Post-CCSD(T) contributions to total atomization energies in multireference systems

Amir Karton

Post-CCSD(T) contributions to total atomization energies in multireference systems

Amir Karton[a)]

School of Molecular Sciences, The University of Western Australia, Perth, Western Australia 6009, Australia

We examine the magnitude and the basis set convergence of post-coupled-cluster with single, double, and perturbative triple excitations (CCSD(T)) contributions (up to CCSDTQ567) for a wide and diverse set of 21 first- and second-row molecules with up to four non-hydrogen atoms. We focus on multireference systems for which post-CCSD(T) effects are particularly pronounced. The considered molecules are $BN(^1\sum^+)$, $C_2(^1\sum^+)$, O_2, FO, P_2, S_2, ClO, N_2O, NO_2, O_3, FNO, FO_2, F_2O, S_2O, S_3, $ClNO$, $ClOO$, Cl_2O, N_2C_2, P_4, and S_4. This set spans the gamut from molecules dominated by moderate nondynamical correlation (e.g., FO, ClO, NO_2, S_2O, N_2C_2, and P_4) to systems dominated by strong nondynamical correlation (e.g., BN, C_2, FO_2, O_3, $ClOO$, and S_4). We examine the basis set convergence of the CCSDT, CCSDT(Q), CCSDTQ, CCSDTQ(5), CCSDTQ5, CCSDTQ5(6), CCSDTQ56, CCSDTQ56(7), and CCSDTQ567 methods. The largest basis sets employed in each category are cc-pV6Z (CCSDT(Q)), cc-pV5Z (CCSDTQ), cc-pVTZ (CCSDTQ5(6)), and cc-pVDZ (CCSDTQ567). Apart from examining the basis-set convergence of post-CCSD(T) contributions near the one-particle basis-set limit, this work explores cost-effective approaches for obtaining these contributions from fairly small basis sets. We consider both effective basis-set extrapolations and scaling factors. An important finding is that extrapolating the perturbative connected quadruples, (Q), from the cc-pVDZ(4s3p1d) and cc-pVTZ basis sets yields near basis-set limit results and represents a significant improvement relative to cc-pV{D,T}Z extrapolation at no additional computational cost (where cc-pVDZ(4s3p1d) is an extended version of the cc-pVDZ basis set). Combining the (Q)/cc-pV{D(4s3p1d),T}Z extrapolations with the fully iterative connected quadruples, Q–(Q), contribution calculated with the cc-pVDZ (or even the cc-pVDZ(3s2p)) basis set is a cost-effective way for obtaining the connected quadruples component close to the basis-set limit (where cc-pVDZ(3s2p) is a truncated version of the cc-pVDZ basis set). In addition, we show that the (5)/cc-pVDZ(3s2p) and (6)/cc-pVDZ(3s2p) components provide reasonable approximations for the connected quintuple and sextuple components close to the basis-set limit, respectively. *Published by AIP Publishing.* https://doi.org/10.1063/1.5036795

I. INTRODUCTION

Coupled-cluster (CC) theory is one of the most cost-effective methods for approximating the exact solution for the nonrelativistic electronic Schrödinger equation.[1,2] Coupled-cluster theory entails a hierarchy of approximations that can be systematically improved toward the exact quantum mechanical solution, providing a roadmap for the determination of highly accurate and reliable chemical properties. However, due to the inherently slow convergence of the correlation energy with respect to the one-particle basis set size, the only practical way to approximate the exact electronic energy near the one-particle basis-set limit is to use the so-called post-coupled-cluster with single, double, and perturbative triple excitations (CCSD(T)) composite *ab initio* methods.[3–7] In these methods, successively higher cluster expansion terms (i.e., CCSD, (T), T–(T), (Q), Q–(Q), (5), etc.) are extrapolated (or calculated) near the basis-set limit using successively smaller basis sets. Indeed, the

only reason that composite *ab initio* methods like W4 and HEAT-456QP can be carried out at a realistic computational cost for systems with multiple non-hydrogen atoms hinges on the faster basis-set convergence of the higher cluster expansion terms.[3,8,9]

A fairly large number of studies have been dedicated to studying the basis-set convergence of the CCSD and CCSD(T) methods near (or at) the infinite basis-set limit (for a number of representative examples, see Refs. 10–34). Fewer studies were dedicated to a systematic examination of basis set convergence of post-CCSD(T) excitations near the infinite basis-set limit.[8,9,35–42] These studies, however, were limited to considering small systems with 1–2 non-hydrogen atoms and/or did not consider post-CCSD(T) effects with sufficiently large basis sets and basis-set extrapolations. In addition, only a small subset of the considered molecules in these studies are characterized by a strong multireference character.

In the present work, we investigate the basis set convergence of post-CCSD(T) contributions to the total atomization energies (TAEs) for a wider and more diverse set of molecules with up to four non-hydrogen atoms. Furthermore, we focus entirely on systems that are dominated by moderate-to-severe

---

a)E-mail: amir.karton@uwa.edu.au

nondynamical correlation effects, for which post-CCSD(T) effects are expected to be more pronounced. In particular, our set includes the following 21 systems: $BN(^1\sum^+)$, $C_2(^1\sum^+)$, O_2, FO, P_2, S_2, ClO, N_2O, NO_2, O_3, FNO, FO_2, F_2O, S_2O, S_3, ClNO, ClOO, Cl_2O, N_2C_2, P_4, and S_4. The chosen set is composed of five radicals (FO·, ClO·, ONO·, FOO·, and ClOO·), two triplet systems (O_2 and S_2), and 14 closed-shell species. This set evidently spans the gamut from systems dominated by moderate multireference character (e.g., N_2O, NO_2, N_2C_2, and P_4) to systems dominated by severe nondynamical correlation (e.g., C_2, BN, FO_2, O_3, ClOO, and S_4). As such it constitutes an excellent set for analysis of basis set convergence of post-CCSD(T) contributions in multireference systems. For most of the diatomic molecules, we were able to carry out CCSDT(Q)/cc-pV6Z, CCSDTQ/cc-pV5Z, CCSDTQ5(6)/cc-pVTZ, and CCSDTQ567/cc-pVDZ(3s2p) calculations. For most of the triatomic and tetra-atomic molecules, we were able to carry out CCSDT(Q)/cc-pVQZ, CCSDTQ/cc-pVTZ, CCSDTQ(5)/cc-pVDZ(4s3p1d), CCSDTQ5/cc-pVDZ, and CCSDTQ56(7)/cc-pVDZ(3s2p) calculations. Here cc-pVDZ(4s3p1d) and cc-pVDZ(3s2p) denote truncated versions of the cc-pVDZ and cc-pVTZ basis sets, respectively (see Sec. II). Using these basis-set limit values for a large and diverse set of challenging systems, we attempt to answer questions such as the following:

- What is the magnitude of the various post-CCSD(T) contributions (namely, T–(T), (Q), Q–(Q), (5), 5–(5), (6), 6–(6), and 7) to the atomization energies of challenging systems?
- For each of the post-CCSD(T) contributions, what level of accuracy can we expect from a certain basis set or basis set extrapolation?
- What are the most cost-effective approaches for obtaining each of the post-CCSD(T) contributions?

Finally, it should be pointed out that exploring basis-set convergence of higher-order correlation effects outside the equilibrium region is a topic of great interest, albeit beyond the scope of the present investigation. For previous studies on this topic, see, for example, Refs. 43–50.

II. COMPUTATIONAL METHODS

This work represents an extensive computational effort for obtaining basis-set limit values for post-CCSD(T) contributions for systems with 2–4 non-hydrogen atoms. We note that many of the calculations reported here strained our computational resources to the absolute limit. For example, they involved 5–10×10^9 amplitudes and ran for over a month on 20 cores of dual Intel Xeon machines with 256–1024 GB of RAM (see the supplementary material for further details).

All the geometries were optimized at the CCSD(T)/cc-pV(Q+d)Z level of theory and were taken from the W4-17 database[51] (the geometries are given in Table S1 of the supplementary material). This level of theory has been shown to yield geometries that are in close agreement with CCSD(T) geometries near the complete basis-set (CBS) limit, e.g., with mean absolute deviations (MADs) of 0.001 Å from

CCSD(T)/cc-pV(6+d)Z geometries.[52] All calculations were carried out using the MRCC program suite[53,54] with the standard correlation-consistent basis sets of Dunning and co-workers.[55–57] For the sake of brevity, the cc-pVnZ basis sets (n = D, T, Q, 5, 6) are denoted by VnZ. In addition to the standard Dunning basis sets, we use two non-standard basis sets. The first is a truncated version of the cc-pVDZ basis set in which the d function has been omitted (denoted by VDZ(3s2p)). The second is an extended version of the cc-pVDZ basis set in which the sp functions have been replaced with the sp functions from the cc-pVTZ basis set (denoted by VDZ(4s3p1d)). All basis set extrapolations employ the $E(L) = E_\infty + A/L^\alpha$ two-point extrapolation formula, unless otherwise indicated with $\alpha = 3$ (where L is the highest angular momentum represented in the basis sets for the non-hydrogen atoms). Basis set extrapolations using the VnZ and V(n+1)Z basis sets are denoted by V{n, n+1}Z. All calculations are nonrelativistic and are carried out within the frozen-core approximation; i.e., the 1s orbitals for first-row atoms and the 1s, 2s, and 2p orbitals for second-row atoms are constrained to be doubly occupied in all configurations.

III. RESULTS AND DISCUSSION

In the present study, we examine the basis set convergence of post-CCSD(T) contributions to molecular atomization energies near the one-particle basis set limit. We consider iterative and perturbative coupled cluster contributions up to connected septuples (CCSDTQ567) for a set of 21 first- and second-row molecules with up to four non-hydrogen atoms, spanning the gamut from systems dominated by moderate nondynamical correlation (e.g., FO, ClO, NO_2, S_2O, N_2C_2, and P_4) to systems dominated by strong nondynamical correlation (e.g., BN, C_2, FO_2, O_3, ClOO, S_3, and S_4). For easy navigation between Secs. III A and III J, Table I gives an overview of the CC excitations that are discussed in each of the subsections along with the abbreviations that are used.

A. Full-iterative connected triple excitations

For seven diatomic systems (BN, C_2, O_2, FO, P_2, S_2, and ClO), we were able to obtain the T–(T) contribution to the TAEs at the complete basis-set limit from V{5,6}Z extrapolations. These results, along with results for the VnZ basis sets (n = D–6), are given in Table II. Overall, the T–(T) component converges fairly smoothly to the CBS limit. For example, we obtain the following root mean square deviations (RMSDs) relative to the V{5,6}Z reference values: 0.69 (VDZ), 0.22 (VTZ), 0.11 (VQZ), 0.06 (V5Z), and 0.04 (V6Z) kcal mol^{-1}. Remarkably, the V{D,T}Z extrapolation results in a similar overall performance to the V6Z basis set, with an RMSD of merely 0.04 kcal mol^{-1}. Needless to say, this represents a remarkable saving in computer time and demonstrates the effectiveness of using basis set extrapolations for the T–(T) component. Extrapolating from the V{T,Q}Z basis-set pair results in a slightly lower RMSD of 0.03 kcal mol^{-1} when an exponent of $\alpha = 3$ is used. It was suggested that convergence in this basis set size regime is, however, slower than $\alpha = 3$. In particular, the authors of Ref. 38 suggested $\alpha = 2.5$ as an effective exponent. Using $\alpha = 2.5$ cuts the RMSD by

TABLE I. Overview of the coupled-cluster contributions discussed in the present work.

