

# W4-17: A Diverse and High-Confidence Dataset of Atomization Energies for Benchmarking High-Level Electronic Structure Methods

Amir Karton <sup>[a]</sup>, Nitai Sylvetsky,<sup>[b]</sup> and Jan M. L. Martin<sup>\*,[b]</sup>

Atomization reactions are among the most challenging tests for electronic structure methods. We use the first-principles Weizmann-4 (W4) computational thermochemistry protocol to generate the W4-17 dataset of 200 total atomization energies (TAEs) with  $3\sigma$  confidence intervals of  $1 \text{ kJ mol}^{-1}$ . W4-17 is an extension of the earlier W4-11 dataset; it includes first- and second-row molecules and radicals with up to eight non-hydrogen atoms. These cover a broad spectrum of bonding situations and multireference character, and as such are an excellent benchmark for the parameterization and validation of highly accurate *ab initio* methods (e.g., CCSD(T) composite procedures) and double-hybrid density functional theory (DHDFT) methods. The W4-17 dataset contains

two subsets (i) a non-multireference subset of 183 systems characterized by dynamical or moderate nondynamical correlation effects (denoted W4-17-nonMR) and (ii) a highly multireference subset of 17 systems (W4-17-MR). We use these databases to evaluate the performance of a wide range of CCSD(T) composite procedures (e.g., G4, G4(MP2), G4(MP2)-6X, ROG4(MP2)-6X, CBS-QB3, ROCBS-QB3, CBS-APNO, ccCA-PS3, W1, W2, W1-F12, W2-F12, W1X-1, and W2X) and DHDFT methods (e.g., B2-PLYP, B2GP-PLYP, B2K-PLYP, DSD-BLYP, DSD-PBEP86, PWPB95,  $\omega$ B97X-2(LP), and  $\omega$ B97X-2(TQZ)). © 2017 Wiley Periodicals, Inc.

DOI: 10.1002/jcc.24854

## Introduction

It is well established that the performance of approximate theoretical procedures can vary for different types of chemical reactions. In particular, the accuracy of any given approximate method should increase as larger molecular fragments are conserved on the two sides of the reaction, due to an increasing degree of error cancellation between reactants and products.<sup>[1–11]</sup> For example, the performance should improve along the sequence: atomization  $\rightarrow$  isogyric  $\rightarrow$  isodesmic  $\rightarrow$  hypohomodesmotic  $\rightarrow$  homodesmotic  $\rightarrow$  hyperhomodesmotic reactions.<sup>[1–4]</sup> The same trend is expected along a generalization of this hierarchy, namely, the connectivity-based hierarchical (CBH-*n*) reaction scheme, in the sequence: atomization (CBH-0)  $\rightarrow$  atom-centric (CBH-1)  $\rightarrow$  bond-centric (CBH-2), and so forth.<sup>[6–8]</sup>

The performance of theoretical procedures can vary significantly between the various reaction types. In the context of density functional theory (DFT), it has been shown that the performance of a wide range of exchange-correlation functionals for linear alkane  $\rightarrow$  branched alkane isomerizations is significantly better than that for isogyric structural isomerizations.<sup>[10–20]</sup> For example, the root mean square deviations (RMSDs) for a wide range of dispersion-corrected DFT methods for linear  $\rightarrow$  branched alkane isomerizations (i.e., reactions that conserve the number of C atoms in each hybridization state in addition to being isodesmic) vary between  $\sim 0.1$ – $1 \text{ kcal mol}^{-1}$ .<sup>[5]</sup> The RMSDs for isodesmic reactions involving alkanes vary between  $\sim 0.3$ – $3 \text{ kcal mol}^{-1}$ , while for alkane atomization reactions they vary between  $\sim 2$ – $25 \text{ kcal mol}^{-1}$ .<sup>[5]</sup> Similarly, for the atomization energies in the W4-11 database the RMSDs for a wide range of DFT methods range between 3.6 and  $24.9 \text{ kcal mol}^{-1}$ .<sup>[1]</sup> Similar observations can be made when comparing the performance of a wide range of composite *ab initio* methods that

