

# On the accuracy of explicitly correlated coupled-cluster interaction energies — have orbital results been beaten yet?

Cite as: J. Chem. Phys. **137**, 034103 (2012); <https://doi.org/10.1063/1.4734597>

Submitted: 15 May 2012 . Accepted: 25 June 2012 . Published Online: 17 July 2012

Konrad Patkowski



View Online



Export Citation

## ARTICLES YOU MAY BE INTERESTED IN

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

The Journal of Chemical Physics **138**, 154101 (2013); <https://doi.org/10.1063/1.4800981>

[A simple and efficient CCSD\(T\)-F12 approximation](#)

The Journal of Chemical Physics **127**, 221106 (2007); <https://doi.org/10.1063/1.2817618>

[Simplified CCSD\(T\)-F12 methods: Theory and benchmarks](#)

The Journal of Chemical Physics **130**, 054104 (2009); <https://doi.org/10.1063/1.3054300>

Lock-in Amplifiers  
up to 600 MHz



# On the accuracy of explicitly correlated coupled-cluster interaction energies — have orbital results been beaten yet?

Konrad Patkowski

*Department of Chemistry and Biochemistry, Auburn University, Auburn, Alabama 36849, USA*

(Received 15 May 2012; accepted 25 June 2012; published online 17 July 2012)

The basis set convergence of weak interaction energies for dimers of noble gases helium through krypton is studied for six variants of the explicitly correlated, frozen geminal coupled-cluster singles, doubles, and noniterative triples [CCSD(T)-F12] approach: the CCSD(T)-F12a, CCSD(T)-F12b, and CCSD(T)(F12\*) methods with scaled and unscaled triples. These dimers were chosen because CCSD(T) complete-basis-set (CBS) limit benchmarks are available for them to a particularly high precision. The dependence of interaction energies on the auxiliary basis sets has been investigated and it was found that the default resolution-of-identity sets cc-pVXZ/JKFIT are far from adequate in this case. Overall, employing the explicitly correlated approach clearly speeds up the basis set convergence of CCSD(T) interaction energies, however, quite surprisingly, the improvement is not as large as the one achieved by a simple addition of bond functions to the orbital basis set. Bond functions substantially improve the CCSD(T)-F12 interaction energies as well. For small and moderate bases with bond functions, the accuracy delivered by the CCSD(T)-F12 approach cannot be matched by conventional CCSD(T). However, the latter method in the largest available bases still delivers the CBS limit to a better precision than CCSD(T)-F12 in the largest bases available for that approach. Our calculations suggest that the primary reason for the limited accuracy of the large-basis CCSD(T)-F12 treatment are the approximations made at the CCSD-F12 level and the non-explicitly correlated treatment of triples. In contrast, the explicitly correlated second-order Møller-Plesset perturbation theory (MP2-F12) approach is able to pinpoint the complete-basis-set limit MP2 interaction energies of rare gas dimers to a better precision than conventional MP2. Finally, we report and analyze an unexpected failure of the CCSD(T)-F12 method to deliver the core-core and core-valence correlation corrections to interaction energies consistently and accurately. © 2012 American Institute of Physics. [<http://dx.doi.org/10.1063/1.4734597>]

## I. INTRODUCTION

It is well recognized that the slow basis-set convergence of electron correlation effects is a consequence of an inadequate recovery of the interelectronic correlation cusp by linear combinations of one-electron basis functions. The best way to improve convergence is by making basis functions explicitly depend on the interelectronic distance  $r_{12}$ . High-accuracy calculations on atoms and small molecules have employed  $r_{12}$ -dependent bases from the very start<sup>1,2</sup> and continue to do so until today, with the highest accuracy achieved using very long expansions of  $r_{12}$ -dependent many-electron basis functions with parameters obtained by full nonlinear optimization.<sup>3–6</sup> For larger systems such an optimization is not feasible, but the accuracy of a correlation energy calculation can be improved even by inserting a constant, unoptimized  $r_{12}$  dependence into the basis set.<sup>7</sup> This idea, coupled with the application of the density-fitting (DF) techniques to simplify the required integrals<sup>8</sup> and with the choice of a frozen Gaussian geminal-like<sup>9</sup> (F12) form of the  $r_{12}$  dependence,  $f_{12} = \sum_k c_k e^{-\gamma_k r_{12}^2}$ ,<sup>10</sup> has led to the development of the widely available explicitly correlated second-order Møller-Plesset perturbation theory, DF-MP2-F12.<sup>11</sup>

While a frozen  $r_{12}$  factor was incorporated in the coupled-cluster singles and doubles theory (CCSD) as early

as 1992,<sup>12</sup> the actual numerical calculations were too complex for general applications until the recent introduction of DF and F12. The full explicitly correlated frozen-geminal CCSD (CCSD-F12) (Refs. 13 and 14) along with the F12-dependent noniterative contribution (T) of triple excitations<sup>15</sup> as in the coupled-cluster approach with singles, doubles, and noniterative triples [CCSD(T)] have been recently derived and implemented, but the codes are not yet generally available. Instead, several approximations to the full CCSD-F12 have been implemented in general-purpose electronic-structure programs such as MOLPRO<sup>16</sup> and TURBOMOLE<sup>17</sup> and these are the variants of the explicitly correlated CC theory that have gained widespread use—see Refs. 18 and 19 for up-to-date reviews. The approximations include CCSD(F12),<sup>20,21</sup> CCSD-F12a and CCSD-F12b,<sup>22</sup> and CCSD(F12\*), also termed CCSD-F12c.<sup>23</sup> For each of these variants, the triples contribution  $E_{(T)} = E_{\text{CCSD(T)}} - E_{\text{CCSD}}$  can be calculated in a perturbative manner just like in conventional CCSD(T), leading, for example, from CCSD-F12b to CCSD(T)-F12b. While the triples contribution computed in this way depends on the converged CCSD-F12b amplitudes and thus implicitly includes some explicitly correlated contribution, this inclusion is by no means exact and the inaccuracy of triples may dominate<sup>24</sup> the overall error with respect to the CCSD(T) complete basis set (CBS)

limit. To alleviate this problem, a simple scaling approximation has been proposed:<sup>24</sup>

$$E_{(T)}^* = E_{(T)} \cdot \frac{E_{\text{MP2-F12}}^{\text{corr}}}{E_{\text{MP2}}^{\text{corr}}}, \quad (1)$$

where  $E_X^{\text{corr}}$  denotes correlation energy at the X level of theory, i.e.,  $E_X - E_{\text{SCF}}$ . Various CCSD(T)-F12 variants have been extensively benchmarked on molecular energies and properties<sup>21,23,24</sup> and the accuracy of results is quite impressive. Approximate CCSD(T)-F12 is frequently — and rightfully — advertised as providing at least quintuple-zeta accuracy at the triple-zeta basis set level and nearly triple-zeta cost.

Weak intermolecular interactions, especially those dominated by dispersion, present one of the most challenging cases for the *ab initio* theory due to the slow convergence of interaction energies with respect to both the one-electron basis set and the treatment of electron correlation. In particular, MP2 tends to overestimate dispersion-dominated interaction energies quite substantially<sup>25</sup> and one needs to go to the CCSD(T) level of theory to obtain interaction energies accurate to a few percent (the convergence with respect to theory level becomes still slower if significant static correlation is involved<sup>26</sup>). In terms of basis set convergence, dispersion energy requires an adequate description of the interelectronic cusp (and thus, sufficient flexibility of the orbital space) not only in the molecular region, but also outside the interacting molecules. In conventional (not explicitly correlated) electronic structure theory, these requirements are typically satisfied, with significant computational cost, by using large correlation-consistent basis sets cc-pVXZ (Ref. 27) augmented with diffuse functions.<sup>28</sup> An often better way to provide long-range orbital space flexibility and saturate the dispersion energy is by including basis functions centered away from the atoms, in the region between the interacting molecules.<sup>29–32</sup> Such midbond functions significantly speed up basis-set convergence and have become a popular choice for accurate studies of weakly interacting systems. It should be noted that including midbond functions is not an obstacle for performing CBS extrapolations and, while extrapolation does partly alleviate the difference between midbondless and midbond-containing bases, the latter remain more accurate.<sup>33</sup> Typically, the same midbond set is added to all basis sets employed in the extrapolation. To our knowledge, CBS extrapolations using bases where the midbond functions vary with the cardinal number X of the atomic basis have not been investigated so far.

The performance of the F12 approach for weak intermolecular interaction energies has been tested quite extensively both for MP2-F12 (Refs. 34 and 35) and approximate CCSD(T)-F12.<sup>36–42</sup> Marchetti and Werner<sup>37</sup> calculated frozen-core (FC) scaled-triples CCSD(T)-F12a interaction energies for the S22 benchmark set of weakly interacting systems<sup>25</sup> using the aug-cc-pVDZ and (for smaller systems) aug-cc-pVTZ orbital bases (the shorthand notation aVXZ $\equiv$ aug-cc-pVXZ and dVXZ $\equiv$ d-aug-cc-pVXZ will be used from now on). By comparison to orbital benchmarks extrapolated from the CCSD(T)/aVTZ' and CCSD(T)/aVQZ' values (the prime in the basis set symbol denotes the omission of diffuse functions on hydrogen atoms) for smaller systems

and the original benchmark values<sup>25</sup> for larger ones, Marchetti and Werner concluded that CCSD(T)-F12a/aVDZ is more accurate than CCSD(T)/aVQZ' and CCSD(T)-F12a/aVTZ is likely better converged than the benchmark itself. Interestingly, the subsequent improvement of the original S22 interaction energies<sup>43,44</sup> slightly improves the agreement with the CCSD(T)-F12a/aVTZ results of Ref. 37 but worsens the agreement with the CCSD(T)-F12a/aVDZ ones for large dimers where the former are not available. Thus, while the CCSD(T)-F12a/aVTZ interaction energies are certainly in the competition for the most accurate S22 values available, the smaller-basis ones might be slightly less accurate than initially thought.

It is certainly encouraging that explicitly correlated CCSD(T) at the double-zeta basis set level can provide more accurate interaction energies than conventional CCSD(T) at the quadruple-zeta level. Whether explicitly correlated CCSD(T) in the largest available bases can provide a more precise CBS limit than conventional CCSD(T) in the largest available bases is, however, another story. Important progress towards answering this question has been made through the benchmark CCSD(T) and CCSD(T)-F12 studies for H<sub>2</sub>O–H<sub>2</sub>O and three other hydrogen-bonded dimers,<sup>38</sup> several weakly interacting complexes involving CO<sub>2</sub>,<sup>39</sup> and two minimum structures of CO<sub>2</sub>–CO<sub>2</sub>.<sup>40</sup> For the hydrogen-bonded complexes of Ref. 38, CBS-limit CCSD(T)/FC interaction energies were obtained by extrapolations from (counterpoise-uncorrected) aVXZ and aVXZ' values for X = Q,5. These values were very nicely recovered by both CCSD(T)-F12a and CCSD(T)-F12b in the cc-pVXZ-F12 basis set sequence<sup>45</sup> (see below), with the cc-pVTZ-F12 result being at least as accurate as the conventional aV5Z one. The extensive studies of dimers involving CO<sub>2</sub> (Refs. 39 and 40) included basis sets as large as aV6Z for CCSD(T) and cc-pVQZ-F12 (for CO<sub>2</sub>–CO<sub>2</sub>,<sup>40</sup> also aV5Z) for CCSD(T)-F12b (though no midbond bases were employed). The results showed that at least a part of the significant improvement afforded by the explicitly correlated approach at the double- and triple-zeta level carried on to the largest basis sets available. However, for some of these difficult dispersion-dominated systems (most notably CO<sub>2</sub>–CO<sub>2</sub>) the interaction energy was not fully converged even with the largest X employed, as indicated by the still substantial differences between the values obtained for X and X – 1.

While one can obviously expect CCSD(T)-F12 to be more accurate than CCSD(T) in the same basis with minor additional computational cost, there are several factors that can limit the accuracy attainable by the explicitly correlated approach. First, the accuracy of the a/b/c approximations to CCSD-F12, and of the non-F12 treatment of the triples contribution (whether scaled or unscaled), might be an issue. Second, the F12 methods have not, to our knowledge, been tested with midbond functions, and it is not clear *a priori* whether the improved description of the interelectronic cusp and the correlation hole provided by the explicitly correlated approach will carry to the region between the interacting molecules that is crucial for dispersion. Third, the density fitting required in a CCSD(T)-F12 calculation limits the choice of a one-electron basis set to those for which adequate auxiliary bases are available. The DF and resolution-of-identity

(RI) auxiliary sets that complement the  $aVXZ$  orbital bases have been developed in Refs. 46–48 for  $X = D, T, Q, 5$ . Similar sets for the  $aV6Z$  basis are not generally available, and such sets for the first- and second-row atoms might have to include at least  $k$  functions (as the orbital bases include up to  $i$  functions) which are available in very few electronic structure codes, none of them able to compute CCSD(T)-F12. A class of correlation-consistent basis sets  $cc-pVXZ$ -F12, specifically optimized for F12 calculations, has been recently developed<sup>45</sup> together with the appropriate RI bases.<sup>49</sup> A similar  $cc-pCVXZ$ -F12 sequence containing additional functions optimized for the core-core and core-valence correlation (in the spirit of the  $aug-cc-pCVXZ \equiv aCVXZ$  and  $aug-cc-pwCVXZ \equiv awCVXZ$  basis sets<sup>50</sup>) has also been created.<sup>51</sup> We will test both of these sequences in the present work, denoting them by  $VXZ$ -F12 and  $CVXZ$ -F12 throughout the rest of the text. However, these sequences terminate at  $X = Q$  so it is not clear whether converged interaction energies can be obtained in this way. In any case, larger orbital basis sets, including correlation-consistent sextuple-zeta (sometimes septuple-zeta) bases as well as doubly and triply augmented sets, are presently available for conventional CCSD(T) calculations than for explicitly correlated ones.

In this article, the central question we attempt to answer is whether the approximate CCSD(T)-F12 approach in the largest available bases is presently superior to conventional CCSD(T) in the largest available bases for dispersion-dominated weak interaction energies. The particular case considered will be the rare gas dimers  $He_2$ ,  $Ne_2$ ,  $Ar_2$ , and  $Kr_2$  at their near van der Waals minimum separations. Rare gas dimers are ultimate dispersion-bound systems, even more so than the  $(CO_2)_2$  complex studied in Ref. 40. For this reason, and because of an unusually large triples contribution to the interaction energy, such systems present a particularly difficult case for CCSD(T)-F12. Fortunately, the CBS-limit CCSD(T) interaction energies are known to a very good accuracy for the helium,<sup>5,52,53</sup> neon,<sup>54</sup> and argon dimers.<sup>55,56</sup> These benchmark interaction energies, obtained using a combination of conventional and explicitly correlated approaches ( $He_2$ ) or by CBS extrapolations from very large orbital bases including extensive midbond sets ( $Ne_2$  and  $Ar_2$ ), will enable us to unambiguously determine the accuracy of various conventional and explicitly correlated CCSD(T) results. For  $Kr_2$ , both CCSD(T)/FC and CCSD(T)-F12/FC interaction energies obtained in the present work will (slightly) surpass the accuracy of the previous best available result<sup>57</sup> and we will be able to see which approach provides the CCSD(T)/FC CBS limit to a better precision. We will also examine how the quality of conventional and explicitly correlated CCSD(T) interaction energies varies with the interatomic separation  $R$  and what factors limit the observed CCSD(T)-F12 accuracy. Finally, we will report an unexpectedly failed attempt to accurately recover the core-core and core-valence correlation contributions to the  $Ne_2$  and  $Ar_2$  interaction energies using explicitly correlated CCSD(T). The helium and neon dimers have been studied before using MP2-F12 (Ref. 34) and some variants of CCSD(T)-F12.<sup>35,36</sup> The latter calculations involved the  $aVXZ$  family of bases up to  $X = 6$  ( $He_2$ ) and  $X = 5$  ( $Ne_2$ ) and included the counterpoise (CP) correction for basis set super-

position error (BSSE).<sup>58,59</sup> The studies of Refs. 35 and 36 did find that the CCSD(T)-F12 interaction energies converge substantially faster than the conventional CCSD(T) ones, however, as will be shown below, the improvement brought about by the F12 approach is smaller than the one provided by inclusion of midbond functions and/or a second augmentation of the atomic basis set — see Secs. III A and III B for further discussion.