Subsection	Name	Definition	Abbreviation
III A	Full-iterative connected triples	CCSDT–CCSD(T)	T–(T)
III B	Noniterative connected quadruples	CCSDT(Q)–CCSDT	(Q)
III C	Full-iterative connected quadruples	CCSDTQ–CCSDT(Q)	Q–(Q)
III D	Connected quadruples as a whole	CCSDTQ–CCSDT	Q
III E	Noniterative quintuples	CCSDTQ(5)–CCSDTQ	(5)
III F	Full-iterative connected quintuples	CCSDTQ5–CCSDTQ(5)	5–(5)
III G	Connected quintuples as a whole	CCSDTQ5–CCSDTQ	5
III H	Noniterative sextuples	CCSDTQ5(6)–CCSDTQ5	(6)
III H	Full-iterative connected sextuples	CCSDTQ56–CCSDTQ5(6)	6–(6)
III H	Connected sextuples as a whole	CCSDTQ56–CCSDTQ5	6
III I	Noniterative septuples	CCSDTQ56(7)–CCSDTQ56	(7)
III I	Full-iterative connected septuples	CCSDTQ567–CCSDTQ56(7)	7–(7)
III I	Connected septuples as a whole	CCSDTQ567–CCSDTQ56	7
III J	Post-CCSD(T) as a whole	CCSDTQ567–CCSD(T)	N/A

over 50% to 0.013 kcal mol^{-1}, and using an optimized exponent for our set of $\alpha = 2.35$ results in an RMSD of merely 0.009 kcal mol^{-1}.

Since the V$\{$T,Q$\}$Z extrapolation with $\alpha = 2.35$ results in a near-zero RMSD of 0.009 kcal mol^{-1} relative to the V$\{$5,6$\}$Z reference values, it is worthwhile evaluating the smaller basis sets against our best V$\{$T,Q$\}$Z values, which we were able to obtain for the entire set of 21 molecules. These results are presented in Table III. We start by noting that for most molecules in our set the T–(T) contribution to the TAE is negative and large. The two exceptions for which the T–(T) contribution is positive are FO and ClO (namely, it is +0.25 and +0.03 kcal mol^{-1}, respectively). For 12 out of the 21 systems, the T–(T) contribution approaches or exceeds the -1 kcal mol^{-1} mark, and for four systems it even exceeds -2 kcal mol^{-1}, namely, C_2 (-2.31), BN (-2.67), P_4 (-3.13), and S_4 (-3.17 kcal mol^{-1}).

As expected, the double-ζ basis sets result in poor performance, with RMSDs of 1.44 (VDZ(3s2p)), 1.31 (VDZ), and 1.06 (VDZ(4s3p1d)) kcal mol^{-1}. The VTZ basis set still

results in an unacceptably large RMSD of 0.45 kcal mol^{-1}. This RMSD is reduced to 0.11 kcal mol^{-1} for the V$\{$D,T$\}$Z extrapolation with $\alpha = 3.0$. Optimizing the exponent results in an RMSD of 0.09 kcal mol^{-1} with $\alpha = 2.7$. The largest deviations of +0.17 and -0.29 kcal mol^{-1} are obtained for S_4 and P_4, respectively. Exclusion of these two systems results in an RMSD of 0.05 kcal mol^{-1}.

B. Perturbative, noniterative connected quadruple excitations

For the seven diatomic molecules in our set, we were able to obtain basis-set limit values from V$\{$5,6$\}$Z extrapolations. Comparison of these V$\{$5,6$\}$Z CBS limits with V$\{$Q,5$\}$Z data reveals that the V$\{$Q,5$\}$Z extrapolations are practically at the basis set limit, the largest deviations being 0.012 kcal mol^{-1} for BN and C_2, followed by 0.005 kcal mol^{-1} for FO and ClO. Over the set of seven diatomics, the V$\{$Q,5$\}$Z extrapolation results in an RMSD of 0.007 kcal mol^{-1} and a MAD of 0.002 kcal mol^{-1}.

TABLE II. Convergence of the full-iterative connected triples contribution, CCSDT–CCSD(T), to the total atomization energy for the set of seven diatomic molecules for which we were able to obtain basis-set limit values from V$\{$5,6$\}$Z extrapolations. The V$\{$5,6$\}$Z reference TAEs are listed in the last column, and the tabulated values in the other columns are deviations relative to these basis-set limit values (in kcal mol^{-1}).

Basis set α	VDZ	VTZ	VQZ	V5Z	V6Z	V$\{$D,T$\}$Z 3.0	V$\{$T,Q$\}$Z 3.0	V$\{$T,Q$\}$Z 2.35	V$\{$Q,5$\}$Z 3.0	V$\{$5,6$\}$Z 3.0
BN	0.695	0.229	0.122	0.064	0.037	0.033	0.043	0.010	0.004	-2.676
C_2	1.048	0.379	0.185	0.097	0.056	0.097	0.043	-0.016	0.004	-2.291
O_2	0.507	0.139	0.077	0.040	0.023	-0.015	0.031	0.012	0.002	-0.528
FO	0.355	0.098	0.053	0.030	0.018	-0.010	0.021	0.007	0.006	0.240
P_2	0.906	0.252	0.132	0.082	0.048	-0.023	0.044	0.007	0.030	-0.974
S_2	0.652	0.200	0.100	0.063	0.036	0.010	0.026	-0.004	0.024	-0.508
ClO	0.301	0.092	0.048	0.030	0.018	0.005	0.015	0.002	0.012	0.031
RMSD[a,b]	0.687	0.220	0.112	0.063	0.036	0.040	0.034	0.009	0.016	
MAD[a,b]	0.638	0.199	0.102	0.058	0.034	0.027	0.032	0.008	0.012	
MSD[a,b]	0.638	0.199	0.102	0.058	0.034	0.014	0.032	0.002	0.012	

[a]RMSD = root mean square deviation, MAD = mean absolute deviation, MSD = mean signed deviation, deviations are taken as [approximate value] − [reference value].
[b]Error statistics with respect to the V$\{$5,6$\}$Z reference values.

TABLE III. Convergence of the full-iterative connected triples contribution, CCSDT–CCSD(T), to the total atomization energy for the set of 21 molecules. The V{T,Q}Z reference TAEs (with $\alpha = 2.35$) are listed in the last column, and the tabulated values in the other columns are deviations relative to these basis-set limit values (in kcal mol^{-1}).

Basis set					V{D,T}Z	V{D,T}Z	V{T,Q}Z
α	VDZ(3s2p)	VDZ	VDZ(4s3p1d)	VTZ	3.0	2.7	2.35
BN	−0.370	0.685	0.323	0.219	0.023	−0.015	−2.666
C_2	0.818	1.065	0.724	0.395	0.113	0.058	−2.307
O_2	0.460	0.494	0.244	0.127	−0.028	−0.058	−0.516
FO	0.410	0.348	0.198	0.091	−0.017	−0.038	0.247
P_2	1.048	0.899	0.811	0.246	−0.029	−0.083	−0.967
S_2	0.647	0.656	0.591	0.205	0.014	−0.023	−0.512
ClO	0.327	0.299	0.260	0.091	0.003	−0.014	0.033
N_2O	0.949	1.085	0.666	0.369	0.068	0.009	−1.571
NO_2	1.093	1.103	0.683	0.385	0.082	0.024	−0.963
O_3	1.148	1.327	0.777	0.468	0.107	0.037	−1.462
FNO	1.169	1.065	0.631	0.362	0.066	0.009	−0.731
FO_2	0.921	1.113	0.676	0.397	0.096	0.037	−0.181
F_2O	0.701	0.768	0.418	0.213	−0.021	−0.067	−0.581
S_2O	1.649	1.504	1.193	0.536	0.129	0.050	−1.482
S_3	2.039	1.714	1.569	0.617	0.156	0.066	−1.529
ClNO	1.412	1.218	0.876	0.420	0.085	0.019	−0.744
ClOO	1.144	1.401	1.039	0.540	0.178	0.107	−0.079
Cl_2O	1.032	0.936	0.744	0.314	0.051	0.001	−0.975
N_2C_2	1.577	1.394	0.896	0.452	0.056	−0.021	−1.956
P_4	3.392	2.354	2.177	0.594	−0.147	−0.291	−3.131
S_4	3.100	2.955	2.742	1.104	0.324	0.173	−3.175
RMSD[a,b]	1.444	1.312	1.062	0.447	0.112	0.087	
MAD[a,b]	1.210	1.161	0.868	0.388	0.085	0.057	
MSD[a,b]	1.175	1.161	0.868	0.388	0.062	−0.001	

[a]See footnote a to Table II.
[b]Error statistics with respect to the V{T,Q}Z reference values.

Even basis-set limit results extrapolated from V{T,Q}Z agree well with the V{5,6}Z results, with an RMSD and MAD of 0.014 and 0.004 kcal mol^{-1}, respectively. The largest deviations being 0.017, 0.018, and 0.022 kcal mol^{-1} for P_2, BN, and C_2, respectively. This suggests that the V{T,Q}Z results, which we were able to obtain for all 21 systems, could be used as reference values for evaluating the performance of the smaller basis sets. These results are presented in Table IV.

For all the molecules in our set, the (Q)/V{T,Q}Z contribution to the TAE is positive and large. Namely, it ranges between 1 and 6 kcal mol^{-1}. For seven molecules, this contribution exceeds the 3 kcal mol^{-1} mark, namely, P_4 (3.15), BN (3.28), C_2 (3.45), FO_2 (3.61), ClOO (3.85), O_3 (4.79), and S_4 (5.82 kcal mol^{-1}).

It has previously been shown that the VDZ basis set overshoots the (Q) contribution in highly polar systems (such as HF and H_2O) and severely undershoots the (Q) contribution in second-row molecules.[38] For the set of molecules in Table IV, the VDZ basis set systematically underestimates the V{T,Q}Z results. For the first-row systems, underestimations ranging between 0.08 (O_2) and 0.82 (BN) kcal mol^{-1} are obtained. Whilst for systems comprising of solely second-row systems, higher underestimations are obtained, namely, they range between 0.48 (S_2) and 2.01 (S_4) kcal mol^{-1}. Overall, for the set of 21 systems in Table IV, an RMSD of 0.77 kcal mol^{-1} is obtained relative to the

V{T,Q}Z results. Scaling the (Q)/VDZ component by 1.1, as is done in the W4lite composite method,[8] reduces the RMSD to 0.59 kcal mol^{-1}. Whilst an optimal scaling factor of 1.25 results in a lower RMSD of 0.40 kcal mol^{-1}. Nevertheless, underestimations of about 1 kcal mol^{-1} are still obtained for P_4 (−0.94) and S_4 (−1.06) kcal mol^{-1}. Thus, the VDZ basis set is not recommended for highly accurate applications involving second-row and/or highly polar species.