attempt to approximate the CCSD(T) energy (e.g., the Gaussian-*n*, CBS, and Weizmann-*n* protocols). Namely, RMSDs in the following ranges are obtained:  $0.2$ – $1.0 \text{ kcal mol}^{-1}$  (structural isomerizations) and  $0.9$ – $2.2 \text{ kcal mol}^{-1}$  (atomization reactions).<sup>[1]</sup> We note that in the evaluation of these accurate composite procedures FCI/CBS reference data from W4 theory were used.<sup>[1,21–23]</sup>

Atomization reactions are therefore of special interest in the evaluation and development of approximate quantum chemical methods as they serve as the ultimate test for these procedures. This is particularly true for high-level quantum chemical methods such as composite *ab initio* protocols (e.g., Gaussian-*n*, CBS, ccCA-PS3, and Weizmann-1 theories) as these procedures are expected to yield chemical (or even sub-chemical) accuracy for atomization energies of species which are not characterized by multireference effects (chemical accuracy is arbitrarily defined as  $1 \text{ kcal mol}^{-1}$ ).<sup>[24,25]</sup> It is also worth mentioning that a large dataset of atomization energies (such as the W4-11 database) can be

[a] A. Karton

School of Molecular Sciences, The University of Western Australia, Perth, Western Australia 6009, Australia  
E-mail: amir.karton@uwa.edu.au

[b] N. Sylvetsky, J. M. L. Martin

Department of Organic Chemistry, Weizmann Institute of Science, Rehovot 76100, Israel  
E-mail: gershom@weizmann.ac.il

Contract grant sponsors: Australian Research Council (ARC) Discovery Early Career Researcher Award (DECRA, Project No. DE140100311; to A.K.), National Computational Infrastructure (NCI) National Facility grand (to A.K.), Feinberg Graduate School Fellowship (to N.S.); Contract grant sponsor: Israel Science Foundation (Research at Weizmann); Contract grant number: 1358/15; Contract grant sponsor: Minerva Foundation, Lise Meitner-Minerva Center for Computational Quantum Chemistry.

© 2017 Wiley Periodicals, Inc.

used to generate a database of hundreds<sup>[1]</sup> or even thousands<sup>[26]</sup> of chemical reactions. These extensive databases of chemical reactions can be used for testing and parameterization of DFT methods.<sup>[1,26–30]</sup>

The present work generates a dataset of 200 total atomization energies (TAEs) with  $3\sigma$  confidence intervals (CIs) of  $1 \text{ kJ mol}^{-1}$ . The W4-17 database is an extension of the smaller W4-11 dataset.<sup>[1]</sup> This extended dataset includes first/second-row molecules and radicals with up to eight non-hydrogen atoms. The species in the W4-17 database cover a broad spectrum of bonding situations and multireference character, and as such they constitute an excellent benchmark set for the parameterization and validation of highly accurate *ab initio* methods (e.g., CCSD(T) composite procedures) and double-hybrid density functional theory (DHDFT) methods. Most of the TAEs in the W4-17 database (167) were obtained at the CCSDTQ5/CBS level of theory (from W4 and W4.2 theories) or the CCSDTQ56/CBS level of theory (from W4.3 and W4.4 theories).<sup>[21–23]</sup> The TAEs for 33 larger systems (e.g.,  $\text{C}_2\text{F}_6$ ,  $\text{C}_2\text{Cl}_6$ ,  $\text{SF}_6$ , and  $\text{C}_6\text{H}_6$ ) were obtained at the CCSDT(Q)/CBS level of theory by means of the W4lite thermochemical protocol.<sup>[21,22]</sup>