The structure of the rest of this article is as follows. In Sec. II we will present all the computational details including the fine points of the explicitly correlated calculations. The numerical results are presented and discussed in Sec. III. Finally, Sec. IV contains conclusions.

## II. DETAILS OF THE COMPUTATIONAL PROCEDURE

All numerical calculations have been performed using the MOLPRO2010.1 program.<sup>16</sup> Most of the interaction energies have been obtained for the near van der Waals minima distances:  $R = 5.6$  bohr for  $He_2$ ,<sup>52</sup>  $R = 3.1$  Å for  $Ne_2$ ,<sup>54</sup>  $R = 3.75$  Å for  $Ar_2$ ,<sup>56</sup> and  $R = 4.06$  Å for  $Kr_2$ .<sup>57</sup> The atomic basis set families included singly and doubly augmented  $cc-pVXZ$  (Ref. 27) and  $cc-pV(X+d)Z$  (Ref. 60) as well as  $VXZ$ -F12 (Ref. 45) in FC calculations and, additionally, the  $aCVXZ$ ,  $awCVXZ$ ,<sup>50</sup> and  $CVXZ$ -F12 (Ref. 51) sequences in all-electron (AE) calculations.

Unless otherwise specified, the calculations employed default,  $aVXZ$ /MP2FIT and  $awCVXZ$ /MP2FIT (Refs. 46 and 48) auxiliary bases in the DF of the MP2-F12 part (MOLPRO keyword `DF_BASIS`). The calculations in AE-optimized  $aCVXZ$ ,  $awCVXZ$ , and  $CVXZ$ -F12 orbital sets employed the  $awCVXZ$ /MP2FIT auxiliary bases and all others used  $aVXZ$ /MP2FIT. The cardinal number  $X$  was the same as for the orbital basis except for the  $aV6Z$  orbital sets for which  $aV6Z$ /MP2FIT is not available and  $aV5Z$ /MP2FIT was used instead. It will be shown in Sec. III B that the default  $cc-pVXZ$ /JKFIT bases<sup>47</sup> are not adequate for the RI approximation, i.e., as a complementary auxiliary basis set (CABS).<sup>61</sup> While the default  $cc-pVXZ$ /JKFIT sets are more appropriate for fitting the Fock matrix used in MP2-F12 and in the CABS singles expression (keyword `DF_BASIS_EXCH`), larger auxiliary bases are helpful in that case as well. In this work, the  $aV6Z$ -RI auxiliary basis set developed in Ref. 62, with the  $k$  functions removed because of MOLPRO limitations, was used for both RI and density-fitted Fock matrices for all orbital bases. This set will be denoted  $aV6Z$ -RI\* throughout the rest of this work. It should be noted that the auxiliary sets described above cannot be fully adequate in case of the  $aV6Z$  orbital bases as they do not include large enough angular momenta. Nevertheless, it will be shown below that the auxiliary bases employed here lead to very reasonable  $aV6Z$  results in all cases.

While several midbond bases were used in conventional CCSD(T) calculations, for CCSD(T)-F12 we had to choose midbond sets for which suitable auxiliary bases are available. We used two such sets in the present work, one that does not change with  $X$  and one that varies along with the atomic basis set cardinal number. The constant set had a  $(3s3p2d2f)$  composition with exponents equal to

(0.9,0.3,0.1) for  $sp$  and (0.6,0.2) for  $df$  functions. The corresponding auxiliary basis was the same in all DF and RI contexts and consisted of exponents (1.8,1.2,0.6,0.4,0.2) for  $spd$ , (1.5,0.9,0.5,0.3) for  $f$ , and (1.5,0.9,0.3) for  $g$  functions. This combination of orbital and auxiliary bases has been used quite extensively<sup>63,64</sup> in density-fitted symmetry-adapted perturbation theory with density functional theory description of monomers [SAPT(DFT)] (Refs. 65 and 66) calculations. The midbond set that varied with  $X$  was simply the hydrogen-atom basis set selected from the same orbital and auxiliary bases as employed for the atoms. We will adopt the notation (atomic basis)+(midbond). For example, the midbond part of the aVTZ+(aVTZ) basis involves a hydrogenic aVTZ orbital set, hydrogenic aVTZ/MP2FIT and aV6Z-RI\* sets for DF, and the hydrogenic aV6Z-RI\* set for RI.

The default, diagonal fixed-amplitude Ansatz 3C(FIX) (Ref. 67) has been used throughout this work. Aside from its simplicity, its important virtues for interaction energy calculations are its size consistency and the absence of geminal basis set superposition error.<sup>34,68</sup> On the other hand, this Ansatz is known to be relatively sensitive<sup>69,70</sup> to the choice of the exponent  $\beta$  in the expression  $F(r_{12}) = -\frac{1}{\beta} \exp(-\beta r_{12})$  for the F12 correlation factor. Therefore, throughout this work we employed the values of  $\beta$  recommended previously:<sup>45</sup>  $1.0 a_0^{-1}$  and  $1.1 a_0^{-1}$  for VTZ-F12 and VQZ-F12, respectively, and 1.1, 1.2, 1.4, and  $1.4 a_0^{-1}$  for aVDZ, aVTZ, aVQZ, and aV5Z, respectively (for the aV6Z basis,  $\beta$  was set to  $1.4 a_0^{-1}$  as well). It should be noted that slightly different values of  $\beta$  have been recommended recently specifically for the 3C(FIX) Ansatz,<sup>70</sup> however, the differences between the recommended values do not exceed  $0.1 a_0^{-1}$  and we chose to stick to the ones of Ref. 45 which were employed in several other investigations of weakly interacting systems. For the AE-optimized basis sets that will be used for calculating the core-core and core-valence contribution to the interaction energy, the optimal values of  $\beta$  were established by Hill *et al.*<sup>51</sup> who recommended to use them in both FC and AE calculations. These optimal  $\beta$  amount to  $1.4 a_0^{-1}$  and  $1.5 a_0^{-1}$  for CVTZ-F12 and CVQZ-F12 and 1.5, 1.6, and  $1.5 a_0^{-1}$  for aCVTZ, aCVQZ, and aCV5Z, respectively. It has recently become possible to perform CCSD(T)-F12 calculations with different geminal exponents  $\beta$  for the core-core, core-valence, and valence-valence pairs.<sup>71</sup> We have tested this approach for Ar<sub>2</sub>, choosing exponents that minimize the core-core, core-valence, and valence-valence contributions to the MP2-F12 correlation energy of the dimer. However, the resulting core corrections to the interaction energy (not shown) did not display any consistent improvement with respect to the results obtained with a single value of  $\beta$ .

We tested all three flavors of the CCSD(T)-F12 approximation available in MOLPRO: the CCSD(T)-F12a and CCSD(T)-F12b theories of Ref. 22 and the CCSD(F12\*) method of Ref. 23, denoted CCSD-F12c in MOLPRO and in this paper. We investigated the performance of these approximate CCSD(T)-F12 approaches both with and without the scaling of the triples contribution according to Eq. (1). If triples were scaled, the scaling factor determined for the dimer was also used for the monomers to ensure size consistency.<sup>37</sup>

The counterpoise correction for BSSE (Refs. 58 and 59) has been applied in all calculations. Some authors argue that CP-uncorrected wave function methods can deliver comparable accuracy even for rare gas dimers.<sup>72</sup> Indeed, the CCSD(T)/FC and CCSD(T)-F12b/FC interaction energies in the aVXZ bases are actually more accurate without the CP correction than with it for all four dimers studied, cf. Tables SI–SII and Figs. S1–S4 in the supplementary material.<sup>73</sup> However, it is well known that an addition of extra diffuse functions to the basis set increases the BSSE and can decrease the accuracy of CP-uncorrected interaction energies<sup>74</sup> (this observation has led to a family of “calendar” sets where diffusion functions are progressively removed from the basis set to make CP-uncorrected calculations more economical while preserving accuracy<sup>75,76</sup>). Such behavior is clearly seen in Tables SI–SII and Figs. S1–S4 (supplementary material<sup>73</sup>) where the addition of a second augmentation, while substantially improving the CP-corrected interaction energies, worsens the CP-uncorrected results dramatically. Even more importantly, midbond bases are incompatible with CP-uncorrected calculations. For midbond-containing bases, CP-uncorrected interaction energies are wildly inaccurate regardless of whether the basis set used for a monomer contains this monomer-centered functions only or this monomer-centered functions and midbond functions (one could think of the second alternative as a “half-CP-corrected” scheme). As our calculations in Sec. III show that the addition of midbond functions is the single most effective way of improving the accuracy of both conventional and explicitly correlated results (this fact is clearly seen in Figs. S1–S4),<sup>73</sup> employing the CP correction is the only reasonable choice.

The finite-basis interaction energies were extrapolated to the CBS limit using the most popular  $X^{-3}$  scheme. Specifically, the correlation part of interaction energy calculated in a basis set with the cardinal number  $X$ ,  $E_{\text{corr}}(X)$ , was assumed to approach its CBS limit  $E_{\text{corr}}(\infty)$  as

$$E_{\text{corr}}(X) = E_{\text{corr}}(\infty) + AX^{-3}, \quad (2)$$

where  $A$  is some constant. This assumption is supported by rigorous analytical results for atomic MP2 correlation energies<sup>77–79</sup> as well as by a multitude of numerical studies pioneered by Helgaker and collaborators.<sup>80,81</sup> While the explicitly correlated correlation energies formally exhibit much faster ( $l_{\text{max}}^{-7}$  instead of  $l_{\text{max}}^{-3}$ ) convergence with the maximum angular momentum  $l_{\text{max}}$  in the atomic partial wave expansion,<sup>7,82</sup> the  $X^{-7}$  convergence is virtually impossible to attain in practice due to the incompleteness of the orbital basis set for each  $l$ , auxiliary basis set incompleteness, and the non-F12 treatment of triples. As the SCF part of the interaction energy converges to its CBS limit in a much faster (exponential) manner, this part was not extrapolated: instead, the value computed in a larger of the two bases was adopted as the CBS estimate (in F12 calculations, the SCF part of the interaction energy included the CABS singles correction<sup>22</sup>). We will use the  $(X-1, X)$  notation for a CBS extrapolation from bases with cardinal numbers  $X-1$  and  $X$ . For example, the CCSD(T) interaction energy (T,Q)-extrapolated from the aVXZ family is the sum of the SCF part computed in the

aVQZ basis and the correlation part extrapolated from the aVTZ and aVQZ results using the  $X^{-3}$  scheme. Various alternative extrapolation schemes have been proposed.<sup>33,83,84</sup> In particular, Schwenke<sup>84</sup> has advocated the use of different correlation factors  $F_{X_1, X_2}^c$  for CBS extrapolations of different energy contributions  $E^c$  [where  $c$  can be, in our case, SCF, the correlation part of CCSD, and/or (T)] according to the formula

$$E^c(\infty) = (E^c(X_2) - E^c(X_1))F_{X_1, X_2}^c + E^c(X_1). \quad (3)$$

The correlation factors  $F_{X_1, X_2}^c$  are basis-set dependent: their values for the aVXZ sequence have been optimized for both conventional CCSD(T) [through  $(X_1, X_2) = (5, 6)$ ] (Ref. 84) and CCSD(T)-F12b [through  $(X_1, X_2) = (Q, 5)$ ].<sup>70</sup> An optimization of similar correlation factors for other basis sets (in particular, for the ones including midbond functions) is outside the scope of this paper. For the aVXZ family, our tests do not show any systematic improvement of the Schwenke-style extrapolated energies (using preoptimized correlation factors from Refs. 70 and 84) over the  $X^{-3}$ -extrapolated ones at either conventional CCSD(T) or CCSD(T)-F12b so the  $X^{-3}$  extrapolation will be the only one considered further.

### III. NUMERICAL RESULTS AND DISCUSSION

#### A. He<sub>2</sub>

The most accurate estimates of the CBS-limit CCSD(T) interaction energy for the helium dimer come from a combination of orbital and explicitly correlated treatments.<sup>52</sup> At the near-minimum interatomic distance  $R = 5.6$  bohr, the CCSD(T) limit obtained in Ref. 52 was  $-10.6853 \pm 0.0013$  K (the conversion factor 1 hartree = 315 774.65 K is employed in the present work). This limit was computed as a sum of the CCSD interaction energy calculated<sup>33,85</sup> in a fully optimized explicitly correlated Gaussian-type geminal basis<sup>86,87</sup> and of the triples contribution extrapolated from correlation consistent orbital basis sets up to dV7Z (Refs. 88–90) supplemented by an extensive (6s6p6d3f3g3h) set of midbond functions<sup>91</sup> (denoted b135 in Ref. 52). It should be noted that the exact (FCI) nonrelativistic Born-Oppenheimer interaction energy for He<sub>2</sub> at  $R = 5.6$  bohr is known to a better precision ( $E_{\text{FCI}} = -11.0006 \pm 0.0002$  K) than the CCSD(T) limit thanks to the four-electron explicitly correlated variational calculations of Ref. 5. This result can be combined with the value of the beyond-CCSD(T) term  $\delta E_{\text{FCI}} = E_{\text{FCI}} - E_{\text{CCSD(T)}}$  extrapolated from large-basis orbital calculations. The  $\delta E_{\text{FCI}}$  estimate obtained in this way amounts to  $-0.3183 \pm 0.0028$  K using bases up to dV6Z and aV6Z+(6s6p6d3f1g1h) midbond,<sup>52</sup> or  $-0.31625 \pm 0.00095$  K using also bases dV7Z and aV7Z+(6s6p6d3f1g1h).<sup>53</sup> Using the latter results and adding the uncertainties quadratically, one arrives at the estimate  $E_{\text{CCSD(T)}} = -10.6844 \pm 0.0010$  K. This last estimate is somewhat more precise than the one of Ref. 52 and we adopt it as the benchmark value of the CCSD(T) interaction energy.

The CCSD(T) interaction energies for He<sub>2</sub> at  $R = 5.6$  bohr computed and CBS-extrapolated using various basis sets are listed in Tables I (conventional results) and

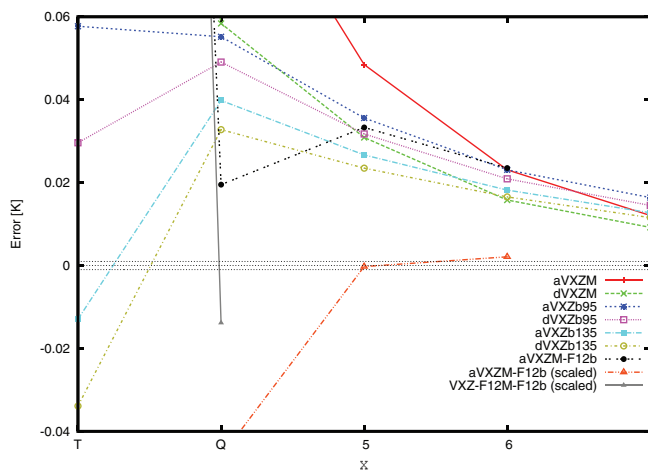


FIG. 1. Absolute errors (in Kelvin) for CP-corrected CCSD(T) interaction energies for the helium dimer at  $R = 5.6$  bohr as a function of the cardinal number  $X$  of the basis set. Both conventional CCSD(T) and CCSD(T)-F12b results are presented. The letter M in the basis set symbol stands for the hydrogenic set of midbond functions from the same basis, and b95 and b135 stand for the (6s6p6d3f1g1h) and (6s6p6d3f3g3h) midbond sets, respectively. No other CCSD(T)-F12b result (in particular, no result without midbond) fits in the range shown.