The (Q)/VTZ level of theory results in an RMSD of 0.31 kcal mol^{-1} and largest underestimations of 0.41 (O_3 and S_3), 0.50 (P_4), and 0.69 (S_4). Two ways of improving the performance of the (Q)/VTZ results with no increase in computational cost are scaling the (Q)/VTZ results by 1.1 as used in the W4 composite method[8] or extrapolating from the V{D,T}Z basis set pair. Scaling by 1.1 results in significantly better performance with an RMSD of 0.09 kcal mol^{-1}, compared to an RMSD of 0.16 kcal mol^{-1} for the V{D,T}Z extrapolation with $\alpha = 3$. We note that using an optimal extrapolation exponent (of 2.4 instead of 3) in the (Q)/V{D,T}Z extrapolation still results in inferior performance to scaling with an RMSD of 0.13 kcal mol^{-1}. Furthermore, inspection of the largest deviations reveals that scaling is a better approach. In particular, the (Q)/V{D,T}Z extrapolation (with $\alpha = 2.4$) results in 11 deviations larger than 0.1 kcal mol^{-1} (in absolute value), whereas

TABLE IV. Convergence of the perturbative, noniterative connected quadruples contribution, CCSDT(Q)–CCSDT, to the total atomization energy (in kcal mol^{-1}). The last five columns list TAEs, and the tabulated values in the other columns are deviations relative to basis-set limit values from V{T,Q}Z extrapolations (in kcal mol^{-1}).

Basis set α	VDZ	1.25 × VDZ	VTZ	1.1 × VTZ	VQZ	V{D,T}Z 3.0	V{D,T}Z 2.4	V{T(4s3p1d),T}Z 2.3	V{T,Q}Z 3.0	V5Z	V6Z	V{Q,5}Z 3.0	V{5,6}Z 3.0
BN	−0.818	−0.202	−0.255	0.048	−0.108	−0.018	0.084	−0.024	3.281	3.223	3.240	3.275	3.263
C_2	−0.793	−0.129	−0.228	0.094	−0.096	0.010	0.111	0.041	3.448	3.393	3.407	3.437	3.426
O_2	−0.082	0.199	−0.111	−0.001	−0.047	−0.123	−0.128	−0.032	1.204	1.179	1.187	1.202	1.198
FO	−0.097	0.065	−0.102	−0.037	−0.043	−0.104	−0.105	−0.027	0.747	0.725	0.733	0.748	0.743
P_2	−0.626	−0.366	−0.235	−0.092	−0.099	−0.071	0.000	0.003	1.666	1.608	1.625	1.651	1.649
S_2	−0.475	−0.350	−0.179	−0.099	−0.075	−0.054	0.000	0.016	0.975	0.939	0.956	0.980	0.980
ClO	−0.324	−0.210	−0.154	−0.092	−0.065	−0.083	−0.053	−0.019	0.780	0.756	0.772	0.799	0.794
N_2O	−0.267	0.284	−0.209	0.017	−0.088	−0.184	−0.174	−0.075	2.468				
NO_2	−0.308	0.206	−0.238	−0.025	−0.100	−0.208	−0.196	−0.075	2.361				
O_3	−0.663	0.370	−0.412	0.026	−0.174	−0.306	−0.261	−0.132	4.795				
FNO	−0.216	0.207	−0.196	−0.025	−0.083	−0.188	−0.185	−0.066	1.910				
FO_2	−0.605	0.146	−0.370	−0.046	−0.156	−0.271	−0.229	−0.138	3.610				
F_2O	−0.186	0.185	−0.198	−0.052	−0.084	−0.204	−0.206	−0.060	1.667				
S_2O	−0.604	−0.140	−0.270	−0.050	−0.114	−0.129	−0.069	0.011	2.463				
S_3	−1.146	−0.707	−0.410	−0.160	−0.173	−0.100	0.033	0.050	2.901				
ClNO	−0.570	−0.124	−0.271	−0.062	−0.114	−0.145	−0.090	−0.037	2.354				
ClOO	−0.968	−0.247	−0.448	−0.107	−0.189	−0.229	−0.135	−0.069	3.852				
Cl_2O	−0.444	−0.201	−0.209	−0.089	−0.088	−0.111	−0.068	0.004	1.417				
N_2C_2	−0.496	0.045	−0.245	−0.004	−0.103	−0.140	−0.094	−0.062	2.657				
P_4	−1.381	−0.937	−0.497	−0.231	−0.210	−0.125	0.035	0.030	3.155				
S_4	−2.013	−1.060	−0.688	−0.175	−0.290	−0.131	0.108	0.136	5.823				
RMSD[a,b]	0.769	0.402	0.314	0.093	0.133	0.158	0.133	0.066					
MAD[a,b]	0.623	0.304	0.282	0.073	0.119	0.140	0.113	0.053					
MSD[a,b]	−0.623	−0.141	−0.282	−0.055	−0.119	−0.139	−0.077	−0.025					

[a]See footnote a to Table II.

[b]Error statistics with respect to the V{T,Q}Z reference values for the 21 systems above.

the 1.1 × (Q)/VTZ approach results in only four deviations larger than 0.1 kcal mol^{-1}.

As pointed out previously,[8,38,39] the poor performance of the V{D,T}Z extrapolations is due to the anemic character of the cc-pVDZ basis set. One way to overcome this problem is to extrapolate using the cc-pVTZ basis set in conjunction with an spd quality basis set in which the s and p shells are more saturated than in the cc-pVDZ basis set. Indeed, extrapolating the (Q) component from the VDZ(4s3p1d) and VTZ basis sets results in an RMSD of merely 0.07 kcal mol^{-1} and only three deviations larger than 0.1 kcal mol^{-1}, namely, −0.14 (FO_2), −0.13 (O_3), and +0.14 (S_4) kcal mol^{-1}. A near-zero mean signed deviation (MSD) of −0.03 kcal mol^{-1} suggests that this extrapolation is free of systematic bias. This approach has the same computational cost as the 1.1 × (Q)/VTZ and (Q)/V{D,T}Z levels of theory but results in better performance and therefore recommended for the (Q) component.

C. Full-iterative connected quadruple excitations

For three systems (BN, C_2, and P_2), we were able to obtain the Q–(Q) contribution to the TAEs at the complete basis-set limit from V{Q,5}Z extrapolations. The V{T,Q}Z extrapolation essentially reproduces these basis-set limit values spot on with all deviations being smaller than 0.005 kcal mol^{-1}. We can therefore use the V{T,Q}Z values, which we have for the seven diatomics, for evaluating the performance of the

smaller basis sets. These results are presented in Table V. The VTZ basis set results in a respectable RMSD of 0.037 kcal mol^{-1} and a largest deviation of 0.046 kcal mol^{-1} obtained for BN. Extrapolating from the VDZ and VTZ basis sets with an exponent of α = 3 significantly improves the performance and results in an RMSD of 0.023 kcal mol^{-1} and a largest deviation of 0.030 kcal mol^{-1} obtained for C_2. Nevertheless, calculating the CCSDTQ energy in conjunction with the VTZ is prohibitively expensive for molecules with more than two non-hydrogen atoms. The largest basis set with which we were able to calculate the Q–(Q) corrections for all 21 systems is the VDZ(4s3p1d) basis set. Relative to the V{T,Q}Z CBS values, this basis set attains an RMSD of 0.051 kcal mol^{-1} and a largest deviation of 0.072 kcal mol^{-1} obtained for P_2 (Table V). Scaling the VDZ(4s3p1d) results by a scaling factor of 1.05 slightly improves the performance and reduces the RMSD to 0.043 kcal mol^{-1}. We note, however, that this does not represent an improvement over scaling the VDZ results with a scaling factor of 1.1, which results in essentially the same error statistics (Table V). Scaling the VDZ results by 1.1 has been found to be an efficient and cost-effective approach for obtaining the Q–(Q) correction and is used in the W4 composite method.[8,38]

Table VI gives the Q–(Q) contribution to the TAEs for the 21 systems in our set calculated with basis sets of up to VDZ(4s3p1d) and error statistics for the VDZ(3s2p) and VDZ with respect to the best reference values (i.e.,

TABLE V. Convergence of the full-iterative connected quadruples contribution, CCSDTQ–CCSDT(Q), to the total atomization energy for the set of seven diatomic molecules for which we were able to obtain basis-set limit values from V{T,Q}Z extrapolations. The last two columns list TAEs, and the tabulated values in the other columns are deviations relative to basis-set limit values from V{T,Q}Z extrapolations (in kcal mol^{-1}).

	VDZ (3s2p)	VDZ	1.1 × VDZ	VDZ (4s3p1d)	1.05 × VDZ (4s3p1d)	VTZ	VQZ	V{D,T}Z	V{T,Q}Z	V{Q,5}Z
BN	0.391	0.115	0.006	0.016	−0.043	0.046	0.019	0.017	−1.201	−1.200
C$_2$	−0.147	0.080	−0.027	0.062	0.007	0.045	0.019	0.030	−1.147	−1.151
O$_2$	0.071	0.034	0.021	0.051	0.045	0.031	0.013	0.029	−0.159	
FO	0.082	0.057	0.042	0.036	0.028	0.022	0.009	0.008	−0.207	
P$_2$	0.039	0.078	0.067	0.072	0.066	0.040	0.017	0.024	−0.185	−0.189
S$_2$	0.124	0.038	0.033	0.035	0.033	0.028	0.012	0.023	−0.088	
ClO	0.128	0.080	0.073	0.060	0.056	0.039	0.017	0.022	−0.149	
RMSD[a,b]	0.177	0.074	0.045	0.051	0.043	0.037	0.015	0.023		
MAD[a,b]	0.140	0.069	0.039	0.047	0.040	0.036	0.015	0.022		
MSD[a,b]	0.098	0.069	0.031	0.047	0.027	0.036	0.015	0.022		

[a]See footnote a to Table II.
[b]Error statistics with respect to the V{T,Q}Z reference values.

1.05 × VDZ(4s3p1d)). Let us begin by noting that the magnitude of the Q–(Q) contribution to the TAE spans a wide range from −0.06 (S$_2$) to −1.25 (BN) kcal mol^{-1}. Perhaps it is not surprising that the largest Q–(Q) contributions are obtained for the strongly multireference systems: ClOO (−0.62), FO$_2$ (−0.68), S$_4$ (−0.79), O$_3$ (−0.96), C$_2$ (−1.14), and BN (−1.25 kcal mol^{-1}).

