

Basis Set Convergence of the Post-CCSD(T) Contribution to Noncovalent Interaction Energies

Daniel G. A. Smith

Department of Chemistry and Biochemistry, Auburn University, Auburn, Alabama 36849, United States

Piotr Jankowski

Department of Quantum Chemistry, Faculty of Chemistry, Nicolaus Copernicus University, Gagarina 7, 87-100 Toruń, Poland

Michał Ślawik and Henryk A. Witek

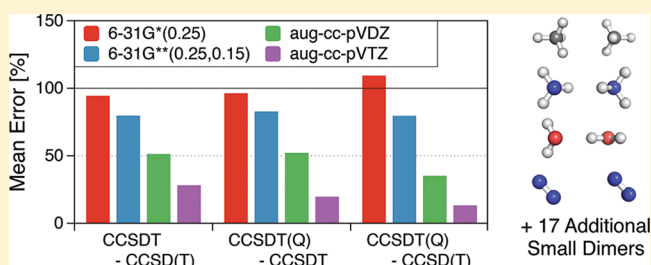
Department of Applied Chemistry and Institute of Molecular Science, National Chiao Tung University, 1001 Ta-Hsueh Road, Hsinchu 30010, Taiwan

Konrad Patkowski*

Department of Chemistry and Biochemistry, Auburn University, Auburn, Alabama 36849, United States

Supporting Information

ABSTRACT: We investigated the basis set convergence of high-order coupled-cluster interaction energy contributions for 21 small weakly bound complexes. By performing CCSDT(Q) calculations in at least the aug-cc-pVTZ basis set, and CCSDT calculations in at least aug-cc-pVQZ (aug-cc-pVTZ for one system), we found the convergence to be quite slow. In particular, the 6-31G*(0.25) and 6-31G**(0.25,0.15) bases advocated by Hobza et al. (*J. Chem. Theory Comput.* **2013**, *9*, 2151; *ibid.* **2013**, *9*, 3420) are unsuitable for the post-CCSD(T) effects, with average errors for the CCSDT(Q)-CCSD(T) interaction energy contribution of about 80% for 6-31G**(0.25,0.15) and 110% for 6-31G*(0.25). Upgrading the basis set to aug-cc-pVDZ reduces the average error to about 35% and extremely demanding CCSDT(Q)/aug-cc-pVTZ calculations are necessary for further improvement in accuracy. An error cancellation between basis set incompleteness effects at the CCSDT-CCSD(T) and CCSDT(Q)-CCSDT levels occurs for most (but not all) complexes, making it unproductive to carry out CCSDT calculations in a larger basis set than the more demanding CCSDT(Q) calculations. We also found that the frozen natural orbital approximation at the CCSDT and CCSDT(Q) levels works well only if the thresholds for discarding least occupied natural orbitals are very tight (significantly tighter than the thresholds recommended for molecular correlation energies in the original work of Rolić and Kállay, *J. Chem. Phys.* **2011**, *134*, 124111), making the performance gains quite limited. The interaction energy contributions through CCSDT(Q) are both a necessity and a bottleneck in the construction of top-accuracy interaction potentials and further improvements in the efficiency of high-order coupled-cluster calculations will be of great help.



I. INTRODUCTION

The “gold-standard” coupled-cluster approach^{1,2} with single, double, and noniterative triple excitations [CCSD(T)]³ has proven immensely successful in generating highly accurate noncovalent interaction energies. In particular, the databases of CCSD(T)-level interaction energies are an extremely valuable tool for the benchmarking and tuning of more approximate methods.^{4–8} Due to the increase of the available computational power and the advances in overcoming basis set incompleteness effects at the CCSD(T) level (such as complete-basis-set (CBS) extrapolations,^{9,10} bond functions,¹¹ and explicitly correlated F12

approaches^{12,13}), the CCSD(T)/CBS interaction energies can be obtained so precisely that higher-order coupled-cluster excitations can become the leading contribution to the residual errors of interaction energies (along with the corrections for core–core and core–valence correlation, relativistic effects, and post-Born–Oppenheimer terms). Therefore, higher-order coupled-cluster interaction energy contributions, obtained using CCSDT,¹⁴ CCSDT(Q),^{15,16} and/or CCSDTQ,^{17,18} have

Received: April 22, 2014



Table 1. Interaction Energy Contributions (in cm^{-1}) for the He–He, He–H₂, and H₂–H₂ Complexes in Their Respective van der Waals Minimum Geometries

basis set	CCSD(T)	δ_T	$\delta_{(Q)}$	$\delta_{T+(Q)}$	δ_Q
He–He					
6-31G*(0.25)	6.7387	−0.0027	0.0000	−0.0027	0.0000
6-31G**(0.25,0.15)	−1.5310	−0.1273	−0.0073	−0.1346	−0.0036
aDZ	−2.9223	−0.1891	−0.0042	−0.1933	−0.0021
aTZ	−5.7698	−0.2124	−0.0066	−0.2190	−0.0031
aQZ	−6.4344	−0.2124	−0.0074	−0.2198	−0.0034
aSZ	−6.8306	−0.2108	−0.0077	−0.2185	−0.0032
a6Z	−7.1316	−0.2117	−0.0080	−0.2197	−0.0034
He–H ₂					
6-31G*(0.25)	10.9419	−0.0120	0.0010	−0.0110	0.0007
6-31G**(0.25,0.15)	−6.2201	−0.2826	−0.0020	−0.2846	0.0008
aDZ	−6.8620	−0.4035	−0.0096	−0.4130	−0.0027
aTZ	−9.1984	−0.4040	−0.0140	−0.4180	−0.0034
aQZ	−9.9053	−0.3947	−0.0159	−0.4106	−0.0036
aSZ	−10.2463	−0.3870	−0.0165	−0.4035	−0.0037
a6Z	−10.4263	−0.3830			
H ₂ –H ₂					
6-31G*(0.25)	28.2768	−0.0747	−0.0238	−0.0985	−0.0144
6-31G**(0.25,0.15)	−22.8972	−0.9264	−0.0736	−1.0000	−0.0342
aDZ	−24.3163	−1.1896	−0.0901	−1.2797	−0.0333
aTZ	−35.0745	−1.2392	−0.1006	−1.3398	−0.0306
aQZ	−37.3259	−1.1885	−0.1059	−1.2944	−0.0307
aSZ	−37.8178	−1.1550	−0.1072	−1.2622	
a6Z	−38.0588	−1.1364			

attracted significant attention in the last year, culminated by the extensive and illuminating studies of Hobza and collaborators.^{19–21} These authors have constructed the A24 database of 24 small complexes for which the benchmark interaction energies contain, in addition to an estimate of the CCSD(T)/CBS interaction energy, the corrections due to core–core and core–valence correlation, relativistic effects, and post-CCSD(T) terms. The latter correction was calculated²⁰ at the CCSDT(Q) level using a 6-31G**(0.25,0.15) basis set. Furthermore, by performing calculations up to CCSDTQP (for the smallest dimers, even full CI), Hobza et al. have shown²¹ that the CCSDT(Q) interaction energies are already converged to about 0.5 cm^{-1} or better. This conclusion is in line with earlier observations of a remarkable agreement between the interaction energies predicted by CCSDT(Q) and CCSDTQ.^{22–24} On the other hand, the inclusion of full, iterative (CCSDT) triples alone does not provide a consistent improvement over CCSD(T); the inclusion of quadruple excitations at least at the noniterative (Q) level is crucial.¹⁹

The work of Hobza and collaborators used a small 6-31G**(0.25,0.15) basis for the post-CCSD(T) corrections to make the calculations feasible (only ref 19 presented results in the slightly larger aDZ \equiv aug-cc-pVDZ basis as well). Due to the steep computational scaling of the methods involved (N^8 for CCSDT and N^9 for CCSDT(Q)), it is not possible to obtain fully converged CBS estimates of the post-CCSD(T) contributions except for small four-electron systems (the latter are, however, not quite representative of larger systems as will be shown below). However, there is some compelling evidence that the basis set convergence of the post-CCSD(T) terms is no faster than for the CCSD(T) interaction energy, so that basis sets such as 6-31G**(0.25,0.15) or aDZ are likely inadequate. In particular, for the argon dimer, the $\delta_T = \text{CCSDT} - \text{CCSD(T)}$ and $\delta_{(Q)} = \text{CCSDT(Q)} - \text{CCSDT}$ contributions to the near-

minimum interaction energy, computed in the aDZ basis set, amount to only −4% (that is, the sign is wrong) and 32%, respectively, of the values computed in the largest basis sets feasible, aug-cc-pV(5 + d)Z and aug-cc-pV(Q + d)Z, respectively.²⁴

In view of the slow basis set convergence of δ_T and $\delta_{(Q)}$ for the argon dimer, and of a similar pattern observed for other weakly interacting dimers investigated in our groups,^{25,26} we suspect that the largest remaining error in the interaction energies of the A24 database might be due to the, possibly severe, basis set incompleteness effects in δ_T and $\delta_{(Q)}$. On a more fundamental level, we would like to gauge the utility of adding a small-basis CCSDT(Q) correction to the interaction energy and to establish the basis set requirements for an accurate description of this correction. Additionally, as for some systems δ_T and $\delta_{(Q)}$ cancel each other to a large extent, we want to study the basis set convergence of the overall CCSDT(Q)–CCSD(T) contribution and compare it with the properties of δ_T and $\delta_{(Q)}$ separately. As our working hypothesis is that small-basis δ_T and $\delta_{(Q)}$ corrections may be qualitatively inaccurate, we restrict our considerations to small systems for which the CCSDT(Q) calculations in at least the aTZ basis are feasible. For some smaller complexes, we will be able to run CCSDT and CCSDT(Q) in bases as large as aSZ and aQZ, respectively, and for the four-electron systems (where CCSDTQ is equivalent to FCI) still larger basis sets can be utilized even at the CCSDTQ level.