In the present work, we will use the W4-17 database to evaluate the performance of a wide range of composite methods that attempt to approximate the coupled-cluster energy with singles, doubles, and quasiperturbative triple excitations (CCSD(T)) at various degrees of basis-set completeness.<sup>[1,24,25,31–38]</sup> The composite procedures that attempt to approximate the CCSD(T) energy close to the complete basis set (CBS) limit include the W1, W1-F12, W1X-1, W1X-2, W2, W2-F12, W2X, and cCA-PS3 methods. The composite procedures that attempt to approximate the CCSD(T) energy further away from the CBS limit include the G4, G3, G3B3, G4(MP2), G4(MP2)-6X, ROG4(MP2)-6X, G3(MP2), G3(MP2)B3, CBS-QB3, and CBS-APNO methods. We will also assess the performance of a range of DHDFT procedures (e.g., B2-PLYP, B2GP-PLYP, B2K-PLYP, DSD-BLYP, DSD-PBEP86, PWPB95,  $\omega$ B97X-2(LP), and  $\omega$ B97X-2(TQZ)).

## Computational Details

All the atomization energies in the W4-17 database have been obtained by means of the W4 family of thermochemical protocols. The computational details of these procedures have been specified and rationalized in great detail in Refs. [21] (W4lite, W4, W4.2, and W4.3), [22] (W4-F12), and [23] (W4.4) (see also Ref. [24] for a recent review). For the purpose of making the article self-contained, we will briefly summarize the steps involved in W4 theory:

- The geometries are optimized at the CCSD(T)/cc-pV(Q+d)Z level of theory. This level of theory has been shown to yield geometries that are in close agreement with CCSD(T)/cc-pV(6+d)Z geometries.<sup>[39]</sup>
- The Hartree–Fock (HF) energies are extrapolated from the aug'-cc-pV(n+d)Z basis sets ( $n = 5, 6$ ) using the Karton–Martin basis set extrapolation.<sup>[40]</sup> The aug' notation indicates the combination of a non-augmented basis set on hydrogen and an augmented basis set on all other elements.<sup>[41–43]</sup>

- The valence CCSD energy is extrapolated from the aug'-cc-pV(n+d)Z basis sets ( $n = 5, 6$ ) where the singlet- and triplet-coupled pair energies are extrapolated separately to the infinite basis-set limit.
- The quasiperturbative triple excitations, (T), are extrapolated from the aug'-cc-pV(n+d)Z basis sets ( $n = Q, 5$ ).
- The higher-order connected triples,  $T_3$ –(T), correlation contribution is extrapolated from the cc-pVnZ basis sets ( $n = D, T$ ).
- The quasiperturbative quadruple excitations, (Q), are calculated with the cc-pVTZ basis set.
- The higher-order connected quadruples,  $T_4$ –(Q), correlation contribution is calculated with the cc-pVDZ basis set.
- The connected quintuple,  $T_5$ , correlation contribution is calculated with the *sp* part of the cc-pVDZ basis set.
- In all the above correlation calculations, the inner-shell orbitals (1s for first-row atoms, and 1s, 2s, and 2p for second-row atoms) are constrained to be doubly occupied in all configurations. The core–valence (CV) correction is extrapolated from the aug'-cc-pwCVnZ basis sets ( $n = T, Q$ ) at the CCSD(T) level.<sup>[44]</sup> We note that the 1s orbitals of second-row atoms remain frozen in the all-electron aug'-cc-pwCVnZ calculations.
- The scalar relativistic contribution (in the second-order Douglas–Kroll–Hess approximation)<sup>[45,46]</sup> is obtained from the difference between nonrelativistic CCSD(T)/aug'-cc-pV(Q+d)Z and relativistic CCSD(T)/aug'-cc-pV(Q+d)Z-DK calculations.<sup>[47]</sup>
- Atomic spin-orbit coupling corrections are obtained as the weighted average of the experimental fine structure levels of the atomic ground states, according to the formula

$$\Delta_{\text{SO}} = \frac{\sum_{|L-S|}^{L+S} (2J+1)\epsilon_J}{\sum_{|L-S|}^{L+S} 2J+1}$$

in which  $L$  and  $S$  stand for the angular and spin term quantum numbers,  $J$  for the  $L$ – $S$  coupling quantum number, and  $\epsilon_J$  for the fine structure level. The required spectroscopic data can be taken from the NIST database.<sup>[48]</sup> For the few molecular species in spatially degenerate ground states, the analogous approach is taken.