The VDZ(3s2p) basis set results in a fairly large RMSD of 0.155 kcal mol^{-1}. This is in part due to the very large deviations of 0.43 and 0.37 kcal mol^{-1} obtained for BN and S$_4$, respectively. Eliminating these two systems results in an RMSD of 0.097 kcal mol^{-1}. Table VI also shows that scaling the VDZ results by a scaling factor of 1.1 results in an RMSD of merely 0.027 kcal mol^{-1} for the set of 21 systems.

TABLE VI. Convergence of the full-iterative connected quadruples contribution, CCSDTQ–CCSDT(Q), to the total atomization energy for the set of 21 molecules for which we were able to obtain this contribution with the 1.05 × VDZ(4s3p1d) basis set. The last two columns list TAEs, and the tabulated values in the other columns are deviations relative to 1.05 × VDZ(4s3p1d) values (in kcal mol^{-1}).

	VDZ(3s2p)	1.05 × VDZ(3s2p)	VDZ	1.1 × VDZ	1.05 × VDZ(4s3p1d)	V{D,T}Z
BN	0.434	0.394	0.158	0.049	−1.245	−1.185
C$_2$	−0.154	−0.219	0.072	−0.035	−1.139	−1.117
O$_2$	0.026	0.022	−0.011	−0.024	−0.114	−0.130
FO	0.055	0.048	0.030	0.015	−0.179	−0.199
P$_2$	−0.027	−0.035	0.012	0.001	−0.119	−0.162
S$_2$	0.092	0.094	0.006	0.001	−0.055	−0.065
ClO	0.072	0.071	0.025	0.018	−0.094	−0.127
N$_2$O	−0.011	−0.034	−0.002	−0.048	−0.453	
NO$_2$	0.033	0.015	0.013	−0.027	−0.404	
O$_3$	0.101	0.059	0.034	−0.058	−0.955	−1.002
FNO	0.022	0.007	0.009	−0.024	−0.332	
FO$_2$	0.122	0.094	0.072	0.011	−0.683	
F$_2$O	0.031	0.022	0.000	−0.021	−0.212	−0.228
S$_2$O	−0.017	−0.038	0.020	−0.017	−0.393	
S$_3$	0.143	0.134	0.028	−0.001	−0.316	−0.348
ClNO	0.057	0.044	0.016	−0.015	−0.320	
ClOO	0.139	0.115	0.052	−0.005	−0.615	
Cl$_2$O	0.074	0.071	0.001	−0.014	−0.152	−0.175
N$_2$C$_2$	0.026	0.009	−0.016	−0.052	−0.349	
P$_4$	0.247	0.242	0.028	−0.005	−0.352	
S$_4$	0.366	0.345	0.062	−0.011	−0.788	
RMSD[a,b]	0.155	0.147	0.047	0.027		
MAD[a,b]	0.107	0.100	0.032	0.021		
MSD[a,b]	0.087	0.069	0.029	−0.012		

[a]See footnote a to Table II.
[b]Error statistics with respect to the 1.05 × VDZ(4s3p1d) reference values for the 21 systems.

TABLE VII. Convergence of the connected quadruples contribution, CCSDTQ–CCSDT, to the total atomization energy for the set of seven diatomic molecules for which we were able to obtain basis-set limit values from V{T,Q}Z extrapolations. The last two columns list TAEs, and the tabulated values in the other columns are deviations relative to basis-set limit values from V{T,Q}Z extrapolations (in kcal mol^{-1}).

α	VDZ (3s2p)	VDZ	1.4 × VDZ	VDZ (4s3p1d)	1.4 × VDZ (4s3p1d)	VTZ	1.1 × VTZ	VQZ	V{D,T}Z 3.0	V{D(4s3p1d),T}Z 2.6	V{T,Q}Z 3.0	V{Q,5}Z 3.0
BN	−1.545	−0.703	−0.153	−0.595	−0.001	−0.209	−0.022	−0.088	0.003	−0.003	2.079	2.075
C$_2$	−1.510	−0.713	−0.078	−0.580	0.108	−0.184	0.028	−0.077	0.044	0.029	2.301	2.286
O$_2$	0.039	−0.048	0.351	−0.181	0.164	−0.080	0.017	−0.034	−0.094	−0.026	1.045	
FO	−0.106	−0.040	0.160	−0.181	−0.037	−0.079	−0.033	−0.034	−0.097	−0.025	0.540	
P$_2$	−0.392	−0.548	−0.175	−0.531	−0.152	−0.196	−0.067	−0.083	−0.044	−0.016	1.481	1.462
S$_2$	−0.388	−0.437	−0.257	−0.444	−0.266	−0.151	−0.078	−0.064	−0.029	0.005	0.887	
ClO	−0.386	−0.244	−0.089	−0.303	−0.172	−0.115	−0.063	−0.049	−0.060	−0.014	0.630	
RMSD[a,b]	0.857	0.472	0.201	0.435	0.153	0.153	0.050	0.065	0.061	0.019		
MAD[a,b]	0.624	0.390	0.180	0.402	0.129	0.145	0.044	0.061	0.053	0.017		
MSD[a,b]	−0.613	−0.390	−0.034	−0.402	−0.051	−0.145	−0.031	−0.061	−0.039	−0.007		

[a]See footnote a to Table II.
[b]Error statistics with respect to the V{T,Q}Z reference values.

D. Connected quadruple excitations as a whole

For three systems, we were able to extrapolate the CCSDTQ–CCSDT contribution to the TAE from the cc-pVQZ and cc-pV5Z basis sets. The differences between these basis-set limit values and cc-pV{T,Q}Z results are 0.005 (BN), 0.015 (C$_2$), and 0.019 (P$_2$) kcal mol^{-1}. We will therefore use the V{T,Q}Z results to evaluate the performance of more cost-effective approaches. These results are given in Table VII. The double-ζ type basis sets result in very large RMSDs of 0.86 (VDZ(3s2p)), 0.47 (VDZ), and 0.44 (VDZ(4s3p1d)) kcal mol^{-1}. Even the VTZ basis set leads to an appreciable RMSD of 0.15 kcal mol^{-1}. Inspection of the MSDs reveals that the VDZ, VDZ(4s3p1d), and VTZ basis sets systematically underestimate the V{T,Q}Z results (Table VII). Thus, scaling significantly improves the performance and leads to RMSDs of 0.20 (1.4 × VDZ), 0.15 (1.4 × VDZ(4s3p1d)), and 0.05 (1.1 × VTZ) kcal mol^{-1}.

Extrapolating the connected quadruple excitations from the VDZ and VTZ basis sets with an exponent of $\alpha = 3.0$ results in an RMSD of 0.06 kcal mol^{-1}, which is higher than that obtained for scaling the VTZ results by a scaling factor of 1.1 (0.05 kcal mol^{-1}, Table VII). We note that optimizing the extrapolation exponent results in essentially the same exponent and RMSD. This indicates that the VDZ might be doing more harm than good as discussed in Sec. III B. However, extrapolating from the VDZ(4s3p1d) and VTZ basis sets with $\alpha = 2.6$, at no additional computational cost, results in a significantly lower RMSD of 0.02 kcal mol^{-1}.

Let us turn our attention to a more cost-effective approach, in which the (Q) contribution is obtained with a triple-ζ-type basis set and the Q–(Q) contribution with a double-ζ-type basis set. In particular, we will consider the following equations:

$$c \times [\text{CCSDT(Q)/VTZ} - \text{CCSDT/VTZ} + \text{CCSDTQ/VDZ} - \text{CCSDT(Q)/VDZ}], \tag{1}$$

$$c \times [\text{CCSDT(Q)/V\{D, T\}Z} - \text{CCSDT/V\{D, T\}Z} + \text{CCSDTQ/VDZ} - \text{CCSDT(Q)/VDZ}], \tag{2}$$

$$c \times [\text{CCSDT(Q)/V\{D(4s3p1d), T\}Z} - \text{CCSDT/V\{D(4s3p1d), T\}Z} + \text{CCSDTQ/VDZ} - \text{CCSDT(Q)/VDZ}], \tag{3}$$

$$c \times [\text{CCSDT(Q)/V\{D(4s3p1d), T\}Z} - \text{CCSDT/V\{D(4s3p1d), T\}Z} + \text{CCSDTQ/VDZ(3s2p)} - \text{CCSDT(Q)/VDZ(3s2p)}], \tag{4}$$

$$\text{CCSDT(Q)/V\{D(4s3p1d), T\}Z} - \text{CCSDT/V\{D(4s3p1d), T\}Z} + c \times [\text{CCSDTQ/VDZ(3s2p)} - \text{CCSDT(Q)/VDZ(3s2p)}]. \tag{5}$$

We begin by noting that Eqs. (1) and (2) have the same computational cost and that the W4 method uses Eq. (1) with $c = 1.10$. For the seven diatomics for which we have V{T,Q}Z basis-set limit results, the W4 approach results in a relatively large RMSD of 0.043 kcal mol^{-1} and largest deviations of +0.067 (C$_2$) and −0.066 (S$_2$) kcal mol^{-1}. These results are shown in Table VIII. We note that optimizing the scaling factor results in $c = 1.09$ and essentially no change in the error statistics. Let us turn our attention to extrapolating the (Q)

contribution from the V{D,T}Z basis set pair and calculating the Q–(Q) component with the VDZ basis set. Using Eq. (2) with $c = 1.06$ and using $\alpha = 6.3$ in the (Q) extrapolation result in an RMSD of 0.037 kcal mol^{-1} and a largest deviation of −0.069 kcal mol^{-1} for S$_2$. This does not represent a significant improvement over the W4 approach (i.e., Eq. (1) with $c = 1.10$). However, extrapolating the (Q) contribution from the V{D(4s3p1d),T}Z basis set pair [via Eq. (3)] with $\alpha = 3.0$ in the extrapolation and using $c = 1.0$ result in an RMSD of

TABLE VIII. Cost-effective approaches for calculating the connected quadruples contribution, CCSDTQ–CCSDT, to the total atomization energy for the set of seven diatomic molecules for which we were able to obtain basis-set limit values from V{T,Q}Z extrapolations. The last column lists the TAEs extrapolated from the V{T,Q}Z basis sets, and the tabulated values in the other columns are deviations relative to these basis-set limit values (in kcal mol^{-1}).