The example dimers considered in this work include nearly all complexes for which post-CCSD(T) interaction energy corrections have been previously computed using any basis set larger than aDZ as well as a number of dimers (e.g., NH₃–NH₃ and CH₄–CH₄) for which such calculations have not been done before. Specifically, we consider the four-electron systems He–He,²⁷ He–H₂,²⁸ and H₂–H₂,²⁹ rare gas dimers Ne–Ne³⁰ and Ar–Ar,^{23,24} the complexes H₂–CO,^{25,31} H₂–HCl,³² He–C₃,²⁶

Table 2. Interaction Energy Contributions (in cm^{-1}) for the He–LiH and LiH–LiH Complexes^a

basis set	CCSD(T)	δ_{T}	$\delta_{(\text{Q})}$	$\delta_{\text{T}+(\text{Q})}$	δ_{Q}
He–LiH					
6-31G*(0.25)	–12.718	–0.522	–0.030	–0.552	–0.007
6-31G**(0.25,0.15)	–31.749	–1.157	–0.102	–1.258	–0.041
aDZ	–98.359	–1.063	–0.054	–1.117	–0.025
aTZ	–139.102	–1.120	–0.066	–1.185	–0.028
aQZ	–147.367	–1.105	–0.071	–1.176	–0.029
aSZ	–148.140	–1.080	–0.072	–1.152	
LiH–LiH					
6-31G*(0.25)	–13872.24	–28.86	–1.08	–29.94	–0.16
6-31G**(0.25,0.15)	–14473.47	–31.54	–0.16	–31.70	0.39
aDZ	–15851.05	–29.48	–1.17	–30.65	–0.21
aTZ	–16689.70	–35.06	–2.51	–37.56	–0.56
aQZ	–16818.09	–33.20	–2.83	–36.03	–0.50
aSZ	–16854.59	–31.94			

^aThe lithium 1s core electrons have not been correlated.

Ar–HF,^{33–35} Ne–HF, and H₂–HF under investigation in our groups, and the N₂–N₂,³⁶ H₂O–H₂O,³⁷ He–LiH, LiH–LiH, Ar–CH₄, HF–HF, NH₃–NH₃, and CH₄–CH₄ dimers. Compared to the post-CCSD(T) interaction energies beyond the aDZ basis available in the literature, to our knowledge, we only exclude the systems with less than four electrons for which FCI calculations are quite straightforward^{38,39} and the alkaline earth metal dimers which are four-electron systems within the frozen-core approximation (this approximation, however, works poorly in this case, and the accuracy of the CCSD(T) interaction energies is particularly low; see ref 40 and references therein).

II. COMPUTATIONAL DETAILS

The majority of the calculations are performed at the global-minimum geometries, which are taken from the literature to facilitate comparison with previous studies. For the H₂O–H₂O and N₂–N₂ complexes, where the global minimum has relatively low symmetry, we have performed additional calculations at a more symmetric configuration. The geometries of all studied complexes are described in Table SI in the Supporting Information (which also lists the corresponding Cartesian coordinates). We have employed the 6-31G*(0.25)^{41,42} and 6-31G**(0.25,0.15) basis sets popularized by Hobza as well as the singly augmented Dunning sequence aXZ, X = D,T,Q,S,6.^{43,44} Note that for the second-row atoms (Cl and Ar) the regular aXZ sets were used, not the aug-cc-pV(X + d)Z sets⁴⁵ as the latter, while slightly larger, have not shown any consistent improvement for interaction energies. Moreover, the aXZ, X = D,T,Q, bases for lithium are the original sets from ref 46 (as listed on the Basis Set Exchange Web site⁴⁷), not the revised sets of ref 48. All interaction energies include the counterpoise (CP) correction for basis set superposition error. Contrary to refs 19–21, which correlated all electrons, our δ_{T} and $\delta_{(\text{Q})}$ corrections will be obtained within the frozen-core approximation. This approach corresponds to treating the corrections to CCSD(T)/CBS for higher-order excitations and for core–core and core–valence correlation as additive, as implicitly assumed in nearly all studies to date. We have, however, tested this assumption on the “worst-case-scenario” examples involving the LiH molecule (the frozen-core approximation generally works poorly for interactions involving alkali or alkaline earth metals⁴⁰) as well as on the much less critical example of Ne–Ne.

All CCSDT(Q) and CCSDTQ calculations have been carried out using the MRCC program,^{49–51} either stand-alone or

interfaced to MOLPRO2012.1.⁵² The largest CCSDT calculations have used the CFOUR code.⁵³ For a few systems, we have carried out additional CCSDT and CCSDT(Q) calculations with the virtual orbital space truncated according to occupations of MP2 natural orbitals;⁵⁴ see section III.B for details.

III. NUMERICAL RESULTS AND DISCUSSION

The basis set convergence of the δ_{T} , $\delta_{(\text{Q})}$, and δ_{Q} = CCSDTQ–CCSDT(Q) contributions to the near-minimum interaction energies of four-electron dimers He–He, He–H₂, and H₂–H₂ is presented in Table 1. The corresponding values of the CCSD(T) interaction energies are shown to illustrate the significance of the post-CCSD(T) contribution.

The first observation from Table 1 affirms that for the four-electron systems quadruple excitations are only of minor importance. At the largest basis sets listed, the $\delta_{(\text{Q})}$ term does not exceed 10% of δ_{T} , and the correction for full CCSDTQ (equivalent to full CI in this case) is still smaller by a factor of at least 2. While the basis set convergence of $\delta_{(\text{Q})}$ might not be of critical importance for these systems, it is quite slow, with the aDZ value amounting to 54–84% of the aSZ result. Fortunately, the dominating δ_{T} contribution exhibits faster basis set convergence, with the aDZ basis reproducing 89–105% of the a6Z value. The 6-31G**(0.25,0.15) basis accidentally gives a very good value of $\delta_{(\text{Q})}$ for He–He, but it is quite inaccurate for the remaining two dimers as well as for the dominating δ_{T} correction, for which the errors are about four times that of aDZ. As could be expected, the still smaller 6-31G*(0.25) basis (equivalent to just 6-31G for hydrogen and helium) is completely unsuitable for dispersion-dominated interactions.

The δ_{T} , $\delta_{(\text{Q})}$, and δ_{Q} results for the He–LiH and LiH–LiH dimers are presented in Table 2. As the 1s lithium electrons were not correlated (we will examine the all-electron δ_{T} and $\delta_{(\text{Q})}$ contributions for these systems in section III.C), these are also four-electron systems and even CCSDTQ \equiv FCI can be computed in reasonably large basis sets. Similar to the other four-electron systems (Table 1), the δ_{T} effect strongly dominates over $\delta_{(\text{Q})}$ and δ_{Q} . However, the overall post-CCSD(T) effects are very minor (below 1%) for He–LiH and LiH–LiH and the convergence of the δ_{T} contribution is generally quite good even though it becomes monotonic only from the aTZ level on. Even the 6-31G**(0.25,0.15) set, while inadequate for the minor $\delta_{(\text{Q})}$ contribution, recovers 98–107% of the aSZ result for δ_{T} ; the 6-31G*(0.25) basis is not much worse for the LiH–LiH complex.

Table 3. Interaction Energy Contributions (in cm^{-1}) for the Ne–Ne, Ar–Ar, and He–C₃ Complexes in Their Respective Global Minima^a

basis set	CCSD(T)	δ_T	$\delta_{(Q)}$	$\delta_{T+(Q)}$	δ_Q
Ne–Ne					
6-31G*(0.25)	–16.8056	–0.1078	0.1276	0.0199	–0.0129
aDZ	–3.0539	–0.0407	0.0326	–0.0080	–0.0074
aTZ	–16.6713	–0.2694	–0.0269	–0.2963	
aQZ	–22.9583	–0.3728	–0.0704	–0.4432	
aSZ	–25.7525	–0.4072			
Ar–Ar					
6-31G*(0.25)	–12.940	0.005	0.774	0.780	–0.148
aDZ	–24.467	–0.054	–0.547	–0.601	0.104
aTZ	–66.090	0.824	–1.206	–0.382	
aQZ	–81.579	1.052	–1.612	–0.560	
aSZ	–90.097	1.200			
He–C ₃					
6-31G*(0.25)	19.023	0.075	–0.139	–0.064	0.007
6-31G**(0.25,0.15)	–12.309	0.084	–0.295	–0.211	0.029
aDZ	–15.499	0.189	–0.579	–0.390	0.021
aTZ	–22.628	0.370	–0.649	–0.279	
aQZ	–24.485	0.432			

^aThe Ar–Ar and He–C₃ results (except for the 6-31G*(0.25) and 6-31G**(0.25,0.15) ones) are taken from Refs 24 and 26, respectively, and most of the Ne–Ne results have been obtained in Ref 30. The 6-31G*(0.25) and 6-31G**(0.25,0.15) bases are identical when no hydrogen or helium atoms are present.