- The diagonal Born–Oppenheimer correction (DBOC) is calculated at the HF/aug-cc-pVTZ level of theory.

The main differences between W4 theory and the higher order W4.x theories can be briefly described as follows: in W4.2 and W4.3, the CV correction is obtained at the CCSDT level,<sup>[21]</sup> and in W4.4 it is obtained at the CCSDT(Q) level.<sup>[23]</sup> W4.3 and W4.4 theories also use larger basis sets than W4 to calculate the valence post-CCSD(T) contributions, and they include contributions of connected sextuple excitations from CCSDTQ5(6)/DZ or CCSDTQ56/DZ calculations.<sup>[21,23]</sup> Due to these computationally demanding higher-order calculations, W4.x theories are applicable to very small systems with up to 2–3 non-hydrogen atoms.

**Table 1.** Overview of the W4 thermochemical protocols used for obtaining the 200 total atomization energies in the W4-17 database.

Level of theory	Protocol <sup>[a]</sup>	Number of species	Largest species <sup>[b]</sup>
CCSDTQ56/CBS	W4.3 or W4.4	47	SiH <sub>4</sub> , CH <sub>4</sub> , C <sub>2</sub> H <sub>2</sub> , HOF
CCSDTQ5/CBS	W4 or W4.2	120	CCl <sub>4</sub> , SiF <sub>4</sub> , CF <sub>2</sub> Cl <sub>2</sub> , CF <sub>4</sub>
CCSDT(Q)/CBS	W4lite	33	SF <sub>6</sub> , N <sub>2</sub> O <sub>4</sub> , C <sub>6</sub> H <sub>6</sub> , C <sub>2</sub> Cl <sub>6</sub> , C <sub>2</sub> F <sub>2</sub> Cl <sub>2</sub>

[a] The W<sub>n</sub> thermochemical protocols are reviewed in Ref. [24] and are described in detail in Refs. [23] (W4.4) and [21] (W4.3, W4.2, W4, and W4lite).  
[b] These are only representative examples, the W4-17 database includes many other systems of similar sizes (see Table S3 of the Supporting Information for the full list of molecules).

In the computationally more economical W4lite theory, the post-CCSDT(Q) contributions are completely neglected and the (Q) contribution is calculated with the cc-pVDZ basis set. This makes W4lite applicable to larger with up to eight non-hydrogen atoms.

The performance of W4 theories was thoroughly validated against highly accurate experimental atomization energies from the Active Thermochemical Tables (ATcT) thermochemical network.<sup>[49–52]</sup> Specifically, the performance of W4 theory was evaluated against 35 TAEs from ATcT (associated with error bars  $\leq 0.05$  kcal mol<sup>-1</sup>) including both first- and second-row species.<sup>[1,24]\*</sup> Against these highly accurate experimental values, W4 theory attains an RMSD of 0.085 kcal mol<sup>-1</sup>, implying a 2 $\sigma$  CI of 0.170 kcal mol<sup>-1</sup>, and a 3 $\sigma$  CI of 0.255 kcal mol<sup>-1</sup>. The largest deviations (theory–experiment) are obtained for ozone (–0.23) and nitrous acid (+0.20 kcal mol<sup>-1</sup>). For 29 of these systems, TAEs were obtained with the higher W4.x theories (x = 2–4). The RMSD for the higher W4.x theories is reduced to 0.060 kcal mol<sup>-1</sup>, implying a 95% CI of merely 0.120 kcal mol<sup>-1</sup>. In addition, the largest deviations are significantly reduced; only four deviations exceed 0.1 kcal mol<sup>-1</sup>, namely: ozone (–0.14), hydrogen peroxide (–0.12), dichlorine (–0.11), and ethane (+0.11 kcal mol<sup>-1</sup>).

