The Journal of Chemical Physics

# Basis set converged weak interaction energies from conventional and explicitly correlated coupled-cluster approach

Cite as: J. Chem. Phys. **138**, 154101 (2013); https://doi.org/10.1063/1.4800981
Submitted: 11 February 2013 . Accepted: 25 March 2013 . Published Online: 16 April 2013

Konrad Patkowski

# Basis set converged weak interaction energies from conventional and explicitly correlated coupled-cluster approach

Konrad Patkowski[a)]
*Department of Chemistry and Biochemistry, Auburn University, Auburn, Alabama 36849, USA*

(Received 11 February 2013; accepted 25 March 2013; published online 16 April 2013)

Interaction energies for seven weakly bound dimers involving helium, argon, water, and methane are computed using large correlation-consistent basis sets augmented with bond functions. The estimates of the coupled-cluster singles, doubles, and noniterative triples [CCSD(T)] complete basis set limit are obtained using both the conventional approach and several variants of the explicitly correlated CCSD(T)-F12 method. It is shown that both bond functions and the F12 approach significantly speed up the convergence of the CCSD(T)/aug-cc-pV*X*Z interaction energies with the basis set cardinal number *X*. However, the extent of improvement provided by each technique varies with the character of the interactions—the F12 method works best for polar, electrostatics-bound dimers, while for dispersion-dominated complexes the addition of bond functions is more efficient. The convergence rate afforded by different coupled-cluster variants is fairly consistent across the entire attractive region of the potential curve, while the improvement provided by the F12 correction increases along the repulsive wall. The use of large basis sets and the agreement between conventional and explicitly correlated approaches allow us to assess the importance of different residual approximations present in the popular CCSD(T)-F12 implementations. © *2013 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4800981]

## I. INTRODUCTION

It is well recognized that one-electron Gaussian basis sets, despite their immense popularity, afford only a slow convergence of molecular electron correlation energies. The underlying reason is the inability of one-electron functions to reproduce the correct form of the many-electron wavefunction when the distance between any two electrons approaches zero (the interelectronic cusp[1–4]). In order to enable such a reproduction, the many-electron function has to be equipped with terms that explicitly depend on the distances $r_{ij}$ between electrons $i$ and $j$. When such terms are present, that is, when the basis functions include *explicit correlation*, the convergence of atomic and molecular correlation energies improves dramatically. In fact, all the most accurate energies and properties available for few-electron systems (four electrons or less) have been computed with some variant of the explicitly correlated approach—see Ref. 5 for a recent review. In the last decade, an explicitly correlated treatment of larger systems has become possible thanks to a breakdown of the many-electron integrals using the resolution-of-identity (RI) and density-fitting (DF) techniques[6,7] and a fixed-amplitude *Ansatz* with the explicitly correlated terms determined from cusp conditions rather than optimized.[8] Thus, the frozen-geminal (F12) explicitly correlated approaches have made its way into the mainstream of the electronic structure theory, with several different methods and approximations now available in popular quantum-chemistry codes.[9,10] The rapid development of the F12 methodology still continues—some very recent enhancements include analytic energy gradients[11,12] and multireference methods for systems with significant static correlation.[13–15]

It should be noted that the F12 technique, as currently implemented in the "gold-standard" coupled-cluster approach with singles, doubles, and noniterative triples [CCSD(T)] method, involves several approximations that are not present in conventional (not explicitly correlated) CCSD(T). First, a CCSD(T)-F12 calculation relies inherently on density fitting and thus may suffer from auxiliary basis incompleteness errors. In fact, in the most popular and efficient variants, such as CCSD(T)-F12a, CCSD(T)-F12b,[16,17] and CCSD(T)(F12*)[18] (also termed CCSD(T)-F12c), no less than three, in general different, auxiliary basis sets are required: a complementary auxiliary basis set (CABS)[19] for the RI approximation to many-electron integrals, a DF set for the expansion of the MP2-F12 pair functions, and a DF set for fitting the Fock matrix used in MP2-F12 and in the CABS singles expression.[16] Second, the F12 correlation factor (the form of the dependence of pair functions on the interelectronic distance $r_{12}$) is assumed as $F(r_{12}) = -\frac{1}{\beta}\exp(-\beta r_{12})$ and involves a parameter $\beta$ that is chosen *a priori* rather than optimized. On a related note, several different *Ansätze* are possible in the MP2-F12 part of the calculation. The diagonal fixed-amplitude *Ansatz* 3C(FIX) (Ref. 8), assumed in the subsequent CCSD-F12 calculations, has several desirable features including size consistency and lack of geminal basis set superposition error.[20,21] On the other hand, this *Ansatz* is known to be quite sensitive to the choice of the geminal exponent $\beta$.[22,23] Last but not least, CCSD(T)-F12a, CCSD(T)-F12b, and CCSD(T)(F12*) all involve some approximations in the CCSD-F12 part. The (T) correction is in turn computed from the same formula

a)Author to whom correspondence should be addressed. Electronic mail: patkowsk@auburn.edu

 

as in conventional CCSD(T) and contains no F12 contribution except for the indirect effect of the change in the converged singles and doubles amplitudes. While both the exact CCSD-F12 approach[24] and the explicitly correlated variant of the (T) correction[25] have been developed, the computer implementations of these approaches are not yet generally available and the algorithms are computationally more demanding.

The efficiency of the CCSD(T)-F12a/b/c approaches in reproducing complete-basis-set (CBS) limit CCSD(T) molecular energies and properties has been confirmed in a number of numerical studies.[3,4,17,26] The convergence of CCSD(T)-F12 molecular correlation energies with the orbital basis set is remarkable—explicitly correlated results in Dunning's augmented correlation-consistent double-zeta (aug-cc-pVDZ≡ aDZ) basis set[27,28] are as close to the CBS limit as conventional CCSD(T) in aQZ or even a5Z. The results obtained using different variants of CCSD(T)-F12 are very similar so that the a/b/c approximations are believed to be highly accurate. To compensate for the lack of an F12 triples contribution, a scaling approach has been suggested[17]

$$E^{*}_{(\mathrm{T})} = E_{(\mathrm{T})} \cdot \frac{E^{\mathrm{corr}}_{\mathrm{MP2-F12}}}{E^{\mathrm{corr}}_{\mathrm{MP2}}}, \tag{1}$$

where $E^{\mathrm{corr}}_{\mathrm{X}}$ denotes correlation energy at the X level of theory, i.e., $E_{\mathrm{X}} - E_{\mathrm{SCF}}$.

The objective of the present work is to investigate the basis set convergence of CCSD(T) and CCSD(T)-F12 noncovalent interaction energies. Weak noncovalent interactions are one of the hardest challenges for *ab initio* quantum chemistry. Interaction energies, especially those dominated by long-range dynamical correlation (dispersion), are very hard to converge with respect to both the theory level and the basis set. As far as the former is concerned, the failure of standard density functional theories to even qualitatively describe dispersion is well documented[29,30] and so is the substantial overestimation of dispersion energies by the MP2 method.[31,32] Numerous approaches have been proposed to overcome both of these deficiencies,[33–55] resulting, on the average, in substantially improved weak interaction energies. However, no method below CCSD(T) has been demonstrated to maintain high accuracy across the entire spectrum of weakly interacting systems.[56]

As far as the basis set dependence of interaction energies is concerned, the convergence can usually be accelerated by CBS extrapolations[57,58] and/or the inclusion of bond functions,[59] but the description of the interelectronic cusp remains a serious problem. Thus, the explicitly correlated approach has a potential to significantly speed up basis set convergence of weak interaction energies. In fact, the most accurate interaction potentials for few-electron systems, such as $He_2$, have been obtained using the explicitly correlated approach with a large number of optimized Gaussian-type geminals (GTGs).[60–62] In the last few years, multiple CCSD(T)-F12 studies of weak interaction energies have been performed.[63–74] However, most of these studies were restricted to small and moderate basis sets for which, as expected, the F12 approach provides highly superior convergence. The notable exceptions are Refs. 63, 64, 68, and 74 where approximate CCSD(T)-F12 interaction energies for $Ne_2$, $He_2/Ne_2$, $(CO_2)_2$, and $(H_2O)_2$, respectively, have been obtained in bases up to a5Z (or even a6Z[64]). While for the water dimer the performance of CCSD(T)-F12 is, as we will see below, particularly impressive, the remaining studies indicated that the improvement over conventional CCSD(T) is somewhat diminished in large bases (which was attributed to the lack of an F12 treatment of triples[63,64]). However, none of these studies used bond functions.

In a recent study,[72] we pointed out that, for a specific example of rare gas dimers ($He_2$ through $Kr_2$), a larger convergence speedup is provided by a simple addition of bond functions than by the F12 approach. The addition of bond functions improves the F12 results as well. Using virtually converged CCSD(T)/CBS interaction energies[60,61,75,76] as benchmarks, we demonstrated that, surprisingly, the conventional CCSD(T) treatment in the largest available basis sets is able to recover benchmark near-minimum interaction energies to a better precision than approximate CCSD(T)-F12 in the largest basis sets available for that approach. There are two reasons for this. First, the effects of the a/b/c approximations to CCSD-F12 and of the lack of an explicit F12 dependence of triples (scaled or unscaled) become significant. Second, the limited availability of suitable auxiliary basis sets limits the choice of orbital bases for CCSD(T)-F12 (note that auxiliary bases need to extend to higher angular momenta than their parent orbital set).

One may argue that the rare-gas dimers investigated in Ref. 72 present the worst-case scenario for CCSD(T)-F12. Indeed, these systems are bound by dispersion which is sensitive to the behavior of monomer wavefunctions at large distances from the nuclei, in particular, in the region between the interacting molecules. Moreover, the MP2 approach misses an unusually large portion of the interaction energy in this case[76,77] (note that MP2-F12 has been proven to be highly superior to MP2 even for rare gas dimers[20,72]). Thus, it is worthwhile to extend the study of Ref. 72 to other small dimers with a varying importance of dispersion.

In this work, we compute CCSD(T) and CCSD(T)-F12a/b/c interaction energies for seven dimers: He–$H_2O$, Ar–$H_2O$, He–$CH_4$, Ar–$CH_4$, $CH_4$–$CH_4$, $CH_4$–$H_2O$, and $H_2O$–$H_2O$. All complexes are investigated both at their van der Waals minimum geometries and along the radial potential energy curves passing through the minima. While accurate interaction potentials have been previously constructed for most of these systems,[78–82] the CCSD(T)/CBS limit values are not known to an accuracy similar as for the rare-gas dimers (an exception is the water dimer for which a highly accurate study[74] was published when the current project was nearly completed). Therefore, the range of basis sets employed here (up to a6Z for CCSD(T) and a5Z for CCSD(T)-F12, with bond functions included) allows us to pinpoint the CBS limit precisely enough to compare the convergence of different approaches. Additionally, we study the effects of different approximations to CCSD(T)-F12, auxiliary bases, and values of the F12 correlation factor to quantify the residual uncertainties of the CCSD(T)-F12 approaches.

The remainder of this paper is structured as follows. In Sec. II, we explain the methodology and list the pertinent

computational details. In Sec. III, we present and analyze the numerical results. Finally, in Sec. IV we summarize our findings and present conclusions.

## II. DETAILS OF THE COMPUTATIONAL PROCEDURE

All numerical calculations have been performed using the MOLPRO2010.1 program.[9] Most of the interaction energies were obtained for the near van der Waals minima geometries given in the supplementary material.[83] The geometries for the water and methane dimers were taken from Refs. 31 and 32, respectively, and all the other geometries were optimized at the CCSD(T)/aTZ level with frozen monomers. In addition to the minimum geometries, we computed interaction energies along radial cross sections through the potential energy surfaces passing through the minima. In other words, we varied the distance between the central atoms of the monomers with all the angles fixed relative to the line connecting the monomers' central atoms. All interaction energies were corrected for basis set superposition error using the counterpoise (CP) correction of Boys and Bernardi.[84,85] The core electrons were not correlated.

The calculations employed two families of orbital basis sets: aug-cc-pV*X*Z[27,28] and cc-pV*X*Z-F12.[86] We will use short-hand notations a*X*Z and *X*Z-F12, respectively, for these bases. Unless otherwise specified, the default a*X*Z/MP2FIT auxiliary bases[87,88] (also known as a*X*Z-RI[89]), with $X$ the same as for the orbital set, were used to fit the MP2-F12 pair functions. For the RI approximation to many-electron integrals, as well as for the density fitting of the Fock matrix, we employed the same a6Z/MP2FIT* auxiliary basis[90] for all orbital sets (the asterisk denotes that the $k$ functions, present in a6Z/MP2FIT for heavy atoms, were removed because of MOLPRO limitations). We have shown in Ref. 72 that the MOLPRO default selection for the a*X*Z orbital set, the cc-pV*X*Z/JKFIT[91] auxiliary basis, is sometimes far from adequate for the RI approximation. We have carried out similar auxiliary basis tests for two dimers studied here, He–$CH_4$ (in the aDZ, aTZ, aQZ, and a5Z orbital bases) and $CH_4$–$H_2O$ (in the aTZ orbital basis). The DF and RI sets included a*X*Z/OPTRI[92] (with the same $X$ as the orbital basis as the OPTRI set is tailored to a particular a*X*Z basis by construction) as well as the *X*Z/JKFIT[91] and a*X*Z/MP2FIT[87,88] sets with $X$ larger or equal to the orbital basis cardinal number (for He–$CH_4$ in the a5Z orbital basis we additionally tested the QZ/JKFIT and aQZ/MP2FIT sets). The helium *X*Z/JKFIT sets are not available and the a*X*Z/OPTRI sets with the same $X$ were used instead. The resulting unscaled-triples CCSD(T)-F12b interaction energies at the minima are displayed in Figs. S1–S5 of the supplementary material.[83] These figures show that the errors incurred by choosing a suboptimal RI basis set, quite dramatic for the neon dimer,[72] are smaller for He–$CH_4$ and even less pronounced for $CH_4$–$H_2O$. Surprisingly, the errors incurred by choosing a suboptimal DF set for the Fock matrix become more significant. In particular, the a*X*Z/OPTRI sets, while decent in the RI context, give the largest errors when employed for the density fitting of the Fock matrix. On the other hand, the a*X*Z/MP2FIT family exhibits superior basis set convergence in all three contexts.