Table 4. Interaction Energy Contributions (in cm^{-1}) for the Ne–HF, Ar–HF, and Ar–CH₄ Complexes in Their Respective Global Minima

basis set	CCSD(T)	δ_T	$\delta_{(Q)}$	$\delta_{T+(Q)}$	δ_Q
Ne–HF					
6-31G*(0.25)	–49.798	–0.054	0.191	0.137	–0.025
6-31G**(0.25,0.15)	–88.120	–0.123	0.102	–0.021	–0.038
aDZ	–55.225	–0.252	–0.303	–0.555	0.028
aTZ	–72.406	–0.928	–0.610	–1.538	
aQZ	–79.262	–0.938	–0.736	–1.674	
aSZ	–86.125	–0.911			
Ar–HF					
6-31G*(0.25)	–21.707	0.488	0.067	0.555	–0.100
6-31G**(0.25,0.15)	–82.679	–0.083	0.030	–0.052	–0.030
aDZ	–94.051	–1.288	–0.849	–2.137	–0.013
aTZ	–185.643	–0.889	–1.977	–2.866	
aQZ	–207.969	–0.477	–2.331	–2.808	
aSZ	–212.617	–0.238			
Ar–CH ₄					
6-31G*(0.25)	–22.898	–0.136	0.241	0.105	–0.111
6-31G**(0.25,0.15)	–50.559	–0.469	0.454	–0.015	–0.121
aDZ	–82.669	–0.506	–1.020	–1.526	
aTZ	–119.564	0.522	–1.924	–1.402	
aQZ	–132.986	0.880			

As seen in Table 2, the frozen-core He–LiH and LiH–LiH systems are not particularly demanding when it comes to the basis set selection for the post-CCSD(T) corrections. It is, unfortunately, not the case for most of the larger dimers, as illustrated in Tables 3 (Ne–Ne, Ar–Ar, He–C₃), 4 (Ne–HF, Ar–HF, Ar–CH₄), 5 (H₂–HF, H₂–HCl, H₂–CO), 6 (H₂O–H₂O, N₂–N₂), and 7 (HF–HF, NH₃–NH₃, CH₄–CH₄). In the most extreme cases of the H₂O–H₂O and HF–HF complexes, the aDZ basis either fails to capture the correct sign of the $\delta_{T+(Q)} \equiv \delta_T + \delta_{(Q)}$ sum or underestimates $\delta_{T+(Q)}$ several times, and the 6-31G*(0.25) and 6-31G**(0.25,0.15) bases are still worse. Tables 3–7 also show that the $\delta_{(Q)}$ correction tends to converge from above to a negative CBS result, while the δ_T term can have

either sign and mostly approaches its CBS limit from below although the convergence often becomes monotonic only from the aTZ level on.

To facilitate the analysis of the overall performance of different basis sets for δ_T and $\delta_{(Q)}$, we will partition the complexes into four groups as follows.

- *Four-valence-electron dimers:* He–He, He–H₂, H₂–H₂, He–LiH, and LiH–LiH
- *Nonpolar–nonpolar dimers:* Ne–Ne, Ar–Ar, He–C₃, H₂–CO, Ar–CH₄, N₂–N₂ (2 orientations), and CH₄–CH₄
- *Polar–nonpolar dimers:* H₂–HF, H₂–HCl, Ne–HF, Ar–HF

Table 5. Interaction Energy Contributions (in cm⁻¹) for the H₂–HF, H₂–HCl, and H₂–CO Complexes in Their Respective near-Global Minimum Geometries^a

basis set	CCSD(T)	δ_T	$\delta_{(Q)}$	$\delta_{T+(Q)}$	δ_Q
H ₂ –HF					
6-31G*(0.25)	233.239	–0.218	0.242	0.024	–0.128
6-31G**(0.25,0.15)	–93.337	–2.545	0.326	–2.219	–0.105
aDZ	–296.317	–3.612	0.156	–3.456	–0.130
aTZ	–370.023	–3.937	–0.458	–4.395	
aQZ	–387.151	–3.450	–0.646	–4.096	
aSZ	–391.676	–3.181			
H ₂ –HCl					
6-31G*(0.25)	118.346	–0.700	–0.026	–0.726	–0.036
6-31G**(0.25,0.15)	–57.089	–2.460	0.182	–2.278	–0.096
aDZ	–122.888	–2.722	–1.123	–3.845	–0.010
aTZ	–183.349	–2.426	–1.774	–4.200	
aQZ	–199.442	–1.927	–1.991	–3.919	
aSZ	–202.779	–1.683			
H ₂ –CO					
6-31G*(0.25)	–15.706	–0.842	–0.095	–0.937	–0.109
6-31G**(0.25,0.15)	–89.638	–1.536	–0.532	–2.068	0.021
aDZ	–73.160	–1.850	–0.856	–2.706	–0.045
aTZ	–86.503	–1.469	–1.131	–2.600	
aQZ	–89.257	–1.303	–1.205	–2.508	
aSZ	–90.210	–1.207			

^aMost of the H₂–CO results were obtained in the course of the work on refs 31 and 25.**Table 6. Interaction Energy Contributions (in cm⁻¹) for the Water and Nitrogen Dimers^a**

basis set	CCSD(T)	δ_T	$\delta_{(Q)}$	$\delta_{T+(Q)}$	δ_Q
H ₂ O–H ₂ O (global minimum)					
6-31G*(0.25)	–1386.50	5.65	–0.71	4.93	–0.08
6-31G**(0.25,0.15)	–1475.05	5.52	–1.50	4.02	0.11
aDZ	–1530.34	2.36	–3.24	–0.88	
aTZ	–1665.60	1.51	–4.95	–3.44	
aQZ	–1726.71	2.02			
H ₂ O–H ₂ O (C _{2v})					
6-31G*(0.25)	–902.55	3.89	1.04	4.93	–0.18
6-31G**(0.25,0.15)	–1021.62	3.75	0.27	4.03	–0.07
aDZ	–1026.27	1.14	–0.73	0.41	–0.08
aTZ	–1107.16	0.30	–2.26	–1.96	
aQZ	–1141.02	0.55	–2.44 ^b	–1.89 ^b	
N ₂ –N ₂ (global minimum)					
6-31G*(0.25)	–85.401	2.479	–4.448	–1.969	1.462
aDZ	–79.561	2.855	–4.852	–1.997	
aTZ	–97.162	3.622	–5.456	–1.835	
aQZ	–102.468	3.857			
N ₂ –N ₂ (D _{2h})					
6-31G*(0.25)	–50.828	1.877	–3.050	–1.173	1.372
aDZ	–49.207	2.294	–3.540	–1.246	1.122
aTZ	–69.069	3.267	–4.391	–1.124	
aQZ	–72.826	3.531			

^aEach complex is presented in two geometries: the global-minimum one (less symmetric) and a highly symmetric structure that is only a radial minimum. The 6-31G*(0.25) and 6-31G**(0.25,0.15) bases are identical when no hydrogen or helium atoms are present. ^bResult obtained in the FNO approximation (section III.B).

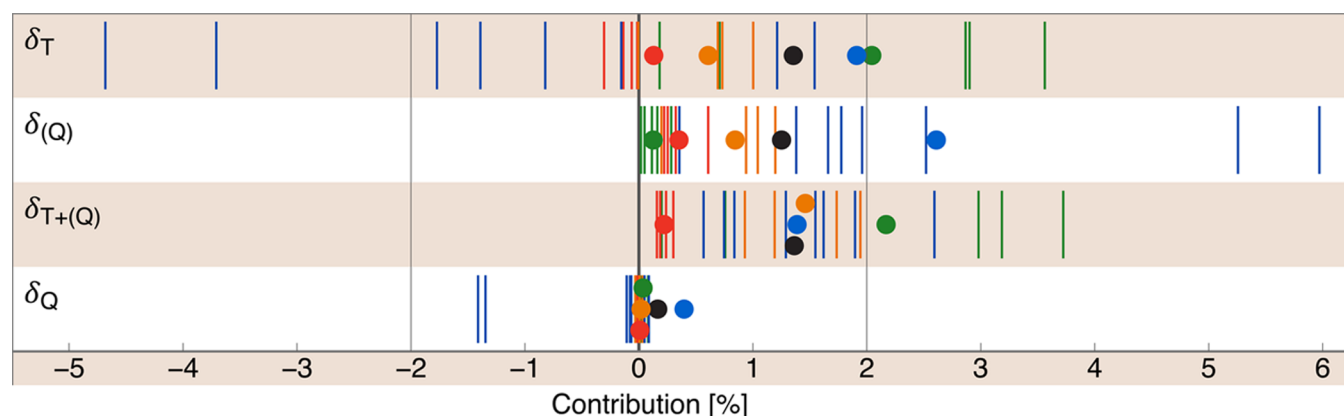
- *Polar–polar dimers:* H₂O–H₂O (2 orientations), HF–HF, and NH₃–NH₃