Table 1 gives an overview of the W4 protocols that have been used for obtaining the reference values in the W4-17 database. Most of the atomization energies in the database have been obtained at the CCSDTQ5/CBS level of theory by means of the W4 or W4.2 procedure. For a subset of 47 species with up to two non-hydrogen atoms, we were able to obtain the TAEs at the CCSDTQ56/CBS level of theory by means of the W4.3 or W4.4 procedure. We note, however, that for nearly all of these systems the connected sextuple contribution is below 0.01 kcal mol<sup>-1</sup>. The main exceptions are the pathologically multireference singlet states of BN and C<sub>2</sub> for which the T<sub>6</sub> contribution amounts to 0.04 and 0.07 kcal mol<sup>-1</sup>, respectively. The geometries and reference TAEs for the molecules in the W4-17 database are given in Table S1 of

\*The ATcT benchmark set included the following 35 molecules: CH, CH<sub>2</sub>, CH<sub>3</sub>, CH<sub>4</sub>, C<sub>2</sub>H<sub>2</sub>, C<sub>2</sub>H<sub>4</sub>, C<sub>2</sub>H<sub>6</sub>, H<sub>2</sub>CO, H<sub>2</sub>C=C=O, CH<sub>3</sub>OH, NH, NH<sub>2</sub>, NH<sub>3</sub>, OH, H<sub>2</sub>O, HO<sub>2</sub>, H<sub>2</sub>O<sub>2</sub>, HCN, HNO, trans-HONO, CO, CO<sub>2</sub>, NO, NO<sub>2</sub>, N<sub>2</sub>O, O<sub>2</sub>, O<sub>3</sub>, SO, SO<sub>2</sub>, H<sub>2</sub>, HF, HCl, N<sub>2</sub>, F<sub>2</sub>, and Cl<sub>2</sub>.

the Supporting Information, while the component breakdown of the 200 TAEs is given in Supporting Information Table S2.

In the present work we will evaluate the performance of a large variety of composite *ab initio* methods that approximate the CCSD(T) energy. In particular, we will consider composite methods that (i) attempt to approximate the CCSD(T) energy in conjunction with a triple- $\zeta$ -quality basis set, such as the popular Gaussian-*n*<sup>[31,53–56]</sup> and CBS<sup>[57,58]</sup> family of methods, and (ii) attempt to approximate the CCSD(T) energy closer to the infinite basis-set limit, such as the Weizman-*n* theories (W1, W2, W1-F12, W2-F12),<sup>[38,59,60]</sup> the modified W<sub>n</sub> methods (W1X-1, W1X-2, and W2X),<sup>[61–63]</sup> and the correlation-consistent Composite Approach (ccCA) methods.<sup>[32,64–67]</sup> The reference values are all-electron, relativistic, DBOC-inclusive TAEs at the bottom of the well, that is, excluding the zero-point vibrational energy (ZPVE). The reference values used for evaluating the CCSD(T) composite procedures and DHDFT methods are given in Table S1 of the Supporting Information. We note that Table S1 of the Supporting Information also lists the all-electron, relativistic, DBOC-inclusive TAEs at 0 K (TAE<sub>0</sub>). The ZPVEs for the 140 systems in the W4-11 database are taken from experiment or high-level *ab initio* anharmonic force field calculations and are deemed to be sufficiently accurate to be included in the W4-17 database (for further details see Ref. [1]). Most of the remaining systems in the W4-17 database are too large for obtaining anharmonic ZPVEs from high-level *ab initio* calculations. The ZPVEs for these systems are obtained from scaling harmonic DSD-PBEP86-D3BJ/aug'-cc-pV(Q+d)Z frequencies by 0.9831 as recommended in Ref. [68]. It has been shown that this approach can yield anharmonic ZPVEs with accuracies approaching those obtained from quartic force fields. Nevertheless, for the purpose of benchmarking composite *ab initio* methods we prefer to use our ZPVE-exclusive reference values due to the potential errors in our scaled harmonic ZPVEs for some of the larger systems in the W4-17 database.