When midbond functions were used, the additional orbital basis functions, centered halfway between the central atoms of the monomers, were chosen as hydrogenic functions from the same a*X*Z or *X*Z-F12 set as for all atoms, and the additional auxiliary basis functions for DF and RI were chosen as described above. In other words, the midbond set varies with $X$ in accordance with the atomic basis sets. Other midbond sets could also be used although the choice is limited by the availability of auxiliary bases.[72] The interaction energies have been shown to be fairly insensitive to the exact placement of the midbond center[93] as long as it does not lie too close to either monomer.[94] We will add "M" to the basis set symbol to indicate that midbond functions are included. The addition of midbond increases the basis set size by 8%–22% for the systems studied here, much less than the difference between a*X*Z and a($X$ + 1)Z.

The standard diagonal fixed-amplitude *Ansatz* 3C(FIX)[8] was employed throughout the present work, and the approximate variants investigated included CCSD(T)-F12a, CCSD(T)-F12b,[16,17] and CCSD(T)(F12*)≡CCSD(T)-F12c.[18] We will present interaction energies obtained both with and without the scaling of the triples contribution, Eq. (1). In the former case, the scaling factor determined for the dimer was also used for both monomers to maintain size consistency, as suggested in Ref. 65. Unless stated otherwise, the parameter $\beta$ (GEM_BETA in MOLPRO) in the F12 correlation factor was set to the values recommended in Ref. 86, that is, 0.9, 1.0, and 1.1 $a_0^{-1}$ for DZ-F12, TZ-F12, and QZ-F12, respectively, and 1.1, 1.2, 1.4, and 1.4 $a_0^{-1}$ for aDZ, aTZ, aQZ, and a5Z, respectively.

The finite-basis results, both conventional and explicitly correlated, were extrapolated to the CBS limit using the standard $X^{-3}$ formula.[57,95] Specifically, the interaction energy extrapolated from basis sets a($X$ − 1)Z and a*X*Z (this extrapolation will be denoted as a($X$ − 1, $X$)Z) is a sum of the self-consistent field (SCF) contribution $E_{\rm int}^{\rm SCF}$ computed in the larger a*X*Z set and the correlation contribution $E_{\rm int}^{\rm corr}$(CBS) obtained from the computed correlation energies $E_{\rm int}^{\rm corr}$(a($X$ − 1)Z) and $E_{\rm int}^{\rm corr}$(a*X*Z) as

$$E_{\rm int}^{\rm corr}({\rm CBS}) = E_{\rm int}^{\rm corr}({\rm a}XZ) + \frac{\left(1-\frac{1}{X}\right)^3}{1-\left(1-\frac{1}{X}\right)^3} \cdot \left(E_{\rm int}^{\rm corr}({\rm a}XZ) - E_{\rm int}^{\rm corr}({\rm a}(X-1)Z)\right). \quad (2)$$

While the $X^{-3}$ formula has been extensively tested in conventional CCSD(T) calculations, it might not be the best choice for CCSD(T)-F12 as, under favorable circumstances, the latter can be expected to converge as quickly as $X^{-7}$.[96,97] Unfortunately, the assumptions leading to the $X^{-7}$ convergence (the completeness of the orbital set for each angular momentum $l$ included in the basis, the completeness of all auxiliary sets, and a genuine F12 treatment of triples) are far from satisfied in our practical calculations. Therefore, we chose to stick to the $X^{-3}$ scheme, treating it as an empirically justified way to improve basis set convergence.

TABLE I. CCSD(T)/a*XZ* and CCSD(T)-F12b/a*XZ* interaction energies (in $cm^{-1}$) for the minimum geometry of the He–$H_2O$ complex as functions of the basis set cardinal number *X*. The extrapolated value (rows "ext.") in the *X* column is computed using interaction energies in bases a($X-1$)Z and a*XZ*. The midbond functions are chosen as hydrogenic functions from the same a*XZ* orbital basis set. The benchmark CCSD(T)/CBS interaction energy amounts to $-34.34 \pm 0.07$ $cm^{-1}$.

| | *X* | | | | |
|---|---|---|---|---|---|
| Basis | D | T | Q | 5 | 6 |
| | No midbond | | | | |
| CCSD(T) | −21.17 | −29.72 | −32.55 | −33.56 | −33.98 |
| ext. | | −33.39 | −34.35 | −34.49 | −34.51 |
| CCSD(T)-F12b | −27.75 | −32.24 | −33.35 | −33.91 | |
| ext. | | −34.13 | −34.11 | −34.48 | |
| CCSD(T)-F12b (scaled) | −29.04 | −32.82 | −33.62 | −34.05 | |
| ext. | | −34.41 | −34.16 | −34.49 | |
| | With midbond | | | | |
| CCSD(T) | −27.82 | −33.55 | −34.12 | −34.23 | −34.28 |
| ext. | | −36.11 | −34.56 | −34.35 | −34.34 |
| CCSD(T)-F12b | −30.43 | −33.91 | −34.19 | −34.27 | |
| ext. | | −35.40 | −34.39 | −34.35 | |
| CCSD(T)-F12b (scaled) | −31.94 | −34.52 | −34.47 | −34.41 | |
| ext. | | −35.63 | −34.42 | −34.36 | |

TABLE III. CCSD(T)/a*XZ* and CCSD(T)-F12b/a*XZ* interaction energies (in $cm^{-1}$) for the minimum geometry of the He–$CH_4$ complex as functions of the basis set cardinal number *X*. The extrapolated value (rows "ext.") in the *X* column is computed using interaction energies in bases a($X-1$)Z and a*XZ*. The midbond functions are chosen as hydrogenic functions from the same a*XZ* orbital basis set. The benchmark CCSD(T)/CBS interaction energy amounts to $-29.43 \pm 0.08$ $cm^{-1}$.

| | *X* | | | | |
|---|---|---|---|---|---|
| Basis | D | T | Q | 5 | 6 |
| | No midbond | | | | |
| CCSD(T) | −19.61 | −26.01 | −27.94 | −28.71 | −29.08 |
| ext. | | −28.59 | −29.40 | −29.52 | −29.60 |
| CCSD(T)-F12b | −24.59 | −27.60 | −28.53 | −29.01 | |
| ext. | | −28.84 | −29.20 | −29.52 | |
| CCSD(T)-F12b (scaled) | −25.53 | −27.99 | −28.70 | −29.11 | |
| ext. | | −28.99 | −29.22 | −29.52 | |
| | With midbond | | | | |
| CCSD(T) | −24.20 | −28.54 | −29.15 | −29.29 | −29.35 |
| ext. | | −30.71 | −29.56 | −29.42 | −29.43 |
| CCSD(T)-F12b | −27.20 | −28.87 | −29.24 | −29.33 | |
| ext. | | −29.61 | −29.50 | −29.43 | |
| CCSD(T)-F12b (scaled) | −28.24 | −29.28 | −29.42 | −29.43 | |
| ext. | | −29.75 | −29.51 | −29.43 | |

## III. NUMERICAL RESULTS AND DISCUSSION

The near-minimum CCSD(T) and CCSD(T)-F12b interaction energies for the seven dimers considered in this work, computed in the a*XZ* basis sets with and without midbond, are presented in Tables I–VII. The analogous CCSD(T)-F12a/c results are given in Tables SI–SVII in the supplementary material[83] which also contains graphical representations of the CCSD(T) and CCSD(T)-F12b data (Figs. S6–S12). Tables I–VII display conventional CCSD(T) results for $X = $ D–6 and CCSD(T)-F12b results, with and without the scaling of triples [Eq. (1)], for $X = $ D–5.

The relative performance of CCSD(T) and CCSD(T)-F12 for the four dimers containing a rare gas atom, Tables I–IV, is not much different for the one observed for rare gas

TABLE II. CCSD(T)/a*XZ* and CCSD(T)-F12b/a*XZ* interaction energies (in $cm^{-1}$) for the minimum geometry of the Ar–$H_2O$ complex as functions of the basis set cardinal number *X*. The extrapolated value (rows "ext.") in the *X* column is computed using interaction energies in bases a($X-1$)Z and a*XZ*. The midbond functions are chosen as hydrogenic functions from the same a*XZ* orbital basis set. The benchmark CCSD(T)/CBS interaction energy amounts to $-139.52 \pm 0.12$ $cm^{-1}$.

| | *X* | | | | |
|---|---|---|---|---|---|
| Basis | D | T | Q | 5 | 6 |
| | No midbond | | | | |
| CCSD(T) | −79.53 | −121.24 | −134.51 | −137.35 | −138.39 |
| ext. | | −137.15 | −140.57 | −139.68 | −139.62 |
| CCSD(T)-F12b | −110.74 | −132.04 | −136.99 | −138.34 | |
| ext. | | −140.58 | −140.18 | −139.68 | |
| CCSD(T)-F12b (scaled) | −119.66 | −135.95 | −138.88 | −139.49 | |
| ext. | | −142.39 | −140.60 | −140.05 | |
| | With midbond | | | | |
| CCSD(T) | −99.68 | −134.80 | −138.20 | −139.23 | −139.40 |
| ext. | | −145.47 | −139.55 | −139.91 | −139.52 |
| CCSD(T)-F12b | −121.71 | −137.93 | −138.64 | −139.11 | |
| ext. | | −144.14 | −139.03 | −139.60 | |
| CCSD(T)-F12b (scaled) | −131.99 | −142.01 | −140.53 | −140.24 | |
| ext. | | −145.60 | −139.32 | −139.93 | |

TABLE IV. CCSD(T)/a*XZ* and CCSD(T)-F12b/a*XZ* interaction energies (in $cm^{-1}$) for the minimum geometry of the Ar–$CH_4$ complex as functions of the basis set cardinal number *X*. The extrapolated value (rows "ext.") in the *X* column is computed using interaction energies in bases a($X-1$)Z and a*XZ*. The midbond functions are chosen as hydrogenic functions from the same a*XZ* orbital basis set. The benchmark CCSD(T)/CBS interaction energy amounts to $-141.16 \pm 0.41$ $cm^{-1}$.

| | *X* | | | | |
|---|---|---|---|---|---|
| Basis | D | T | Q | 5 | 6 |
| | No midbond | | | | |
| CCSD(T) | −89.40 | −121.97 | −133.81 | −137.92 | −139.58 |
| ext. | | −136.38 | −142.83 | −142.17 | −141.85 |
| CCSD(T)-F12b | −110.05 | −130.13 | −136.75 | −139.45 | |
| ext. | | −138.62 | −141.61 | −142.28 | |
| CCSD(T)-F12b (scaled) | −118.02 | −133.58 | −138.45 | −140.53 | |
| ext. | | −140.17 | −142.02 | −142.71 | |
| | With midbond | | | | |
| CCSD(T) | −108.32 | −135.75 | −139.51 | −140.46 | −140.76 |
| ext. | | −147.99 | −142.31 | −141.41 | −141.16 |
| CCSD(T)-F12b | −119.05 | −137.73 | −140.13 | −140.70 | |
| ext. | | −145.71 | −141.87 | −141.29 | |
| CCSD(T)-F12b (scaled) | −127.99 | −141.39 | −141.83 | −141.76 | |
| ext. | | −147.15 | −142.15 | −141.67 | |

TABLE V. CCSD(T)/a*XZ* and CCSD(T)-F12b/a*XZ* interaction energies (in $cm^{-1}$) for the minimum geometry of the $H_2O$–$H_2O$ complex as functions of the basis set cardinal number *X*. The extrapolated value (rows "ext.") in the *X* column is computed using interaction energies in bases a($X-1$)Z and a*XZ*. The midbond functions are chosen as hydrogenic functions from the same a*XZ* orbital basis set. The benchmark CCSD(T)/CBS interaction energy amounts to $-1745.0 \pm 1.2$ $cm^{-1}$.

| | *X* | | | | |
|---|---|---|---|---|---|
| Basis | D | T | Q | 5 | 6 |
| | | | No midbond | | |
| CCSD(T) | −1515.0 | −1653.9 | −1716.5 | −1730.7 | −1736.7 |
| ext. | | −1715.3 | −1752.9 | −1745.8 | −1744.8 |
| CCSD(T)-F12b | −1678.6 | −1728.6 | −1741.4 | −1743.3 | |
| ext. | | −1750.2 | −1751.3 | −1745.8 | |
| CCSD(T)-F12b (scaled) | −1698.4 | −1737.2 | −1745.5 | −1745.5 | |
| ext. | | −1754.2 | −1752.0 | −1745.9 | |
| | | | With midbond | | |
| CCSD(T) | −1580.2 | −1687.5 | −1723.7 | −1734.0 | −1738.5 |
| ext. | | −1730.3 | −1749.3 | −1744.5 | −1744.8 |
| CCSD(T)-F12b | −1701.6 | −1736.3 | −1743.1 | −1743.8 | |
| ext. | | −1751.1 | −1749.2 | −1745.0 | |
| CCSD(T)-F12b (scaled) | −1723.5 | −1745.0 | −1747.1 | −1745.9 | |
| ext. | | −1754.3 | −1749.7 | −1745.1 | |