The first important factor that differentiates between complexes belonging to different groups is the relative importance of the δ_T and $\delta_{(Q)}$ interaction energy contributions compared to the leading CCSD(T) term. There exists ample literature concerning the methodology (basis sets, counterpoise

correction or lack thereof, CBS extrapolations, and different explicitly correlated CCSD(T)-F12 variants) of obtaining precise and reliable CCSD(T)/CBS interaction energy estimates,^{55–59} and it is not our aim to generate ultra-accurate values of the CCSD(T)/CBS term. Therefore, we will select the most accurate CCSD(T)/CBS benchmark values from the literature if available for a given geometry: if not, we will obtain

Table 7. Interaction Energy Contributions (in cm^{-1}) for the HF, NH_3 , and CH_4 Dimers in Their Respective near-Minimum Configurations

basis set	CCSD(T)	δ_T	$\delta_{(Q)}$	$\delta_{T+(Q)}$	δ_Q
HF–HF					
6-31G*(0.25)	–1316.83	6.17	–0.98	5.18	0.15
6-31G**(0.25,0.15)	–1247.58	9.23	–3.21	6.02	0.32
aDZ	–1391.01	4.52	–2.15	2.37	
aTZ	–1504.08	0.91	–3.48	–2.57	
aQZ	–1576.91	0.52			
NH_3 – NH_3					
6-31G*(0.25)	–791.26	1.13	–1.01	0.12	0.00
6-31G**(0.25,0.15)	–927.82	–0.24	–1.25	–1.49	0.00
aDZ	–928.12	–0.47	–3.72	–4.19	
aTZ	–1050.72	1.56	–5.80	–4.24	
aQZ	–1082.74	2.60			
CH_4 – CH_4					
6-31G*(0.25)	–23.384	–0.124	–0.992	–1.116	–0.043
6-31G**(0.25,0.15)	–102.161	–1.008	–1.170	–2.178	0.124
aDZ	–143.081	–1.403	–1.888	–3.292	
aTZ	–174.241	–0.212	–2.893	–3.105	
aQZ	–181.931				

**Figure 1.** δ_T , $\delta_{(Q)}$, $\delta_{T+(Q)}$, and δ_Q interaction energy contributions as percentages of the CCSD(T)/CBS interaction energy. The four-valence-electron, nonpolar–nonpolar, polar–nonpolar, and polar–polar dimers are represented by green, blue, orange, and red symbols, respectively. The circles represent the unsigned averages of the individual values, with the black circle showing the unsigned average for all complexes.

CCSD(T)/CBS estimates by the standard X^{-3} extrapolation⁹ of the results computed in the aQZ and aSZ bases (aSZ and a6Z for a few systems). The complete set of benchmark CCSD(T)/CBS interaction energies is listed in Table SI in the Supporting Information.

The values of the δ_T and $\delta_{(Q)}$ corrections, as well as of their $\delta_{T+(Q)}$ sum and of the post-CCSDT(Q) term δ_Q , as percentages of the CCSD(T)/CBS interaction energy are presented in Figure 1. In this figure (its design is inspired by refs 58 and 60), the four-valence-electron, nonpolar–nonpolar, polar–nonpolar, and polar–polar dimers (see the partitioning above) are represented by green, blue, orange, and red symbols, respectively. Note that, as the CCSD(T)/CBS interaction energy is negative, a positive percentage means a negative interaction energy contribution. For individual systems, the actual percentages (positive or negative) are shown; however, all averages pertain to unsigned percentages (in other words, we show the mean unsigned relative deviations of, for example, CCSD(T) + δ_T relative to CCSD(T)). In Figure 1, the values of δ_T and $\delta_{(Q)}$ are chosen as (supposedly) the most accurate estimates of these corrections available from Tables 1–7, that is, the results of the X^{-3} extrapolation using the two largest- X aXZ values (note that the largest X available is typically

higher for δ_T than for $\delta_{(Q)}$). In the particular case of the a(D,T)Z extrapolation, it was observed for the CCSD(T)–MP2 interaction energy contribution⁶¹ that the inadequacy of the aDZ basis often makes the extrapolated result inferior to the calculated aTZ value. In order to check whether the similar observation holds for δ_T and $\delta_{(Q)}$, we compared the performance of the a(D,T)Z and aTZ estimates of these corrections to the benchmark values obtained by the a(T,Q)Z extrapolation or higher: this comparison was done on 18 systems (all but CH_4 – CH_4 , HF–HF, and Ar–HF, excluded for the reasons described below) for δ_T and 13 systems for $\delta_{(Q)}$. For δ_T , the overall accuracy of the a(D,T)Z and aTZ results was about the same (a mean unsigned relative error (MURE) of 25.3% versus 28.3%), however, the a(D,T)Z approach performed much better on the nonpolar–nonpolar dimers (a MURE of 10.6% versus 30.0% for bare aTZ) so we chose the δ_T /a(D,T)Z result as benchmark for the methane dimer (the only complex for which CCSDT/aQZ could not be calculated). In the case of $\delta_{(Q)}$, the extrapolated a(D,T)Z estimates performed significantly better (a MURE of 10.8% versus 24.5% for aTZ) so we adopted the $\delta_{(Q)}$ /a(D,T)Z values as benchmarks when larger-basis results are not available. All benchmark estimates of the post-CCSD(T) corrections are

listed in Table SII in the Supporting Information. On the average, the δ_T and $\delta_{(Q)}$ effects amount to, respectively, 2.04% and 0.13% for four-valence-electron dimers, 1.91% and 2.61% for nonpolar–nonpolar dimers, 0.61% and 0.84% for polar–nonpolar dimers, 0.13% and 0.35% for polar–polar dimers, and 1.36% and 1.25% overall. The magnitude of the post-CCSD(T) effects can be contrasted with the $\delta_{(T)} = \text{CCSD(T)} - \text{CCSD}$ interaction energy contribution which amounts, on the average, to 17.4% of the CCSD(T) benchmark (the MP2, CCSD, and $\delta_{(T)}$ interaction energy terms are given in Table SII for completeness). As evident from Figure 1, for the nonpolar–nonpolar dimers the δ_T and $\delta_{(Q)}$ contributions cancel out partially but systematically while for the polar–nonpolar and polar–polar dimers no such cancellation is present. As mentioned above, the δ_T contribution strongly dominates for the four-valence-electron complexes. The benchmark values for the nonperturbative quadruples correction δ_Q were computed in the largest bases available in Tables 1–7 without any extrapolation. As shown in Figure 1, the δ_Q correction is generally very small. The only exceptions are the two geometries of the nitrogen dimer: interactions between triply bonded molecules such as N_2 or CO are known to require particularly high orders of Møller–Plesset perturbation theory (MP n) or high-order coupled-cluster excitations.⁶² One should note that the same two N_2 – N_2 structures are responsible for the largest percentage contributions to δ_T (the two leftmost lines in Figure 1) and $\delta_{(Q)}$ (the two rightmost lines); however, these two terms cancel out to a large extent.

When the δ_T term can be computed in a larger basis than $\delta_{(Q)}$ (which is the case for all complexes but four, cf. Tables 1–7), there are two sensible ways of estimating the benchmark value of the overall $\delta_{T+(Q)}$ contribution: as a sum of the δ_T and $\delta_{(Q)}$ benchmarks (obtained in separate extrapolations) or via a single extrapolation of the $\delta_{T+(Q)}$ term computed in the two largest basis sets for which the $\delta_{(Q)}$ value is available. While the first approach is formally closer to CBS (it includes the δ_T contribution in larger basis sets), the second approach is preferable if any cancellation between the basis set incompleteness errors at the δ_T and $\delta_{(Q)}$ levels occurs. For the 12 complexes for which the CCSDT/a5Z and CCSDT(Q)/aQZ calculations are possible, the single-extrapolation $\delta_{T+(Q)}/a(T,Q)Z$ estimate differs from the separate-extrapolations $\delta_T/a(Q,5)Z + \delta_{(Q)}/a(T,Q)Z$ one by an average of 5.2%. As the results in Tables 1–7 show that the $\delta_{T+(Q)}$ sum exhibits faster basis set convergence than its δ_T and $\delta_{(Q)}$ components for most but not all dimers (in other words, error cancellation is likely but not at all systematic), it is not clear which one of these estimates is more accurate but the difference is minor. For smaller bases, an additional factor is the poor performance of the $\delta_{T+(Q)}/a(D,T)Z$ extrapolation—comparison with the $a(T,Q)Z$ (or higher) benchmark data for 13 complexes shows a MURE of 19.9% for $\delta_{T+(Q)}/a(D,T)Z$ and 13.5% for $\delta_{T+(Q)}/aTZ$. Therefore, computing the δ_T part from the $a(T,Q)Z$ extrapolation should be highly preferred over either $a(D,T)Z$ or aTZ : not only the basis set is larger but the extrapolation is clearly beneficial. Consequently, we decided to compute our $\delta_{T+(Q)}$ benchmark values, presented in Figure 1, using the separate-extrapolations approach whenever possible (for the methane dimer, we chose the $\delta_{T+(Q)}/a(D,T)Z$ value as benchmark).