We will also evaluate the performance of a variety DHDFT procedures against all-electron, non-relativistic, DBOC-exclusive TAEs at the bottom of the well from the W4-17 database. We will consider the following DHDFT procedures: B2-PLYP,<sup>[69]</sup> B2GP-PLYP,<sup>[70]</sup> B2K-PLYP,<sup>[71]</sup> B2T-PLYP,<sup>[71]</sup> DSD-BLYP,<sup>[72]</sup> DSD-PBEP86,<sup>[73,74]</sup> PWPB95,<sup>[75]</sup>  $\omega$ B97X-2(LP), and  $\omega$ B97X-2(TQZ).<sup>[76]</sup> Unless otherwise indicated, the DHDFT calculations have been carried out in conjunction with the aug'-cc-pV(5+d)Z basis set with the Gaussian09 and Q-Chem program suites.<sup>[77,78]</sup> Empirical D3 dispersion corrections are included using the Becke–Johnson damping potential (denoted by the suffix -D3).<sup>[79–83]</sup> The empirical D3 corrections were calculated using the DFT-D3 program by Grimme and co-workers<sup>[80,81]</sup>; the  $s_6$ ,  $s_8$ , and  $a_2$  parameters needed for calculating the D3 correction for B2K-PLYP were taken from the Supporting Information to Ref. [84]. We also consider the performance of the MP2, SCS-MP2,<sup>[85]</sup> MP2.5,<sup>[86]</sup> MP3, and SCS-MP3 methods.<sup>[87]</sup> All the HF, CCSD, and CCSD(T) calculations involved in the W4 protocol have been performed with the Molpro 2012.1 program suite.<sup>[88,89]</sup> All the post-CCSD(T) calculations were carried out with the MRCC program suite.<sup>[90,91]</sup>

Table 2. Overview of the species in the W4-17 database.

		Number of species
Number of Non-hydrogen atoms	≤2	85
	3	57
	4	31
	5	16
	6	8
	7	1
	8	2
Containing only first-row atoms		129
Containing only second-row atoms		19
Mixed first- and second-row atoms		52
Closed-shell species <sup>[a]</sup>		160
Open-shell species <sup>[b]</sup>		40

[a] Nine of the closed-shell systems are singlet biradicals. [b] Of the open-shell systems, 33 are radicals and 7 are triplet systems.

## Results and Discussion

### Overview of the species in the W4-17 database

Table 2 gives an overview of the types of species in the W4-17 database, while the complete list of molecules in the database is given in Table S3 of the Supporting Information. In terms of the molecular size distribution the database includes:

- 3 systems with 7–8 non-hydrogen atoms ( $C_2F_6$ ,  $C_2Cl_6$ ,  $SF_6$ )
- 8 systems with 6 non-hydrogen atoms (e.g.,  $C_6H_6$ ,  $N_2O_4$ ,  $PF_5$ ,  $C_2Cl_4$ ,  $ClF_5$ )
- 16 systems with 5 non-hydrogen atoms (e.g., pentane, thiophene,  $\beta$ -lactim,  $HClO_4$ ,  $CCl_4$ )
- 31 systems with 4 non-hydrogen atoms (e.g., butane, tetrahedrane, acetic acid, chloroform)
- 57 systems with 3 non-hydrogen atoms (e.g., propane, acetaldehyde, formic acid, ethanol)
- 85 systems with less than 3 non-hydrogen atoms (e.g.,  $CH_4$ ,  $CH_3OH$ ,  $NH_2OH$ ,  $N_2H_4$ ,  $CClH_3$ )

The resulting database includes a total of 200 neutral species. From one perspective, they consist of 160 closed shell, 33 radicals, and 7 triplet systems (we note that nine of the closed shell species are singlet carbenes); from another, they comprise 129 purely first-row systems, 52 mixed first/second-row systems, and 19 purely second-row systems.