TABLE VII. CCSD(T)/a*XZ* and CCSD(T)-F12b/a*XZ* interaction energies (in $cm^{-1}$) for the minimum geometry of the $CH_4$–$CH_4$ complex as functions of the basis set cardinal number *X*. The extrapolated value (rows "ext.") in the *X* column is computed using interaction energies in bases a($X-1$)Z and a*XZ*. The midbond functions are chosen as hydrogenic functions from the same a*XZ* orbital basis set. The benchmark CCSD(T)/CBS interaction energy amounts to $-187.30 \pm 0.30$ $cm^{-1}$.

| | *X* | | | | |
|---|---|---|---|---|---|
| Basis | D | T | Q | 5 | 6 |
| | | | No midbond | | |
| CCSD(T) | −149.60 | −176.80 | −183.64 | −185.87 | −186.60 |
| ext. | | −189.33 | −188.53 | −188.04 | −187.60 |
| CCSD(T)-F12b | −164.90 | −182.07 | −185.53 | −186.68 | |
| ext. | | −189.35 | −187.98 | −187.84 | |
| CCSD(T)-F12b (scaled) | −174.61 | −185.71 | −187.13 | −187.52 | |
| ext. | | −190.43 | −188.10 | −187.88 | |
| | | | With midbond | | |
| CCSD(T) | −163.71 | −183.65 | −186.11 | −186.76 | −186.99 |
| ext. | | −193.04 | −187.81 | −187.33 | −187.29 |
| CCSD(T)-F12b | −171.80 | −185.25 | −186.66 | −187.00 | |
| ext. | | −190.92 | −187.61 | −187.31 | |
| CCSD(T)-F12b (scaled) | −181.95 | −188.94 | −188.25 | −187.83 | |
| ext. | | −191.88 | −187.67 | −187.35 | |

dimers.[72] The results, both conventional and explicitly correlated, obtained in midbond-containing bases are much better converged than the results computed in the corresponding midbondless bases. The convergence improvement resulting from the addition of midbond functions compares favorably to the improvement due to adopting the unscaled-triples F12b approach. The scaling of the triples correction does improve the F12b results further, and for the Ar–$H_2O$ and Ar–$CH_4$ dimers the scaled-triples CCSD(T)-F12b/a*XZ* results are just as good or slightly better than the conventional CCSD(T)/a*XZM* values (for dimers containing helium the latter results are better converged except for $X = D$). The benefits of midbond functions and of the F12b treatment can be combined, leading to the most accurate CBS-limit estimates for a given cardinal number *X*. When midbond functions are present, the scaling of triples improves results for $X = D,T$ but mostly overshoots for $X = Q,5$, similar to what has been observed for noble gas dimers.[72]

TABLE VI. CCSD(T)/a*XZ* and CCSD(T)-F12b/a*XZ* interaction energies (in $cm^{-1}$) for the minimum geometry of the $CH_4$–$H_2O$ complex as functions of the basis set cardinal number *X*. The extrapolated value (rows "ext.") in the *X* column is computed using interaction energies in bases a($X-1$)Z and a*XZ*. The midbond functions are chosen as hydrogenic functions from the same a*XZ* orbital basis set. The benchmark CCSD(T)/CBS interaction energy amounts to $-354.8 \pm 0.5$ $cm^{-1}$.

| | *X* | | | | |
|---|---|---|---|---|---|
| Basis | D | T | Q | 5 | 6 |
| | | | No midbond | | |
| CCSD(T) | −230.6 | −333.3 | −347.8 | −351.3 | −352.8 |
| ext. | | −365.9 | −358.4 | −355.1 | −354.9 |
| CCSD(T)-F12b | −308.7 | −349.0 | −353.3 | −354.1 | |
| ext. | | −364.7 | −356.7 | −355.0 | |
| CCSD(T)-F12b (scaled) | −322.9 | −354.9 | −356.0 | −355.5 | |
| ext. | | −367.1 | −357.0 | −355.0 | |
| | | | With midbond | | |
| CCSD(T) | −287.0 | −339.9 | −349.7 | −352.3 | −353.3 |
| ext. | | −362.6 | −356.8 | −354.9 | −354.8 |
| CCSD(T)-F12b | −327.1 | −351.3 | −353.9 | −354.3 | |
| ext. | | −361.8 | −356.0 | −354.8 | |
| CCSD(T)-F12b (scaled) | −342.4 | −357.2 | −356.6 | −355.7 | |
| ext. | | −363.8 | −356.3 | −354.9 | |

The hierarchy of different CCSD(T) approaches is completely different for the $H_2O$–$H_2O$ and $CH_4$–$H_2O$ dimers (Tables V and VI, respectively). For these systems, the improvement provided by the F12b method is remarkable: the CCSD(T)-F12b/aTZ water-water interaction energy is virtually as accurate as the conventional CCSD(T)/a5Z result, or even the CCSD(T)/a6Z one if the triples correction is scaled. The addition of midbond functions also improves the convergence for all variants but the improvement is quite minimal. On the other hand, the CCSD(T) and CCSD(T)-F12b basis set convergence for the methane dimer, Table VII, is more similar to the complexes containing a noble gas atom. Again, midbond functions improve the interaction energy more than the unscaled-triples F12b approach (the scaled-triples F12b treatment brings a*XZ* results very close to the CBS limit but overcorrects the a*XZM* energies). Nevertheless, the improvement brought about by explicit correlation is also substantial, with the largest-basis CCSD(T)/a6ZM and CCSD(T)-F12b/a5ZM results providing virtually the same accuracy.

Tables I–VII and Figs. S6–S12 also contain the results extrapolated to the CBS limit using the standard $X^{-3}$

algorithm. It is evident that the extrapolation improves the convergence of both CCSD(T) and CCSD(T)-F12b interaction energies. The latter success of the $X^{-3}$ extrapolation comes despite the fact that the F12 approach formally exhibits different convergence of correlation energy (up to $X^{-7}$ (Refs. 96 and 97)). The best illustration of this success is the excellent agreement between the CCSD(T) and unscaled-triples CCSD(T)-F12b results extrapolated from the a*X*ZM basis sets, leading to remarkably consistent conventional and explicitly correlated estimates of the CBS limit—see below. The midbondless a*X*Z bases lead to somewhat less accurate and more erratic extrapolations. It is also worth noting that the unscaled- and scaled-triples CCSD(T)-F12b interaction energies become nearly identical upon the CBS extrapolation despite the fact that the corresponding computed results approach the limit from two different sides.

As the CCSD(T) and CCSD(T)-F12 results obtained using bond functions are clearly better converged than those without bond functions, we decided to base our benchmark CCSD(T)/CBS values on two results $E_1$, $E_2$: the conventional CCSD(T)/a(5,6)ZM value and the CCSD(T)-F12b/a(Q,5)ZM result with unscaled triples. The uncertainty $\sigma_i$ of each $E_i$, $i = 1, 2$, was taken as the absolute difference between the extrapolated value and the one calculated using the larger of the two bases (CCSD(T)/a6ZM or CCSD(T)-F12b/a5ZM). Next, the two estimates were combined and the benchmark $(E - \sigma, E + \sigma)$ interval for the CCSD(T)/CBS interaction energy was taken as the intersection of the intervals $(E_i - \sigma_i, E_i + \sigma_i)$ for $i = 1, 2$. The resulting CBS-limit estimates $E$ and their uncertainties $\sigma$ are given in captions to Tables I–VII. The CCSD(T) and unscaled-triples CCSD(T)-F12b estimates are consistent for all systems. The same could not be said about the scaled-triples CCSD(T)-F12b estimates (for Ar–$CH_4$, the interval predicted by scaled-triples CCSD(T)-F12b does not intersect with the CCSD(T) interval)—as we have stated above, scaling overshoots for bases containing midbond. It is worth noting that for the dispersion-bound dimers (He–$H_2O$, Ar–$H_2O$, He–$CH_4$, Ar–$CH_4$, and $CH_4$–$CH_4$) the CCSD(T)/a(5,6)ZM treatment provides a slightly more precise CBS limit (a narrower interval) than the CCSD(T)-F12b/a(Q,5)ZM one. In other words, while the F12b method obviously improves convergence, a larger improvement can be achieved by utilizing bigger orbital bases (and possibly bigger midbond bases, as shown in Ref. 72) available for conventional CCSD(T). For the more strongly bound $H_2O$–$H_2O$ and $CH_4$–$H_2O$ complexes, the CCSD(T)-F12b/a(Q,5)ZM estimate is several times more precise than the CCSD(T)/a(5,6)ZM one. For these systems, conventional CCSD(T) cannot match the CCSD(T)-F12b/a(Q,5)ZM accuracy in any feasible basis set.

The benchmark interaction energies obtained as described above allow us to compare the accuracy of different methods and basis sets in reproducing the CCSD(T)/CBS limits for the seven van der Waals minima. To this end, we considered the mean unsigned errors (MUE, in $cm^{-1}$) and mean unsigned relative errors (MURE, in percent). The values of MUE and MURE obtained for different approaches (conventional or F12), different F12 variants, and different basis sets (a*X*Z with and without midbond and/or CBS extrapolation) are listed in Tables VIII and IX, respectively. Analogous statistical measures for the *X*Z-F12 and *X*Z-F12M basis set families are given in Table SVIII in the supplementary material[83] and show that these basis sets are, on the average, substantially inferior to the a*X*Z and a*X*ZM sets for the same *X*. As the *X*Z-F12 set extends only to $X = Q$, it cannot be used to converge CCSD(T)-F12 interaction energies to an extent comparable to the best results from the CCSD(T)-F12/a*X*Z and/or CCSD(T)-F12/a*X*ZM treatments.

While all six variants of the F12 method (a/b/c approximations to CCSD-F12, scaled or unscaled triples) provide improvement over conventional CCSD(T), the differences between variants are unexpectedly large. For example, the scaled-triples CCSD(T)-F12a variant is by far the best one at the aDZM level but the *least accurate* one

TABLE VIII. Mean unsigned errors (MUE, in $cm^{-1}$) of different CCSD(T)/CCSD(T)-F12 variants and basis sets. The errors are averaged over the van der Waals minimum geometries for the seven dimers considered in this work. The benchmark CCSD(T)/CBS interaction energies have been obtained as described in Sec. III.

| | | Computed results, $X =$ | | | | | Extrapolated results | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Method | Basis | D | T | Q | 5 | 6 | (D,T) | (T,Q) | (Q,5) | (5,6) |
| CCSD(T) | a*X*Z | 75.24 | 24.09 | 7.83 | 3.73 | 2.06 | 7.40 | 2.21 | 0.47 | 0.25 |
| CCSD(T) | a*X*ZM | 48.67 | 12.56 | 4.45 | 2.20 | 1.28 | 6.30 | 1.19 | 0.18 | 0.04 |
| CCSD(T)-F12a | a*X*Z | 22.11 | 3.77 | 1.23 | 0.59 | | 4.26 | 2.04 | 0.73 | |
| CCSD(T)-F12a | a*X*ZM | 12.04 | 0.46 | 0.68 | 0.42 | | 5.11 | 1.34 | 0.24 | |
| CCSD(T)-F12a (scaled) | a*X*Z | 13.14 | 2.08 | 1.91 | 1.16 | | 5.50 | 2.31 | 0.88 | |
| CCSD(T)-F12a (scaled) | a*X*ZM | 2.53 | 3.89 | 2.40 | 1.37 | | 6.45 | 1.47 | 0.35 | |
| CCSD(T)-F12b | a*X*Z | 29.46 | 7.14 | 2.24 | 0.97 | | 3.08 | 1.49 | 0.43 | |
| CCSD(T)-F12b | a*X*ZM | 18.97 | 2.90 | 0.81 | 0.43 | | 3.87 | 1.01 | 0.03 | |
| CCSD(T)-F12b (scaled) | a*X*Z | 20.49 | 3.37 | 0.95 | 0.37 | | 4.13 | 1.76 | 0.57 | |
| CCSD(T)-F12b (scaled) | a*X*ZM | 9.07 | 1.01 | 0.94 | 0.52 | | 5.21 | 1.12 | 0.16 | |
| CCSD(T)-F12c | a*X*Z | 29.00 | 7.35 | 2.75 | 1.33 | | 2.89 | 0.77 | 0.24 | |
| CCSD(T)-F12c | a*X*ZM | 18.76 | 3.33 | 1.33 | 0.78 | | 3.17 | 0.46 | 0.14 | |
| CCSD(T)-F12c (scaled) | a*X*Z | 20.27 | 3.60 | 1.06 | 0.38 | | 3.95 | 1.05 | 0.38 | |
| CCSD(T)-F12c (scaled) | a*X*ZM | 9.14 | 0.91 | 0.43 | 0.18 | | 4.56 | 0.58 | 0.18 | |