Equation	(1)	(2)	(3)	(4)	(5)	
(Q)	VTZ	V{D,T}Z	V{D(4s3p1d),T}Z	V{D(4s3p1d),T}Z	V{D(4s3p1d),T}Z	V{T,Q}Z
Q–(Q)	VDZ	VDZ	VDZ	VDZ(3s2p)	VDZ(3s2p)	V{T,Q}Z
c	1.10	1.06	1.0	1.0	0.96	N/A
α[a]	N/A	6.3	3.0	3.0	3.0	3.0
BN	0.054	0.026	0.010	0.286	−0.041	2.079
C_2	0.067	0.031	0.026	−0.201	0.051	2.301
O_2	0.020	−0.021	−0.026	0.012	−0.005	1.045
FO	0.005	−0.015	0.004	0.029	−0.010	0.540
P_2	−0.025	−0.043	−0.003	−0.042	−0.004	1.481
S_2	−0.066	−0.069	−0.014	0.072	−0.015	0.887
ClO	−0.019	−0.026	0.014	0.061	−0.003	0.630
RMSD[b,c]	0.043	0.037	0.016	0.138	0.026	
MAD[b,c]	0.036	0.033	0.014	0.100	0.019	
MSD[b,c]	0.005	−0.017	0.001	0.031	−0.004	

[a]Exponent used in the extrapolation of the (Q) component.
[b]See footnote a to Table II.
[c]Error statistics with respect to the V{T,Q}Z reference values.

merely 0.016 kcal mol^{-1} and largest deviations of +0.026 (C_2) and −0.026 (O_2) kcal mol^{-1}. This approach does not only outperform Eqs. (1) and (2) but also requires no scaling factor (i.e., $c = 1.0$) and no optimization of the extrapolation exponent (i.e., $\alpha = 3.0$). We note that optimization of either or both of these parameters does not lead to an improvement in performance.

Equation (3) still requires calculating the CCSDTQ energy with the VDZ basis set. Using the VDZ(3s2p) basis set instead of the VDZ basis set would represent a significant saving in computational cost. Equation (4) extrapolates the (Q) component from the VDZ(4s3p1d) and VTZ basis sets and calculates the Q–(Q) contribution with the VDZ(3s2p) basis set. With $c = 1.0$ and $\alpha = 3.0$, this results in a significant deterioration in performance relative to Eq. (3). In particular, the RMSD is increased by an order of magnitude, from 0.016 [Eq. (3)] to 0.138 [Eq. (4)] kcal mol^{-1}. Optimizing the scaling factor and extrapolation exponent does not alleviate the situation and results only in a minor improvement in performance (i.e., RMSD = 0.128 kcal mol^{-1}). On the other hand, using Eq. (5), which only scales the Q–(Q) component, results in a respectable RMSD of 0.026 kcal mol^{-1} at a significantly reduced computational cost. We note that this approach uses only one adjustable parameter (i.e., a scaling factor of $c = 0.96$).

Having established that Eq. (3) with $c = 1.0$ and $\alpha = 3.0$ leads to an RMSD of merely 0.016 kcal mol^{-1} and a MAD of 0.014 kcal mol^{-1} relative to V{T,Q}Z results, we can use these reference values to evaluate the performance of more approximate approaches. These results are presented in Table IX for our set of 21 challenging systems. We start by noting that for all the systems the overall CCSDTQ–CCSDT contribution to the TAE is fairly large. Apart from FO, ClO, and S_2, this contribution exceeds 1 kcal mol^{-1} for all the systems and even

exceeds 2 kcal mol^{-1} for as much as ten systems. Particularly large contributions of over 3 kcal mol^{-1} are obtained for ClOO (3.09), O_3 (3.64), and S_4 (4.94 kcal mol^{-1}).

Let us turn our attention to the performance of cost-effective approximations for the CCSDTQ–CCSDT component to the TAE for our set of 21 molecules (Table IX). Calculating the CCSDTQ–CCSDT contribution in conjunction with the VDZ(3s2p) basis set and scaling by 1.4 result in poor performance with an unacceptably high RMSD of 0.72 kcal mol^{-1}. As may be expected, particularly large underestimations are obtained for systems containing multiple second-row atoms, e.g., 1.83 (S_4), 1.04 (S_3), and 0.91 (P_4) kcal mol^{-1}. Calculating the (Q) part with the larger VDZ basis set significantly improves the performance and results in an overall RMSD of 0.34 kcal mol^{-1}. However, fairly large underestimations are still obtained for systems containing multiple second-row atoms, e.g., 0.54 (S_4) and 0.51 (P_4) kcal mol^{-1}, and fairly large overestimations are obtained for highly multireference systems, e.g., 0.62 (O_3) kcal mol^{-1}. Further increasing the basis set size in the (Q) part to the VDZ(4s3p1d) basis set reduces the RMSD to 0.25 kcal mol^{-1} and is therefore recommended for large systems. We note, however, that fairly large deviations are obtained for systems containing multiple second-row atoms and/or strongly multireference systems. For example, for S_4, P_4, and S_3, underestimations of 0.4–0.5 kcal mol^{-1} are obtained, whilst for BN and O_3 overestimations of 0.3–0.4 kcal mol^{-1} are obtained.

E. Perturbative, noniterative connected quintuple excitations

For the set of seven diatomics, we were able to extrapolate the CCSDTQ(5)–CCSDTQ contribution from the VDZ and VTZ basis sets. These results are presented in Table X along with the results for the VDZ(3s2p), VDZ, VDZ(4s3p1d), and

TABLE IX. Performance of cost-effective approaches for obtaining the connected quadruples contribution, CCSDTQ–CCSDT, to the total atomization energy for the set of 21 molecules for which we were able to obtain reference values via Eq. (3) (i.e., extrapolating the (Q) contribution from the VDZ(4s3p1d) and VTZ basis sets and calculating the Q–(Q) contribution with the VDZ basis set). The reference TAEs are listed in the last column, and the other columns list deviations from these reference values (in kcal mol^{-1}).

(Q)	VDZ(3s2p)	VDZ	VDZ(4s3p1d)	V{D(4s3p1d),T}Z
Q–(Q)	VDZ(3s2p)	VDZ(3s2p)	VDZ(3s2p)	VDZ
c	1.4	1.3	1.3	
BN	−1.342	0.059	0.328	2.089
C$_2$	−1.220	−0.557	−0.361	2.327
O$_2$	0.498	0.326	0.131	1.019
FO	0.064	0.139	−0.017	0.544
P$_2$	0.047	−0.316	−0.287	1.478
S$_2$	−0.175	−0.176	−0.180	0.873
ClO	−0.302	−0.079	−0.130	0.644
N$_2$O	0.768	0.368	0.174	1.891
NO$_2$	0.351	0.350	0.114	1.838
O$_3$	0.574	0.618	0.383	3.644
FNO	0.314	0.324	0.088	1.474
FO$_2$	0.164	0.399	0.239	2.778
F$_2$O	0.285	0.343	0.049	1.346
S$_2$O	−0.139	−0.120	−0.248	2.003
S$_3$	−1.044	−0.445	−0.409	2.502
ClNO	−0.232	0.046	−0.033	1.931
ClOO	−0.588	0.044	−0.038	3.086
Cl$_2$O	−0.288	−0.032	−0.155	1.196
N$_2$C$_2$	0.404	0.223	0.182	2.166
P$_4$	−0.907	−0.506	−0.413	2.676
S$_4$	−1.825	−0.539	−0.467	4.943
RMSD[a,b]	0.718	0.339	0.251	
MAD[a,b]	0.549	0.286	0.211	
MSD[a,b]	−0.219	0.022	−0.050	

[a]See footnote a to Table II.
[b]Error statistics with respect to reference values obtained via Eq. (3) for the 21 systems above.

TABLE X. Convergence of the noniterative connected quintuples contribution, CCSDTQ(5)–CCSDTQ, to the total atomization energy for the set of seven diatomic molecules for which we were able to obtain basis-set limit values from V{D,T}Z extrapolations. The reference TAEs are listed in the last column, and the other columns list deviations from these reference values (in kcal mol^{-1}).

	VDZ(3s2p)	VDZ	VDZ(4s3p1d)	VTZ	V{D,T}Z
BN	0.233	0.066	−0.034	0.019	−0.175
C$_2$	0.031	0.094	−0.003	0.028	0.371
O$_2$	−0.025	−0.011	−0.010	−0.003	0.119
FO	0.018	0.018	0.008	0.005	0.024
P$_2$	0.025	0.023	0.022	0.007	0.081
S$_2$	−0.012	0.012	0.014	0.004	0.037
ClO	−0.009	0.007	0.002	0.002	0.032
RMSD[a,b]	0.090	0.045	0.017	0.013	
MAD[a,b]	0.050	0.033	0.013	0.010	
MSD[a,b]	0.037	0.030	0.000	0.009	

[a]See footnote a to Table II.
[b]Error statistics with respect to the V{D,T}Z reference values.

The results in Table X show that, in conjunction with the VDZ(4s3p1d) basis set, the (5) contribution to the TAE is almost converged to the CBS limit. We were able to obtain the (5) contribution to the TAE in conjunction with the VDZ(4s3p1d) basis set for 17 molecules (namely, all the molecules in Table IX apart from FO$_2$, ClOO, P$_4$, and S$_4$). These results are presented in Table S2 of the supplementary material. The (5)/VDZ(4s3p1d) contribution to the TAE is positive, with the exception of BN and FNO for which it is −0.208 and −0.038 kcal mol^{-1}, respectively. For the other systems, the (5)/VDZ(4s3p1d) contribution to the TAE ranges between 0.009 (ClNO) and 0.451 (O$_3$) kcal mol^{-1}. However for five systems, the (5) contribution to the TAE exceeds 0.1 kcal mol^{-1}, namely, P$_2$ (0.103), O$_2$ (0.109), N$_2$C$_2$ (0.205), C$_2$ (0.368), and O$_3$ (0.451 kcal mol^{-1}).

Relative to the VDZ(4s3p1d) results for the abovementioned 17 systems, the VDZ(3s2p) basis set attains an RMSD of 0.079 kcal mol^{-1}. Exclusion of the BN system from the error statistics results in an RMSD of 0.047 kcal mol^{-1}. Apart from BN, all the deviations are smaller or equal to 0.1 kcal mol^{-1}. The largest deviations being 0.10 (NO$_2$), 0.09 (S$_2$O), and 0.07 (FNO and ClNO) kcal mol^{-1}.

F. Full-iterative connected quintuple excitations

For our set of seven diatomics, we were able to extrapolate the CCSDTQ5–CCSDTQ(5) contribution from the VDZ and VTZ basis sets. These results are presented in Table XI along with results for the VDZ(3s2p), VDZ, VDZ(4s3p1d), and VTZ basis sets. We begin with evaluating the performance of the rather small VDZ(3s2p) basis set, which would be the only realistic choice for systems with more than four non-hydrogen atoms. For the seven diatomic systems in Table XI, the VDZ(3s2p) basis set attains a rather large RMSD of 0.118 kcal mol^{-1}. This large RMSD is a result of very large deviations obtained for BN and C$_2$ (Table XI). Elimination of these two challenging systems results in an RMSD of 0.051 kcal mol^{-1}. The overall RMSD for the VDZ basis set

VTZ basis sets. Consistent with previous studies,[8,38] the (5) contribution to the TAE converges fairly rapidly to the basis-set limit. We begin with evaluating the performance of the rather small VDZ(3s2p) basis set and note that for systems with more than four non-hydrogen atoms this would normally be the only feasible option with current mainstream technology. For the seven diatomic systems in Table X, the VDZ(3s2p) basis set attains a fairly large RMSD of 0.090 kcal mol^{-1}. We note, however, that the VDZ(3s2p) basis set gives fairly good results for all the systems apart from the highly multireference BN system. Once this system is eliminated, the RMSD is reduced to merely 0.022 kcal mol^{-1}. The error for the BN system is reduced from 0.233 kcal mol^{-1} for the VDZ(3s2p) basis set to 0.066 kcal mol^{-1} for the VDZ basis set. Thus, for strongly multireference and polar systems we recommend using the VDZ basis set. The overall RMSD for the VDZ basis set is 50% of that for the VDZ(3s2p) basis set, namely, it is 0.045 kcal mol^{-1}. Increasing the basis set size to VDZ(4s3p1d) results in another significant reduction in the RMSD to merely 0.017 kcal mol^{-1}, and the VTZ basis set leads to an RMSD of 0.013 kcal mol^{-1}.