II (explicitly correlated results). The corresponding errors with respect to benchmark are plotted in Fig. 1 for conventional CCSD(T) as well as the CCSD(T)-F12b explicitly correlated variant with and without the scaling of triples. We employed singly- and doubly-augmented Dunning-type cc-pVXZ basis sets with or without midbond functions. In non-F12 calculations we employed both constant midbond sets (6s6p6d3f1g1h) and (6s6p6d3f3g3h) and hydrogenic midbond sets that change in concert with the atomic part of the basis like in aVXZ+(aVXZ). As an aV7Z basis is not available for hydrogen, we used the aug-mcc-pV7Z set of Ref. 92 instead and added a second augmentation to it (to form a d-aug-mcc-pV7Z set) in an even-tempered manner. Because of the limited choice of auxiliary bases, the explicitly correlated calculations included only hydrogenic midbond sets that match the orbital basis, e.g., aVTZ+(aVTZ).

The first thing that is evident from Tables I–II and Fig. 1 is how crucial midbond functions are for saturating a dispersion-dominated interaction energy. A midbondless basis set as large as aV7Z is only slightly more accurate than aVTZ+(aVTZ) and even dV7Z is no match for aV5Z+(aV5Z). Moreover, while the CCSD(T)-F12b results are obviously better than conventional CCSD(T) in the same basis set, the improvement is not nearly as dramatic as observed for most other systems: for  $X = Q, 5, 6$ , CCSD(T)-F12b/aVXZ is roughly as accurate as CCSD(T)/aV(X+1)Z. Interestingly, CCSD(T)-F12b in the aV5Z basis is less accurate than conventional CCSD(T) in the smallest midbond-containing basis presented here, aVTZ+(aVTZ), for which the calculation takes only about 6% of the time needed for CCSD(T)-F12b/aV5Z. While the scaling of the triples contribution to CCSD(T)-F12b as described in Sec. II does bring about a minor improvement in accuracy, explicitly correlated calculations without midbond remain inferior to conventional calculations with midbond. Our results for the aVXZ

TABLE I. CP-corrected CCSD(T) interaction energies for the helium dimer at  $R = 5.6$  bohr calculated and extrapolated using various orbital bases. The rows marked “ext.” contain the CBS-extrapolated results, with the value in the “X” column obtained from the energies calculated in XZ and  $(X - 1)Z$  basis sets. The energy unit is 1 K. The benchmark CCSD(T)/CBS value of the interaction energy is  $-10.6844 \pm 0.0010$  K (see text).

Basis set	X = T	X = Q	X = 5	X = 6	X = 7
aVXZ	-8.3015	-9.2577	-9.8276	-10.2607	-10.4355
ext.		-9.9919	-10.4368	-10.8448	-10.7291
dVXZ	-8.9225	-10.0028	-10.4057	-10.5781	-10.6176
ext.		-10.8353	-10.8149	-10.8021	-10.6846
aVXZ+(aVXZ)	-10.3133	-10.5802	-10.6359	-10.6613	-10.6723
ext.		-10.7528	-10.6882	-10.6929	-10.6909
dVXZ+(dVXZ)	-10.3925	-10.6260	-10.6535	-10.6686	-10.6752
ext.		-10.7768	-10.6915	-10.6878	-10.6866
aVXZ+(6s6p6d3f1g1h)	-10.6267	-10.6292	-10.6488	-10.6613	-10.6680
ext.		-10.6339	-10.6656	-10.6757	-10.6792
dVXZ+(6s6p6d3f1g1h)	-10.6548	-10.6353	-10.6527	-10.6635	-10.6699
ext.		-10.6408	-10.6685	-10.6756	-10.6803
aVXZ+(6s6p6d3f3g3h)	-10.6974	-10.6446	-10.6577	-10.6662	-10.6717
ext.		-10.5961	-10.6672	-10.6751	-10.6805
dVXZ+(6s6p6d3f3g3h)	-10.7183	-10.6516	-10.6609	-10.6679	-10.6728
ext.		-10.6066	-10.6695	-10.6764	-10.6808

sequence, Table II, can be compared to the ones of Tew *et al.*<sup>36</sup> using two, less approximate, variants of CCSD(T)(F12) and very large CABS in the RI approximation (there was no density fitting in the MP2-F12 component in Ref. 36, and the

interatomic distance was 5.62 bohr instead of 5.6 bohr). The results of Tew *et al.* are in fairly good agreement with the ones in Table II — the differences are of the same order as those between the a/b/c variants of CCSD(T)-F12. Still, those

TABLE II. CP-corrected CCSD(T)-F12 interaction energies for the helium dimer at  $R = 5.6$  bohr calculated and extrapolated using various bases. The rows marked “ext.” contain the CBS-extrapolated results, with the value in the “X” column obtained from the energies calculated in XZ and  $(X - 1)Z$  basis sets. The energy unit is 1 K. The benchmark CCSD(T)/CBS value of the interaction energy is  $-10.6844 \pm 0.0010$  K (see text).

Basis set	Unscaled triples				Scaled triples			
	X = T	X = Q	X = 5	X = 6	X = T	X = Q	X = 5	X = 6
CCSD(T)-F12a results								
aVXZ	-9.5956	-9.7590	-10.2185	-10.4873	-9.7283	-9.8208	-10.2512	-10.5089
ext.		-9.8825	-10.7006	-10.8532		-9.8925	-10.7027	-10.8597
aVXZ+(aVXZ)	-9.9557	-10.7027	-10.6701	-10.6720	-10.1032	-10.7683	-10.7037	-10.6934
ext.		-11.2375	-10.6335	-10.6733		-11.2434	-10.6334	-10.6780
VXZ-F12	-8.2742	-9.1815			-8.4019	-9.2945		
ext.		-9.8468				-9.9491		
VXZ-F12+(VXZ-F12)	-9.5694	-10.6558			-9.7480	-10.7865		
ext.		-11.4473				-11.5430		
CCSD(T)-F12b results								
aVXZ	-9.4578	-9.6990	-10.1808	-10.4603	-9.5905	-9.7608	-10.2135	-10.4819
ext.		-9.8792	-10.6863	-10.8410		-9.8893	-10.6884	-10.8475
aVXZ+(aVXZ)	-9.8622	-10.6649	-10.6511	-10.6609	-10.0097	-10.7305	-10.6847	-10.6823
ext.		-11.2404	-10.6342	-10.6731		-11.2463	-10.6341	-10.6778
VXZ-F12	-8.1404	-9.0771			-8.2681	-9.1902		
ext.		-9.7639				-9.8662		
VXZ-F12+(VXZ-F12)	-9.4521	-10.5676			-9.6307	-10.6982		
ext.		-11.3801				-11.4759		
CCSD(T)-F12c results								
aVXZ	-9.4196	-9.6776	-10.1682	-10.4528	-9.5519	-9.7394	-10.2009	-10.4744
ext.		-9.8702	-10.6829	-10.8406		-9.8804	-10.6850	-10.8470
aVXZ+(aVXZ)	-9.8170	-10.6422	-10.6388	-10.6529	-9.9638	-10.7076	-10.6723	-10.6743
ext.		-11.2341	-10.6329	-10.6711		-11.2402	-10.6329	-10.6758
VXZ-F12	-7.9486	-8.9951			-8.0759	-9.1077		
ext.		-9.7619				-9.8637		
VXZ-F12+(VXZ-F12)	-9.3993	-10.5171			-9.5770	-10.6472		
ext.		-11.3314				-11.4268		

differences are substantial (up to 0.1 K at the aV6Z level) which shows that the accuracy of the *a/b/c* approximations may be a major limiting factor for the accuracy of CCSD(T)-F12 interaction energies.

Explicitly correlated calculations employing bases with midbond have attracted little attention so far. This might be somewhat surprising since the explicitly correlated approach and the use of midbond bases are designed to address two distinct effects that are hard to saturate — the short-range dynamical correlation (the interelectronic cusp conditions) and the long-range dynamical correlation (dispersion), respectively. One could expect that it should be possible to combine the strengths of both approaches without any adverse interference. As the results in Table II and Fig. 1 show, this is indeed the case. The CCSD(T)-F12b results in bases aVXZ+(aVXZ),  $X = Q, 5, 6$ , and VQZ-F12+(VQZ-F12) match or exceed the accuracy of orbital results including midbond functions at the same  $X$ . As visible in Fig. 1, when the triples correction is scaled, the accuracy of the CCSD(T)-F12b/aV5Z+(aV5Z) approach becomes amazing — it provides the only computed (nonextrapolated) result that is within the tight error bars of the benchmark (no orbital result is that accurate). The scaled-triples CCSD(T)-F12b/aV6Z+(aV6Z) result comes a close second but is outside the error bars of the benchmark. As all the midbond sets utilized here are much larger than typically used, it is apparent how hard it is to saturate the dispersion-dominated interaction energy even using midbond functions. The (6s6p6d3f3g3h) set is clearly superior to (6s6p6d3f1g1h): interestingly, additional midbond functions prove to be more important than additional diffuse functions as indicated by the results in the aVXZ+(6s6p6d3f3g3h) bases being more accurate than those in the (larger) dVXZ+(6s6p6d3f1g1h) bases. As expected, the aVXZ+(aVXZ) and dVXZ+(dVXZ) results become more accurate than the ones involving constant midbond for sufficiently large  $X$ . Still, even the largest, 714-term dV7Z+(dV7Z) set gives a CCSD(T) energy that is 0.0092 K off the benchmark value and CBS extrapolation is necessary to improve the accuracy further.

The standard  $X^{-3}$  CBS extrapolation clearly improves the agreement of orbital results with the benchmark. This fact is best illustrated in Fig. 2 where all different midbond-containing bases are compared in CBS extrapolations. Interestingly, the extrapolated constant-midbond and variable-midbond results approach the benchmark from two different sides, in accordance with the different behavior of the computed results as a function of  $X$ . In fact, the undershooting of the constant-midbond extrapolations and the overshooting of the variable-midbond ones means that the optimal exponent  $\alpha$  in an  $X^{-\alpha}$  fit to the computed results is less than 3 (i.e., slower convergence) for constant midbond and more than 3 for variable midbond. It should be noted that in the former case [precisely, for the aVXZ+(6s6p6d3f1g1h) family] an analogous conclusion has been reached by an extensive analysis of extrapolation patterns for the correlation part of the CCSD interaction energy and for its dominant dispersion contribution, with the optimal exponent  $\alpha$  equal to about 2.3 for the (6,7) CCSD extrapolation.<sup>93</sup> Regardless of the overshooting/undershooting, all midbond-containing bases in the (6,7) extrapolation give orbital results that are within 0.007 K

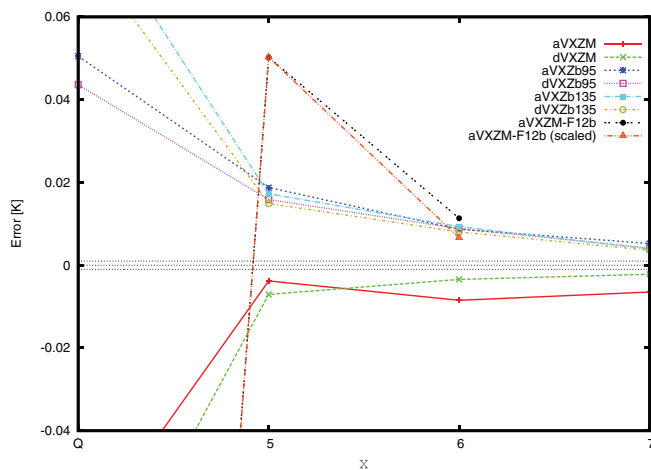


FIG. 2. Absolute errors (in Kelvin) for CP-corrected CCSD(T) interaction energies for the helium dimer at  $R = 5.6$  bohr, extrapolated to the CBS limit using values computed in  $(X - 1)Z$  and  $XZ$  basis sets, as a function of the cardinal number  $X$ . Both conventional CCSD(T) and CCSD(T)-F12b results are presented. The letter M in the basis set symbol stands for the hydrogenic set of midbond functions from the same basis, and b95 and b135 stand for the (6s6p6d3f1g1h) and (6s6p6d3f3g3h) midbond sets, respectively. The midbondless CCSD(T)-F12b results extrapolated from bases aVQZ and aV5Z, unlike the results extrapolated from aV5Z and aV6Z, would accidentally fit in the range shown. All other CCSD(T)-F12b results extrapolated from midbondless bases and all CCSD(T)-F12b results extrapolated from the VXZ-F12 bases (with or without midbond) are outside the range.

of the benchmark, with the most accurate result, 0.0022 K below the benchmark, provided by the largest dVXZ+(dVXZ) sequence.

The (T,Q) extrapolated CCSD(T)-F12b results are fairly inaccurate. This includes the VXZ-F12+(VXZ-F12) value where the  $X = Q$  result is much more accurate than the  $X = T$  one so that any extrapolation is deemed to overshoot. At the (Q,5) level, the extrapolated CCSD(T)-F12b energies are much more accurate — the errors are  $-0.002$  K and  $0.050$  K for the aVXZ and aVXZ+(aVXZ) sequences, respectively. The first result is clearly fortuitous and worsens substantially (to  $-0.157$  K) at the (5,6) level where the aVXZ+(aVXZ) error is reduced to 0.011 K. The latter accuracy is not matched by any orbital results in bases without midbond except for the (6,7) extrapolation from the dVXZ sequence. However, extrapolations from orbital results with midbond surpass this accuracy already at the (Q,5) (variable midbond) and (5,6) (constant midbond) level.

An interesting observation based on the results in Table II is that the differences between unscaled-triples and scaled-triples results are greatly diminished upon CBS extrapolation [at least for the aVXZ and aVXZ+(aVXZ) families]. Extrapolation also reduces the dependence of the explicitly correlated results on the particular version of the CCSD-F12 approximation (a, b, or c).

Based on the numerical data in Table II and Figs. 1 and 2, the VXZ-F12 basis set family cannot be recommended for such a dispersion-dominated system even when supplemented by midbond functions. While, in the latter case, occasional accurate interaction energies are obtained, there is no consistency with respect to the scaling or not scaling of triples and to the *a/b/c* variant of the CCSD-F12 approximation.



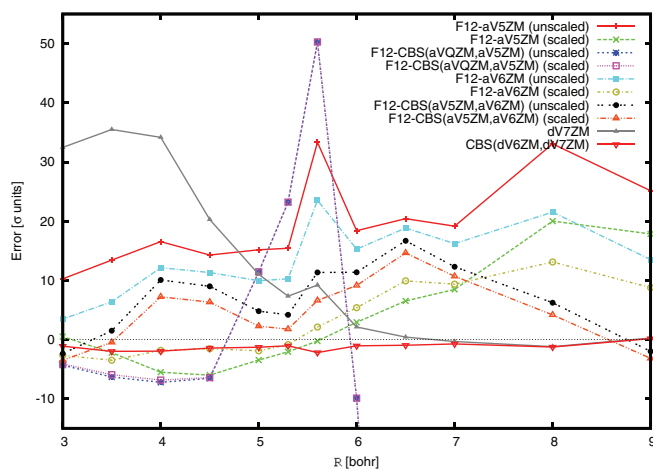


FIG. 3. Accuracy of the He-He potential energy curve calculated and extrapolated at the CCSD(T)-F12b/aVXZ+(aVXZ) level of theory. The numbers shown are differences between the calculated result and the benchmark divided by the uncertainty of the benchmark for a given  $R$ : that is, the benchmark value corresponds to  $0 \pm 1$  at each  $R$ . The best available conventional CCSD(T) results, computed in the dV7Z+(dV7Z) basis and CBS-extrapolated using that basis and the dV6Z+(dV6Z) one, are also included for comparison. The benchmark CCSD(T)/CBS interaction energies have been taken from Table 9 of Ref. 52 except for the value at  $R = 5.6$  bohr which was obtained as described in the text. Note that the benchmark uncertainty is smaller at  $R = 5.6$  bohr than at the surrounding distances: thus the local maxima of the relative errors at  $R = 5.6$  bohr. The letter M in the basis set symbol stands for the hydrogenic set of midbond functions from the same basis.