While, as expected, the overall $\delta_{T+(Q)}$ interaction energy correction is quite minor (0.16% to 3.72% of the CCSD(T)/CBS value, or a maximum of 2.68% if four-valence-electron dimers are neglected), it nevertheless becomes important in top-accuracy calculations. Therefore, one should ask how well the δ_T , $\delta_{(Q)}$, and $\delta_{T+(Q)}$ corrections are recovered (relative to the benchmark

values of these terms from Figure 1) when computed in a smaller basis set. To answer this question, we computed the MURE values of the δ_T correction obtained in the 6-31G*(0.25), 6-31G**(0.25,0.15), aDZ, aTZ, and aQZ basis sets, and of the $\delta_{(Q)}$ and $\delta_{T+(Q)}$ terms computed in the 6-31G*(0.25), 6-31G**(0.25,0.15), aDZ, and aTZ basis sets, relative to the benchmark values from Figure 1. The results are presented in Figures 2, 3,

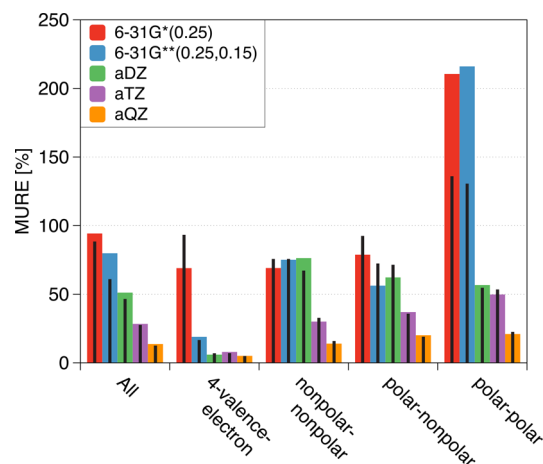


Figure 2. Performance of different basis sets in the recovery of the δ_T benchmark (extrapolated from the two largest-basis results in Tables 1–7) displayed as mean unsigned relative errors (MURE). The black lines represent the median unsigned relative errors (MeURE). The CH_4 – CH_4 , HF – HF , and Ar – HF complexes are excluded from the MURE and MeURE calculation due to the accidental smallness of the benchmark value (see text for details).

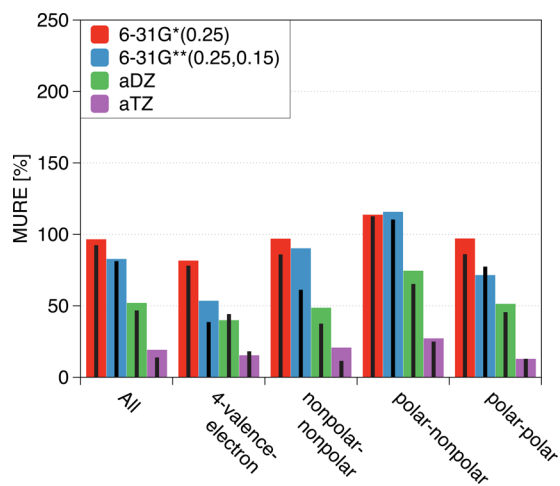


Figure 3. Performance of different basis sets in the recovery of the $\delta_{(Q)}$ benchmark (extrapolated from the two largest-basis results in Tables 1–7) displayed as mean unsigned relative errors (MURE). The black lines represent the median unsigned relative errors (MeURE).

and 4 for δ_T , $\delta_{(Q)}$, and $\delta_{T+(Q)}$, respectively. Note that a MURE of over 100%, obtained for some classes of systems at the 6-31G*(0.25) and 6-31G**(0.25,0.15) basis sets, means that it is better to neglect a contribution completely (a 100% error) than to compute it using such a small basis set. It should also be noted that even our largest-basis estimates are in general not converged, as indicated by the overall MURE values of 13.7, 19.2, and 15.8% obtained for δ_T/aQZ , $\delta_{(Q)}/aTZ$, and $\delta_{T+(Q)}/aTZ$, respectively.

There are three systems for which the benchmark δ_T value is particularly close to zero so that the relative errors are greatly

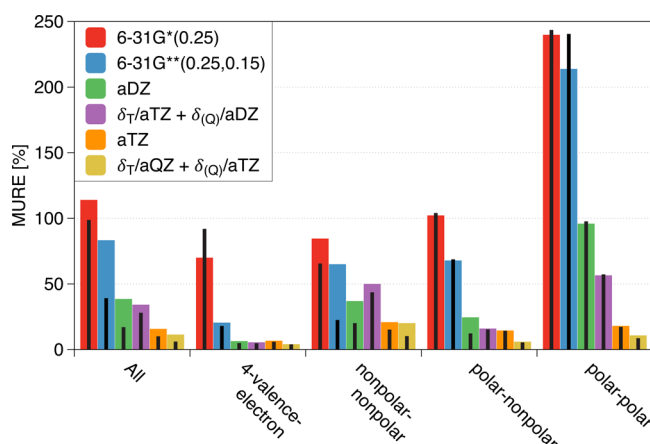


Figure 4. Performance of different basis sets in the recovery of the $\delta_{T+(Q)}$ benchmark (computed as a sum of the δ_T and $\delta_{(Q)}$ benchmarks obtained in separate extrapolations) displayed as mean unsigned relative errors (MURE). The black lines represent the median unsigned relative errors (MeURE).

exaggerated (no similar issues arise for $\delta_{(Q)}$ and $\delta_{T+(Q)}$). These systems are $\text{CH}_4\text{--CH}_4$, HF--HF , and Ar--HF , with the two largest-basis computed results, $\delta_T/a(X-1)Z$ and δ_T/aXZ , and the extrapolated $\delta_T/(a(X-1)Z, aXZ)$ value amounting to $(-1.403, -0.212, 0.289) \text{ cm}^{-1}$, respectively, for $\text{CH}_4\text{--CH}_4$ (note that the correction is particularly poorly converged in this case because the aQZ value is not available), $(0.91, 0.52, 0.24) \text{ cm}^{-1}$ for HF--HF , and $(-0.477, -0.238, 0.013) \text{ cm}^{-1}$ for Ar--HF . Thus, the unsigned relative errors of even the largest-basis computed results (with respect to the extrapolated value) exceed 100% for these three systems. Therefore, to make the statistics in Figure 2 meaningful, the $\text{CH}_4\text{--CH}_4$, HF--HF , and Ar--HF complexes had to be excluded. Even after this exclusion, the 6-31G*(0.25) and 6-31G**(0.25,0.15) bases lead to δ_T errors of over 50% in all cases except the four-valence-electron dimers in the latter set. With the same exception of the four-valence-electron systems, the aDZ basis set is also inadequate for the δ_T term, with the largest errors (excluding the $\text{CH}_4\text{--CH}_4$, HF--HF , and Ar--HF complexes) of 144.3, 113.9, and 104.0% obtained for Ar--CH_4 , $\text{NH}_3\text{--NH}_3$, and Ar--Ar , respectively. The percentage errors are particularly large for the polar–polar dimers: fortunately, these are also the systems where the post-CCSD(T) corrections are least important in relative terms, cf. Figure 1.

For the $\delta_{(Q)}$ contribution (Figure 3), the 6-31G*(0.25) and 6-31G**(0.25,0.15) basis sets give average errors close to 100% while the errors in the aDZ set are around 50%. Notably, the errors of $\delta_{(Q)}$ are quite similar across all classes of systems despite the vastly different relative importance of this correction (cf. Figure 1), making it particularly critical to converge this term for the nonpolar–nonpolar and polar–nonpolar dimers. The largest errors at the $\delta_{(Q)}/aDZ$ level amount to 132.0, 119.9, and 71.6% for Ne--Ne , $\text{H}_2\text{--HF}$, and $\text{H}_2\text{O--H}_2\text{O}$ (C_{2v}), respectively.

For most of the systems, if the $\delta_{(Q)}$ calculation is feasible up to the aXZ basis, the δ_T contribution can be obtained also in the $a(X+1)Z$ set. However, the statistical errors in Figure 4 indicate that the improvement of δ_T is not really beneficial unless one improves $\delta_{(Q)}$ at the same time. In fact, the $\delta_T/aTZ + \delta_{(Q)}/aDZ$ estimates perform only a little bit better than $\delta_{T+(Q)}/aDZ$ (a MURE of 34.2% versus 38.6%), and $\delta_T/aQZ + \delta_{(Q)}/aTZ$ compared to $\delta_{T+(Q)}/aTZ$ exhibits the same trend (a MURE of 11.4% versus 15.7% excluding the $\text{CH}_4\text{--CH}_4$ complex). Thus, a separate δ_T calculation using a basis set beyond the $\delta_{(Q)}$

capabilities is usually not worthwhile. The largest errors at the $\delta_{T+(Q)}/aDZ$ level amount to 162.3, 122.2, and 98.5% for HF--HF , $\text{H}_2\text{O--H}_2\text{O}$ (C_{2v}), and Ne--Ne , respectively, and clearly show that extreme caution is needed if one wishes to include post-CCSD(T) interaction energy contributions using basis sets of this size.