TABLE XI. Convergence of the full-iterative connected quintuples contribution, CCSDTQ5–CCSDTQ(5), to the total atomization energy for the set of seven diatomic molecules for which we were able to obtain basis-set limit values from V{D,T}Z extrapolations. The reference TAEs are listed in the last column, and the other columns list deviations from these reference values (in kcal mol^{-1}).

	VDZ(3s2p)	VDZ	VDZ(4s3p1d)	VTZ	V{D,T}Z
BN	−0.254	−0.105	−0.019	−0.031	0.373
C$_2$	−0.144	−0.120	−0.049	−0.035	−0.026
O$_2$	0.001	−0.002	−0.002	−0.001	−0.003
FO	−0.028	−0.024	−0.013	−0.007	0.015
P$_2$	−0.096	−0.071	−0.068	−0.021	0.091
S$_2$	−0.045	−0.043	−0.042	−0.013	0.050
ClO	−0.027	−0.024	−0.018	−0.007	0.020
RMSD[a,b]	0.118	0.069	0.037	0.020	
MAD[a,b]	0.085	0.055	0.030	0.016	
MSD[a,b]	−0.085	−0.055	−0.030	−0.016	

[a]See footnote a to Table II.
[b]Error statistics with respect to the V{D,T}Z reference values.

is about 50% of that for the VDZ(3s2p) basis set, namely, it is 0.069 kcal mol^{-1} (elimination of BN and C$_2$ reduces the RMSD to 0.040 kcal mol^{-1}). Increasing the basis set size to VDZ(4s3p1d) results in another significant reduction in the RMSD to merely 0.037 kcal mol^{-1}. The VTZ basis set results in an RMSD of 0.020 kcal mol^{-1}.

For all 21 systems, we were able to obtain the 5–(5) contribution to the TAE in conjunction with the VDZ(3s2p) basis set, and for 16 systems we were able to obtain this contribution with the VDZ basis set. These results are presented in Table S3 of the supplementary material. As indicated above, the VDZ values are not sufficiently close to the basis-set limit to be used as reference values, nevertheless, it is worthwhile comparing between the VDZ and VDZ(3s2p) results. Relative to the VDZ values, the VDZ(3s2p) basis set attains an RMSD 0.035 kcal mol^{-1} (once BN is excluded, *vide supra*). Inspection of the VDZ(3s2p) results reveals that for about half of the systems,

the 5–(5) contribution to the TAE is negative and for about half it is positive. The largest negative contributions to the TAE are obtained for C$_2$ (−0.170), O$_3$ (−0.080), ClOO (−0.051), and FO$_2$ (−0.042 kcal mol^{-1}). The largest positive contributions to the TAE are obtained for N$_2$O (0.159), BN (0.120), S$_4$ (0.083), FNO (0.058), and ClNO (0.039 kcal mol^{-1}).

G. Connected quintuple excitations as a whole

In Subsection III E, we examined the basis set convergence of the (5) component relative to (5)/V{D,T}Z reference values. It is also of interest to evaluate the performance of the (5) component calculated with the VDZ(3s2p), VDZ, VDZ(4s3p1d), and VTZ basis sets relative to the full connected quintuple, CCSDTQ5–CCSDTQ, component. For our set of seven diatomics, we were able to extrapolate the CCSDTQ5–CCSDTQ component to the basis-set limit from the VDZ and VTZ basis sets. These results are presented in Table XII. Relative to the 5/V{D,T}Z reference values, the (5)/VDZ(3s2p) results in a surprisingly small RMSD of 0.068 kcal mol^{-1}. This is due to error cancellation between the neglect of the fully iterative quintuple excitations and basis set incompleteness errors. Indeed, the performance deteriorates significantly when larger basis sets are used, namely, the RMSDs for the larger basis sets range between 0.128 (VDZ) and 0.157 (VDZ(4s3p1d)) kcal mol^{-1}.

What about calculating the CCSDTQ5–CCSDTQ component in conjunction with small basis sets? The VDZ(3s2p) basis set results in an RMSD of 0.058 kcal mol^{-1} and seems to be a fairly robust and cost-effective option. Adding the d function from the VDZ basis set, however, cuts the RMSD by 50% to 0.029 kcal mol^{-1}; thus, this would be the recommended option in case it is computationally affordable. The VDZ(4s3p1d) basis set results in a similar performance and is therefore not recommended. Finally, we note that the VTZ basis set gives results that are essentially identical to the V{D,T}Z results, i.e., RMSD = 0.009 kcal mol^{-1} and the deviations (all underestimations) are smaller or equal to 0.01 kcal mol^{-1}.

TABLE XII. Performance of the CCSDTQ(5)–CCSDTQ and CCSDTQ5–CCSDTQ contributions to the total atomization energy for a set of seven diatomic molecules relative to the CCSDTQ5–CCSDTQ contribution extrapolated from the V{D,T}Z basis set pair. The reference TAEs are listed in the last column, and the other columns list deviations from these reference values (in kcal mol^{-1}).

	(5) VDZ(3s2p)	(5) VDZ	(5) VDZ(4s3p1d)	(5) VTZ	5 VDZ(3s2p)	5 VDZ	5 VDZ(4s3p1d)	5 VTZ	5 V{D,T}Z
BN	−0.141	−0.308	−0.407	−0.354	−0.021	−0.039	−0.053	−0.012	0.199
C$_2$	0.058	0.121	0.023	0.054	−0.112	−0.025	−0.052	−0.007	0.345
O$_2$	−0.023	−0.008	−0.007	0.000	−0.024	−0.013	−0.012	−0.004	0.116
FO	0.003	0.004	−0.007	−0.009	−0.011	−0.005	−0.005	−0.002	0.039
P$_2$	−0.066	−0.068	−0.069	−0.084	−0.072	−0.047	−0.045	−0.014	0.172
S$_2$	−0.062	−0.038	−0.037	−0.047	−0.057	−0.031	−0.028	−0.009	0.088
ClO	−0.030	−0.014	−0.019	−0.018	−0.036	−0.017	−0.016	−0.005	0.053
RMSD[a,b]	0.068	0.128	0.157	0.140	0.058	0.029	0.035	0.009	
MAD[a,b]	0.055	0.080	0.081	0.081	0.048	0.025	0.030	0.008	
MSD[a,b]	−0.037	−0.044	−0.075	−0.066	−0.048	−0.025	−0.030	−0.008	

[a]See footnote a to Table II.
[b]Error statistics with respect to the CCSDTQ5/V{D,T}Z–CCSDTQ/V{D,T}Z reference values for the seven systems above.

For a set of 15 di- and triatomics, we were able to obtain the CCSDTQ5–CCSDTQ contribution to the TAE in conjunction with the VDZ basis set. Although these values are not sufficiently close to the basis-set limit to be used as reference values, it is worthwhile comparing between the VDZ and VDZ(3s2p) results. For the set of 15 systems, we obtain an overall RMSD of 0.030 kcal mol^{-1}, with only one deviation (for C_2) above 0.04 kcal mol^{-1}. This is another indication that calculating the CCSDTQ5–CCSDTQ contribution in conjunction with the VDZ(3s2p) is a cost-effective option.

Finally, a note is due on the magnitude of the CCSDTQ5–CCSDTQ contribution in conjunction with the VDZ(3s2p) basis set. The CCSDTQ5–CCSDTQ contribution to the TAE [in conjunction with the VDZ(3s2p) basis set] spans a wide interval, ranging from 0.017 (ClO) to 0.395 (O_3) kcal mol^{-1}. For most systems (13 out of 21), the CCSDTQ5–CCSDTQ contribution exceeds 0.1 kcal mol^{-1}, and for six of these it exceeds 0.2 kcal mol^{-1}, namely, for ClOO (0.222), N_2C_2 (0.229), FO_2 (0.230), C_2 (0.232), S_4 (0.291), and O_3 (0.395 kcal mol^{-1}).

H. Connected sextuple excitations

Contributions beyond CCSDTQ5 can still be chemically significant for systems dominated by moderate-to-severe multireference effects. Table XIII gathers the CCSDTQ5(6)–CCSDTQ5 contribution to the TAE calculated in conjunction with the VDZ(3s2p), VDZ, VDZ(4s3p1d), VTZ, and V{D,T}Z basis sets, as well as the CCSDTQ56–CCSDTQ5(6) contribution to the TAE calculated in conjunction with the VDZ(3s2p) basis set. Let us begin by considering the basis set

convergence of the CCSDTQ5(6)–CCSDTQ5 component. For three systems (BN, C_2, and P_2), we were able to obtain the (6) contribution with the V{D,T}Z basis set pair. The highly multireference BN diatomic exhibits a somewhat slow basis-set convergence. At the (6)/V{D,T}Z level of theory, we obtain a value of −0.011 kcal mol^{-1}. Both the VDZ(4s3p1d) and VTZ basis sets overshoot this value by about 0.008 kcal mol^{-1}. The VDZ and VDZ(3s2p) basis sets overshoot this value by as much as 0.025 and 0.067 kcal mol^{-1}, respectively. Nevertheless, for the other two systems for which we have (6)/V{D,T}Z basis-set limit values (C_2 and P_2), the VDZ and VTZ basis sets are practically spot on the (6)/V{D,T}Z values, and even the VDZ(3s2p) basis set results in deviations smaller or equal to 0.01 kcal mol^{-1}.

For the seven diatomics in our set, we were able to calculate the (6) contribution in conjunction with the VDZ(4s3p1d) basis set. Relative to these values, the VDZ and VDZ(3s2p) basis sets attain RMSDs of 0.002 and 0.005 kcal mol^{-1}, respectively (after exclusion of the problematic BN system). Thus, it seems that the (6)/VDZ(3s2p) contribution is fairly close to the basis-set limit. With this in mind, it is useful to examine the magnitude of the (6)/VDZ(3s2p) contribution for our set of 21 systems (Table XIII). For most of the systems, the (6) contribution ranges between 0.00 and 0.03 kcal mol^{-1}; however, for three strongly multireference systems (BN, C_2, and O_3), it ranges between 0.05 and 0.07 kcal mol^{-1}.