TABLE IX. Mean unsigned relative errors (MURE, in percent) of different CCSD(T)/CCSD(T)-F12 variants and basis sets. The errors are averaged over the van der Waals minimum geometries for the seven dimers considered in this work. The benchmark CCSD(T)/CBS interaction energies have been obtained as described in Sec. III.

| | | Computed results, $X =$ | | | | | Extrapolated results | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Method | Basis | D | T | Q | 5 | 6 | (D,T) | (T,Q) | (Q,5) | (5,6) |
| CCSD(T) | a*X*Z | 31.39 | 9.81 | 3.52 | 1.59 | 0.80 | 2.37 | 0.59 | 0.30 | 0.26 |
| CCSD(T) | a*X*ZM | 18.53 | 3.14 | 1.00 | 0.45 | 0.26 | 3.53 | 0.43 | 0.08 | 0.00 |
| CCSD(T)-F12a | a*X*Z | 11.85 | 2.96 | 1.16 | 0.44 | | 1.39 | 0.63 | 0.37 | |
| CCSD(T)-F12a | a*X*ZM | 6.08 | 0.35 | 0.20 | 0.12 | | 2.47 | 0.33 | 0.05 | |
| CCSD(T)-F12a (scaled) | a*X*Z | 7.66 | 1.73 | 0.94 | 0.46 | | 1.72 | 0.71 | 0.46 | |
| CCSD(T)-F12a (scaled) | a*X*ZM | 1.62 | 1.77 | 0.97 | 0.56 | | 3.11 | 0.38 | 0.12 | |
| CCSD(T)-F12b | a*X*Z | 15.29 | 4.41 | 1.78 | 0.77 | | 1.33 | 0.50 | 0.28 | |
| CCSD(T)-F12b | a*X*ZM | 9.43 | 1.33 | 0.45 | 0.22 | | 2.07 | 0.29 | 0.02 | |
| CCSD(T)-F12b (scaled) | a*X*Z | 11.10 | 2.65 | 1.06 | 0.39 | | 1.44 | 0.58 | 0.38 | |
| CCSD(T)-F12b (scaled) | a*X*ZM | 4.76 | 0.64 | 0.38 | 0.25 | | 2.71 | 0.32 | 0.11 | |
| CCSD(T)-F12c | a*X*Z | 16.15 | 4.91 | 2.07 | 0.95 | | 1.30 | 0.40 | 0.22 | |
| CCSD(T)-F12c | a*X*ZM | 10.21 | 1.90 | 0.74 | 0.39 | | 1.60 | 0.21 | 0.03 | |
| CCSD(T)-F12c (scaled) | a*X*Z | 12.07 | 3.17 | 1.26 | 0.49 | | 1.36 | 0.48 | 0.31 | |
| CCSD(T)-F12c (scaled) | a*X*ZM | 5.67 | 0.47 | 0.15 | 0.10 | | 2.26 | 0.25 | 0.07 | |

in the aQZM and a5ZM bases or if the CBS extrapolation is performed. On the other hand, the unscaled-triples CCSD(T)-F12a/a*X*ZM approach, while not particularly well converged at the $X = \mathrm{D}$ level, is one of the best variants from $X = \mathrm{T}$ on. For the largest, a5ZM basis set, the scaled-triples CCSD(T)-F12c method provides the best accuracy and unscaled-triples CCSD(T)-F12a comes a close second (a virtually the same ranking was observed in our earlier study on rare gas dimers[72]). The presence of midbond functions reduces the MURE of both conventional CCSD(T) and all variants of CCSD(T)-F12 except for the aforementioned scaled-triples CCSD(T)-F12a approach. When the CBS extrapolation is performed, the results are, on the average, closer to the CBS limit and the differences between various F12 variants and between CCSD(T)-F12 and conventional CCSD(T) are substantially diminished. However, the CCSD(T)/CBS estimates obtained from the a*X*ZM sequence are more accurate than those from the a*X*Z one for all (T,Q), (Q,5), and (5,6) extrapolations.

The two lowest values of MUE and MURE in Tables VIII and IX are provided by the CCSD(T)/a(5,6)ZM and unscaled-triples CCSD(T)-F12b/a(Q,5)ZM extrapolations. This result is nothing more than a consequence of how the benchmark interaction energies were established using precisely these two approaches. The fact that the two extrapolations are so close to each other (the MUE of one method with respect to the other one is just 0.07 $\mathrm{cm}^{-1}$, the sum of the two MUE values from Table VIII) is, however, highly encouraging. Obviously, this agreement might be partially accidental and other F12b/c variants might give a(Q,5)ZM extrapolated results that are at least as close to the true CBS limit. Nevertheless, the agreement of the benchmark with the CCSD(T)/a(5,6)ZM, CCSD(T)-F12b/a(Q,5)ZM, and CCSD(T)-F12c/a(Q,5)ZM interaction energies (with or without the scaling of triples) to better than 0.2 $\mathrm{cm}^{-1}$ (MUE) and about 0.1% (MURE) is a strong indicator that the benchmarks adopted by us are also converged to at least 0.2 $\mathrm{cm}^{-1}$ and 0.1%. Thus, all our remarks regarding the relative performance of different variants should be statistically significant.

The properties of the conventional and explicitly correlated approaches can be understood by looking at the CCSD and triples interaction energy contributions separately. Accordingly, Table X presents the MURE for the CCSD/CCSD-F12 interaction energies, the (T) triples interaction energy contributions with and without scaling and, for completeness, the MP2/MP2-F12 results. Just like in Tables VIII and IX, the statistical averaging has been performed over the van der Waals minima for the seven dimers considered here. The benchmark values have been obtained in the same way as for the entire CCSD(T) interaction energy, that is, using the a(5,6)ZM extrapolation from the conventional results and the a(Q,5)ZM one from the F12b results. The confidence intervals provided by these two extrapolations do intersect for all methods and systems. Interestingly, at the MP2 and CCSD levels, the F12 approach provides a more precise CBS estimate, with the F12 confidence interval contained entirely in the conventional interval in all cases but two. At the (T) level, the opposite is true: the conventional confidence intervals are always contained within the unscaled F12a≡F12b ones. If the triples correction were scaled, the resulting confidence intervals would not intersect with the conventional (T) intervals for some systems.

Similar to the case of the rare gas dimers,[72] the basis set convergence of the MP2-F12 approach is highly superior to that of conventional MP2. Extrapolation improves results significantly and alleviates a large part of the MP2 deficiency as compared to MP2-F12. At the CCSD level, the F12 interaction energies are more accurate than the conventional results, but the three variants lead to dramatically different performance. For double-zeta basis sets, CCSD-F12a is the most accurate variant by far, but it is the least accurate variant by far for the aTZM, aQZM, and a5ZM bases (where CCSD-F12b performs best and CCSD-F12c is also very good). Clearly, the CCSD-F12a approach, involving more drastic

TABLE X. Mean unsigned relative errors (MURE, in percent) of the conventional and explicitly correlated MP2, CCSD, and (T) contributions to the interaction energy. The errors are averaged over the van der Waals minimum geometries for the seven dimers considered in this work. The benchmark CBS values have been obtained as described in Sec. III. The F12a and F12b triples corrections are identical.

| | | Computed results, $X =$ | | | | | Extrapolated results | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Method | Basis | D | T | Q | 5 | 6 | (D,T) | (T,Q) | (Q,5) | (5,6) |
| MP2 | a*X*Z | 32.83 | 11.52 | 4.70 | 2.36 | 1.32 | 3.27 | 0.37 | 0.15 | 0.18 |
| MP2 | a*X*ZM | 20.48 | 4.86 | 1.94 | 0.99 | 0.60 | 2.21 | 0.18 | 0.13 | 0.08 |
| MP2-F12 | a*X*Z | 10.04 | 3.05 | 1.49 | 0.64 | | 1.16 | 0.44 | 0.23 | |
| MP2-F12 | a*X*ZM | 5.65 | 0.89 | 0.39 | 0.19 | | 1.31 | 0.12 | 0.01 | |
| CCSD | a*X*Z | 31.65 | 9.88 | 3.47 | 1.55 | 0.76 | 2.51 | 0.62 | 0.29 | 0.28 |
| CCSD | a*X*ZM | 17.44 | 2.59 | 0.76 | 0.34 | 0.20 | 3.85 | 0.41 | 0.05 | 0.01 |
| CCSD-F12a | a*X*Z | 6.71 | 1.51 | 0.89 | 0.44 | | 1.57 | 0.69 | 0.40 | |
| CCSD-F12a | a*X*ZM | 1.57 | 1.47 | 0.77 | 0.40 | | 2.67 | 0.31 | 0.07 | |
| CCSD-F12b | a*X*Z | 11.09 | 2.89 | 1.18 | 0.45 | | 1.26 | 0.52 | 0.29 | |
| CCSD-F12b | a*X*ZM | 5.65 | 0.37 | 0.13 | 0.06 | | 2.16 | 0.26 | 0.02 | |
| CCSD-F12c | a*X*Z | 11.81 | 3.31 | 1.42 | 0.61 | | 1.24 | 0.43 | 0.26 | |
| CCSD-F12c | a*X*ZM | 6.24 | 0.64 | 0.24 | 0.16 | | 1.76 | 0.20 | 0.08 | |
| (T) | a*X*Z | 32.12 | 9.91 | 3.78 | 1.75 | 0.93 | 1.83 | 0.70 | 0.37 | 0.21 |
| (T) | a*X*ZM | 24.00 | 5.29 | 1.86 | 0.85 | 0.49 | 2.58 | 0.64 | 0.21 | 0.00 |
| (T)-F12ab | a*X*Z | 34.33 | 11.03 | 4.36 | 2.10 | | 2.04 | 0.52 | 0.28 | |
| (T)-F12ab | a*X*ZM | 26.01 | 6.30 | 2.37 | 1.12 | | 1.99 | 0.51 | 0.18 | |
| (T)-F12ab (scaled) | a*X*Z | 12.38 | 2.59 | 1.19 | 0.62 | | 2.82 | 1.24 | 0.69 | |
| (T)-F12ab (scaled) | a*X*ZM | 2.81 | 3.31 | 1.97 | 1.27 | | 5.37 | 0.99 | 0.54 | |
| (T)-F12c | a*X*Z | 35.94 | 11.99 | 4.85 | 2.37 | | 2.43 | 0.43 | 0.23 | |
| (T)-F12c | a*X*ZM | 27.84 | 7.32 | 2.85 | 1.39 | | 1.32 | 0.44 | 0.15 | |
| (T)-F12c (scaled) | a*X*Z | 14.54 | 3.08 | 1.12 | 0.53 | | 2.50 | 1.13 | 0.65 | |
| (T)-F12c (scaled) | a*X*ZM | 4.05 | 2.19 | 1.46 | 0.99 | | 4.80 | 0.92 | 0.51 | |

approximations than CCSD-F12b and CCSD-F12c, benefits from some cancellation of errors in small basis sets. This cancellation seems to be quite systematic across different systems, as indicated by the surprisingly good performance of the CCSD(T)-F12a/aDZ approach found in earlier studies.[17,73]

The lack of a proper explicitly correlated contribution to triples[25] is the obvious reason why the relative performance of various variants of the (T) correction is entirely different than for MP2 and CCSD. The F12a/b/c corrections to doubles amplitudes are not designed to improve the basis set saturation of the perturbative triples expression, so they do not improve it—in fact, as the results in Table X indicate, they worsen it slightly but systematically. As demonstrated previously,[65] the scaling in Eq. (1) results in a significant improvement of the (T) contribution to the interaction energy, at least for the a*X*Z basis sets. For the a*X*ZM bases, scaling does make the results more accurate but the large-*X* improvement is limited as the procedure overshoots. This leads to a surprising situation where the scaled large-basis triples corrections are more accurate *without* midbond. However, the accuracy gain afforded by midbond functions at the CCSD-F12 level more than makes up for the worsening of triples. Finally, it is worth noting that the scaling of triples does not work well for the extrapolated results—CBS extrapolation improves unscaled-triples corrections dramatically, while scaled-triples corrections are hardly improved or even worsened.

In an earlier analysis,[72] we came to a conclusion that the residual deviations of large-basis CCSD(T)-F12 interaction energies from the CBS limit are mainly due to the F12a/b/c approximations to CCSD and the lack of a rigorous explicitly correlated treatment of triples. The results of Table X generally support this conclusion. However, if one stays away from the least accurate CCSD-F12a approximation, the inaccuracy of triples, with or without scaling, dominates over the inaccuracy of CCSD-F12. This behavior is seen from the MURE values in Table X and, even more clearly, from the corresponding absolute errors (MUE) in Table SIX in the supplementary material.[83] The mean errors at the CCSD-F12b and CCSD-F12c levels are actually quite similar to those of MP2-F12 and to each other, indicating that the accuracy of the F12b and F12c approximations is high and systematic. While the triples contribution, with or without scaling, is the least converged part of the CCSD(T)-F12 interaction energy, the relative performance of different variants is strongly dependent on the degree to which the errors between different contributions cancel each other. This cancellation is best exemplified by the formally most approximate unscaled-triples CCSD(T)-F12a approach, cf. Tables VIII–IX.