Moreover, the accurate computed results are badly worsened by the CBS extrapolation. The only basis set family for which consistently accurate CCSD(T)-F12 results are obtained is the aVXZ+(aVXZ) one at  $X = 5, 6$ , and the presence of midbond functions is nearly as beneficial as for orbital calculations. As mentioned before, the scaled-triples CCSD(T)-F12b/aV5Z+(aV5Z) energy is the only result, either computed or extrapolated, that actually lies within the error bars of the benchmark. Thus, it is interesting to see whether the very good performance of the CCSD(T)-F12b/aV5Z+(aV5Z) approach at  $R = 5.6$  bohr is consistent for the whole potential energy curve. As the results plotted in Fig. 3 show, this is unfortunately not the case. The accuracy of the scaled-triples result at  $R = 5.6$  bohr is clearly accidental as for other distances the same method gives results that are off the benchmark by up to 20 times the uncertainty. The unscaled-triples results as well as the extrapolated energies are even worse in terms of the quality of the whole potential energy curve. Going from  $X = 5$  to  $X = 6$  improves both the computed results (slightly) and the extrapolated ones (very significantly at large  $R$ ). Still, any explicitly correlated approach, with or without CBS extrapolation, leads to interaction energies that are off the benchmark by more than 10 times its uncertainty for some values of  $R$ . Similar conclusions hold for the other variants of the explicitly correlated approach, CCSD(T)-F12a and CCSD(T)-F12c. It is interesting to note that while the (Q,5) extrapolation of the CCSD(T)-F12b/aVXZ+(aVXZ) results clearly overshoots at large  $R$  (due to the  $X = Q$  results being highly inaccurate), an analogous (5,6) extrapolation actually undershoots. Thus, the optimal exponent  $\alpha$  for the  $X^{-\alpha}$  extrapolation is less than

3 at the (5,6) level, a surprising result in view of the formal  $l_{\max}^{-7}$  convergence of the F12 correlation energies.<sup>7,82</sup>

Even though CCSD(T)-F12 cannot match the accuracy of the fully optimized Gaussian geminal benchmark, it does surpass the accuracy of any conventional CCSD(T) result for the repulsive part of the potential energy curve. This is clearly visible in Fig. 3 where the largest-basis orbital results, the dV7Z+(dV7Z) ones, are plotted for comparison. While these results are much better than the CCSD(T)-F12 ones in the asymptotic region and deliver more or less comparable accuracy around the van der Waals minimum, for smaller  $R$  the orbital results are inferior to any flavor of the CCSD(T)-F12 method, with and without the scaling of the triples term and/or CBS extrapolation. This further substantiates our hypothesis that while CCSD(T)-F12 does an excellent job of describing the interelectronic cusp conditions and thus the short-range dynamical correlation, it is not nearly as successful in recovering long-range dynamical correlation, that is, dispersion energy at large  $R$ . Finally, it should be noted that the dVXZ+(dVXZ) orbital results benefit immensely from the  $X^{-3}$  extrapolation — the (6,7) extrapolated results are the most accurate ones throughout the whole range of  $R$ . As the  $X^{-3}$  extrapolation did not perform that well in the extensive study involving constant-midbond bases,<sup>93</sup> this might suggest that enlarging the midbond set in line with the atom-centered part of the basis is crucial for maintaining the  $X^{-3}$  decay of correlation energy with the cardinal number  $X$ .

## B. Ne<sub>2</sub>

The most accurate calculations of the neon-neon interaction energy to date are those of Hellmann *et al.*<sup>54</sup> and include coupled-cluster contributions all the way to the singles, doubles, triples, and connected quadruples [CCSDT(Q)] level<sup>94,95</sup> as well as corrections for core-core and core-valence correlation and relativistic effects. At the frozen-core CCSD(T) level, the CBS limit of  $-41.375$  K at the near-minimum interatomic distance  $R = 3.1$  Å was established by CP-corrected calculations in bases t-aug-cc-pV5Z+(4s4p3d3f2g) and t-aug-cc-pV6Z+(4s4p3d3f2g) followed by an  $X^{-3}$  extrapolation of the correlation energy. The CCSD(T) correction for core-core and core-valence correlation was computed using the dwCV5Z basis set<sup>50</sup> and amounts to  $+0.068$  K at the same  $R$ . The benchmark uncertainty has not been given in Ref. 54 but the difference between the extrapolated CCSD(T)/FC interaction energy and the result computed in the largest, t-aug-cc-pV6Z+(4s4p3d3f2g) basis can be taken as a trustworthy, albeit conservative, uncertainty estimate. In this way, we arrive at the benchmark CCSD(T)/FC value of  $-41.375 \pm 0.013$  K at  $R = 3.1$  Å.

The convergence of computed and extrapolated frozen-core CCSD(T) Ne-Ne interaction energies is presented in Tables III (orbital calculations) and IV (explicitly correlated calculations). The errors of these interaction energies with respect to the benchmark are plotted in Fig. 4. As far as the orbital calculations are concerned, Table III shows that the aVXZ sequence leads to quite inaccurate results — even at  $X = 6$  the error is

TABLE III. CP-corrected CCSD(T)/FC interaction energies for the neon dimer at  $R = 3.1 \text{ \AA}$  calculated and extrapolated using various orbital bases. The rows marked “ext.” contain the CBS-extrapolated results, with the value in the “X” column obtained from the energies calculated in XZ and  $(X - 1)Z$  basis sets. The energy unit is 1 K. The benchmark CCSD(T)/CBS value of the interaction energy is  $-41.375 \pm 0.013 \text{ K}$  (see text).

Basis set	$X = T$	$X = Q$	$X = 5$	$X = 6$
aVXZ	-23.995	-33.032	-37.052	-38.795
ext.		-39.330	-41.773	-41.165
dVXZ	-35.716	-40.199	-40.343	-40.746
ext.		-43.533	-40.680	-41.263
aVXZ+(aVXZ)	-38.704	-40.847	-41.253	-41.338
ext.		-42.462	-41.703	-41.469
dVXZ+(dVXZ)	-41.103	-41.341	-41.341	-41.384
ext.		-41.779	-41.362	-41.446
aVXZ+(6s6p6d3f3g3h)	-42.474	-41.477	-41.342	-41.344
ext.		-40.903	-41.173	-41.347
dVXZ+(6s6p6d3f3g3h)	-42.588	-41.513	-41.375	-41.374
ext.		-40.956	-41.210	-41.375

about 2.6 K or 6% of the interaction energy. A second augmentation helps a lot, but the single most effective way to improve basis set convergence of the interaction energy is to include midbond functions. In fact, the accuracy of the midbondless dV6Z basis is surpassed by the aVXZ+(aVXZ) and

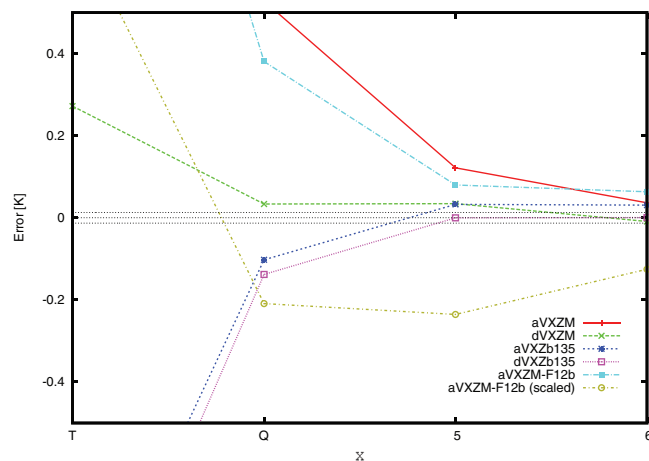


FIG. 4. Absolute errors (in Kelvin) for the CP-corrected CCSD(T)/FC interaction energies for the neon dimer at  $R = 3.1 \text{ \AA}$  as a function of the cardinal number  $X$  of the basis set. Both conventional CCSD(T) and CCSD(T)-F12b results are presented. The letter M in the basis set symbol stands for the hydrogenic set of midbond functions from the same basis and b135 stands for the  $(6s6p6d3f3g3h)$  midbond set. All other CP-corrected CCSD(T)-F12b results (scaled and unscaled triples) are outside the range shown.

dVXZ+(dVXZ) sequences already at the  $X = Q$  and  $X = T$  levels, respectively.

Among the explicitly correlated CCSD(T) interaction energies, the VXZ-F12 ones are grossly inaccurate without

TABLE IV. CP-corrected CCSD(T)-F12/FC interaction energies for the neon dimer at  $R = 3.1 \text{ \AA}$  calculated and extrapolated using various bases. The rows marked “ext.” contain the CBS-extrapolated results, with the value in the “X” column obtained from the energies calculated in XZ and  $(X - 1)Z$  basis sets. The energy unit is 1 K. The benchmark CCSD(T)/CBS value of the interaction energy is  $-41.375 \pm 0.013 \text{ K}$  (see text).

Basis set	Unscaled triples				Scaled triples			
	$X = T$	$X = Q$	$X = 5$	$X = 6$	$X = T$	$X = Q$	$X = 5$	$X = 6$
CCSD(T)-F12a results								
aVXZ	-34.278	-37.055	-39.325	-40.268	-35.166	-37.571	-39.623	-40.452
ext.		-38.998	-41.750	-41.557		-39.243	-41.819	-41.583
aVXZ+(aVXZ)	-40.157	-41.295	-41.455	-41.405	-41.348	-41.885	-41.770	-41.594
ext.		-42.190	-41.649	-41.332		-42.342	-41.677	-41.347
VXZ-F12	-31.568	-33.504			-32.075	-33.853		
ext.		-34.973				-35.207		
VXZ-F12+(VXZ-F12)	-37.164	-40.051			-38.045	-40.532		
ext.		-42.175				-42.364		
CCSD(T)-F12b results								
aVXZ	-34.223	-36.932	-39.221	-40.187	-35.112	-37.448	-39.519	-40.370
ext.		-38.826	-41.666	-41.505		-39.070	-41.735	-41.532
aVXZ+(aVXZ)	-39.453	-40.994	-41.295	-41.312	-40.644	-41.584	-41.610	-41.500
ext.		-42.183	-41.638	-41.329		-42.335	-41.665	-41.344
VXZ-F12	-33.600	-34.428			-34.106	-34.777		
ext.		-35.088				-35.322		
VXZ-F12+(VXZ-F12)	-36.760	-39.823			-37.640	-40.304		
ext.		-42.077				-42.265		
CCSD(T)-F12c results								
aVXZ	-33.037	-36.548	-39.027	-40.072	-33.916	-37.061	-39.324	-40.256
ext.		-39.027	-41.670	-41.501		-39.273	-41.740	-41.528
aVXZ+(aVXZ)	-39.053	-40.755	-41.157	-41.225	-40.227	-41.341	-41.471	-41.413
ext.		-42.062	-41.605	-41.313		-42.219	-41.634	-41.328
VXZ-F12	-29.146	-32.571			-29.652	-32.919		
ext.		-35.126				-35.358		
VXZ-F12+(VXZ-F12)	-36.501	-39.675			-37.372	-40.154		
ext.		-42.010				-42.201		

midbond functions and better, but still not competitive, when midbond functions are added. Apparently, as suggested in the original work,<sup>45</sup> the VXZ-F12 bases are not suitable for dispersion-dominated interactions due to their lack of higher-angular-momentum diffuse functions. The aVXZ family performs better — CCSD(T)-F12b is more accurate than conventional CCSD(T) using the same basis set. However, just like for the helium dimer, much larger improvement is brought about by a simple inclusion of midbond functions than by the explicitly correlated treatment. Midbond functions are also beneficial for the explicitly correlated energies and, as should have been expected, CCSD(T)-F12b is (slightly) superior to CCSD(T) in the same aVXZ+(aVXZ) basis set. However, a larger improvement is achieved by adding a second set of diffuse functions to the atomic and midbond bases. The scaled-triples aVTZ+(aVTZ) result is substantially better than the unscaled one, but for large  $X$  the scaling of triples improves results only slightly or not at all. Our aVXZ results are in the same range as the ones obtained using two different variants of CCSD(T)(F12).<sup>36</sup> They also are in good agreement with the CCSD(T)-F12a and CCSD(T)-F12b values previously reported by Marchetti and Werner,<sup>35</sup> the main reason for the discrepancy being different values of the exponent  $\beta$  in the F12 correlation factor (Sec. II).

As shown in Fig. 4, three different computed results are within the error bars of the benchmark. All of them are conventional CCSD(T): the dVXZ+(6s6p6d3f3g3h) results at  $X = 5, 6$  and the dVXZ+(dVXZ) one for  $X = 6$ . The most accurate explicitly correlated result is, accidentally, the scaled-triples CCSD(T)-F12a/aVTZ+(aVTZ) one with an error of 0.027 K. For larger  $X$ , the most accurate interaction energies seem to be provided by the unscaled-triples CCSD(T)-F12a and scaled-triples CCSD(T)-F12c variants. All six CCSD(T)-F12 approaches lead to aV5Z+(aV5Z) and aV6Z+(aV6Z) interaction energies that are within 0.4 K and 0.22 K of the benchmark, respectively.

As for the helium dimer, CBS extrapolation makes orbital results without midbond more competitive, although still inferior, to the ones with midbond (the same thing happens to the CCSD(T)-F12 results), and the convergence patterns of constant-midbond and variable-midbond sequences are quite different. Interestingly, extrapolation again reduces the dependence of explicitly correlated results on the particular variant employed (CCSD(T)-F12a/b/c, unscaled or scaled triples). The (Q,5) extrapolated CCSD(T)-F12/aVXZ+(aVXZ) energies are all accurate to about 0.2–0.3 K while the (5,6) extrapolated ones are accurate to 0.03–0.06 K. Thus, once again the aVXZ+(aVXZ) sequence is the recommended one for explicitly correlated CCSD(T) calculations.

We mentioned in Sec. II that the default cc-pVXZ/JKFIT auxiliary sets are not adequate for the resolution of identity and, to a lesser extent, for the density fitting of the Fock matrix. This fact is clearly illustrated in Table SIII in the supplementary material<sup>73</sup> which contains the CCSD(T)-F12b/FC interaction energies (unscaled triples) computed using the aVTZ orbital basis and various choices of auxiliary bases. It should be stressed that Table SIII of the supplementary material<sup>73</sup> represents the worst-case scenario where both the orbital basis and the default RI basis are so small that their

union is not at all adequate for performing the resolution of identity. Larger orbital bases as well as bases involving midbond functions are less dependent on the choice of the RI basis because the default cc-pVXZ/JKFIT set does a better job. In the aVTZ case, the dependence of results on the RI basis quality is dramatic and the default cc-pVTZ/JKFIT set overestimates the interaction energy by a striking 10 K. On the other hand, the aVXZ-RI sequence<sup>48</sup> exhibits smooth convergence of interaction energies and the aV6Z-RI\* auxiliary set (aV6Z-RI with the  $k$  functions omitted) should be adequate for all our calculations. We have tested even larger RI bases obtained from aV6Z-RI\* by an extension of the exponent range for each angular momentum  $l$  in an even-tempered manner. The sets aV6Z-RI\*+ (one large and one small exponent added for each  $l$ ) and aV6Z-RI\*++ (two large and two small exponents added for each  $l$ ) lead to a further increase of the CCSD(T)-F12b/FC interaction energy by about 0.9 K, showing how hard it is to converge results with respect to the RI basis. As the RI-basis dependence diminishes for larger orbital sets, we decided that this further increase does not warrant the associated increase in the basis set size and stuck to the aV6Z-RI\* set for all production runs. The effect of an improvement to the DF basis used to approximate the Fock matrix is much smaller (about 0.24 K) but we included it nonetheless, choosing the same aV6Z-RI\* set for all orbital bases. Finally, the effect of improving the DF(MP2-F12) basis seems to be negligible (no more than 0.05 K) and we stuck to the default aVXZ/MP2FIT choice. For comparison, an extensive set of explicitly correlated results using default auxiliary bases has been included in the supplementary material<sup>73</sup> (Tables SIV–SVII).