With the exception of FO_2, ClOO, P_4, and S_4, we were able to obtain the CCSDTQ56–CCSDTQ5 contribution in conjunction with the VDZ(3s2p) basis set for all the molecules in our set (Table XIII). For most of the systems, this contribution is equal to or smaller than 0.01 kcal mol^{-1}. Nevertheless,

TABLE XIII. Overview of basis set convergence of the CCSDTQ5(6)–CCSDTQ5 contribution to the total atomization energy and the magnitude of the CCSDTQ56–CCSDTQ5(6) contribution calculated with the VDZ(3s2p) basis set (in kcal mol^{-1}).

	(6) VDZ(3s2p)	(6) VDZ	(6) VDZ(4s3p1d)	(6) VTZ	(6) V{D,T}Z	6–(6) VDZ(3s2p)	6 VDZ(3s2p)
BN	0.056	0.014	−0.004	−0.003	−0.011	−0.026	0.030
C_2	0.065	0.061	0.057	0.060	0.060	0.003	0.069
O_2	0.012	0.011	0.010			0.002	0.014
FO	0.003	0.003	0.001			0.000	0.003
P_2	0.020	0.010	0.011	0.010	0.010	0.006	0.026
S_2	0.003	0.004	0.004			0.001	0.004
ClO	0.001	0.003	0.002			0.000	0.001
N_2O	0.010					0.015	0.025
NO_2	0.017					0.003	0.020
O_3	0.055					−0.001	0.053
FNO	0.003					0.006	0.009
FO_2	0.026						
F_2O	0.010					0.001	0.010
S_2O	0.015					0.003	0.018
S_3	0.011					0.003	0.014
ClNO	0.007					0.004	0.012
ClOO	0.023						
Cl_2O	0.006					0.001	0.007
N_2C_2	0.023					0.003	0.026
P_4	0.020						
S_4	0.030						

TABLE XIV. Overview of the magnitude of the CCSDTQ56(7)–CCSDTQ56, CCSDTQ567–CCSDTQ56(7), and CCSDTQ567–CCSDTQ56 contributions to the total atomization energy calculated with the VDZ(3s2p) basis set (in kcal mol^{-1}).

	(7)	7–(7)	7
BN	−0.003	0.005	0.003
C_2	0.003	0.001	0.004
O_2	0.001	0.000	0.001
FO	0.000	0.000	0.000
P_2	0.001	0.000	0.002
S_2	0.000	0.000	0.000
ClO	0.000	0.000	0.000
N_2O	0.000		
NO_2	0.000		
O_3	0.003	0.001	0.005
F_2O	0.001		
S_2O	0.000		
S_3	0.000	0.001	0.001
ClNO	0.000		
Cl_2O	0.000		

we find that for six systems the CCSDTQ56–CCSDTQ5 contribution ranges between 0.02 and 0.03 kcal mol^{-1}. These systems are S_2O (0.018), NO_2 (0.020), N_2O (0.025), P_2 (0.026), N_2C_2 (0.026), and BN (0.030 kcal mol^{-1}). Whilst for the strongly multireference molecules O_3 and C_2, the CCSDTQ56–CCSDTQ5 contribution is 0.053 and 0.069 kcal mol^{-1}, respectively.

I. Connected septuple excitations

For the sake of completeness, we have calculated the CCSDTQ56(7)–CCSDTQ56, CCSDTQ567–CCSDTQ56(7), and CCSDTQ567–CCSDTQ56 contributions to the TAE for a subset of 15 systems. These results are presented in Table XIV. As expected, these contributions are practically nil for all the systems including the strongly multi-

reference systems. In particular, the largest CCSDTQ567–CCSDTQ56 contributions to the TAE are obtained for BN (0.003), C_2 (0.004), and O_3 (0.005 kcal mol^{-1}).

J. Post-CCSD(T) excitations as a whole

Finally, it is of interest to examine the basis set convergence of post-CCSD(T) excitations as a whole. Table XV lists the contribution of post-CCSD(T) contributions (up to CCSDTQ5(6)) to the total atomization energy for the set of seven diatomic molecules. The reference values are taken as [T–(T)]/V{5,6}Z + (Q)/V{Q,5}Z + [Q–(Q)]/V{T,Q}Z + (5)/V{D,T}Z + [5–(5)]/V{D,T}Z + (6)/VDZ(4s3p1d). Relative to these CBS values, the VDZ(3s2p) basis set attains a large RMSD of 0.83 kcal mol^{-1}. Inclusion of the d functions in the VDZ basis set reduces the RMSD by 66% to 0.28 kcal mol^{-1}. Thus, it is clear that calculating the post-CCSD(T) contributions as a whole benefits from a significant degree of error cancellation. For example, the VDZ basis set gives RMSDs close to 1 kcal mol^{-1} for the T–(T) and (Q) components (Tables II–IV). Addition of s and p functions in the VDZ(4s3p1d) basis set further reduces the RMSD to 0.17 kcal mol^{-1}. However, in order to obtain an RMSD smaller than 0.1 kcal mol^{-1} with a single basis set, one has to calculate all the post-CCSD(T) excitations with the VTZ basis set. This is of course not a viable option for systems with more than 1–2 non-hydrogen atoms and illustrates the for using efficient post-CCSD(T) composite approaches.[3–8,38] Sections III A–III I suggest that calculating the post-CCSD(T) in the following way would be a cost effective approach,

$$[T–(T)]/V\{D,T\}Z + (Q)/V\{VDZ(4s3p1d),T\}Z$$
$$+ 1.1 \times [Q–(Q)]/VDZ + (5)/VDZ(3s2p)$$
$$+ [5–(5)]/VDZ(3s2p) + (6)/VDZ(3s2p),$$

TABLE XV. Convergence of post-CCSD(T) contributions as a whole (up to CCSDTQ5(6)) to the total atomization energy for the set of seven diatomic molecules. The reference TAEs (see footnote a) are listed in the last column, and the tabulated values in the other columns are deviations relative to these basis-set limit values (in kcal mol^{-1}).

Basis set	VDZ(3s2p)	VDZ	VDZ(4s3p1d)	VTZ	V{D,T}Z	Composite[a]	Reference[b]
BN	−1.860	−0.023	−0.309	0.014	0.022	0.022	−0.408
C_2	−0.801	0.326	0.086	0.198	0.174	−0.037	0.401
O_2	0.492	0.449	0.065	0.058	−0.018	−0.076	0.641
FO	0.300	0.310	0.019	0.016	−0.041	−0.026	0.821
P_2	0.615	0.324	0.256	0.057	0.006	−0.053	0.675
S_2	0.191	0.179	0.109	0.034	0.006	−0.041	0.476
ClO	−0.114	0.021	−0.077	−0.047	−0.060	−0.015	0.736
RMSD[c,d]	0.833	0.278	0.165	0.084	0.072	0.043	
MAD[c,d]	0.625	0.233	0.131	0.061	0.047	0.039	
MSD[c,d]	−0.168	0.226	0.021	0.047	0.013	−0.032	

[a]These values are taken as [T–(T)]/V{D,T}Z + (Q)/V{VDZ(4s3p1d),T}Z + 1.1 × [Q–(Q)]/VDZ + (5)/VDZ(3s2p) + [5–(5)]/VDZ(3s2p) + (6)/VDZ(3s2p), where the T–(T) component is extrapolated with α = 2.7 and the (Q) component is extrapolated with α = 2.3.
[b]The post-CCSD(T)/CBS reference vales are taken as [T–(T)]/V{5,6}Z + (Q)/V{Q,5}Z + [Q–(Q)]/V{T,Q}Z + (5)/V{D,T}Z + [5–(5)]/V{D,T}Z + (6)/VDZ(4s3p1d).
[c]See footnote a to Table II.
[d]Error statistics with respect to the reference values listed in the last column.

where the T–(T) and (Q) components are extrapolated with α = 2.7 and 2.3, respectively. This approach results in an overall RMSD of 0.04 kcal mol^{-1} at a significantly reduced computational cost. For comparison, calculating all the post-CCSD(T) contributions with the VTZ basis set results in an RMSD of 0.08 kcal mol^{-1}. Finally, we note that extrapolating the post-CCSD(T) contribution as a whole from the VDZ and VTZ basis sets (with an effective exponent of α = 4.5) results in an RMSD of 0.07 kcal mol^{-1} and does not represent a significant improvement over the VTZ results.

IV. CONCLUSIONS

We have examined the basis set convergence of post-CCSD(T) contributions to the total atomization energies near (or at) the one-particle basis-set limit for a diverse set of 21 challenging molecules. We considered iterative coupled cluster contributions up to connected septuples, CCSDTQ567. The set of 21 diatomic, triatomic, and tetra-atomic molecules spans the gamut from molecules dominated by moderate nondynamical correlation (e.g., FO, ClO, NO_2, S_2O, N_2C_2, and P_4) to systems dominated by strong nondynamical correlation (e.g., BN, C_2, FO_2, O_3, ClOO, and S_4). Our main conclusions can be summarized as follows:

- **The CCSDT–CCSD(T) component.** The V{T,Q}Z extrapolation with an effective exponent of α = 2.35 provides near basis-set limit results, with an RMSD of merely 0.009 kcal mol^{-1} relative to V{5,6}Z extrapolations. The computationally more economical V{D,T}Z extrapolation with an effective exponent of α = 2.7 seems to be the most cost effective approach, resulting in an RMSD of 0.09 kcal mol^{-1} relative to V{T,Q}Z extrapolations for a wide range of 21 challenging systems. Nevertheless, the V{D,T}Z extrapolation should be used with caution for systems containing multiple second-row atoms, for which large deviations may be obtained. For example, deviations of +0.17 and −0.29 kcal mol^{-1} are obtained for S_4 and P_4, respectively. Exclusion of these two systems from the error statistics results in an RMSD of 0.05 kcal mol^{-1} for a set of 19 challenging systems.

- **The CCSDT(Q)–CCSDT component.** Similar to the T–(T) component, the V{T,Q}Z extrapolation (with α = 3.0) agrees very well with basis-set limit V{5,6}Z results, with an RMSD of 0.014 kcal mol^{-1}. Using the V{T,Q}Z results, which we were able to obtain for the entire set of 21 molecules, we obtain a somewhat disappointing RMSD of 0.16 kcal mol^{-1} for the V{D,T}Z extrapolation (with α = 3.0). Optimizing the exponent results in α = 2.4 and an RMSD of 0.13 kcal mol^{-1}. However, this performance can be significantly improved at no additional computational cost by replacing the VDZ basis set with the VDZ(4s3p1d) basis set. In particular, the V{D(4s3p1d),T}Z extrapolation (with an effective exponent of α = 2.3) results in an RMSD of 0.07 kcal mol^{-1}. The use of the V{D(4s3p1d),T}Z extrapolation instead of V{D,T}Z is particularly recommended for systems containing multiple second-row atoms.