The benchmark interaction energies obtained here enable us to revisit the choice of the exponent $\alpha = 3.0$ in the $X^{-\alpha}$ CBS extrapolations, both for the total CCSD(T)-F12 interaction energy and for its CCSD-F12 and (T) contributions (which may exhibit different convergence[23,58]). To this end, we found the exponents $\alpha_{\rm opt}$ that, when employed in the extrapolations of the correlation parts of the CCSD(T)-F12 and CCSD-F12 interaction energies, and of the (T) interaction energy contribution, minimize the MURE for the seven van der Waals minima. The optimal exponents and the resulting

MURE values are shown in Table SX in the supplementary material.[83] In a few cases $\alpha_{opt}$ turned out to be infinite, that is, the nonextrapolated results are more accurate than any extrapolation. The origin of the unusually large and/or infinite optimal exponents is typically the nonextrapolated results crossing onto the other side of the CBS limit value. This happens mostly for the (D,T) extrapolations as the computed interaction energies become more monotonic for larger $X$.

Table SX shows that the unscaled (T) corrections in CCSD(T)-F12 interaction energies exhibit near-$X^{-3}$ convergence in all cases (the extrapolated scaled (T) corrections are substantially worse and the associated optimal exponents are quite erratic). On the other hand, $\alpha_{opt}$ for CCSD-F12 and CCSD(T)-F12 are quite similar, especially for the (Q,5) extrapolations. These exponents increase with $X$ for the a$X$Z bases and decrease with $X$ for the a$X$ZM ones. In particular, at the (Q,5) level, the optimal extrapolation exponents for the a$X$ZM bases are close to 3.0 (although the MURE are so small that the uncertainty of the benchmark itself might be a crucial factor in this case), while for the a$X$Z sets they are somewhat larger, around 3.8. The overall conclusion of Table SX is that, once the initial nonmonotonic convergence has passed, the optimal extrapolation exponents are all fairly close to 3.0. Thus, while one could put more effort into a careful, basis set-specific and midbond-specific optimization of these exponents, the expected improvement is less than for molecular correlation energies (for which $\alpha_{opt}$ exhibit more diversity[23]) and we chose to stick to the well established $X^{-3}$ scheme.

### A. Potential energy curves

The largest calculations that we have carried out for the van der Waals minima (Tables I–VII) would be too computationally demanding to extend to the entire potential energy curves. For the sake of consistency, we restricted the potential curve calculations for all systems to the a$X$Z and a$X$ZM basis sets with $X$ = D,T,Q. Therefore, we need to determine how to select benchmark interaction energies for points other than the minima. Tables VIII and IX show that, if one restricts the values of $X$ to D, T, Q, (D,T), and (T,Q), the lowest MUE and MURE for the seven minima amount to 0.43 cm$^{-1}$ and 0.15%, respectively, and they are provided by the same CCSD(T)-F12c/aQZM approach with scaled triples. Thus, the scaled-triples CCSD(T)-F12c/aQZM interaction energies will be used as benchmarks for the entire potential energy curves.

The performance of the CCSD(T) and CCSD(T)-F12 variants for the He–$CH_4$, $CH_4$–$CH_4$, and $H_2O$–$H_2O$ dimer curves is presented in Figs. 1–3. The analogous graphs for the remaining four complexes are given in the supplementary material.[83] The purpose of these figures is to illustrate the relative performance of different approaches as a function of the intermolecular distance $R$. Such an illustration for the entire range of $R$ is not provided by either absolute errors (which decay quickly with $R$) or relative errors (which are greatly enhanced close to the point where the interaction energy crosses zero). It is more meaningful to arbitrarily select one of the approximate methods as a reference and express all interaction energy errors (with respect to the scaled-triples CCSD(T)-F12c/aQZM benchmark) relative to the errors of the reference approach. In our case, we selected the conventional CCSD(T)/aQZ treatment (which lies in the middle of the pack among the methods tested) as the reference and scaled the errors in Figs. 1–3 and S13–S16 by the reference errors at the same $R$.

The relative CCSD(T) and CCSD(T)-F12 performance across different regions of the interaction varies with the type of the complex. For the dispersion-bound He–$CH_4$ system, the most important conclusion from Fig. 1 is the same as for the He–He curve investigated in Ref. 72: the advantage of the F12 approach over conventional CCSD(T) substantially increases for short intermolecular distances where the unscaled-triples CCSD(T)-F12b treatment is always better than conventional CCSD(T) in the same basis set and the scaled-triples version is still better. On the other hand, the ordering of CCSD(T) and unscaled-triples CCSD(T)-F12b is in most cases reversed in the asymptotic region. The scaled-triples CCSD(T)-F12b approach remains superior to conventional CCSD(T) in that regime but its advantage is not nearly as large as for small $R$.

The CBS-extrapolated He–$CH_4$ potential energy curves, presented in the lower panel of Fig. 1, show that the (D,T) extrapolations, both conventional and explicitly correlated, are fairly inaccurate. On the other hand, all (T,Q) extrapolations perform well and provide a substantial improvement over the reference CCSD(T)/aQZ method across the entire range of $R$. In fact, the conventional and explicitly correlated results provide very similar accuracy when the (T,Q) extrapolation is performed. Moreover, similar to what has been found for the minima, the unscaled-triples and scaled-triples extrapolations give nearly identical values. The relative deviations of potential energy curves for the methane dimer, Fig. 2, are fairly similar to the helium-methane case except for the fact that the CCSD(T)-F12b/a(D,T)Z and CCSD(T)-F12b/a(D,T)ZM extrapolations no longer suffer from a dramatic breakdown at large $R$.

In the case of the water dimer (Fig. 3), the errors of different approaches relative to the CCSD(T)/aQZ error are fairly constant. Thus, the advantage of CCSD(T)-F12b over CCSD(T) is just as striking as for the minimum: the most accurate conventional results, CCSD(T)/aQZM, are (slightly) worse than the least accurate F12 results, unscaled-triples CCSD(T)-F12b/aTZ. The benefits of adding midbond functions, although undeniable for both CCSD(T) and CCSD(T)-F12, are smaller than for the less polar dimers so that the results in the aQZ basis are better than the aTZM ones. The CBS extrapolation makes conventional CCSD(T) results competitive to CCSD(T)-F12, but only at the (T,Q) level. The potential energy curves for the remaining four dimers, shown in Figs. S13–S16 in the supplementary material,[83] are fairly similar to either the He–$CH_4$ and $CH_4$–$CH_4$ case (He–$H_2O$, Ar–$H_2O$, and Ar–$CH_4$) or to the $H_2O$–$H_2O$ case ($CH_4$–$H_2O$). The occasional larger errors at the largest $R$ are just an artifact of the reference CCSD(T)/aQZ calculations being particularly close to the benchmark in this case.

As mentioned earlier, neither MUE nor MURE is a good statistical measure of the performance of different approaches for the entire potential energy curves. Therefore, we followed the idea of Ref. 98 and computed the *median* unsigned

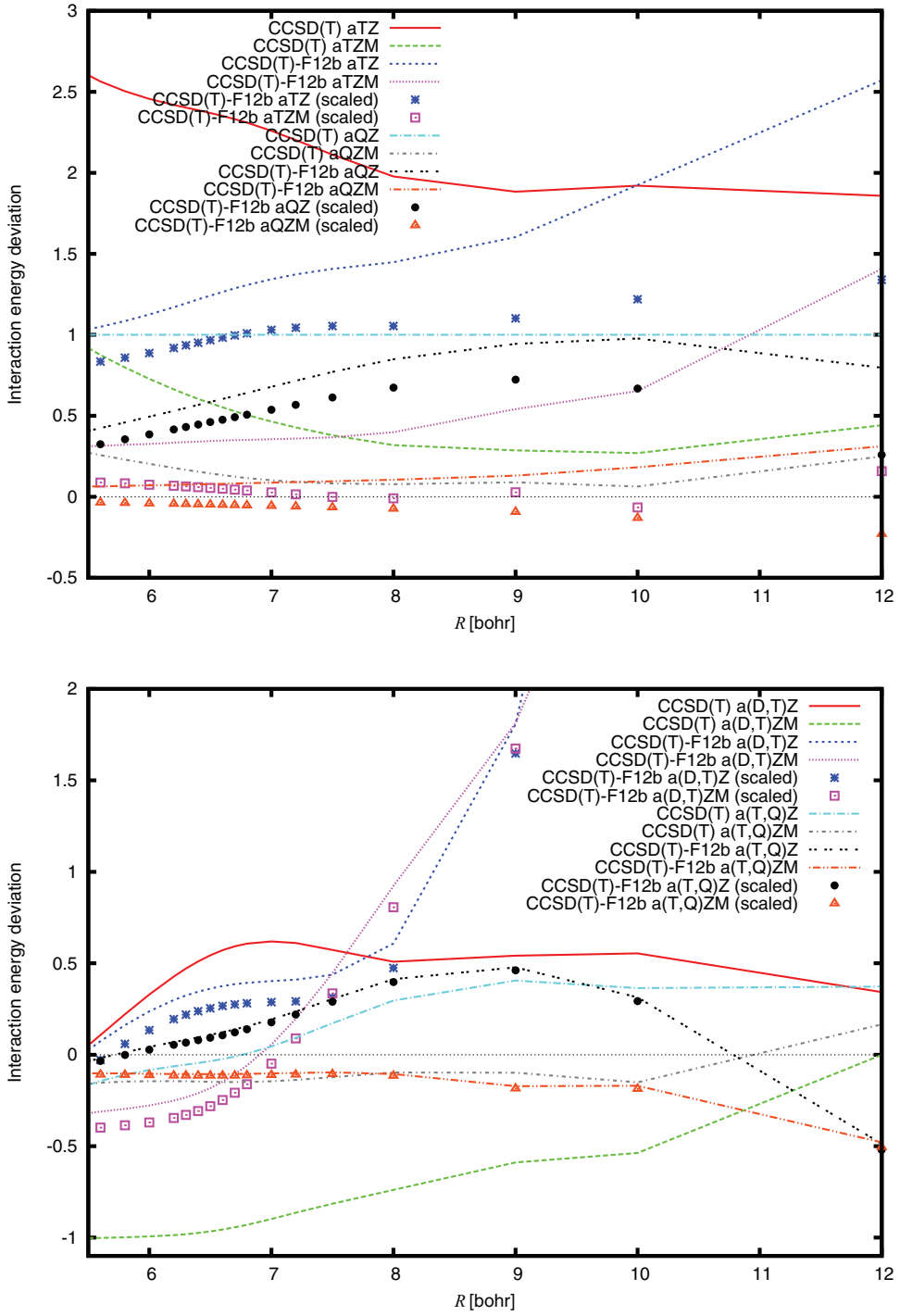


FIG. 1. Performance of the CCSD(T) and CCSD(T)-F12b approaches for the radial potential energy curve of the He–$CH_4$ complex passing through the global minimum. The values displayed are relative to the benchmark interaction energy (estimated from the scaled-triples CCSD(T)-F12c/aQZM calculation) and normalized by the absolute deviation of the CCSD(T)/aQZ interaction energy from the benchmark at a given intermolecular distance $R$.

relative errors (which will be denoted by MeURE to distinguish them from the regular MURE) for all methods and basis sets (one should note that other, properly weighted modifications of MURE would also be suitable[99]). The results are collected in Table XI. The values of MeURE are highly similar to the MURE for the seven van der Waals minima (Table IX) despite comparing to a different, less precise benchmark. In other words, the relative performance of various CCSD(T)/CBS approximations at the minima is a reliable indicator of the relative performance for other intermonomer separations. The scaled-triples CCSD(T)-F12a/a$X$ZM treatment is again the best one for $X$ = D and the worst out of the six F12 variants for $X$ = Q. On the other hand, the CCSD(T)-F12b and CCSD(T)-F12c approaches perform quite similar to each other and in general different from CCSD(T)-F12a. This observation is consistent with both the formal hierarchy of approximations to CCSD-F12[16–18] and the behavior of the CCSD-F12 interaction energies shown in Table X.