While the effects of the core-core and core-valence correlation are nearly negligible for the neon dimer, it is instructive to examine how the CCSD(T) core correction  $E_{\text{int}}^{\text{CCSD(T)}/\text{AE}} - E_{\text{int}}^{\text{CCSD(T)}/\text{FC}}$  is recovered in orbital and explicitly correlated calculations in various bases. The relevant numerical results are presented in Table V. In the explicitly correlated case, CCSD(T)-F12b results with unscaled triples have been presented. The scaling of the triples term actually worsens most of the explicitly correlated results. The AE-optimized aCVXZ, awCVXZ,<sup>50</sup> and CVXZ-F12 (Ref. 51) sequences are included in Table V in addition to the FC-optimized bases. As the results in this table show, all orbital basis sequences smoothly converge to the vicinity of the benchmark result. As expected, the aCVXZ and awCVXZ sequences perform very similarly and they both are better than the purely FC-optimized aVXZ and dVXZ bases at the same  $X$ , especially when midbond functions are present [the (3s3p2d2f1g) midbond set<sup>57</sup> employed here has exponents (0.9,0.3,0.1) for  $sp$ , (0.6, 0.2) for  $df$ , and 0.35 for  $g$  functions]. However, the aV6Z and dV6Z values are also well converged. On the other hand, the CCSD(T)-F12b core corrections are somewhat erratic. The explicitly correlated results in the aVXZ and aVXZ+(aVXZ) bases are actually worse than their conventional CCSD(T) counterparts except for the aV6Z+(aV6Z) case where the excellent accuracy of the CCSD(T)-F12b result may be accidental. The aCVXZ and aCVXZ+(aVXZ) results are also less accurate at CCSD(T)-F12b than at CCSD(T) although they are better than the aVXZ

TABLE V. Values of the core correction  $E_{\text{int}}^{\text{CCSD(T)/AE}} - E_{\text{int}}^{\text{CCSD(T)/FC}}$  for the neon dimer at  $R = 3.1 \text{ \AA}$  calculated in various orbital and explicitly correlated bases. Counterpoise correction and, in case of the explicitly correlated values, the CCSD(T)-F12b variant (unscaled triples) have been applied. The energy unit is 1 K. The benchmark value of this correction is 0.068 K.<sup>54</sup>

Basis set	$X = T$	$X = Q$	$X = 5$	$X = 6$
Orbital results				
aVXZ	-0.063	0.002	0.037	0.057
aVXZ+(3s3p2d2f1g)	-1.490	-0.042	0.043	0.063
aVXZ+(aVXZ)	-0.096	-0.017	0.026	0.048
dVXZ	-0.150	0.023	0.042	0.064
aCVXZ	-0.019	0.028	0.051	0.062
aCVXZ+(3s3p2d2f1g)	0.042	0.065	0.071	0.073
aCVXZ+(aVXZ)	0.038	0.064	0.071	
awCVXZ	-0.019	0.029	0.052	
awCVXZ+(3s3p2d2f1g)	0.050	0.067	0.072	
awCVXZ+(aVXZ)	0.042	0.065	0.072	
CCSD(T)-F12b results				
aVXZ	0.739	0.245	0.159	0.112
aVXZ+(aVXZ)	0.985	0.199	0.152	0.068
aCVXZ	-0.022	0.031	-0.005	-0.706
aCVXZ+(aVXZ)	0.110	0.047	-0.038	
VXZ-F12	0.129	-0.035		
VXZ-F12+(VXZ-F12)	0.295	0.078		
CVXZ-F12	-0.119	-0.033		
CVXZ-F12+(VXZ-F12)	0.061	0.071		

and aVXZ+(aVXZ) values at the same  $X$ . The CCSD(T)-F12b core correction in the aCV6Z basis set, obtained by adding diffuse functions (in an even-tempered manner) to the cc-pCV6Z set,<sup>96</sup> is far from accurate due to the lack of an adequate DF basis set (the awCV5Z/MP2FIT set was used). Note that aCV6Z performs very well in conventional CCSD(T). The VXZ-F12 and CVXZ-F12 core corrections are quite erratic and mostly of wrong sign. However, inclusion of mid-bond functions makes these two sequences acceptably accurate, with CVXZ-F12+(VXZ-F12) having an advantage over VXZ-F12+(VXZ-F12). For the former sequence, with the core correction estimate at  $X = Q$  amounting to 0.071 K, the results are actually just as good or better than the orbital, aCVXZ+(aVXZ) and awCVXZ+(aVXZ) ones at the same  $X$ . However, the latter results are also available at  $X = 5$ .

One may ask a question whether the erratic recovery of the core correction by CCSD(T)-F12b is a consequence of the fact that auxiliary bases that are appropriate for the FC calculations (see the discussion of Table SIII)<sup>73</sup> are not complete enough for the AE calculations. Our tests show that at the “worst-case-scenario” aVTZ level, the auxiliary basis sets employed here [aVTZ/MP2FIT for DF(MP2-F12) and aV6Z-RI\* for DF(Fock) and RI] are indeed not complete enough to get a fully converged CCSD(T)-F12b core correction. The value 0.739 K obtained in this way changes to 0.841 K if the RI basis is upgraded to aV6Z-RI\*++ and to 0.622 K if the DF(MP2-F12) basis is upgraded to aV5Z/MP2FIT. As the two auxiliary basis set incompleteness effects act in opposite directions, a simultaneous change of the DF(MP2-F12) basis to aV5Z/MP2FIT and of the RI basis to aV6Z-RI\*++ alters the core correction only slightly (the value is 0.727 K). However, for larger orbital sets the auxiliary bases employed by

us are quite adequate. For the aV5Z orbital basis, our choice of aV5Z/MP2FIT for the DF(MP2-F12) basis and aV6Z-RI\* for the DF(Fock) and RI bases leads to a CCSD(T)-F12b core correction of 0.159 K. Any upgrade of the DF(MP2-F12) basis (to awCV5Z/MP2FIT or awCV5Z/MP2FIT+, an even-tempered extension of the former set), of the RI basis (to aV6Z-RI\*+ or aV6Z-RI\*++), or of both, leads to a value of the core correction in the range 0.157–0.169 K. These results suggest that the auxiliary basis set incompleteness is only a minor factor as far as the errors of the CCSD(T)-F12b core correction are concerned.

### C. Ar<sub>2</sub>

The most accurate argon-argon interaction potential to date was obtained in Ref. 56 using theory levels up to CCSDTQ (see Refs. 55, 57, and 97 for earlier work of nearly the same accuracy). The CCSD(T)/FC part of the potential was obtained by CBS extrapolations using four different basis set families, all of them including diffuse and midbond functions. Extensive study of the basis set convergence of conventional CCSD(T)/FC has been performed in Ref. 56 and we will not repeat it here, moving right on to the explicitly correlated results. The benchmark CCSD(T)/FC argon-argon interaction energy at the near-minimum distance  $R = 3.75 \text{ \AA}$  amounts to  $-97.445 \pm 0.063 \text{ cm}^{-1}$ , and the CCSD(T) core correction at the same  $R$  is equal to  $-0.817 \pm 0.024 \text{ cm}^{-1}$ .

The CP-corrected computed and extrapolated CCSD(T)/FC argon-argon interaction energies at  $R = 3.75 \text{ \AA}$  are presented in Tables VI [conventional CCSD(T)] and VII [CCSD(T)-F12], and the corresponding deviations from benchmark are plotted in Fig. 5. Table VI is far from

TABLE VI. CP-corrected conventional CCSD(T)/FC interaction energies for the argon dimer at  $R = 3.75 \text{ \AA}$  calculated and extrapolated using various bases. The rows marked “ext.” contain the CBS-extrapolated results, with the value in the “ $X$ ” column obtained from the energies calculated in XZ and ( $X - 1$ )Z basis sets. The energy unit is  $1 \text{ cm}^{-1}$ . The benchmark CCSD(T)/CBS value of the interaction energy is  $-97.445 \pm 0.063 \text{ cm}^{-1}$ .<sup>56</sup> Table IV in Ref. 56 contains CCSD(T)/FC interaction energies in many more basis sets: this table is meant only as a supplement and provides orbital results in basis sets employed at the CCSD(T)-F12 level.

Basis set	$X = T$	$X = Q$	$X = 5$	$X = 6$
aVXZ	-66.068	-81.572	-90.097	-93.616
ext.		-93.830	-98.845	-98.289
aV( $X + d$ )Z	-65.808	-81.400	-89.981	-93.482
ext.		-93.788	-98.865	-98.151
dVXZ	-80.637	-90.643	-94.147	-95.872
ext.		-98.419	-97.633	-98.161
dV( $X + d$ )Z	-80.462	-90.476	-94.048	-95.765
ext.		-98.310	-97.682	-98.060
aVXZ+(aVXZ)	-91.122	-95.713	-96.961	-97.265
ext.		-98.985	-98.196	-97.673
aV( $X + d$ )Z+(aVXZ)	-91.025	-95.586	-96.914	-97.219
ext.		-98.900	-98.315	-97.637
aVXZ+(3s3p2d2f)	-96.189	-95.815	-96.566	-96.911
ext.		-95.699	-97.250	-97.370
aV( $X + d$ )Z+(3s3p2d2f)	-96.064	-95.705	-96.537	-96.870
ext.		-95.647	-97.380	-97.325

TABLE VII. CP-corrected CCSD(T)-F12/FC interaction energies for the argon dimer at  $R = 3.75 \text{ \AA}$  calculated and extrapolated using various bases. The rows marked “ext.” contain the CBS-extrapolated results, with the value in the “X” column obtained from the energies calculated in XZ and  $(X - 1)Z$  basis sets. The energy unit is  $1 \text{ cm}^{-1}$ . The benchmark CCSD(T)/CBS value of the interaction energy is  $-97.445 \pm 0.063 \text{ cm}^{-1}$ .<sup>56</sup>

Basis set	Unscaled triples				Scaled triples			
	X = T	X = Q	X = 5	X = 6	X = T	X = Q	X = 5	X = 6
CCSD(T)-F12a results								
aVXZ	-82.661	-89.354	-94.582	-96.356	-85.771	-91.017	-95.759	-97.086
ext.		-94.396	-100.116	-98.733		-95.003	-100.782	-98.849
aVXZ+(3s3p2d2f)	-99.466	-98.076	-97.847	-97.754	-103.158	-99.846	-99.043	-98.487
ext.		-97.363	-97.524	-97.586		-97.729	-98.118	-97.685
aVXZ+(aVXZ)	-96.545	-97.983	-98.057	-97.694	-100.180	-99.752	-99.246	-98.420
ext.		-99.127	-98.113	-97.169		-99.535	-98.694	-97.259
VXZ-F12	-74.930	-78.260			-78.039	-80.080		
ext.		-80.734				-81.613		
VXZ-F12+(3s3p2d2f)	-96.758	-96.363			-100.942	-98.574		
ext.		-96.013				-96.784		
VXZ-F12+(VXZ-F12)	-91.524	-93.694			-95.528	-95.864		
ext.		-95.262				-96.095		
CCSD(T)-F12b results								
aVXZ	-79.334	-87.452	-93.373	-95.609	-82.444	-89.115	-94.550	-96.338
ext.		-93.533	-99.634	-98.620		-94.141	-100.300	-98.735
aVXZ+(3s3p2d2f)	-96.391	-96.319	-96.854	-97.177	-100.082	-98.088	-98.049	-97.910
ext.		-96.567	-97.333	-97.581		-96.934	-97.926	-97.680
aVXZ+(aVXZ)	-93.499	-96.247	-97.100	-97.169	-97.135	-98.016	-98.289	-97.895
ext.		-98.347	-97.974	-97.236		-98.754	-98.554	-97.326
VXZ-F12	-72.264	-76.692			-75.372	-78.512		
ext.		-79.967				-80.847		
VXZ-F12+(3s3p2d2f)	-94.493	-95.091			-98.678	-97.302		
ext.		-95.466				-96.236		
VXZ-F12+(VXZ-F12)	-89.352	-92.468			-93.356	-94.639		
ext.		-94.728				-95.561		
CCSD(T)-F12c results								
aVXZ	-78.237	-86.836	-92.955	-95.287	-81.291	-88.483	-94.125	-96.014
ext.		-93.270	-99.423	-98.429		-93.889	-100.093	-98.548
aVXZ+(3s3p2d2f)	-95.225	-95.759	-96.441	-96.874	-98.855	-97.512	-97.630	-97.606
ext.		-96.450	-97.073	-97.431		-96.832	-97.671	-97.533
aVXZ+(aVXZ)	-92.342	-95.686	-96.688	-96.875	-95.915	-97.438	-97.871	-97.599
ext.		-98.222	-97.718	-97.106		-98.645	-98.303	-97.200
VXZ-F12	-70.127	-75.591			-73.178	-77.394		
ext.		-79.621				-80.514		
VXZ-F12+(3s3p2d2f)	-93.176	-94.432			-97.306	-96.627		
ext.		-95.288				-96.070		
VXZ-F12+(VXZ-F12)	-88.007	-91.829			-91.958	-93.984		
ext.		-94.603				-95.447		

complete — it supplements the extensive set of CCSD(T)/FC interaction energies from Table IV in Ref. 56 with orbital results in basis sets used for explicitly correlated calculations in this work. In addition to the aVXZ family, we have included in Table VI results calculated in the aug-cc-pV(X + d)Z $\equiv$ aV(X + d)Z sequence,<sup>60</sup> a recommended replacement to aVXZ for second-row atoms. The corresponding explicitly correlated results are given in Table SVIII in the supplementary material.<sup>73</sup> All auxiliary basis sets accompanying the aV(X+d)Z basis were taken identical as for the corresponding aVXZ set. The explicitly correlated calculations have been performed with two different midbond sets: the hydrogenic set from the same basis, and the (3s3p2d2f) set

for which a corresponding auxiliary basis has been defined.<sup>98</sup> Interestingly, the change from aVXZ to aV(X+d)Z slightly worsens most interaction energies, both the conventional and F12 ones as well as with and without midbond. Therefore, we elected to show only results in standard aVXZ bases in Fig. 5.