A. Anisotropy of the Post-CCSD(T) Contributions.

Except for the high-symmetry configurations of the $\text{H}_2\text{O--H}_2\text{O}$ and $\text{N}_2\text{--N}_2$ complexes, the discussion so far concerned only geometries near the van der Waals minima. While the importance, and slow basis set convergence, of post-CCSD(T) effects at near-minimum geometries has been clearly illustrated, one could hope that these effects exhibit low anisotropy and effectively provide a nearly spherical contribution that has minimal effects on, say, spectral lines. Unfortunately, our benchmark calculations for several different angular configurations of LiH--LiH (Table SIII in the Supporting Information), $\text{H}_2\text{--CO}$ (Table 8), and Ar--HF (Table 9) (the geometries

Table 8. Anisotropy of Various Interaction Energy Components (in cm^{-1}) for the $\text{H}_2\text{--CO}$ Complex and Two Basis Sets^a

	(0°,0°,0°)	(0°,90°,0°)	(0°,180°,0°)	(45°,45°,45°)
	aDZ			
δ_T	−0.119	−0.291	−1.736	−0.226
$\delta_{(Q)}$	−0.174	−0.309	−0.794	−0.181
$\delta_{T+(Q)}$	−0.293	−0.600	−2.531	−0.406
CCSD(T)	−48.358	−16.553	−74.655	−20.352
	aTZ			
δ_T	−0.185	−0.214	−1.368	−0.154
$\delta_{(Q)}$	−0.219	−0.373	−1.055	−0.268
$\delta_{T+(Q)}$	−0.403	−0.587	−2.423	−0.421
CCSD(T)	−53.221	−18.475	−86.638	−23.799

^aThe intermolecular distance R is set to $R = 8.0$ bohr, whereas the intramolecular ones, r_{HH} and r_{CO} , are equal to 1.4487 bohr and 2.1399 bohr, respectively. The three angles are defined in the same way as in ref 68: the Cartesian coordinates for all four configurations are given in the Supporting Information.

for all configurations are given in the Supporting Information) indicate that the post-CCSD(T) contributions can have very strong anisotropy. For different orientations presented in Tables 8, 9 and SIII, the largest-basis (nonextrapolated) $\delta_{T+(Q)}$

Table 9. Interaction Energy Components (in cm^{-1}) for the Ar--HF Complex As Functions of the Angle θ between the Lines Going from the HF Center of Mass to the Ar and H Atoms^a

θ	0°	90°	180°
	aDZ		
δ_T	−1.279	−0.267	−0.253
$\delta_{(Q)}$	−0.815	0.087	0.167
$\delta_{T+(Q)}$	−2.094	−0.180	−0.086
CCSD(T)	−98.690	−10.148	−43.114
	aTZ		
δ_T	−0.877	−0.233	0.172
$\delta_{(Q)}$	−1.920	−0.409	−0.296
$\delta_{T+(Q)}$	−2.797	−0.642	−0.124
CCSD(T)	−186.560	−53.046	−75.358

^aThe intermolecular distance R is set to $R = 6.5$ bohr and the intramolecular one r_{HF} is equal to 1.7629 bohr.

contributions constitute 0.2–0.9%, 0.8–3.2%, and 0.2–1.5% of the frozen-core CCSD(T) interaction energy for LiH–LiH, H₂–CO, and Ar–HF, respectively. Moreover, the basis set convergence varies significantly with geometry: for three of the H₂–CO configurations in Table 8, the aDZ basis set recovers 96–104% of the aTZ value for $\delta_{T+(Q)}$; however, for the remaining configuration, this percentage is 73%. The agreement between the two bases for three orientations is actually a consequence of error cancellation between the δ_T and $\delta_{(Q)}$ terms (for which the aDZ results constitute 64–147% and 68–83%, respectively, of the aTZ ones). Such a cancellation cannot be taken for granted and it does not occur for the fourth geometry. An even more striking example of error cancellation is the $\theta = 0^\circ$ orientation of the Ar–HF complex. In this case, when the basis set is changed from aDZ to aTZ, the value of δ_T decreases by 31%, the value of $\delta_{(Q)}$ increases by 136%, whereas the total $\delta_{T+(Q)}$ term increases by 34%. For $\theta = 90^\circ$, such cancellation is not present and the value of $\delta_{T+(Q)}$ for aTZ is over three times larger than for aDZ. For the test complexes and geometries presented in Tables 8, 9, and SIII, there does not appear to be a pattern of change in the anisotropy when the quality of the basis set is increased. Both absolute values of $\delta_{T+(Q)}$ and its relative values with respect to the CCSD(T) interaction energy can increase or decrease for different angular orientations of the interacting subsystems.

The results in Tables 8 and 9, and in Table SIII in the Supporting Information, show that the inclusion of a properly converged post-CCSD(T) interaction energy contribution is even more critical for the anisotropy of the potential energy surface than for the near-minimum interaction energy. The post-CCSD(T) contributions can be especially important for the relative depths of global and local minima or for the heights of energy barriers. For instance, for H₂–CO the difference of the $\delta_{T+(Q)}$ values between the (0°, 0°, 0°) and (0°, 180°, 0°) orientations is equal to 2.0 cm^{−1} while the difference of the CCSD(T) interaction energies for these geometries amounts to 33.4 cm^{−1}. Thus, the inclusion of the $\delta_{T+(Q)}$ interaction energy term changes the relative energy by 6%. Not surprisingly, the inclusion of interaction energy terms up to CCSDT(Q) proved essential for the recovery and assignment of the experimental high-resolution infrared spectrum of *ortho*-H₂–CO.^{25,31}

B. MP2 Frozen Natural Orbital Approximation. In view of the highly unfavorable scaling of high-order coupled-cluster methods with the number of virtual orbitals it is desirable to introduce approximations that reduce this number without a significant adverse effect on the interaction energies. Several approaches to restrict the virtual space have been proposed.^{63,64} In this work, we utilize the MP2 frozen natural orbital (FNO) approach⁶⁵ as implemented in CCSDT and CCSDT(Q) by Rolik and Kállay.⁵⁴ In the MP2 FNO method, the natural orbitals are obtained as eigenvectors of the first-order Møller–Plesset density matrix. The corresponding eigenvalues, that is, the natural orbital occupation numbers, are then sorted and the orbitals with sufficiently low occupations can be removed from the virtual space with little error. In the implementation of ref 54, a cumulative threshold ϵ_{FNO} is employed.⁶⁶ Specifically, natural orbitals are added to the virtual space, in the order of decreasing occupation numbers, until the cumulative occupation of all included orbitals (occupied and virtual) exceeds ϵ_{FNO} times the number of electrons. The remaining natural orbitals are removed from further consideration (care is taken to avoid splitting degenerate sets of orbitals). Rolik and Kállay⁵⁴ investigated the accuracy of the δ_T and $\delta_{(Q)}$ contributions to molecular energies and heats of formation as a function of ϵ_{FNO} and recommended a

threshold of 0.975 for an optimal combination of accuracy and efficiency. Unfortunately, as we will show below, setting $\epsilon_{\text{FNO}} = 0.975$ is not accurate enough for the small δ_T and $\delta_{(Q)}$ contributions to noncovalent interaction energies.

The convergence of the δ_T and $\delta_{(Q)}$ interaction energy contributions for the C_{2v} geometry of the water dimer as a function of the ϵ_{FNO} threshold is shown in Table 10. This

Table 10. Interaction Energy Contributions (in cm^{−1}) for the H₂O–H₂O (C_{2v}) Complex Calculated at Different FNO Thresholds ϵ_{FNO} ^a

ϵ_{FNO}	Nbf	CCSD(T)	δ_T	$\delta_{(Q)}$	$\delta_{T+(Q)}$
aDZ					
0.99	50	−979.489	2.524	−0.966	1.558
0.999	64	−1021.655	1.465	−1.542	−0.077
0.9999	74	−1021.808	1.181	−0.611	0.570
0.99999	78	−1026.659	1.136	−0.738	0.398
1	82	−1026.272	1.137	−0.730	0.407
aTZ					
0.99	79	−963.346	1.953	−2.006	−0.053
0.999	126	−1078.499	0.999	−2.190	−1.191
0.9999	156	−1101.515	0.479	−2.124	−1.645
0.99999	171	−1107.233	0.334	−2.249	−1.915
1	184	−1107.160	0.302	−2.264	−1.962
aQZ					
0.99	99	−962.430	1.926	−2.390	−0.464
0.999	197	−1143.072	1.431	−2.458	−1.027
0.9999	271	−1136.522	0.696	−2.441	−1.745
1	344	−1141.024	0.552		