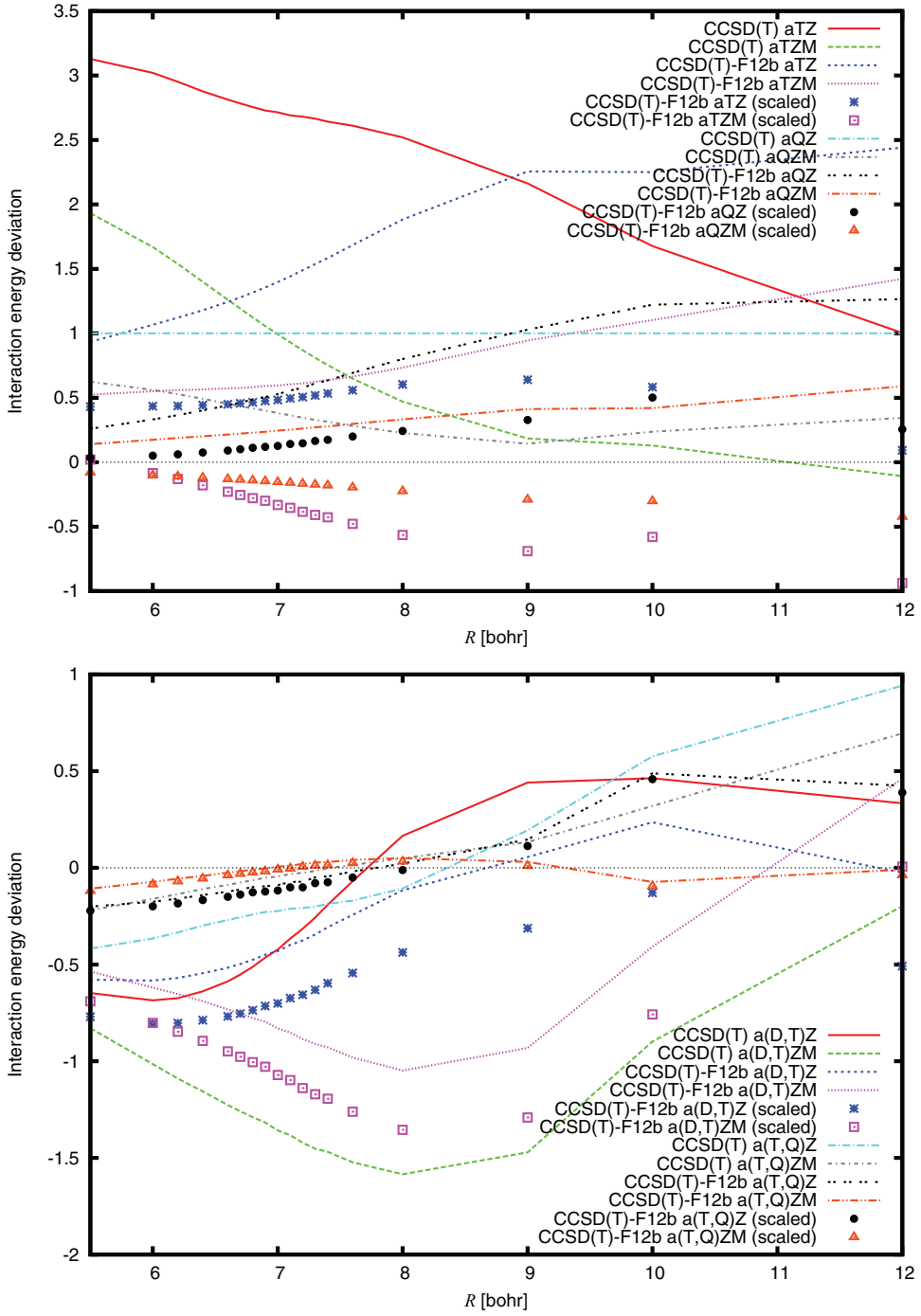


FIG. 2. Performance of the CCSD(T) and CCSD(T)-F12b approaches for the radial potential energy curve of the $CH_4$–$CH_4$ complex passing through the global minimum. The values displayed are relative to the benchmark interaction energy (estimated from the scaled-triples CCSD(T)-F12c/aQZM calculation) and normalized by the absolute deviation of the CCSD(T)/aQZ interaction energy from the benchmark at a given intermolecular distance $R$.

### B. Sensitivity of interaction energies to the geminal exponent $\beta$

In this subsection we investigate how the CCSD(T)-F12 interaction energies depend on the geminal exponent $\beta$ in the F12 correlation factor. The values of $\beta$ recommended for different basis sets[23,86] have been obtained by variationally minimizing the MP2-F12 correlation energy for a set of model atoms and molecules. The value of $\beta$ determines the size of the correlation hole modeled by the geminal[100]—the larger the value of $\beta$, the tighter the correlation hole. The correlation hole associated with dispersion interactions is likely highly diffuse and smaller values of $\beta$ might be preferable. Therefore, we decided to investigate how the choice of $\beta$ influences the MP2-F12 and CCSD(T)-F12b interaction energies for two model systems, He–$CH_4$ and $CH_4$–$H_2O$. The pertinent results are displayed in Fig. 4. As observed before,[8,101,102] the dependence of correlation energy on $\beta$ is particularly strong for small basis sets for which the orbital basis leaves a lot of room for improvement. Larger basis sets, especially the

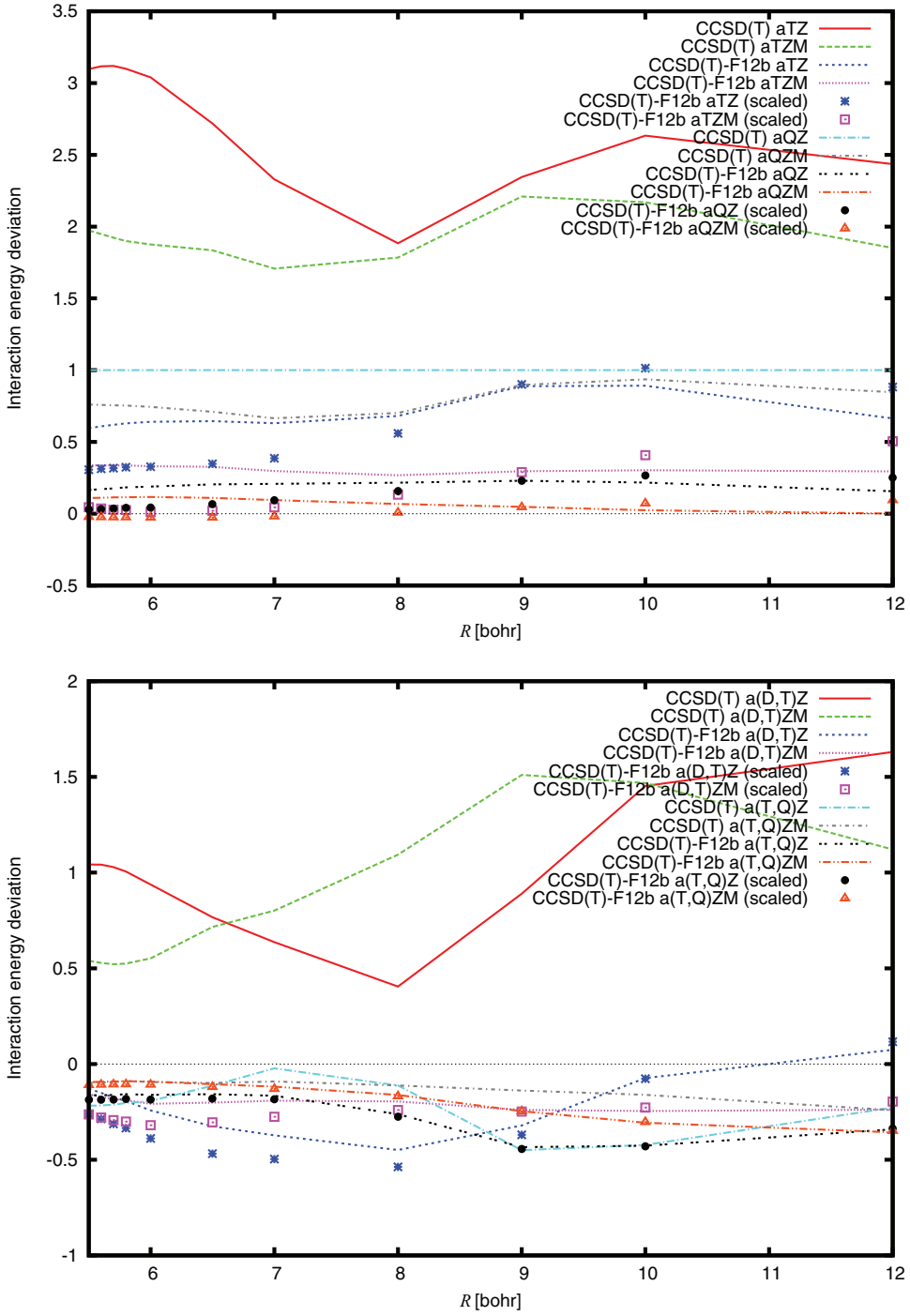


FIG. 3. Performance of the CCSD(T) and CCSD(T)-F12b approaches for the radial potential energy curve of the $H_2O$–$H_2O$ complex passing through the global minimum. The values displayed are relative to the benchmark interaction energy (estimated from the scaled-triples CCSD(T)-F12c/aQZM calculation) and normalized by the absolute deviation of the CCSD(T)/aQZ interaction energy from the benchmark at a given intermolecular distance $R$.

ones containing midbond functions, lead to nearly constant interaction energies over a wide range of $\beta$. The comparison of the left (MP2-F12) and right (CCSD(T)-F12b) panels of Fig. 4 shows that the two approaches result in very similar $\beta$ dependence which validates the optimization of $\beta$ at the MP2-F12 level.[23,86] The values of $\beta$ that minimize the MP2-F12 interaction energy are indeed lower than the commonly recommended values: in fact, they are close to 1.0 $a_0^{-1}$ for all systems and basis sets displayed in Fig. 4.

It should be noted that, in the most popular implementations of MP2-F12 and CCSD(T)-F12, the Slater-type correlation factor is represented by a fit to a fixed number $N_{\mathrm{GTG}}$ of GTGs,[21]

$$e^{-\beta r_{12}} \approx \sum_{i=1}^{N_{\mathrm{GTG}}} c_i e^{-\xi_i r_{12}^2}, \tag{3}$$

where, in general, both the coefficients $c_i$ and the exponents $\xi_i$ are fitted parameters. This fitting obviously affects the

TABLE XI. Median unsigned relative errors (MeURE, in percent) of different CCSD(T)/CCSD(T)-F12 variants and basis sets for the radial interaction energy curves of all seven complexes. The benchmark CCSD(T)/CBS interaction energies have been computed at the scaled-triples CCSD(T)-F12c/aQZM level.

| | | Computed results, $X =$ | | | Extrapolated results | |
|---|---|---|---|---|---|---|
| Method | Basis | D | T | Q | (D,T) | (T,Q) |
| CCSD(T) | a*X*Z | 26.33 | 8.48 | 3.12 | 1.84 | 0.51 |
| CCSD(T) | a*X*ZM | 16.43 | 2.66 | 0.92 | 2.89 | 0.40 |
| CCSD(T)-F12a | a*X*Z | 12.77 | 3.21 | 1.01 | 0.95 | 0.50 |
| CCSD(T)-F12a | a*X*ZM | 6.04 | 0.34 | 0.07 | 2.39 | 0.32 |
| CCSD(T)-F12a (scaled) | a*X*Z | 8.75 | 1.19 | 0.74 | 1.18 | 0.52 |
| CCSD(T)-F12a (scaled) | a*X*ZM | 1.78 | 1.55 | 0.81 | 2.91 | 0.29 |
| CCSD(T)-F12b | a*X*Z | 15.42 | 4.47 | 1.98 | 1.23 | 0.41 |
| CCSD(T)-F12b | a*X*ZM | 9.07 | 1.32 | 0.46 | 2.00 | 0.25 |
| CCSD(T)-F12b (scaled) | a*X*Z | 11.69 | 2.63 | 0.81 | 1.02 | 0.45 |
| CCSD(T)-F12b (scaled) | a*X*ZM | 4.72 | 0.46 | 0.30 | 2.53 | 0.32 |
| CCSD(T)-F12c | a*X*Z | 16.64 | 5.21 | 2.31 | 1.22 | 0.32 |
| CCSD(T)-F12c | a*X*ZM | 10.16 | 1.86 | 0.76 | 1.51 | 0.20 |
| CCSD(T)-F12c (scaled) | a*X*Z | 12.85 | 3.43 | 1.25 | 1.10 | 0.39 |
| CCSD(T)-F12c (scaled) | a*X*ZM | 5.51 | 0.39 | 0.00 | 2.11 | 0.25 |