The W4-17 database covers a broad range of organic and inorganic systems. Overall, it includes 109 organic systems (broadly defined here as carbon containing species) and 91 inorganic systems (i.e., with no carbon atoms). Table 3 gives an overview of the types of organic and inorganic species in the database. The organic systems include alkanes, alkenes, alkynes, haloalkanes, haloalkenes, haloalkynes, arenes, aromatic heterocycles, nonaromatic heterocycles, alcohols, aldehydes, ketones, anhydrides, carboxylic acids, amines, imines, and nitriles. The inorganic species include halogenated species, boranes, oxides, acids, hydrides, and pure atomic clusters. The organic and inorganic species include noncyclic, cyclic, and cage systems with single and multiple bonds that involve

varying degrees of covalent and ionic characters. This set evidently spans the gamut from systems dominated by a single reference configuration (e.g.,  $CH_4$ ,  $CH_3OH$ ,  $CH_3NH_2$ ), to systems which exhibit appreciable nondynamical correlation effects (e.g.,  $O_2$ ,  $SO_3$ ,  $N_2O_4$ ), and also to systems exhibiting pathological nondynamical correlation effects (e.g.,  $C_2$ ,  $O_3$ ,  $F_2O_2$ ).

### Multireference considerations

A number of diagnostics, which are based on the CCSD  $T_1$  amplitudes, have been proposed for gauging the degree of multireference character in a system, for example, the  $T_1$  and  $D_1$  diagnostics<sup>[92,93]</sup> (see also discussion in Refs. [21] and [94]). However, it has been found that these diagnostics are not necessarily good predictors for the magnitude of post-CCSD(T) contributions to the TAEs.<sup>[1,21,95–99]</sup> Conversely, it has been found that an energy-based diagnostic of the form:

$$\%TAE[(T)] = 100 \times \frac{TAE[CCSD(T)] - TAE[CCSD]}{TAE[CCSD(T)]}$$

provides a reliable *a priori* diagnostic for the importance of post-CCSD(T) contributions to the TAEs (where  $TAE[CCSD]$  and  $TAE[CCSD(T)]$  are the TAEs calculated at the CCSD and CCSD(T) levels, respectively).<sup>[1,21]</sup> In particular, it has been shown that  $\%TAE[(T)] < 5\%$  indicates that post-CCSD(T) contributions to the TAEs should not exceed  $0.5 \text{ kcal mol}^{-1}$ ,  $\%TAE[(T)] < 10\%$  indicates that post-CCSD(T) contributions should generally not exceed  $1.0 \text{ kcal mol}^{-1}$ , and  $\%TAE[(T)] > 10\%$  indicates that post-CCSD(T) can exceed  $1.0 \text{ kcal mol}^{-1}$  by significant amounts.<sup>[24]</sup>

For the purpose of evaluating the performance of CCSD(T) composite procedures and DHFDT methods, it is useful to divide the W4-17 database into two subsets: W4-17-nonMR and W4-17-MR. The W4-17-nonMR subset includes 183 systems for which  $\%TAE[(T)] < 10\%$ . This subset includes systems that are dominated by a single-reference configuration (e.g., pentane and ethanol), mild nondynamical correlation effects (e.g., HNO and  $C_2F_2$ ), or moderate nondynamical correlation effects (e.g., NCCN,  $HO_3\cdot$  and  $N_2O_4$ ).<sup>†</sup> The W4-17-MR subset includes 17 systems for which  $\%TAE[(T)] > 10\%$  and includes systems that are dominated by severe multireference effects (e.g.,  $O_3$ ,  $F_2O_2$ ,  $ClO_3$ , and  $ClF_5$ ). The  $\%TAE[(T)]$  diagnostics for all the molecules in the W4-17 database are given in Supporting Information Table S4.

### Performance of CCSD(T) composite methods for systems dominated by mild to moderate multireference effects

As CCSD(T) composite methods are not recommended for systems with severe multireference character, we will evaluate the performance of these procedures for the 183 atomization reactions in the W4-17-nonMR database. The error statistics for a wide range of CCSD(T) composite procedures are summarized in Table 4.