^aThe FNOs with the largest occupation numbers are retained until their cumulative occupation exceeds ϵ_{FNO} times the number of electrons; the remaining FNOs are discarded. A threshold of one represents the full calculation without any FNO truncation of basis functions. The quantity “Nbf” represents the number of basis functions after the FNO truncation.

convergence turns out to be quite slow: a threshold of 0.99 leads to an overestimation of the δ_T term up to six times. The $\delta_{(Q)}$ contribution is less sensitive to the FNO approximation: nevertheless, the error for $\epsilon_{\text{FNO}} = 0.99$ amounts to 32% in the aDZ basis and 11% in aTZ. The results of Table 10 suggest that the minimum acceptable ϵ_{FNO} threshold for noncovalent interactions is 0.9999 for $\delta_{(Q)}$ (errors up to 16%) and 0.99999 for δ_T (errors up to 11%). Unfortunately, such tight thresholds correspond to a fairly small reduction of the virtual space so that the speedup afforded by the FNO approximation is quite limited. In fact, the C_{2v} water dimer (Table 10) is the only system for which we obtained, with significant computational effort, an FNO result (the $\delta_{(Q)}/\text{aQZ}$ value) for which the corresponding nonapproximate value was out of reach. The $\delta_{(Q)}/\text{aQZ}$ contribution for $\epsilon_{\text{FNO}} = 0.9999$ can be expected to be accurate (the accuracy of the FNO $\delta_{(Q)}$ terms for a given threshold increases with the basis set size, cf. Table 10) and the quadruples contribution converges particularly slow for the water dimer (cf. Table 6), so the availability of the aQZ result thanks to the FNO approximation is quite helpful. However, this situation should be viewed as an exception rather than a rule as the FNO approach using reliable thresholds provides only modest performance gains (a similar $\epsilon_{\text{FNO}} = 0.9999$ calculation would be unfeasible for the less symmetric, global-minimum geometry of the water dimer).

C. Effects of the Frozen Core Approximation. In our calculations so far, only valence electrons were correlated. In contrast, the small-bases studies of Hobza et al.^{19–21} correlated all electrons. Therefore, it is worth checking if the conclusions reached on the basis of the frozen-core calculations still hold when all-electron interaction energies are considered. As the importance of the core–core and core–valence correlation varies widely among the atoms present in our test systems, we decided to investigate the all-electron interaction energies for three dimers that represent the best- and worst-case scenarios: Ne–Ne (where the frozen-core approximation is expected to be very accurate), He–LiH, and LiH–LiH (for which the effects of the 1s lithium correlation can be significant). The all-electron results, in the same format as in Tables 1–7, are presented in Tables SIII (LiH–LiH) and SIV (He–LiH and Ne–Ne) in the Supporting Information. In addition to the 6-31G*(0.25), 6-31G**-(0.25,0.15), and aXZ results, we have listed the interaction energy contributions obtained in the polarized core and valence aug-cc-pCVXZ≡aCVXZ sequence.⁶⁷

The results in Tables SIII and SIV in the Supporting Information indicate that, as expected, the interaction energy correction due to core–core and core–valence correlation amounts to a small fraction of the post-CCSD(T) correction for Ne–Ne, but completely dominates the latter for He–LiH and LiH–LiH. However, virtually all of this correction is recovered at the CCSD(T) level. In the largest aCVXZ basis sets considered, the core–core and core–valence correlation contribution constitutes 0–14% of δ_T and 3–25% of δ_Q . While the all-electron δ_T and δ_Q corrections should formally be computed using the aCVXZ bases, the corresponding aXZ results turn out to provide very reasonable approximations. Overall, the results in Tables SIII and SIV indicate that the standard practice of treating the (CCSD(T)-level) core correlation and (frozen-core) post-CCSD(T) interaction energy corrections as additive^{24,30,37,40} is well justified and that the basis set convergence patterns of the δ_T and δ_Q interaction energy components are very similar with and without the frozen core approximation.

IV. SUMMARY

We have studied the basis set convergence of the post-CCSD(T) coupled-cluster interaction energy contributions for 21 weakly bound dimers including the smallest members of the A24 set.²⁰ By performing CCSDT(Q) calculations in at least the aTZ basis set, and CCSDT calculations in at least aQZ (except for one system), we were able to assess the accuracy of small-basis results. We found that, unfortunately, the 6-31G*(0.25) and 6-31G**-(0.25,0.15) bases suggested for post-CCSD(T) corrections by Hobza et al.^{19–21} provide a very poor description of the CCSDT- and CCSDT(Q)-level effects, with mean unsigned relative errors for the δ_{T+Q} sum on the order of 80% for 6-31G**-(0.25,0.15) and 110% for 6-31G*(0.25) (thus, it is often better to neglect the post-CCSD(T) terms completely than to estimate them using these small basis sets). Upgrading the basis set to aDZ reduces the average error to about 35%.

The overall importance of the post-CCSD(T) interaction energy contributions varies dramatically with the size and polarity of the monomers. In agreement with the findings of refs 19 and 21, and of earlier studies for individual complexes,^{22,24} we observe that the full quadruples contribution δ_Q is negligible for all dimers except for N₂–N₂. However, the CCSDT and CCSDT(Q) corrections are generally of similar magnitude: the neglect of δ_Q is a viable approximation only for the four-electron dimers. The total δ_{T+Q} effect amounts to about 1–2% of the

CCSD(T) interaction energy (less for the polar–polar dimers) and can contribute even more to the interaction energy anisotropy, making it critical to go up to CCSDT(Q) in the calculations of potential energy surfaces for high-resolution spectroscopic applications. On the average, the sum δ_{T+Q} converges a little faster (in relative terms) than the δ_T and δ_Q terms separately; however, the error cancellation between δ_T and δ_Q is by no means systematic.

Overall, the basis set convergence of the δ_T and δ_Q interaction energy contributions is somewhat disappointing. Even at the aTZ basis set level, the obtained corrections seem to be quite far from converging (we estimate the mean accuracy of the δ_{T+Q} /aTZ estimate to be about 15%). Obtaining the benchmark δ_{T+Q} /CBS value to within a few percent is only feasible for four-electron dimers (for which the convergence of the coupled-cluster expansion is particularly fast so that δ_Q is much smaller than δ_T). Thus, as observed in some recent potential energy surface studies,^{24–26} the δ_{T+Q} terms, even computed in the aTZ basis or larger, remain one of the largest sources of residual uncertainty in the final potential as long as highly accurate CCSD(T)/CBS limits are obtained using state-of-the-art approaches. Thus, further research into the techniques that make high-order CC calculations more efficient is highly desired. One of such techniques, the MP2 frozen natural orbital approximation, was tested in this work with limited success. While this approximation provided the only way to obtain the δ_Q /aQZ result for the C_{2v} water dimer, the required FNO thresholds were too tight to afford a significant decrease in the size of the virtual space. As the inclusion of interaction energy contributions through CCSDT(Q) has been integral to the development of top-accuracy interaction potentials in recent years, we are in a place where the δ_T and δ_Q calculations are both a necessity and a bottleneck. Thus, further improvements in the accuracy of potential energy surfaces hinge on the design of new accurate approximations, new basis sets specifically optimized for the post-CCSD(T) corrections, or both.

■ ASSOCIATED CONTENT

Supporting Information

Geometries and CCSD(T)/CBS interaction energies for all systems (Table SI), benchmark MP2, δ_{CCSD} , δ_T , δ_Q , and δ_Q interaction energy contributions for all systems (Table SII), all-electron post-CCSD(T) contributions for selected systems (Tables SIII–SIV), and Cartesian coordinates for all complexes investigated in this work. This material is available free of charge via the Internet at <http://pubs.acs.org/>.

■ AUTHOR INFORMATION

Corresponding Author

*Email: patkowski@auburn.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Dr. Krzysztof Szalewicz for his inspiration and his help in establishing the collaboration that led to this work. D.G.A.S. and K.P. are supported by the U.S. National Science Foundation CAREER Award No. CHE-1351978 and by the startup funding from Auburn University. H.A.W. and M.S. acknowledge financial support from the National Science Council of Taiwan (NSC 102-2113-M-009-015-MY3) and from the Ministry of Education of Taiwan (MOE-ATU project).