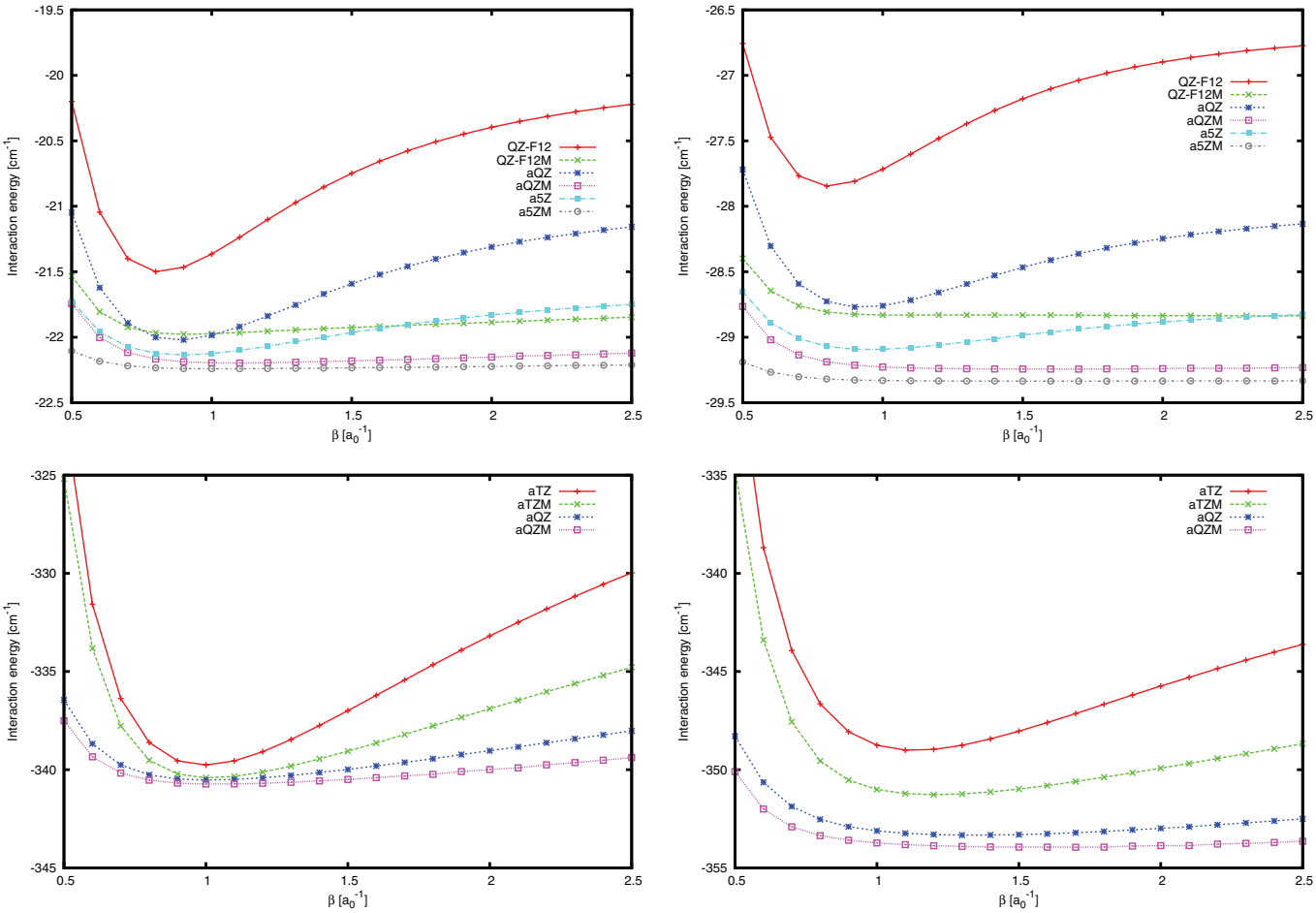


FIG. 4. Dependence of the MP2-F12 (left panels) and CCSD(T)-F12b (right panels) interaction energies on the geminal exponent $\beta$ for the He–$CH_4$ (top) and $CH_4$–$H_2O$ (bottom) dimers. The MP2-F12 calculations use the 3C(FIX) *Ansatz* just like the CCSD(T)-F12b approach. The triples term in CCSD(T)-F12b is not scaled.

description of the interelectronic cusp—in fact, the right-hand side of Eq. (3) behaves for small $r_{12}$ like $C + O(r_{12}^2)$ and is incapable of recovering the leading linear term for any $c_i$ and $\xi_i$. Nevertheless, in the diagonal *Ansatz*,[8] the explicitly correlated pair function amplitudes are fixed to the values that would have reproduced the exact linear terms[1,2] had the Slater-type correlation factor not been expanded in GTGs.

It is not our intention to criticize the Gaussian geminal fit of Eq. (3). Extensive numerical experience with the F12 methods shows that it is the improved description of the entire near-coalescence regions of the many-electron space, rather than forcing the exact behavior of the wavefunction in the coalescence limit, that is the key to the superior basis set convergence of correlation energy. Thus, a reproduction of the exact linear terms by the diagonal *Ansatz* is not a requirement for obtaining accurate correlation energies and one can obtain the coefficients $c_i$ and exponents $\xi_i$ in a manner other than fitting to a Slater-type formula. Therefore, we explored the possibility of generating a GTG *Ansatz* through a straightforward interaction energy minimization. Specifically, we selected the He–$CH_4$ dimer, set $N_{GTG} = 6$, and assumed the exponents $\xi_1, \ldots, \xi_6$ to form an even-tempered expansion with the center at 1.0 and the ratio of 3.0. The corresponding coefficients $c_1, \ldots, c_6$ were then chosen to minimize the value of the (non-density-fitted) MP2-F12 interaction energy in the 3C(FIX) *Ansatz*. The actual minimization was performed using the Powell algorithm as implemented in the *SciPy* library[103] with calls to MOLPRO[9] to compute MP2-F12 interaction energies. Separate minimizations were carried out for the a*X*Z and a*X*ZM basis sets for $X$ = D,T,Q. It should be noted that the idea of optimizing GTG coefficients in the context of MP2-F12 is neither original (Valeev[102] has already observed that short optimized GTG expansions provide superior MP2-F12 correlation energies compared to a single Slater-type factor) nor computationally sensible for each dimer and each basis set separately. Nevertheless, such an optimization provides a useful indication of how much the interaction energies can be improved by varying the F12 correlation factor.

The comparison of the MP2-F12 and CCSD(T)-F12b interaction energies in the He–$CH_4$ complex, obtained using different correlation factors, is presented in Table XII. We listed three selections of $\beta$ for the Slater-type factor: the MOLPRO default choice $\beta = 1.0\ a_0^{-1}$, the value recommended in Ref. 86 (1.1, 1.2, 1.4, and 1.4 $a_0^{-1}$ for aDZ, aTZ, aQZ, and a5Z, respectively), and the value that minimizes the density-fitted MP2-F12 interaction energy (Fig. 4). Additionally, we have listed the results obtained using an optimized contraction of six GTGs, generated as described in the preceding paragraph. Due to computational complexity, we have not performed a GTG optimization for the a5Z and a5ZM basis sets.

The MP2-F12 results in Table XII confirm that the optimal exponent $\beta$ in the Slater-type correlation factor is very close to 1.0 $a_0^{-1}$ regardless of the basis set. Consequently, the optimization of $\beta$ provides little improvement over simply setting $\beta$ to 1.0 $a_0^{-1}$. The use of the $\beta$ exponents recommended based on molecular correlation energy calculations[86] increases the errors somewhat (from 1.59 to 2.45 $cm^{-1}$ for aDZ and from 0.61 to 1.03 $cm^{-1}$ for aTZ). Apparently, the treatment of dispersion energy requires a more diffuse correlation hole than the calculations of molecular correlation energies. It should be noted, however, that the selection of $\beta$ matters only for small basis sets. For the aTZM, aQZM, and a5ZM bases the three choices of $\beta$ do not lead to a difference of more than 0.02 $cm^{-1}$. Moreover, the differences become even less pronounced at the CCSD(T)-F12b level of theory. Therefore, while our study indicates that the simple $\beta = 1.0\ a_0^{-1}$ choice should be the recommended one for weak interaction energy calculations, the expected improvement over the values from Ref. 86 employed throughout the rest of this work does not warrant a recalculation of all results. As the results in the last column of Table XII indicate, the GTG contraction that is fully optimized for interaction energy does provide some improvement over the contractions obtained by fitting the Slater-type correlation factor. However, the improvement quickly diminishes when one goes to larger

TABLE XII. MP2-F12 and CCSD(T)-F12b interaction energies (in $cm^{-1}$) for the near-minimum geometry of the He–$CH_4$ complex, computed using several different choices of the correlation factor. The calculations in the first three columns employed a single Slater-type correlation factor fitted to six GTGs. The geminal exponent $\beta$ was fixed at 1.0 $a_0^{-1}$ (first column), taken as the recommended value from Ref. 86 (like in most calculations in this work, second column), or chosen to minimize the MP2-F12 interaction energy (third column). The calculations in the column marked "Optimized GTGs" utilized six GTGs with even-tempered exponents and coefficients obtained by minimizing the MP2-F12 interaction energy as described in the text. The diagonal 3C(FIX) *Ansatz* and the same DF/RI bases as in the rest of this work were used. The numbers in parentheses are the values of $\beta$. The benchmark MP2/CBS and CCSD(T)/CBS interaction energies amount to −22.283 and −29.425 $cm^{-1}$, respectively.

| Basis | $\beta = 1.0\ a_0^{-1}$ | $\beta$ from Ref. 86 | | Optimal $\beta$ | | Optimized GTGs |
|---|---|---|---|---|---|---|
| | | | MP2-F12 | | | |
| aDZ | −20.694 | −19.838 | (1.1) | −21.135 | (0.90) | −21.794 |
| aDZM | −22.465 | −21.784 | (1.1) | −22.885 | (0.89) | −23.559 |
| aTZ | −21.676 | −21.250 | (1.2) | −21.749 | (0.91) | −21.950 |
| aTZM | −22.028 | −22.011 | (1.2) | −22.032 | (1.05) | −22.111 |
| aQZ | −21.985 | −21.669 | (1.4) | −22.021 | (0.88) | −22.108 |
| aQZM | −22.195 | −22.185 | (1.4) | −22.197 | (1.08) | −22.208 |
| a5Z | −22.124 | −22.001 | (1.4) | −22.135 | (0.91) | |
| a5ZM | −22.241 | −22.236 | (1.4) | −22.241 | (1.04) | |
| | | | CCSD(T)-F12b (unscaled triples) | | | |
| aDZ | −25.248 | −24.590 | (1.1) | −25.416 | (0.90) | −26.090 |
| aDZM | −27.708 | −27.198 | (1.1) | −27.871 | (0.89) | −28.242 |
| aTZ | −27.908 | −27.602 | (1.2) | −27.913 | (0.91) | −28.123 |
| aTZM | −28.844 | −28.868 | (1.2) | −28.859 | (1.05) | −28.972 |
| aQZ | −28.759 | −28.528 | (1.4) | −28.766 | (0.88) | −28.877 |
| aQZM | −29.227 | −29.242 | (1.4) | −29.234 | (1.08) | −29.262 |
| a5Z | −29.093 | −29.012 | (1.4) | −29.094 | (0.91) | |
| a5ZM | −29.331 | −29.335 | (1.4) | −29.332 | (1.04) | |
| | | | CCSD(T)-F12b (scaled triples) | | | |
| aDZ | −26.195 | −25.535 | (1.1) | −26.356 | (0.90) | −27.026 |
| aDZM | −28.756 | −28.242 | (1.1) | −28.909 | (0.89) | −29.214 |
| aTZ | −28.292 | −27.989 | (1.2) | −28.293 | (0.91) | −28.498 |
| aTZM | −29.248 | −29.275 | (1.2) | −29.264 | (1.05) | −29.375 |
| aQZ | −28.934 | −28.704 | (1.4) | −28.939 | (0.88) | −29.045 |
| aQZM | −29.405 | −29.421 | (1.4) | −29.412 | (1.08) | −29.436 |
| a5Z | −29.186 | −29.105 | (1.4) | −29.186 | (0.91) | |
| a5ZM | −29.423 | −29.428 | (1.4) | −29.425 | (1.04) | |

basis sets, and it never constitutes a significant fraction of the error remaining at the CCSD(T)-F12b level of theory. Thus, our tests indicate that the Slater-type correlation factor with a properly chosen parameter $\beta$ is close to optimal and further optimization of this factor does little to overcome the remaining basis set incompleteness effects.

## IV. SUMMARY

We have investigated the basis set convergence of weak interaction energies at the "gold-standard" CCSD(T) level of theory, employing both conventional CCSD(T) and three explicitly correlated approaches: CCSD(T)-F12a, CCSD(T)-F12b, and CCSD(T)(F12*)≡CCSD(T)-F12c. This work extends our earlier study of noble gas dimers[72] to seven complexes of varying polarity and interaction strength: He–$H_2O$, Ar–$H_2O$, He–$CH_4$, Ar–$CH_4$, $H_2O$–$H_2O$, $CH_4$–$H_2O$, and $CH_4$–$CH_4$. We established benchmark CCSD(T)/CBS interaction energies for the van der Waals minima using conventional CCSD(T) in bases up to a6ZM and CCSD(T)-F12 in bases up to a5ZM. The two approaches displayed remarkable consistency giving a mean absolute deviation from each other equal to just 0.07 $cm^{-1}$.

The seven dimers considered in this work can be divided into two classes. For the strongly bound water-water and methane-water dimers, the advantage of CCSD(T)-F12 over conventional CCSD(T) is tremendous. The situation for the remaining five dimers is substantially different: while the F12 treatment clearly improves the results, the improvement is in general smaller than the one afforded by a simple addition of midbond functions. The scaling of the (T) triples contribution [Eq. (1)] does improve convergence in most cases, however, the gains for the midbond-containing basis sets are limited as the scaling overshoots.

Our numerical investigations shed some light on the residual deviations of CCSD(T)-F12 weak interaction energies from the CCSD(T)/CBS limit. They also provide some recommendations how to maximize the performance of F12 interaction energy calculations. First, midbond functions consistently provide a large boost to the basis set convergence. Second, the value of 1.0 $a_0^{-1}$ for the geminal exponent $\beta$ is close to optimal although the dependence of interaction energies on $\beta$ is not strong. Finally, one should stay away from the CCSD(T)-F12a approach. While its scaled-triples version does give very reasonable interaction energies in double-zeta basis sets, it does not provide a systematic improvement when the orbital basis is enlarged and/or the CBS extrapolation is performed. On the other hand, the unscaled-triples CCSD(T)-F12a approach strongly benefits from an error cancellation between the CCSD and (T) contributions. Such a cancellation is much less of a critical factor for the scaled-triples CCSD(T)-F12b and CCSD(T)-F12c approaches and these variants are the best choices to maximize the accuracy of computed interaction energies. If the auxiliary basis sets employed are sufficiently complete, the leading source of residual error in large-basis CCSD(T)-F12b and CCSD(T)-F12c interaction energies is the inaccuracy of the triples term (either scaled or unscaled).