The results in Tables VI–VII and Fig. 5 once again show the importance of midbond functions for an accurate recovery of dispersion energy. No result in a midbondless basis, whether conventional or explicitly correlated, attains any acceptable accuracy. An inclusion of a moderate midbond set in the one-electron basis provides a much larger improvement than the F12 treatment, although the latter clearly improves the results in all bases tested. The size of

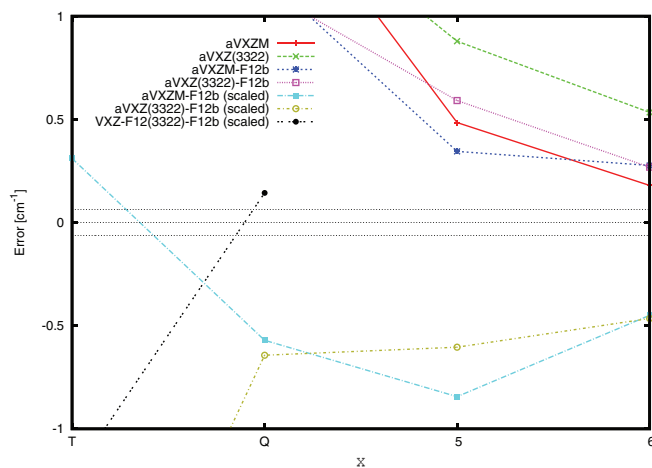


FIG. 5. Absolute errors (in  $\text{cm}^{-1}$ ) for the CP-corrected CCSD(T)/FC interaction energies for the argon dimer at  $R = 3.75 \text{ \AA}$  as a function of the cardinal number  $X$  of the basis set. Both conventional CCSD(T) and CCSD(T)-F12b results are presented. The letter M in the basis set symbol stands for the hydrogenic set of midbond functions from the same basis and (3322) stands for the ( $3s3p2d2f$ ) midbond set. The graph displays all CP-corrected CCSD(T)-F12b results using aVXZ or VXZ-F12 bases on atoms (scaled and unscaled triples) that fit within the range shown.

the midbond set also plays an important role: the variable-midbond aVXZ+(aVXZ) results become more accurate than the constant-midbond aVXZ+( $3s3p2d2f$ ) ones at  $X \geq 5$  (actually, converging the benchmark value to a  $0.063 \text{ cm}^{-1}$  uncertainty would not be possible without using very large midbond sets on top of nearly complete atom-centered bases — see Ref. 56 for details). The unscaled-triples CCSD(T)-F12a and CCSD(T)-F12b results in the aV6Z+(aV6Z) basis, with errors of  $-0.25$  and  $0.28 \text{ cm}^{-1}$  with respect to benchmark, respectively, are slightly inferior to the aV6Z+(aV6Z) and aV(6+d)Z+(aV6Z) orbital results but outperform all other  $X = 6$  orbital bases in Table VI and most of the bases from Table IV in Ref. 56 except for the doubly augmented sets accompanied by large midbond. Similar accuracy (errors of  $-0.31$  and  $0.27 \text{ cm}^{-1}$ , respectively) is attained for the aV6Z+( $3s3p2d2f$ ) basis. The scaled-triples CCSD(T)-F12c interaction energies, with errors of  $-0.15 \text{ cm}^{-1}$  for aV6Z+(aV6Z) and  $-0.16 \text{ cm}^{-1}$  for aV6Z+( $3s3p2d2f$ ), are even better (for the a and b variants, the scaling of triples actually worsens the results). A few lower- $X$  explicitly correlated results are, quite accidentally, more accurate, with the best value (an error of only  $0.007 \text{ cm}^{-1}$ ) provided by the scaled-triples CCSD(T)-F12c/aVQZ+(aVQZ) method. The VXZ-F12+( $3s3p2d2f$ ) [but not VXZ-F12+(VXZ-F12)] results are also occasionally quite accurate but no consistent pattern is observed as far as the CCSD(T)-F12 variant (a/b/c with scaled or unscaled triples) is concerned so these bases cannot be recommended for high-accuracy calculations. As the (5,6) extrapolation employing the CCSD(T)-F12b/aVXZ+(aVXZ) results undershoots by  $0.21 \text{ cm}^{-1}$ , one can conclude that the explicitly correlated CCSD(T) treatment in presently available bases can pinpoint the CBS limit to about  $0.2 \text{ cm}^{-1}$ . This accuracy is competitive to the  $X = 6$  orbital results but significantly inferior to the best available orbital results employing large-midbond bases with  $X$  up to 7.<sup>56</sup>

The values of the CCSD(T) core correction for the argon dimer have been listed in Table SVI in the supplementary material<sup>73</sup> (for orbital bases, this table is meant to complement, rather than replace, Table V in Ref. 56). The orbital basis sets that include functions optimized for the core-core and core-valence correlation like the awCVXZ family converge fairly quickly and smoothly to the benchmark result. The frozen-core-optimized aVXZ and aV( $X+d$ )Z families behave somewhat erratically, but the results at  $X = 6$  are within about  $0.2 \text{ cm}^{-1}$  of the benchmark. At the CCSD(T)-F12b level, unfortunately, most of the core correction values are quite inaccurate. The aVXZ and aV( $X+d$ )Z sequences, with or without midbond, do not even predict the correct sign of the core correction until  $X = 6$ , and the differences between the aVXZ and aV( $X+d$ )Z bases are unexpectedly high. The aCVXZ sequence performs better but the results are still inferior to conventional CCSD(T) in AE-optimized bases at the same  $X$ , cf. Table V in Ref. 56. The cc-pCV6Z basis set for argon has been recently optimized,<sup>51</sup> but its corresponding augmented version turned out to be too close to linear dependency to converge either CCSD or CCSD-F12 iterations. The VXZ-F12 results are superior to aVXZ but worse than aCVXZ at a given  $X$ , and the CVXZ-F12 core corrections are occasionally very accurate but erratic. Overall, the accuracy of the best orbital bases for the CCSD(T) core correction cannot be matched at the explicitly correlated level.

The calculations on all systems so far suggest that the CCSD(T)-F12 approaches using largest available basis sets including midbond functions are capable of recovering dispersion-dominated interaction energies quite accurately but they are no match for conventional CCSD(T) calculations using still larger basis sets. At this point, the question about the sources of remaining errors in the CCSD(T)-F12 interaction energies becomes increasingly important. One hypothesis is that the F12 approach does not improve the description of long-range dynamical correlation nearly as significantly as it improves its short-range, interelectronic-cusp counterpart. Other possible sources of residual error are the approximate character of the CCSD(T)-F12 theory [the a/b/c approximations to CCSD-F12 and a purely orbital character of the (T) correction] and the incompleteness of the auxiliary bases used for density fitting and/or resolution of identity [although, as the tests for  $\text{Ne}_2$  (Table SIII of the supplementary material<sup>73</sup>) suggest, the auxiliary bases used in this work are of high quality]. To investigate these issues, we decided to look at the MP2-F12 interaction energies for which no approximations are needed except for the auxiliary basis set incompleteness. The corresponding benchmark MP2/FC interaction energy at  $R = 3.75 \text{ \AA}$  was obtained as a byproduct of the CCSD(T) benchmark in Ref. 56 (but, mistakenly, not reported in that reference) and amounts to  $-114.241 \pm 0.394 \text{ cm}^{-1}$ . Note that MP2 converges slower with the basis set than CCSD(T): thus the larger uncertainty of the benchmark, and thus the popular “CBS-extrapolated MP2 plus small-basis CCSD(T)” approach<sup>25,99</sup> performs poorly for the argon dimer. A similar behavior is observed for the neon dimer where the benchmark MP2/CBS interaction energy, (5,6)-extrapolated from the same t-aug-cc-pVXZ+( $4s4p3d3f2g$ ) bases as the benchmark CCSD(T) energy from Ref. 54, equals  $-26.174 \pm 0.152 \text{ K}$ .

TABLE VIII. MP2/FC and MP2-F12/FC interaction energies for the near-minima geometries of He<sub>2</sub>, Ne<sub>2</sub>, and Ar<sub>2</sub>. The numbers listed are deviations with respect to benchmark CBS-limit MP2/FC interaction energies (see text). The middle third of the table presents raw MP2-F12 values, and in the lower third these values are corrected for the residual density fitting error by adding a difference between conventional MP2 and DF-MP2 calculated in the same orbital and auxiliary bases. The energy unit is 1 K for He<sub>2</sub> and Ne<sub>2</sub> and 1 cm<sup>-1</sup> for Ar<sub>2</sub>.

System	Basis set	X = T	X = Q	X = 5	X = 6
Conventional MP2					
He <sub>2</sub>	aVXZ	2.09	1.26	0.77	0.41
He <sub>2</sub>	aVXZ+(aVXZ)	0.43	0.14	0.075	0.038
Ne <sub>2</sub>	aVXZ	12.7	6.24	3.51	2.16
Ne <sub>2</sub>	aVXZ+(aVXZ)	2.71	0.77	0.32	0.15
Ar <sub>2</sub>	aVXZ	31.8	16.8	8.3	4.6
Ar <sub>2</sub>	aVXZ+(aVXZ)	10.0	3.72	1.55	0.76
Raw MP2-F12					
He <sub>2</sub>	aVXZ	0.65	0.68	0.32	0.12
He <sub>2</sub>	aVXZ+(aVXZ)	0.70	-0.025	0.014	-0.012
Ne <sub>2</sub>	aVXZ	2.92	2.14	1.11	0.53
Ne <sub>2</sub>	aVXZ+(aVXZ)	0.85	0.12	-0.023	-0.097
Ar <sub>2</sub>	aVXZ	14.4	8.40	3.05	1.02
Ar <sub>2</sub>	aVXZ+(aVXZ)	4.02	1.18	0.04	-0.37
MP2-F12 corrected for DF error					
He <sub>2</sub>	aVXZ	0.64	0.70	0.35	0.15
He <sub>2</sub>	aVXZ+(aVXZ)	0.71	-0.020	0.020	0.006
Ne <sub>2</sub>	aVXZ	2.83	1.77	1.16	0.59
Ne <sub>2</sub>	aVXZ+(aVXZ)	0.86	0.11	-0.030	-0.016
Ar <sub>2</sub>	aVXZ	13.8	8.88	3.23	1.28
Ar <sub>2</sub>	aVXZ+(aVXZ)	4.32	1.24	0.06	0.005

For the helium dimer, a Gaussian-type geminal benchmark is available<sup>52</sup> and amounts to  $-6.7920 \pm 0.0001$  K.

Convergence of the MP2/FC and MP2-F12/FC interaction energies in bases aVXZ and aVXZ+(aVXZ) for the helium, neon, and argon dimers has been illustrated in Table VIII. The same 3C(FIX) Ansatz as in the CCSD(T)-F12 calculations and the same auxiliary bases have been employed. Table VIII shows that the basis set convergence of MP2-F12 is far superior to conventional MP2 at all X for both midbondless and midbond-containing bases. An addition of midbond functions still improves the interaction energies more than the F12 treatment, but the advantage of midbond over F12 is not as dramatic as for CCSD(T). At large X, the MP2-F12 results in aVXZ+(aVXZ) bases, although extremely accurate, seem to be cutting right through the benchmark value and converging to a point that is below the benchmark. The origin of this behavior can be traced down to the incompleteness of the auxiliary basis in DF-MP2-F12 which, however small, becomes important at this level of accuracy. Therefore, we attempted to correct the MP2-F12 results for auxiliary basis set incompleteness by adding the difference between the conventional MP2 and DF-MP2 values employing the same orbital and auxiliary bases. The “corrected” MP2-F12 interaction energies obtained in this way, listed in Table VIII, exhibit impressive convergence towards the benchmark values as long as the basis contains midbond functions. For the argon dimer, the MP2-F12 interaction energy is within 0.06

cm<sup>-1</sup> of the benchmark already at the aV5Z+(aV5Z) level. Note that all the orbital results in basis sets with up to X = 7 including large midbond bases were not able to pin down the MP2/CBS limit to any better than 0.4 cm<sup>-1</sup>. The aVQZ+(aVQZ) MP2-F12 result is already within 1.3 cm<sup>-1</sup> of the benchmark, similar to conventional MP2 in a basis as large as aV6Z+(3s3p2d2f) and only slightly worse than conventional MP2 in aV6Z+(aV6Z) (an error of 0.76 cm<sup>-1</sup>). Without midbond, the aV6Z basis gives an MP2-F12 error of 1.3 cm<sup>-1</sup>, much better than for conventional MP2 (4.6 cm<sup>-1</sup>) but not nearly as good as for the aVXZ+(aVXZ) family. The accuracy of MP2-F12 for midbond-including bases is just as impressive for other rare gas dimers presented in Table VIII. In general, the MP2-F12/aV5Z+(aV5Z) treatment, with a correction for the DF basis set incompleteness, is superior to conventional MP2 in *any* available basis for Ne<sub>2</sub> and Ar<sub>2</sub> and is just as good as the best conventional result [dV7Z+(dV7Z), an error of 0.019 K] for He<sub>2</sub>. By comparison, the poorer performance of CCSD(T)-F12 suggests that the a/b/c approximations to the exact CCSD-F12 as well as the approximate treatment of triples are at least one of the limiting factors for the accuracy of available CCSD(T)-F12 interaction energies.

#### D. Kr<sub>2</sub>

The CP-corrected CCSD(T)/FC interaction energies, both conventional and explicitly correlated, for a near-minimum geometry of the krypton dimer ( $R = 4.06$  Å)

TABLE IX. CP-corrected CCSD(T)/FC and CCSD(T)-F12b/FC (unscaled triples) interaction energies for the krypton dimer at  $R = 4.06$  Å calculated and extrapolated using various bases. The rows marked “ext.” contain the CBS-extrapolated results, with the value in the “X” column obtained from the energies calculated in XZ and (X - 1)Z basis sets. The energy unit is 1 cm<sup>-1</sup>.

Basis set	X = D	X = T	X = Q	X = 5
Conventional results				
aVXZ	-30.06	-89.35	-116.39	-125.95
ext.		-116.01	-136.43	-135.86
aVXZ+(aVXZ)	-78.71	-125.12	-131.25	-132.91
ext.		-144.49	-135.70	-134.64
aVXZ+(3s3p2d2f1g)	-145.75	-137.16	-132.17	-132.76
ext.		-134.50	-128.53	-133.36
aVXZ+(6s6p6d3f3g3h)	-147.01	-137.03	-133.33	-133.42
ext.		-134.06	-130.62	-133.52
dVXZ	-43.01	-107.37	-124.02	-129.67
ext.		-135.93	-136.26	-135.54
dVXZ+(dVXZ)	-85.25	-128.71	-132.28	-133.31
ext.		-146.64	-134.87	-134.39
dVXZ+(3s3p2d2f1g)	-148.13	-138.08	-132.82	-133.15
ext.		-134.57	-129.00	-133.48
dVXZ+(6s6p6d3f3g3h)	-147.77	-136.80	-133.43	-133.54
ext.		-133.26	-130.97	-133.65
CCSD(T)-F12b results				
aVXZ	-71.30	-107.57	-122.39	-129.36
ext.		-123.61	-133.24	-136.63
aVXZ+(aVXZ)	-96.15	-127.73	-131.56	-132.86
ext.		-141.38	-134.35	-134.23
aVXZ+(3s3p2d2f)	-124.26	-131.62	-131.85	-132.67
ext.		-135.49	-132.02	-133.53

have been presented in Table IX. For this dimer, no highly accurate benchmark results are available. In fact, the most accurate nonrelativistic CCSD(T)/FC interaction energy to date is the  $aV5Z+(3s3p2d2f1g)$  result of Ref. 57 which amounts to  $-132.8\text{ cm}^{-1}$  at this  $R$ . As both our conventional and F12 results slightly surpass the accuracy of Ref. 57, we will compare how precise estimates of the CBS-limit CCSD(T)/FC interaction energy can be obtained using each approach. Moreover, the  $VXZ$ -F12 basis sets have not been created for krypton, so we will restrict ourselves to the  $aVXZ$  family (for  $X = D, T, Q, 5$  as the  $aV6Z$  krypton set is not yet available). To converge the conventional CCSD(T)/FC interaction energy as well as possible, we generated the doubly augmented  $dVXZ$  bases by adding the second diffuse function for each  $l$  in an even-tempered manner in addition to using several large midbond sets.

For the  $aV5Z$  set, the default  $aV5Z/MP2FIT$  density fitting basis does not exist. Instead, we tried two other auxiliary bases:  $aVQZ/MP2FIT$  (Ref. 46) and  $aV5Z/JKFIT$ . Both of these choices lead to very reasonable CCSD(T)-F12b/FC interaction energies, and the numerical results in the  $aV5Z$  and  $aV5Z+(aV5Z)$  bases agree to  $0.03\text{ cm}^{-1}$  and  $0.24\text{ cm}^{-1}$ , respectively. We conclude that both auxiliary bases are fairly adequate for our purposes and present the results obtained using  $aVQZ/MP2FIT$  in Table IX. As the largest  $aVXZ$ -RI set available for krypton is the  $X = Q$  one, we have to revisit our choice of auxiliary basis for the resolution of identity and for the density fitting of the Fock matrix (we will use the same set for both purposes and refer to it simply as the RI set). For the  $aV5Z$  orbital basis and the  $aVQZ/MP2FIT$  density-fitting basis in MP2-F12, the RI set  $aVQZ$ -RI gives a CCSD(T)-F12b/FC interaction energy of  $-130.84\text{ cm}^{-1}$ . This value is still not fully converged, as indicated by the results with the RI set extended in an even-tempered manner (in exactly the same way in which the  $aV6Z$ -RI\*+ and  $aV6Z$ -RI\*++ sets of Table SIII of the supplementary material<sup>73</sup> were obtained). The use of RI sets  $aVQZ$ -RI+ and  $aVQZ$ -RI++ leads to interaction energies of  $-129.77$  and  $-129.27\text{ cm}^{-1}$ , respectively. On the other hand, the  $cc$ - $pV5Z/JKFIT$  RI set, the default for this orbital basis, gives an interaction energy of  $-129.36\text{ cm}^{-1}$ , and this energy barely changes (to  $-129.29$  and  $-129.28\text{ cm}^{-1}$  for  $cc$ - $pV5Z/JKFIT$ + and  $cc$ - $pV5Z/JKFIT$ ++, respectively) after an even-tempered extension. We conclude that the  $cc$ - $pV5Z/JKFIT$  RI set is a reasonable and well converged choice and we have used it as the krypton part of the RI basis set in all calculations. One should note that, just like for  $Ne_2$  (Table SIII of the supplementary material<sup>73</sup>), the default  $cc$ - $pVXZ/JKFIT$  RI sets are inadequate for the  $aVXZ$  orbital bases for  $X < 5$ . For example, the  $aVTZ$  orbital set combined with the  $aVTZ/MP2FIT$  DF set for MP2-F12 leads to CCSD(T)-F12b/FC interaction energies equal to  $-110.99$ ,  $-107.57$ , and  $-105.76\text{ cm}^{-1}$  for the RI bases  $cc$ - $pVTZ/JKFIT$ ,  $cc$ - $pV5Z/JKFIT$ , and  $cc$ - $pV5Z/JKFIT$ ++, respectively. The last result is virtually converged with respect to the RI approximation so that the  $cc$ - $pV5Z/JKFIT$  choice is significantly better than the default  $cc$ - $pVTZ/JKFIT$  one.