- **The CCSDTQ–CCSDT component as a whole.** Extrapolating the CCSDTQ–CCSDT component from the V{D(4s3p1d),T}Z basis set pair results in an RMSD of 0.019 kcal mol^{-1}. However, this approach requires running the expensive CCSDTQ/VTZ calculation. A more cost-effective approach is to use Eq. (3), which requires running the CCSDT(Q)/VTZ and CCSDTQ/VDZ calculations. This yields an RMSD of merely 0.016 kcal mol^{-1}. An additional advantage of this approach is that it is completely free from adjustable parameters [i.e., Eq. (3) uses c = 1.0 and α = 3]. Reducing the size of the basis set in the CCS-DTQ calculation to VDZ(3s2p) and using Eq. (5) with c = 0.96 result in a slight deterioration in performance at a significantly lower computational cost. In particular, this approach yields an RMSD of 0.026 kcal mol^{-1}. For larger systems where the CCSDT(Q)/VTZ calculation is prohibitive, we recommend calculating the (Q) contribution with the VDZ(4s3p1d) basis set and the Q–(Q) contribution with the VDZ(3s2p) basis set and scaling this result by 1.3. This approach yields an RMSD of 0.251 kcal mol^{-1} for a set of 21 challenging systems. However, it must be used with caution for systems containing multiple second-row atoms (e.g., S_3, P_4, and S_4).

- **The CCSDTQ(5)–CCSDTQ component.** The VDZ(3s2p) basis set is a very cost-effective basis set for calculating the (5) component. For a set of 17 systems for which we were able to obtain near basis-set-limit results with the VDZ(4s3p1d) basis set, the VDZ(3s2p) attains an RMSD of 0.047 kcal mol^{-1}, after the strongly multireference BN system is excluded. The (5)/VDZ(3s2p) component also shows fairly good performance (RMSD = 0.068 kcal mol^{-1}) relative to the CCSDTQ5–CCSDTQ component. This is due to error cancellation between the neglect of the fully iterative quintuple excitations and basis set incompleteness errors. Nevertheless, for multireference systems, the VDZ(3s2p) basis set can lead to fairly large errors; in these cases, the VDZ basis set is recommended.

- **The CCSDTQ56–CCSDTQ5 component.** The connected sextuple excitations are still chemically significant and can reach up to 0.07 kcal mol^{-1} for strongly multireference systems. This contribution to the TAE is captured fairly accurately at the (6)/VDZ(3s2p) level.

- **The CCSDTQ567–CCSDTQ56 component.** As expected, these contributions are practically nil for all the systems, including the strongly multireference ones. The largest CCSDTQ567–CCSDTQ56 contributions to the TAE are obtained for BN (0.003), C_2 (0.004), and O_3 (0.005 kcal mol^{-1}).

SUPPLEMENTARY MATERIAL

See supplementary material for the optimized geometries for all the structures (Table S1); an overview of the magnitude and basis-set convergence of the CCSDTQ(5)–CCSDTQ contribution to the total atomization energy for the set of 21 molecules (Table S2); an overview of the magnitude and

basis-set convergence of the CCSDTQ5–CCSDTQ(5) contribution to the total atomization energy for the set of 21 molecules (Table S3); and central processing unit (CPU)-times and the number of amplitudes involved in some of the larger calculations reported in the present study (Table S4).

ACKNOWLEDGMENTS

[1] I. Shavitt and R. J. Bartlett, in *Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory*, Cambridge Molecular Science Series (Cambridge University Press, Cambridge, 2009).

[2] R. J. Bartlett and M. Musial, Rev. Mod. Phys. **79**, 291 (2007).

[3] A. Karton, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **6**, 292 (2016).

[4] K. A. Peterson, D. Feller, and D. A. Dixon, Theor. Chem. Acc. **131**, 1079 (2012).

[5] A. Karton, S. Daon, and J. M. L. Martin, Chem. Phys. Lett. **510**, 165 (2011).

[6] D. Feller, K. A. Peterson, and D. A. Dixon, J. Chem. Phys. **129**, 204105 (2008).

[7] T. Helgaker, W. Klopper, and D. P. Tew, Mol. Phys. **106**, 2107 (2008).

[8] A. Karton, E. Rabinovich, J. M. L. Martin, and B. Ruscic, J. Chem. Phys. **125**, 144108 (2006).

[9] M. E. Harding, J. Vázquez, B. Ruscic, A. K. Wilson, J. Gauss, and J. F. Stanton, J. Chem. Phys. **128**, 114111 (2008).

[10] A. J. C. Varandas, Annu. Rev. Phys. Chem. **69**, 177 (2018).

[11] F. N. N. Pansini and A. J. C. Varandas, Chem. Phys. Lett. **631-632**, 70 (2015).

[12] F. N. N. Pansini, A. C. Neto, and A. J. C. Varandas, Chem. Phys. Lett. **641**, 90 (2015).

[13] A. J. C. Varandas and F. N. N. Pansini, J. Chem. Phys. **141**, 224113 (2014).

[14] P. R. Spackman and A. Karton, AIP Adv. **5**, 057148 (2015).

[15] D. Feller, J. Chem. Phys. **138**, 074103 (2013).

[16] D. S. Ranasinghe and G. A. Petersson, J. Chem. Phys. **138**, 144104 (2013).

[17] A. Karton and J. M. L. Martin, J. Chem. Phys. **136**, 124114 (2012).

[18] D. Feller, K. A. Peterson, and J. G. Hill, J. Chem. Phys. **135**, 044102 (2011).

[19] D. Feller, K. A. Peterson, and D. A. Dixon, J. Phys. Chem. A **114**, 613 (2010).

[20] J. G. Hill, K. A. Peterson, G. Knizia, and H.-J. Werner, J. Chem. Phys. **131**, 194105 (2009).

[21] N. DeYonker, T. R. Cundari, and A. K. Wilson, in *Advances in the Theory of Atomic and Molecular Systems*, Progress in Theoretical Chemistry and Physics, edited by P. Piecuch, J. Maruani, G. Delgado-Barrio, and S. Wilson (Springer Netherlands, Dordrecht, 2009), Vol. 19, pp. 197–224.

[22] E. C. Barnes, G. A. Petersson, D. Feller, and K. A. Peterson, J. Chem. Phys. **129**, 194115 (2008).

[23] D. Feller and K. A. Peterson, J. Chem. Phys. **126**, 114105 (2007).

[24] K. Kahn, A. A. Granovsky, and J. Noga, J. Comput. Chem. **28**, 547 (2007).

[25] D. Feller, K. A. Peterson, and T. D. Crawford, J. Chem. Phys. **124**, 054107 (2006).

[26] D. W. Schwenke, J. Chem. Phys. **122**, 014107 (2005).

[27] J. M. L. Martin, "Computational thermochemistry: A brief overview of quantum mechanical approaches," Annu. Rep. Comput. Chem. **1**, 31 (2005).

[28] E. F. Valeev, W. D. Allen, R. Hernandez, C. D. Sherrill, and H. F. Schaefer, J. Chem. Phys. **118**, 8594 (2003).

[29] D. Feller and D. A. Dixon, J. Chem. Phys. **115**, 3484 (2001).

[30] T. Helgaker, W. Klopper, K. L. Bak, A. Halkier, P. Jørgensen, and J. Olsen, "Highly accurate *ab initio* computation of thermochemical data," in *Quantum-Mechanical Prediction of Thermochemical Data*, Understanding Chemical Reactivity, edited by J. Cioslowski (Kluwer, Dordrecht, 2001), Vol. 22, pp. 1–30.

[31] J. M. L. Martin and S. Parthiban, "W1 and W2 theory and their variants: Thermochemistry in the kJ/mol accuracy range," in *Quantum-mechanical Prediction of Thermochemical Data*, Understanding Chemical Reactivity, edited by J. Cioslowski (Kluwer, Dordrecht, 2001), Vol. 22, pp. 31–65.

[32] W. Klopper, Mol. Phys. **99**, 481 (2001).

[33] K. L. Bak, P. Jørgensen, J. Olsen, T. Helgaker, and W. Klopper, J. Chem. Phys. **112**, 9229 (2000).

[34] J. M. L. Martin and G. de Oliveira, J. Chem. Phys. **111**, 1843 (1999).

[35] N. Sylvetsky, K. A. Peterson, A. Karton, and J. M. L. Martin, J. Chem. Phys. **144**, 214101 (2016).

[36] D. Feller, K. A. Peterson, and B. Ruscic, Theor. Chem. Acc. **133**, 1407 (2014).

[37] A. Karton, S. Parthiban, and J. M. L. Martin, J. Phys. Chem. A **113**, 4802 (2009).

[38] A. Karton, P. R. Taylor, and J. M. L. Martin, J. Chem. Phys. **127**, 064104 (2007).

[39] A. D. Boese, M. Oren, O. Atasoylu, J. M. L. Martin, M. Kállay, and J. Gauss, J. Chem. Phys. **120**, 4129 (2004).

[40] A. Tajti, P. G. Szalay, A. G. Császár, M. Kállay, J. Gauss, E. F. Valeev, B. A. Flowers, J. Vázquez, and J. F. Stanton, J. Chem. Phys. **121**, 11599 (2004).

[41] Y. J. Bomble, J. Vázquez, M. Kállay, C. Michauk, P. G. Szalay, A. G. Császár, J. Gauss, and J. F. Stanton, J. Chem. Phys. **125**, 064108 (2006).

[42] T. A. Ruden, T. Helgaker, P. Jørgensen, and J. Olsen, Chem. Phys. Lett. **371**, 62 (2003).

[43] M. Musial and R. J. Bartlett, J. Chem. Phys. **133**, 104102 (2010).

[44] A. J. C. Varandas, J. Chem. Phys. **131**, 124128 (2009).

[45] P. Piecuch, M. Włoch, and A. J. C. Varandas, Theor. Chem. Acc. **120**, 59 (2008).

[46] P. Piecuch, I. S. O. Pimienta, P.-D. Fan, and K. Kowalski, "New alternatives for accurate electronic structure calculations of potential energy surfaces involving bond breaking," ACS Symp. Ser. **958**, 37–73 (2007).

[47] C. D. Sherrill and P. Piecuch, J. Chem. Phys. **122**, 124104 (2005).

[48] M. L. Abrams and C. D. Sherrill, J. Chem. Phys. **121**, 9211 (2004).

[49] A. Dutta and C. D. Sherrill, J. Chem. Phys. **118**, 1610 (2003).

[50] P. Piecuch, K. Kowalski, I. S. O. Pimienta, and M. J. McGuire, Int. Rev. Phys. Chem. **21**, 527 (2002).

[51] A. Karton, N. Sylvetsky, and J. M. L. Martin, J. Comput. Chem. **38**, 2063 (2017).

[52] P. R. Spackman, D. Jayatilaka, and A. Karton, J. Chem. Phys. **145**, 104101 (2016).

[53] MRCC, A Quantum Chemical Program Suite written by M. Kállay *et al.*, See also: http://www.mrcc.hu.

[54] Z. Rolik, L. Szegedy, I. Ladjanszki, B. Ladoczki, and M. Kállay, J. Chem. Phys. **139**, 094105–094117 (2013).

[55] T. H. Dunning, J. Chem. Phys. **90**, 1007 (1989).

[56] R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).

[57] T. H. Dunning, K. A. Peterson, and A. K. Wilson, J. Chem. Phys. **114**, 9244 (2001).