We are well aware that the subtle differences between various large-basis CBS estimates studied in this work are often smaller than the interaction energy contributions neglected at the frozen-core CCSD(T) level, in particular, the effects of higher order coupled-cluster excitations. The increase of computer power and the algorithmic and implementation improvements have made it possible to include the latter effects for small dimers using, for example, a moderate-basis CCSDT(Q) treatment.[104,105] As a result, residual errors in the frozen-core CCSD(T) interaction energy become once again one of the leading contributions to the overall interaction energy uncertainty even when very large basis sets are used.[74–76,106,107] Thus, the benchmark investigations presented here are likely to aid in the further development of ultra-accurate weak interaction potentials for small dimers of experimental and theoretical interest.

## ACKNOWLEDGMENTS

This work has been supported by the Auburn University startup funding.

[1] T. Kato, Commun. Pure Appl. Math. **10**, 151 (1957).
[2] R. T Pack and W. Byers Brown, J. Chem. Phys. **45**, 556 (1966).
[3] C. Hättig, W. Klopper, A. Köhn, and D. P. Tew, Chem. Rev. **112**, 4 (2012).
[4] L. Kong, F. A. Bischoff, and E. F. Valeev, Chem. Rev. **112**, 75 (2012).
[5] J. Mitroy, S. Bubin, K. Varga, W. Horiuchi, Y. Suzuki, L. Adamowicz, W. Cencek, K. Szalewicz, J. Komasa, and D. Blume, "Theory and application of explicitly correlated Gaussians," Rev. Mod. Phys. (submitted).
[6] F. R. Manby, J. Chem. Phys. **119**, 4607 (2003).
[7] A. J. May, E. Valeev, R. Polly, and F. R. Manby, Phys. Chem. Chem. Phys. **7**, 2710 (2005).
[8] S. Ten-no, Chem. Phys. Lett. **398**, 56 (2004).
[9] H.-J. Werner, P. J. Knowles, F. R. Manby, M. Schütz *et al.*, MOLPRO, version 2010.1, a package of *ab initio* programs, 2010, see http://www.molpro.net.
[10] TURBOMOLE V6.3 2011, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989–2007, turbomole GmbH, since 2007, see http://www.turbomole.com.
[11] S. Höfener, C. Hättig, and W. Klopper, Z. Phys. Chem. **224**, 695 (2010).
[12] S. Höfener and W. Klopper, Mol. Phys. **108**, 1783 (2010).
[13] S. Ten-no, Chem. Phys. Lett. **447**, 175 (2007).
[14] T. Shiozaki and H.-J. Werner, J. Chem. Phys. **133**, 141103 (2010).
[15] S. Kedžuch, O. Demel, J. Pittner, S. Ten-no, and J. Noga, Chem. Phys. Lett. **511**, 418 (2011).
[16] T. B. Adler, G. Knizia, and H.-J. Werner, J. Chem. Phys. **127**, 221106 (2007).
[17] G. Knizia, T. B. Adler, and H.-J. Werner, J. Chem. Phys. **130**, 054104 (2009).
[18] C. Hättig, D. P. Tew, and A. Köhn, J. Chem. Phys. **132**, 231102 (2010).
[19] E. F. Valeev, Chem. Phys. Lett. **395**, 190 (2004).
[20] D. P. Tew and W. Klopper, J. Chem. Phys. **125**, 094302 (2006).
[21] H.-J. Werner, T. B. Adler, and F. R. Manby, J. Chem. Phys. **126**, 164102 (2007).
[22] S. Höfener, D. P. Tew, W. Klopper, and T. Helgaker, Chem. Phys. **356**, 25 (2009).
[23] J. G. Hill, K. A. Peterson, G. Knizia, and H.-J. Werner, J. Chem. Phys. **131**, 194105 (2009).
[24] T. Shiozaki, M. Kamiya, S. Hirata, and E. F. Valeev, J. Chem. Phys. **129**, 071101 (2008).
[25] A. Köhn, J. Chem. Phys. **130**, 131101 (2009).
[26] D. P. Tew, W. Klopper, C. Neiss, and C. Hättig, Phys. Chem. Chem. Phys. **9**, 1921 (2007).
[27] T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989).
[28] R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).
[29] J. M. Peréz-Jordá and A. D. Becke, Chem. Phys. Lett. **233**, 134 (1995).

[30]X. Wu, M. C. Vargas, S. Nayak, V. F. Lotrich, and G. Scoles, J. Chem. Phys. **115**, 8748 (2001).
[31]P. Jurečka, J. Šponer, J. Černý, and P. Hobza, Phys. Chem. Chem. Phys. **8**, 1985 (2006).
[32]C. D. Sherrill, T. Takatani, and E. G. Hohenstein, J. Phys. Chem. A **113**, 10146 (2009).
[33]Q. Wu and W. Yang, J. Chem. Phys. **116**, 515 (2002).
[34]S. Grimme, J. Chem. Phys. **118**, 9095 (2003).
[35]A. Hesselmann, G. Jansen, and M. Schütz, J. Chem. Phys. **122**, 014103 (2005).
[36]A. J. Misquitta, R. Podeszwa, B. Jeziorski, and K. Szalewicz, J. Chem. Phys. **123**, 214103 (2005).
[37]Y. Zhao, N. E. Schultz, and D. G. Truhlar, J. Chem. Theory Comput. **2**, 364 (2006).
[38]S. Grimme, J. Comput. Chem. **27**, 1787 (2006).
[39]S. Grimme, J. Chem. Phys. **124**, 034108 (2006).
[40]A. D. Becke and E. R. Johnson, J. Chem. Phys. **127**, 124108 (2007).
[41]J. G. Hill and J. A. Platts, J. Chem. Theory Comput. **3**, 80 (2007).
[42]R. A. DiStasio, Jr. and M. Head-Gordon, Mol. Phys. **105**, 1073 (2007).
[43]Y. Zhao and D. G. Truhlar, Theor. Chem. Acc. **120**, 215 (2008).
[44]J.-D. Chai and M. Head-Gordon, Phys. Chem. Chem. Phys. **10**, 6615 (2008).
[45]A. Hesselmann, J. Chem. Phys. **128**, 144112 (2008).
[46]T. Takatani, E. G. Hohenstein, and C. D. Sherrill, J. Chem. Phys. **128**, 124111 (2008).
[47]A. Tkatchenko and M. Scheffler, Phys. Rev. Lett. **102**, 073005 (2009).
[48]K. Pernal, R. Podeszwa, K. Patkowski, and K. Szalewicz, Phys. Rev. Lett. **103**, 263201 (2009).
[49]Y. Zhang, X. Xu, and W. A. Goddard III, Proc. Natl. Acad. Sci. U.S.A. **106**, 4963 (2009).
[50]M. Pitoňák, P. Neogrady, J. Černý, S. Grimme, and P. Hobza, ChemPhysChem **10**, 282 (2009).
[51]S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, J. Chem. Phys. **132**, 154104 (2010).
[52]S. N. Steinmann and C. Corminboeuf, J. Chem. Theory Comput. **6**, 1990 (2010).
[53]O. A. Vydrov and T. Van Voorhis, J. Chem. Phys. **133**, 244103 (2010).
[54]Ł. Rajchel, P. S. Żuchowski, M. Hapka, M. Modrzejewski, M. M. Szczęśniak, and G. Chałasiński, Phys. Chem. Chem. Phys. **12**, 14686 (2010).
[55]A. Hesselmann, J. Phys. Chem. A **115**, 11321 (2011).
[56]L. A. Burns, A. Vazquez-Mayagoitia, B. G. Sumpter, and C. D. Sherrill, J. Chem. Phys. **134**, 084107 (2011).
[57]A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen, and A. K. Wilson, Chem. Phys. Lett. **286**, 243 (1998).
[58]D. W. Schwenke, J. Chem. Phys. **122**, 014107 (2005).
[59]F.-M. Tao and Y.-K. Pan, J. Phys. Chem. **95**, 3582 (1991).
[60]M. Jeziorska, W. Cencek, K. Patkowski, B. Jeziorski, and K. Szalewicz, J. Chem. Phys. **127**, 124303 (2007).
[61]W. Cencek and K. Szalewicz, Int. J. Quantum Chem. **108**, 2191 (2008).
[62]W. Cencek, M. Przybytek, J. Komasa, J. B. Mehl, B. Jeziorski, and K. Szalewicz, J. Chem. Phys. **136**, 224303 (2012).
[63]O. Marchetti and H.-J. Werner, Phys. Chem. Chem. Phys. **10**, 3400 (2008).
[64]D. P. Tew, W. Klopper, and C. Hättig, Chem. Phys. Lett. **452**, 326 (2008).
[65]O. Marchetti and H.-J. Werner, J. Phys. Chem. A **113**, 11580 (2009).
[66]J. R. Lane and H. G. Kjaergaard, J. Chem. Phys. **131**, 034307 (2009).
[67]K. M. de Lange and J. R. Lane, J. Chem. Phys. **134**, 034301 (2011).
[68]J. D. McMahon and J. R. Lane, J. Chem. Phys. **135**, 154309 (2011).
[69]W. Skomorowski, F. Pawłowski, T. Korona, R. Moszyński, P. S. Żuchowski, and J. M. Hutson, J. Chem. Phys. **134**, 114109 (2011).
[70]M. S. Marshall and C. D. Sherrill, J. Chem. Theory Comput. **7**, 3978 (2011).
[71]E. J. Carrell, C. M. Thorne, and G. S. Tschumper, J. Chem. Phys. **136**, 014103 (2012).
[72]K. Patkowski, J. Chem. Phys. **137**, 034103 (2012).
[73]D. G. A. Smith and K. Patkowski, J. Chem. Theory Comput. **9**, 370 (2013).
[74]J. R. Lane, J. Chem. Theory Comput. **9**, 316 (2013).
[75]R. Hellmann, E. Bich, and E. Vogel, Mol. Phys. **106**, 133 (2008).
[76]K. Patkowski and K. Szalewicz, J. Chem. Phys. **133**, 094304 (2010).
[77]K. Patkowski, W. Cencek, M. Jeziorska, B. Jeziorski, and K. Szalewicz, J. Phys. Chem. A **111**, 7611 (2007).
[78]J. Makarewicz, J. Chem. Phys. **129**, 184310 (2008).
[79]W. A. Alexander and D. Troya, J. Phys. Chem. A **110**, 10834 (2006).
[80]R. Bukowski, K. Szalewicz, G. C. Groenenboom, and A. van der Avoird, Science **315**, 1249 (2007).
[81]O. Akin-Ojo and K. Szalewicz, J. Chem. Phys. **123**, 134311 (2005).
[82]R. Hellmann, E. Bich, and E. Vogel, J. Chem. Phys. **128**, 214303 (2008).
[83]See supplementary material at http://dx.doi.org/10.1063/1.4800981 for additional tables and figures.
[84]S. F. Boys and F. Bernardi, Mol. Phys. **19**, 553 (1970).
[85]F. B. van Duijneveldt, J. G. C. M. van Duijneveldt-van de Rijdt, and J. H. van Lenthe, Chem. Rev. **94**, 1873 (1994).
[86]K. A. Peterson, T. B. Adler, and H.-J. Werner, J. Chem. Phys. **128**, 084102 (2008).
[87]F. Weigend, A. Köhn, and C. Hättig, J. Chem. Phys. **116**, 3175 (2002).
[88]C. Hättig, Phys. Chem. Chem. Phys. **7**, 59 (2005).
[89]K. L. Schuchardt, B. T. Didier, T. Elsethagen, L. Sun, V. Gurumoorthi, J. Chase, J. Li, and T. L. Windus, J. Chem. Inf. Model. **47**, 1045 (2007).
[90]C. Hättig, unpublished results.
[91]F. Weigend, Phys. Chem. Chem. Phys. **4**, 4285 (2002).
[92]K. E. Yousaf and K. A. Peterson, Chem. Phys. Lett. **476**, 303 (2009).
[93]F.-M. Tao, J. Chem. Phys. **100**, 4947 (1994).
[94]O. Akin-Ojo, R. Bukowski, and K. Szalewicz, J. Chem. Phys. **119**, 8379 (2003).
[95]T. Helgaker, W. Klopper, H. Koch, and J. Noga, J. Chem. Phys. **106**, 9639 (1997).
[96]W. Kutzelnigg and W. Klopper, J. Chem. Phys. **94**, 1985 (1991).
[97]W. Klopper and C. C. M. Samson, J. Chem. Phys. **116**, 6397 (2002).
[98]R. Podeszwa and K. Szalewicz, J. Chem. Phys. **136**, 161102 (2012).
[99]K. S. Thanthiriwatte, E. G. Hohenstein, L. A. Burns, and C. D. Sherrill, J. Chem. Theory Comput. **7**, 88 (2011).
[100]H.-J. Werner, G. Knizia, and F. R. Manby, Mol. Phys. **109**, 407 (2011).
[101]D. P. Tew and W. Klopper, J. Chem. Phys. **123**, 074101 (2005).
[102]E. F. Valeev, J. Chem. Phys. **125**, 244106 (2006).
[103]E. Jones, T. Oliphant, and P. Peterson, SciPy: Open source scientific tools for Python, 2001, see http://www.scipy.org/.
[104]S. A. Kucharski and R. J. Bartlett, J. Chem. Phys. **108**, 9221 (1998).
[105]Y. J. Bomble, J. F. Stanton, M. Kállay, and J. Gauss, J. Chem. Phys. **123**, 054101 (2005).
[106]P. Jankowski, A. R. W. McKellar, and K. Szalewicz, Science **336**, 1147 (2012).
[107]J. Řezáč, L. Šimová, and P. Hobza, J. Chem. Theory Comput. **9**, 364 (2013).