First, let us establish a CBS limit using the conventional nonrelativistic CCSD(T)/FC results in Table IX. Similar to

other rare gas dimers, interaction energies in bases without midbond converge slowly and second augmentation cannot make up for the lack of midbond functions. On the other hand, the results with midbond are quite well converged already at the quadruple-zeta level: the computed interaction energies in Table IX lie in the range  $(-133.5\text{ cm}^{-1}, -131.2\text{ cm}^{-1})$  for  $X = Q$  and  $(-133.6\text{ cm}^{-1}, -132.7\text{ cm}^{-1})$  for  $X = 5$ . The CBS-extrapolated energies clearly show different convergence patterns for constant-midbond and variable-midbond sequences: as discussed in Sec. III A, the  $X^{-3}$  extrapolations employed here are likely to undershoot for the former sequences and overshoot for the latter ones (see also Ref. 93). Assuming that this is indeed the case for the two largest-basis extrapolations, the  $dVXZ+(dVXZ)$  and  $dVXZ+(6s6p6d3f3g3h)$  ones with  $X = (Q, 5)$ , one obtains the  $(-134.39\text{ cm}^{-1}, -133.65\text{ cm}^{-1})$ , or  $-134.02 \pm 0.37\text{ cm}^{-1}$ , range for the CCSD(T)/FC CBS limit obtained from conventional calculations. Note that none of the calculated and/or extrapolated results in bases  $aVXZ$  and  $dVXZ$  lie in this range which shows that midbondless bases are not able to pinpoint the CBS limit to any similar accuracy.

In order to obtain a similar estimate of the CCSD(T)/FC CBS limit using F12 calculations, we employed the  $aVXZ+(aVXZ)$  and  $aVXZ+(3s3p2d2f)$  sequences from Table IX. The midbondless  $aVXZ$  energies, while better than the orbital results in the same sequence, are clearly inferior to the energies obtained in any sequence with midbond, either conventional or explicitly correlated. The CCSD(T)-F12b/FC interaction energies converge monotonically for both the  $aVXZ+(aVXZ)$  and  $aVXZ+(3s3p2d2f)$  families. As expected, the constant-midbond results are more accurate for small  $X$  while at  $X = 5$  the variable-midbond energy becomes slightly superior. It may be argued that, similar to the orbital case, the  $X^{-3}$  extrapolations probably undershoot in constant-midbond sequences and overshoot in variable-midbond ones. Based on this observation, we took the range between the  $(Q, 5)$ -extrapolated  $aVXZ+(aVXZ)$  and  $aVXZ+(3s3p2d2f)$  results,  $(-134.23\text{ cm}^{-1}, -133.53\text{ cm}^{-1})$  or  $-133.88 \pm 0.35\text{ cm}^{-1}$ , as our explicitly correlated estimate of the CBS limit. The agreement with the orbital-basis estimate is remarkable and the uncertainties of both approaches are virtually the same. If triples were scaled in the CCSD(T)-F12b calculation, the corresponding CBS limit estimate, obtained in the same way as above, would amount to  $-134.05 \pm 0.34\text{ cm}^{-1}$ , in agreement with the unscaled result.

Based on the CCSD(T)-F12 results for all four systems, Tables II, IV, VII, and IX, we can attempt to answer the question which of the six flavors of the CCSD(T)-F12 approach ( $a/b/c$  approximations, scaled or unscaled triples) provides the most accurate results for rare gas dimers. For this purpose, we selected all midbond-containing basis sets from  $X = T$  upwards (as midbondless bases are clearly inferior) and computed the mean absolute deviations (MADs) with respect to the benchmark CCSD(T)/FC value for each system (for  $Kr_2$  we took the orbital benchmark obtained here,  $-134.02\text{ cm}^{-1}$ ). The results are collected in Table X. To look at the performance of different variants for the largest basis sets only, we repeated the averaging for all bases with  $X \geq Q$  and  $X \geq 5$ : the results are also listed in Table X. The same general trends are seen for all four systems.



TABLE X. Mean absolute deviations of the CCSD(T)-F12/FC interaction energies (*a/b/c* approximation, scaled or unscaled triples) with respect to the benchmark values. The units are Kelvins for He<sub>2</sub> and Ne<sub>2</sub> and cm<sup>-1</sup> for Ar<sub>2</sub> and Kr<sub>2</sub>. The results obtained in all midbond-containing basis sets listed in Tables II, IV, (VII and SVIII), and IX for He<sub>2</sub>, Ne<sub>2</sub>, Ar<sub>2</sub>, and Kr<sub>2</sub>, respectively, with the cardinal number *X* higher or equal to the value *X*<sub>min</sub>, have been included in the averaging.

<i>X</i> <sub>min</sub>	Unscaled triples			Scaled triples		
	F12a	F12b	F12c	F12a	F12b	F12c
	He <sub>2</sub>					
T	0.32	0.37	0.41	0.29	0.30	0.32
Q	0.018	0.048	0.072	0.054	0.016	0.021
5	0.013	0.028	0.039	0.014	0.001	0.011
	Ne <sub>2</sub>					
T	1.16	1.44	1.65	0.89	1.02	1.09
Q	0.38	0.52	0.67	0.49	0.41	0.35
5	0.055	0.072	0.184	0.307	0.180	0.067
	Ar <sub>2</sub>					
T	1.08	1.83	2.52	2.26	1.12	0.91
Q	0.65	1.09	1.55	1.61	0.76	0.48
5	0.36	0.38	0.73	1.34	0.61	0.25
	Kr <sub>2</sub>					
T	0.74	2.64	3.64	3.33	1.08	0.79
Q	0.12	1.79	2.46	2.32	0.62	0.17
5	0.09	1.25	1.87	1.66	0.43	0.19

CCSD(T)-F12a is consistently the most accurate unscaled-triples variant (followed by F12b), and, except for He<sub>2</sub>, CCSD(T)-F12c is the most accurate scaled-triples variant (followed by F12b). The very small MAD for scaled-triples helium-dimer F12b results with  $X \geq 5$  is clearly accidental, cf. Fig. 3. The top performers, unscaled-triples CCSD(T)-F12a and scaled-triples CCSD(T)-F12c, lead to results of similar accuracy. The results in Table X corroborate the validity of the scaling approximation (1) and are in line with the theoretical superiority of the CCSD(T)-F12c variant over the other two.<sup>23</sup> At the same time, the unscaled-triples CCSD(T)-F12b, and especially CCSD(T)-F12a, interaction energies appear to benefit from a cancellation of errors between the *a/b* approximation and the non-F12 treatment of triples.

A more detailed picture of the performance of various flavors of CCSD(T)-F12 can be obtained by looking at the CCSD and (T) contributions separately. Such contributions for Ar<sub>2</sub> are presented in Tables SX and SXI in the supplementary material,<sup>73</sup> respectively. The MADs for the separate CCSD and (T) terms are listed in Tables SXII and SXIII, respectively, for all four dimers. The benchmark CCSD/FC interaction energies amount to  $-9.1507 \pm 0.0006$  K,<sup>52</sup>  $-32.504 \pm 0.015$  K,  $-69.064 \pm 0.075$  cm<sup>-1</sup>, and  $-97.35 \pm 0.22$  cm<sup>-1</sup> for He<sub>2</sub>, Ne<sub>2</sub>, Ar<sub>2</sub>, and Kr<sub>2</sub>, respectively. The latter three benchmarks have been obtained in exactly the same way as their CCSD(T) counterparts. The corresponding benchmark values for the (T) triples contribution to the interaction energy are  $-1.5346 \pm 0.0011$  K,  $-8.871 \pm 0.028$  K,  $-28.386 \pm 0.134$  cm<sup>-1</sup>, and  $-36.67 \pm 0.14$  cm<sup>-1</sup> for He<sub>2</sub>, Ne<sub>2</sub>, Ar<sub>2</sub>, and Kr<sub>2</sub>, respectively. The results in Tables SX–SXIII show that, first of all, CCSD-F12a is inferior to the other two variants for all sufficiently large bases. This could

have been expected as CCSD-F12a is more approximate than CCSD-F12b and is generally recommended for double- and triple-zeta bases only.<sup>24</sup> Interestingly, CCSD-F12c, despite being formally closer to exact CCSD-F12 than CCSD-F12b, does not provide improvement over the latter for rare gas dimers. Due to the differences in the converged doubles amplitudes, the CCSD(T)-F12c triples contributions are different than those for CCSD(T)-F12a/b which in turn are different than for conventional CCSD(T). The F12 contribution to doubles is neither designed nor guaranteed to improve triples so a perturbative approach like the CCSD(2)<sub>F12</sub> method of Valeev *et al.*,<sup>100–102</sup> for which the (T) contribution is identical as in conventional CCSD(T), might be an interesting alternative to try. The scaling of the (T) term consistently improves it for midbondless bases. For bases with midbond, some improvement is also seen but the scaling considerably overshoots. As the MADs in Table SXIII show, scaled CCSD(T)-F12c triples are consistently somewhat more accurate than scaled CCSD(T)-F12a/b triples, however, it is not clear if this holds for other systems. The MADs for the triples term are often larger than the MADs for the entire CCSD(T)-F12 (Table X) so that, in fact, several CCSD(T)-F12 variants including scaled-triples CCSD(T)-F12c benefit from error cancellation between the CCSD and (T) contributions.

The overall MADs in Table X indicate that the CCSD(T)-F12c method with scaled triples should be recommended for calculations of interaction energies in rare gas dimers. If one follows this recommendation and employs the largest, aV6Z+(aV6Z) orbital bases, the resulting estimates of near-minimum CCSD(T)/FC interaction energies for He<sub>2</sub>, Ne<sub>2</sub>, and Ar<sub>2</sub> amount to  $-10.6743 \pm 0.0015$  K,  $-41.41 \pm 0.08$  K, and  $-97.60 \pm 0.40$  cm<sup>-1</sup>, respectively. The uncertainties for these estimates were chosen as differences between the calculated result and the one extrapolated from this basis and the basis with *X* lower by one (note that the extrapolated results are less accurate than the calculated ones for Ne<sub>2</sub> and Ar<sub>2</sub>). Due to an accidental closeness of the aVQZ+(aVQZ) and aV5Z+(aV5Z) results for Kr<sub>2</sub>, the benchmark  $-133.93 \pm 0.01$  cm<sup>-1</sup> computed in this way would have a grossly underestimated uncertainty. A comparison of the results (Q,5)-extrapolated using the aVXZ+(aVXZ) and aVXZ+(3s3p2d2f) families of bases (just like we did earlier in this subsection for the CCSD(T)-F12b approach), gives  $-133.57 \pm 0.35$  cm<sup>-1</sup> as the CBS-limit CCSD(T) interaction energy. Comparing to the benchmark values (of mixed geminal-orbital origin for He<sub>2</sub> and purely orbital origin for all other systems) of  $-10.6844 \pm 0.0010$  K,  $-41.375 \pm 0.013$  K,  $-97.445 \pm 0.063$  cm<sup>-1</sup>, and  $-134.02 \pm 0.37$  cm<sup>-1</sup>, respectively, we find that the accuracy afforded by the explicitly correlated approach is quite competitive. However, for He<sub>2</sub> the two benchmark values are incompatible which clearly indicates that the scaled-triples CCSD(T)-F12c energy in the aV6Z+(aV6Z) basis is not as well converged as it seems. For Ne<sub>2</sub> and Ar<sub>2</sub>, the error bars of the CCSD(T)-F12c and benchmark results overlap but the uncertainties of the latter are several times smaller due to the availability of much larger bases for conventional CCSD(T). For Kr<sub>2</sub>, both values are in perfect agreement and, since no orbital bases with  $X > 5$  are available,

the CCSD(T)-F12c benchmark is just as precise as the orbital one.

#### IV. SUMMARY

The performance of various variants of the explicitly correlated CCSD(T)-F12 approach was studied for the near-minimum geometries of four weakly interacting rare gas dimers: He<sub>2</sub>, Ne<sub>2</sub>, Ar<sub>2</sub>, and Kr<sub>2</sub>. The recent immense popularity of the CCSD(T)-F12 method stems from the ability of the explicitly correlated approach to converge molecular correlation energies using relatively small orbital bases. Preliminary studies have indicated that this improved convergence does translate to an improved convergence of weak intermolecular interaction energies, however, the comparisons to very large orbital calculations are still quite scarce. Rare gas dimers are an extremely demanding case as the interaction is dominated by long-range correlation (dispersion) that is a particularly hard effect to converge with respect to the basis set. At the same time, the accuracy of benchmark rare-gas-dimer interaction energies, at the CCSD(T) level and beyond, is far superior to the accuracy attained for nearly all other systems.

The results presented here confirm that the F12 approach improves the accuracy of CCSD(T) interaction energies quite substantially for basis sets of triple- and quadruple-zeta quality. However, somewhat surprisingly, the improvement in accuracy diminishes for still larger basis sets. Furthermore, a larger improvement is obtained by a simple addition of bond functions than by employing the F12 method. Fortunately, as shown here, the benefits of bond functions and of the F12 approach can be combined to yield the most accurate frozen-core CCSD(T) interaction energies at a given basis set cardinal number  $X$ . The  $aVXZ+(aVXZ)$  basis set sequence, with hydrogenic midbond functions that change in accord with the atom-centered part of the basis, is a particularly good choice for the F12 calculations.

At the present time, the rare-gas-dimer CCSD(T) interaction energies obtained using the conventional approach with the largest available bases remain more accurate than the ones obtained using the F12 method with the largest available bases except for short intermonomer distances corresponding to the repulsive wall of the potential. There are two reasons for that. First, the F12 approach relies inherently on density fitting and the resolution of identity, and the limited availability of auxiliary bases somewhat restricts the orbital basis set choice in this case. Even more importantly, the errors introduced by the  $a/b/c$  approximations in CCSD-F12 and by a non-explicitly correlated treatment of perturbative triples (with and without scaling) become crucial at this level of accuracy. Conversely, the MP2-F12 method, for which no approximations are needed except for the density fitting, leads to interaction energies that are superior to all conventional MP2 results even for the rare gas dimers. A possible third factor limiting the accuracy of the CCSD(T)-F12 interaction energies, the auxiliary basis set incompleteness, has been largely eliminated by a use of extended, nondefault RI basis sets. The default  $cc-pVXZ/JKFIT$  RI bases lead to large errors for the small- $X$   $aVXZ$  orbital bases where the union of the orbital set and the RI set is far from complete.

A comparison of the accuracy of six different CCSD(T)-F12 flavors for rare gas dimers shows that, in general, the unscaled-triples CCSD(T)-F12a and scaled-triples CCSD(T)-F12c methods have a slight advantage over the other variants. The scaled-triples CCSD(T)-F12c method [introduced in Ref. 23 as CCSD(T)(F12\*)] is the one that provides the right answer for the right reason, confirming the validity of the scaling approximation and the superiority of CCSD(T)-F12b/c over CCSD(T)-F12a. Both unscaled-triples CCSD(T)-F12a and scaled-triples CCSD(T)-F12c benefit from a cancellation of errors between the CCSD part and the triples part.