REFERENCES

- (1) Čížek, J. J. *Chem. Phys.* **1966**, 45, 4256–4266.
- (2) Paldus, J.; Čížek, J.; Shavitt, I. *Phys. Rev. A* **1972**, 5, 50–67.
- (3) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, 157, 479–483.
- (4) Jurečka, P.; Šponer, J.; Černý, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006**, 8, 1985–1993.
- (5) Takatani, T.; Hohenstein, E. G.; Malagoli, M.; Marshall, M. S.; Sherrill, C. D. *J. Chem. Phys.* **2010**, 132, 144104.
- (6) Podeszwa, R.; Patkowski, K.; Szalewicz, K. *Phys. Chem. Chem. Phys.* **2010**, 12, 5974–5979.
- (7) Faver, J. C.; Benson, M. L.; He, X.; Roberts, B. P.; Wang, B.; Marshall, M. S.; Kennedy, M. R.; Sherrill, C. D.; Merz, K. M., Jr. *J. Chem. Theory Comput.* **2011**, 7, 790–797.
- (8) Řezáč, J.; Riley, K. E.; Hobza, P. *J. Chem. Theory Comput.* **2011**, 7, 2427–2438.
- (9) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. *Chem. Phys. Lett.* **1998**, 286, 243–252.
- (10) Schwenke, D. W. *J. Chem. Phys.* **2005**, 122, 014107.
- (11) Tao, F.-M.; Pan, Y.-K. *J. Phys. Chem.* **1991**, 95, 3582–3588.
- (12) Hättig, C.; Klopper, W.; Köhn, A.; Tew, D. P. *Chem. Rev.* **2012**, 112, 4–74.
- (13) Kong, L.; Bischoff, F. A.; Valeev, E. F. *Chem. Rev.* **2012**, 112, 75–107.
- (14) Noga, J.; Bartlett, R. J. *J. Chem. Phys.* **1987**, 86, 7041–7050.
- (15) Bomble, Y. J.; Stanton, J. F.; Kállay, M.; Gauss, J. *J. Chem. Phys.* **2005**, 123, 054101.
- (16) Kállay, M.; Gauss, J. *J. Chem. Phys.* **2005**, 123, 214105.
- (17) Oliphant, N.; Adamowicz, L. *J. Chem. Phys.* **1991**, 94, 1229–1235.
- (18) Kucharski, S. A.; Bartlett, R. J. *Theor. Chim. Acta* **1991**, 80, 387–405.
- (19) Řezáč, J.; Šimová, L.; Hobza, P. *J. Chem. Theory Comput.* **2013**, 9, 364–369.
- (20) Řezáč, J.; Hobza, P. *J. Chem. Theory Comput.* **2013**, 9, 2151–2155.
- (21) Šimová, L.; Řezáč, J.; Hobza, P. *J. Chem. Theory Comput.* **2013**, 9, 3420–3428.
- (22) Noga, J.; Kállay, M.; Valiron, P. *Mol. Phys.* **2006**, 104, 2337–2345.
- (23) Jäger, B.; Hellmann, R.; Bich, E.; Vogel, E. *Mol. Phys.* **2009**, 107, 2181–2188.
- (24) Patkowski, K.; Szalewicz, K. *J. Chem. Phys.* **2010**, 133, 094304.
- (25) Jankowski, P.; Surin, L. A.; Potapov, A.; Schlemmer, S.; McKellar, A. R. W.; Szalewicz, K. *J. Chem. Phys.* **2013**, 138, 084307.
- (26) Smith, D. G. A.; Patkowski, K.; Trinh, D.; Balakrishnan, N.; Lee, T.-G.; Forrey, R. C.; Yang, B. H.; Stancil, P. C. *J. Phys. Chem. A* **2014**, DOI: 10.1021/jp412048w.
- (27) Patkowski, K.; Cencek, W.; Jeziorski, B.; Szalewicz, K. *J. Phys. Chem. A* **2007**, 111, 7611–7623.
- (28) Bakr, B. W.; Smith, D. G. A.; Patkowski, K. *J. Chem. Phys.* **2013**, 139, 144305.
- (29) Patkowski, K.; Cencek, W.; Jankowski, P.; Szalewicz, K.; Mehl, J. B.; Garberoglio, G.; Harvey, A. H. *J. Chem. Phys.* **2008**, 129, 094304.
- (30) Hellmann, R.; Bich, E.; Vogel, E. *Mol. Phys.* **2008**, 106, 133–140.
- (31) Jankowski, P.; McKellar, A. R. W.; Szalewicz, K. *Science* **2012**, 336, 1147–1150.
- (32) Ślawik, M.; Li, Y.-T.; Patkowski, K.; Jankowski, P.; Witek, H. A. (to be submitted).
- (33) Jankowski, P. *J. Chem. Phys.* **2004**, 121, 1655–1662.
- (34) Jankowski, P.; Ziolkowski, M. *Mol. Phys.* **2006**, 104, 2293–2302.
- (35) Jankowski, P. *J. Chem. Phys.* **2008**, 128, 154311.
- (36) Hellmann, R. *Mol. Phys.* **2013**, 111, 387–401.
- (37) Lane, J. R. *J. Chem. Theory Comput.* **2013**, 9, 316–323.
- (38) Przybytek, M.; Patkowski, K.; Jeziorski, B. *Collect. Czech. Chem. Commun.* **2004**, 69, 141–176.
- (39) Pieniazek, P. A.; Arnstein, S. A.; Bradforth, S. E.; Krylov, A. I.; Sherrill, C. D. *J. Chem. Phys.* **2007**, 127, 164110.
- (40) Patkowski, K.; Podeszwa, R.; Szalewicz, K. *J. Phys. Chem. A* **2007**, 111, 12822–12838.
- (41) Šponer, J.; Leszczyński, J.; Hobza, P. *J. Phys. Chem.* **1996**, 100, 1965–1974.
- (42) Hobza, P.; Šponer, J. *Chem. Rev.* **1999**, 99, 3247–3276.
- (43) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, 90, 1007–1023.
- (44) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, 96, 6796–6806.
- (45) Dunning, T. H., Jr.; Peterson, K. A.; Wilson, A. K. *J. Chem. Phys.* **2001**, 114, 9244–9253.
- (46) Woon, D. E.; Dunning Jr., T. H. To be published.
- (47) Schuchardt, K. L.; Didier, B. T.; Elsethagen, T.; Sun, L.; Gurumoorthis, V.; Chase, J.; Li, J.; Windus, T. L. *J. Chem. Inf. Model.* **2007**, 47, 1045–1052.
- (48) Prascher, B. P.; Woon, D. E.; Peterson, K. A.; Dunning, T. H., Jr.; Wilson, A. K. *Theor. Chem. Acc.* **2011**, 128, 69–82.
- (49) Kállay, M.; Surján, P. R. *J. Chem. Phys.* **2001**, 115, 2945–2954.
- (50) MRCC, a quantum chemical program suite written by Kállay, M.; Rolik, Z.; Ladjánszki, I.; Szegedy, L.; Ladóczki, B.; Csontos, J.; Kornis, B. (a version from October 2013). See also <http://www.mrcc.hu> (accessed June 2, 2014).
- (51) Rolik, Z.; Szegedy, L.; Ladjánszki, I.; Ladóczki, B.; Kállay, M. *J. Chem. Phys.* **2013**, 139, 094105.
- (52) Werner, H.-J. et al. *MOLPRO, version 2012.1, a package of ab initio programs*, 2012, see <http://www.molpro.net> (accessed June 2, 2014).
- (53) Stanton, J. et al. *CFOUR*, a quantum chemical program package, containing the integral packages MOLECULE (Almlöf, J.; Taylor, P.R.), PROPS (Taylor, P.R.), ABACUS (Helgaker, T.; Jensen, H.J. Aa.; Jørgensen, P.; Olsen, J.), and ECP routines by Mitin, A. V.; van Wüllen, C. For the current version, see <http://www.cfour.de> (accessed June 2, 2014).
- (54) Rolik, Z.; Kállay, M. *J. Chem. Phys.* **2011**, 134, 124111.
- (55) McMahon, J. D.; Lane, J. R. *J. Chem. Phys.* **2011**, 135, 154309.
- (56) Patkowski, K. *J. Chem. Phys.* **2012**, 137, 034103.
- (57) Patkowski, K. *J. Chem. Phys.* **2013**, 138, 154101.
- (58) Burns, L. A.; Marshall, M. S.; Sherrill, C. D. *J. Chem. Theory Comput.* **2014**, 10, 49–57.
- (59) Mentel, L. M.; Baerends, E. J. *J. Chem. Theory Comput.* **2014**, 10, 252–267.
- (60) Burns, L. A.; Vazquez-Mayagoitia, A.; Sumpter, B. G.; Sherrill, C. D. *J. Chem. Phys.* **2011**, 134, 084107.
- (61) Marshall, M. S.; Burns, L. A.; Sherrill, C. D. *J. Chem. Phys.* **2011**, 135, 194102.
- (62) Rode, M.; Sadlej, J.; Moszyński, R.; Wormer, P. E. S.; van der Avoird, A. *Chem. Phys. Lett.* **1999**, 314, 326–332.
- (63) Adamowicz, L.; Bartlett, R. J. *J. Chem. Phys.* **1987**, 86, 6314–6324.
- (64) Neogrady, P.; Pitoňák, M.; Urban, M. *Mol. Phys.* **2005**, 103, 2141–2157.
- (65) Sosa, C.; Geertsen, J.; Trucks, G. W.; Bartlett, R. J.; Franz, J. A. *Chem. Phys. Lett.* **1989**, 159, 148–154.
- (66) Landau, A.; Khistyayev, K.; Dolgikh, S.; Krylov, A. I. *J. Chem. Phys.* **2010**, 132, 014109.
- (67) Peterson, K. A.; Dunning, T. H., Jr. *J. Chem. Phys.* **2002**, 117, 10548–10560.
- (68) Jankowski, P.; Szalewicz, K. *J. Chem. Phys.* **2005**, 123, 104301.