The frozen-core CCSD(T)-F12 method, with proper care taken to select appropriate auxiliary bases, variants, and parameters, does provide interaction energies that are superior to conventional CCSD(T) energies for the same  $X$  and competitive to the best conventional energies available. The same cannot be said, however, about the effects of core-core and core-valence correlation on the CCSD(T) interaction energies. The values of the core correction computed using CCSD(T)-F12 vary widely with the orbital basis set employed. While one of the reasons for the failure of CCSD(T)-F12 to consistently describe core corrections is that the requirements for the density fitting and resolution-of-identity bases are particularly stringent for all-electron calculations, it cannot be excluded that the accuracy of all-electron CCSD(T)-F12 interaction energies is hampered by a more fundamental issue. More research in this direction is needed before the CCSD(T)-F12 approach, despite its immense benefits, can be accepted as the ultimate method for computing top-accuracy intermolecular interaction energies.

#### ACKNOWLEDGMENTS

We thank Dr. Bogumil Jeziorski, Dr. Edward Valeev, Dr. Vincent Ortiz, and Daniel Smith for reading and commenting on the manuscript. This work has been supported by the Auburn University startup funding.

- <sup>1</sup>E. A. Hylleraas, *Z. Phys.* **54**, 347 (1929).
- <sup>2</sup>W. Kołos and L. Wolniewicz, *J. Chem. Phys.* **43**, 2429 (1965).
- <sup>3</sup>C. Schwartz, *Int. J. Mod. Phys. E* **15**, 877 (2006).
- <sup>4</sup>H. Nakatsuji, H. Nakashima, Y. Kurokawa, and A. Ishikawa, *Phys. Rev. Lett.* **99**, 240402 (2007).
- <sup>5</sup>W. Cencek and K. Szalewicz, *Int. J. Quantum Chem.* **108**, 2191 (2008).
- <sup>6</sup>K. Pachucki, *Phys. Rev. A* **82**, 032509 (2010).
- <sup>7</sup>W. Kutzelnigg and W. Klopper, *J. Chem. Phys.* **94**, 1985 (1991).
- <sup>8</sup>F. R. Manby, *J. Chem. Phys.* **119**, 4607 (2003).
- <sup>9</sup>K. Szalewicz and B. Jeziorski, *Mol. Phys.* **108**, 3091 (2010).
- <sup>10</sup>A. J. May and F. R. Manby, *J. Chem. Phys.* **121**, 4479 (2004).
- <sup>11</sup>W. Klopper, F. R. Manby, S. Ten-no, and E. F. Valeev, *Int. Rev. Phys. Chem.* **25**, 427 (2006).
- <sup>12</sup>J. Noga, W. Kutzelnigg, and W. Klopper, *Chem. Phys. Lett.* **199**, 497 (1992).
- <sup>13</sup>T. Shiozaki, M. Kamiya, S. Hirata, and E. F. Valeev, *J. Chem. Phys.* **129**, 071101 (2008).
- <sup>14</sup>A. Köhn, G. W. Richings, and D. P. Tew, *J. Chem. Phys.* **129**, 201103 (2008).
- <sup>15</sup>A. Köhn, *J. Chem. Phys.* **130**, 131101 (2009).
- <sup>16</sup>H.-J. Werner, P. J. Knowles, F. R. Manby, and M. Schütz *et al.*, MOLPRO, version 2010.1, a package of ab initio programs (2010), see <http://www.molpro.net>.
- <sup>17</sup>*TURBOMOLE V6.3 2011, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from <http://www.turbomole.com>.*

- <sup>18</sup>C. Hättig, W. Klopper, A. Köhn, and D. P. Tew, *Chem. Rev.* **112**, 4 (2012).
- <sup>19</sup>L. Kong, F. A. Bischoff, and E. F. Valeev, *Chem. Rev.* **112**, 75 (2012).
- <sup>20</sup>H. Fliegl, W. Klopper, and C. Hättig, *J. Chem. Phys.* **122**, 084107 (2005).
- <sup>21</sup>D. P. Tew, W. Klopper, C. Neiss, and C. Hättig, *Phys. Chem. Chem. Phys.* **9**, 1921 (2007).
- <sup>22</sup>T. B. Adler, G. Knizia, and H.-J. Werner, *J. Chem. Phys.* **127**, 221106 (2007).
- <sup>23</sup>C. Hättig, D. P. Tew, and A. Köhn, *J. Chem. Phys.* **132**, 231102 (2010).
- <sup>24</sup>G. Knizia, T. B. Adler, and H. Werner, *J. Chem. Phys.* **130**, 054104 (2009).
- <sup>25</sup>P. Jurecka, J. Sponer, J. Cerny, and P. Hobza, *Phys. Chem. Chem. Phys.* **8**, 1985 (2006).
- <sup>26</sup>K. Patkowski, R. Podeszwa, and K. Szalewicz, *J. Phys. Chem. A* **111**, 12822 (2007).
- <sup>27</sup>T. H. Dunning, Jr., *J. Chem. Phys.* **90**, 1007 (1989).
- <sup>28</sup>R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, *J. Chem. Phys.* **96**, 6796 (1992).
- <sup>29</sup>F.-M. Tao and Y.-K. Pan, *J. Chem. Phys.* **95**, 3582 (1991).
- <sup>30</sup>R. Burcl, G. Chałasiński, R. Bukowski, and M. M. Szczęśniak, *J. Chem. Phys.* **103**, 1498 (1995).
- <sup>31</sup>H. L. Williams, E. M. Mas, K. Szalewicz, and B. Jeziorski, *J. Chem. Phys.* **103**, 7374 (1995).
- <sup>32</sup>T. van Mourik, R. J. Vos, J. H. van Lenthe, and F. B. van Duijneveldt, *Int. J. Quantum Chem.* **63**, 805 (1997).
- <sup>33</sup>M. Jeziorska, R. Bukowski, W. Cencek, M. Jaszunski, B. Jeziorski, and K. Szalewicz, *Collect. Czech. Chem. Commun.* **68**, 463 (2003).
- <sup>34</sup>D. P. Tew and W. Klopper, *J. Chem. Phys.* **125**, 094302 (2006).
- <sup>35</sup>O. Marchetti and H. Werner, *Phys. Chem. Chem. Phys.* **10**, 3400 (2008).
- <sup>36</sup>D. P. Tew, W. Klopper, and C. Hättig, *Chem. Phys. Lett.* **452**, 326 (2008).
- <sup>37</sup>O. Marchetti and H. Werner, *J. Phys. Chem. A* **113**, 11580 (2009).
- <sup>38</sup>J. R. Lane and H. G. Kjaergaard, *J. Chem. Phys.* **131**, 034307 (2009).
- <sup>39</sup>K. M. de Lange and J. R. Lane, *J. Chem. Phys.* **134**, 034301 (2011).
- <sup>40</sup>J. D. McMahon and J. R. Lane, *J. Chem. Phys.* **135**, 154309 (2011).
- <sup>41</sup>W. Skomorowski, F. Pawłowski, T. Korona, R. Moszyński, P. S. Żuchowski, and J. M. Hutson, *J. Chem. Phys.* **134**, 114109 (2011).
- <sup>42</sup>E. J. Carrell, C. M. Thorne, and G. S. Tschumper, *J. Chem. Phys.* **136**, 014103 (2012).
- <sup>43</sup>T. Takatani, E. G. Hohenstein, M. Malagoli, M. S. Marshall, and C. D. Sherrill, *J. Chem. Phys.* **132**, 144104 (2010).
- <sup>44</sup>R. Podeszwa, K. Patkowski, and K. Szalewicz, *Phys. Chem. Chem. Phys.* **12**, 5974 (2010).
- <sup>45</sup>K. A. Peterson, T. B. Adler, and H.-J. Werner, *J. Chem. Phys.* **128**, 084102 (2008).
- <sup>46</sup>F. Weigend, A. Köhn, and C. Hättig, *J. Chem. Phys.* **116**, 3175 (2002).
- <sup>47</sup>F. Weigend, *Phys. Chem. Chem. Phys.* **4**, 4285 (2002).
- <sup>48</sup>C. Hättig, *Phys. Chem. Chem. Phys.* **7**, 59 (2005).
- <sup>49</sup>K. E. Yousaf and K. A. Peterson, *J. Chem. Phys.* **129**, 184108 (2008).
- <sup>50</sup>K. A. Peterson and T. H. Dunning, Jr., *J. Chem. Phys.* **117**, 10548 (2002).
- <sup>51</sup>J. G. Hill, S. Mazumder, and K. A. Peterson, *J. Chem. Phys.* **132**, 054108 (2010).
- <sup>52</sup>K. Patkowski, W. Cencek, M. Jeziorska, B. Jeziorski, and K. Szalewicz, *J. Phys. Chem. A* **111**, 7611 (2007).
- <sup>53</sup>W. Cencek, K. Patkowski, and K. Szalewicz *et al.* (unpublished).
- <sup>54</sup>R. Hellmann, E. Bich, and E. Vogel, *Mol. Phys.* **106**, 133 (2008).
- <sup>55</sup>B. Jäger, R. Hellmann, E. Bich, and E. Vogel, *Mol. Phys.* **107**, 2181 (2009).
- <sup>56</sup>K. Patkowski and K. Szalewicz, *J. Chem. Phys.* **133**, 094304 (2010).
- <sup>57</sup>P. Slavíček, R. Kalus, P. Paška, I. Odvárková, P. Hobza, and A. Malijevský, *J. Chem. Phys.* **119**, 2102 (2003).
- <sup>58</sup>S. F. Boys and F. Bernardi, *Mol. Phys.* **19**, 553 (1970).
- <sup>59</sup>F. B. van Duijneveldt, J. G. C. M. van Duijneveldt-van de Rijdt, and J. H. van Lenthe, *Chem. Rev.* **94**, 1873 (1994).
- <sup>60</sup>T. H. Dunning, Jr., K. A. Peterson, and A. K. Wilson, *J. Chem. Phys.* **114**, 9244 (2001).
- <sup>61</sup>E. F. Valeev, *Chem. Phys. Lett.* **395**, 190 (2004).
- <sup>62</sup>C. Hättig (unpublished).
- <sup>63</sup>R. Podeszwa, R. Bukowski, and K. Szalewicz, *J. Phys. Chem. A* **110**, 10345 (2006).
- <sup>64</sup>K. Pernal, R. Podeszwa, K. Patkowski, and K. Szalewicz, *Phys. Rev. Lett.* **103**, 263201 (2009).
- <sup>65</sup>A. J. Misquitta, R. Podeszwa, B. Jeziorski, and K. Szalewicz, *J. Chem. Phys.* **123**, 214103 (2005).
- <sup>66</sup>A. Hesselmann, G. Jansen, and M. Schütz, *J. Chem. Phys.* **122**, 014103 (2005).
- <sup>67</sup>S. Ten-no, *Chem. Phys. Lett.* **398**, 56 (2004).
- <sup>68</sup>H.-J. Werner, T. B. Adler, and F. R. Manby, *J. Chem. Phys.* **126**, 164102 (2007).
- <sup>69</sup>S. Höfener, D. P. Tew, W. Klopper, and T. Helgaker, *Chem. Phys.* **356**, 25 (2009).
- <sup>70</sup>J. G. Hill, K. A. Peterson, G. Knizia, and H.-J. Werner, *J. Chem. Phys.* **131**, 194105 (2009).
- <sup>71</sup>H.-J. Werner, G. Knizia, and F. R. Manby, *Mol. Phys.* **109**, 407 (2011).
- <sup>72</sup>A. J. C. Varandas, *J. Phys. Chem. A* **114**, 8505 (2010).
- <sup>73</sup>See supplementary material at <http://dx.doi.org/10.1063/1.4734597> for additional tables and figures.
- <sup>74</sup>E. Papajak, H. R. Leverentz, J. Zheng, and D. G. Truhlar, *J. Chem. Theory Comput.* **5**, 1197 (2009).
- <sup>75</sup>E. Papajak and D. G. Truhlar, *J. Chem. Theory Comput.* **7**, 10 (2011).
- <sup>76</sup>E. Papajak, J. Zheng, X. Xu, H. R. Leverentz, and D. G. Truhlar, *J. Chem. Theory Comput.* **7**, 3027 (2011).
- <sup>77</sup>R. N. Hill, *J. Chem. Phys.* **83**, 1173 (1985).
- <sup>78</sup>W. Kutzelnigg, *Theor. Chem. Acc.* **68**, 445 (1985).
- <sup>79</sup>W. Kutzelnigg and J. D. Morgan III, *J. Chem. Phys.* **96**, 4484 (1992).
- <sup>80</sup>T. Helgaker, W. Klopper, H. Koch, and J. Noga, *J. Chem. Phys.* **106**, 9639 (1997).
- <sup>81</sup>A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen, and A. K. Wilson, *Chem. Phys. Lett.* **286**, 243 (1998).
- <sup>82</sup>W. Klopper and C. C. M. Samson, *J. Chem. Phys.* **116**, 6397 (2002).
- <sup>83</sup>W. Klopper, *Mol. Phys.* **99**, 481 (2001).
- <sup>84</sup>D. W. Schwenke, *J. Chem. Phys.* **122**, 014107 (2005).
- <sup>85</sup>W. Cencek, M. Jeziorska, R. Bukowski, M. Jaszunski, B. Jeziorski, and K. Szalewicz, *J. Phys. Chem. A* **108**, 3211 (2004).
- <sup>86</sup>B. Jeziorski, H. J. Monkhorst, K. Szalewicz, and J. G. Zabolitzky, *J. Chem. Phys.* **81**, 368 (1984).
- <sup>87</sup>R. Bukowski, B. Jeziorski, and K. Szalewicz, *J. Chem. Phys.* **110**, 4165 (1999).
- <sup>88</sup>D. E. Woon and T. H. Dunning, Jr., *J. Chem. Phys.* **100**, 2975 (1994).
- <sup>89</sup>T. van Mourik, A. K. Wilson, and T. H. Dunning, Jr., *Mol. Phys.* **96**, 529 (1999).
- <sup>90</sup>R. J. Gdanitz, *J. Chem. Phys.* **113**, 5145 (2000).
- <sup>91</sup>H. Partridge and C. W. Bauschlicher, *Mol. Phys.* **96**, 705 (1999).
- <sup>92</sup>S. L. Mielke, B. C. Garrett, and K. A. Peterson, *J. Chem. Phys.* **116**, 4142 (2002).
- <sup>93</sup>M. Jeziorska, W. Cencek, K. Patkowski, B. Jeziorski, and K. Szalewicz, *Int. J. Quantum Chem.* **108**, 2053 (2008).
- <sup>94</sup>S. A. Kucharski and R. J. Bartlett, *J. Chem. Phys.* **108**, 9221 (1998).
- <sup>95</sup>Y. J. Bomble, J. F. Stanton, M. Kállay, and J. Gauss, *J. Chem. Phys.* **123**, 054101 (2005).
- <sup>96</sup>A. K. Wilson (unpublished).
- <sup>97</sup>K. Patkowski, G. Murdachaew, C.-M. Fou, and K. Szalewicz, *Mol. Phys.* **103**, 2031 (2005).
- <sup>98</sup>R. Podeszwa, R. Bukowski, and K. Szalewicz, *J. Chem. Theory Comput.* **2**, 400 (2006).
- <sup>99</sup>M. S. Marshall, L. A. Burns, and C. D. Sherrill, *J. Chem. Phys.* **135**, 194102 (2011).
- <sup>100</sup>E. F. Valeev, *Phys. Chem. Chem. Phys.* **10**, 106 (2008).
- <sup>101</sup>M. Torheyden and E. F. Valeev, *Phys. Chem. Chem. Phys.* **10**, 3410 (2008).
- <sup>102</sup>E. F. Valeev and T. D. Crawford, *J. Chem. Phys.* **128**, 244113 (2008).