<sup>†</sup>Note that the W4-17-nonMR subset includes  $F_2$  which has an artificially high  $\%TAE[(T)]$  value (see Refs. [1] and [24] for further details).



Table 3. Overview of the 200 organic and inorganic species in the W4-17 database.

	Number of species	Type	Typical examples
109 organic species	27	Halocarbons	$C_2X_6$ , $C_2X_4$ , $C_2X_2$ , $CX_4$ ( $X = F, Cl$ )
	19	Noncyclic hydrocarbons	pentane, butane, butadiene, propane
	10	Nonaromatic heterocycles	borole, silole, $\beta$ -lactim, 1,3-dioxetan-2-one
	8	Cyclic hydrocarbons	cyclopentadiene, cyclobutane, tetrahydrane
	7	Carbonyls	phosgene, ketene, carbonyl fluoride
	7	Nitriles	cyanogen, cyanomethyl radical, cyanic acid
	6	Aldehydes	formic anhydride, acetaldehyde, glyoxal
	6	Imines	methanimine, isocyanic acid, isofulminic acid
	5	Aromatic	benzene, furan, furazan, pyrrole, thiophene
	4	Alcohols	methanol, ethanol, hydroxymethylene
	4	Amines	methylamine, formamide, aminomethyl radical
	4	Other	methylphosphine, carbon disulfide, dicarbon
	2	Acids	acetic acid, formic acid
	91 inorganic species	28	Diatomics
10		Fluorides	$SF_6$ , $PF_5$ , $ClF_5$ , $SiF_4$ , $BF_3$ , $AlF_3$ , $PF_3$
10		Halogen oxides	$ClO_3$ , $Cl_2O_2$ , $ClO_2$ , $F_2O_2$ , $FO_2$ , $F_2O$
9		Hydrides	$B_2H_6$ , $BH_3$ , $Si_2H_6$ , $SiH_4$ , $AlH_3$ , $PH_3$
8		Acids	$HClO_4$ , $HClO_3$ , $HClO_2$ , $HN_3$ , $HNO_2$
7		Amines	$NH_2OH$ , $NH_2F$ , $NH_2Cl$ , $N_2H_4$ , $NH_3$
7		Oxides	$N_2O_4$ , $NO_2$ , $N_2O$ , $SO_2$ , $SO_3$ , $S_2O$
4		Peroxides	$H_2O_2$ , $HO_2$ , $HO_3$
4		Pure atomic clusters	$S_4$ , $P_4$ , $S_3$ , $O_3$
4		Other	$H_2N_2$ , $AlCl_3$

**Gn, CBS, and ccCA methods.** Let us begin with the performance of the Gn protocols, which represent a good compromise between accuracy and computational cost and can be routinely applied to large systems. For example, the computationally economical G4(MP2) method<sup>[54]</sup> has been recently applied for the calculation of the heat of formation of  $C_{60}$ .<sup>[100]</sup> The G4 procedure shows good overall performance with an RMSD of  $0.95 \text{ kcal mol}^{-1}$  and a MAD of  $0.68 \text{ kcal mol}^{-1}$ . However, it should be noted that for 36 systems, errors larger than  $1 \text{ kcal mol}^{-1}$  (in absolute value) are obtained. Particularly large errors are obtained for systems with highly polar bonds involving O, F, and Cl. In particular, large underestimations ranging between  $1.71$  and  $2.52 \text{ kcal mol}^{-1}$  are obtained for  $HClO_4$ ,  $HClO_3$ ,  $SO_3$ ,  $SO_2$ ,  $S_2O$ ,  $SF_6$ ,  $PF_5$ , and  $BF_3$ . Conversely, the atomization energies of perchlorinated compounds (e.g.,  $CCl_4$ ,  $C_2Cl_2$ , and  $C_2Cl_4$ ) are overestimated by amounts ranging from  $1.69$  to  $2.48 \text{ kcal mol}^{-1}$ . Hexachloroethane ( $C_2Cl_6$ ), for which the TAE is overestimated by as much as  $5.73 \text{ kcal mol}^{-1}$ , represents an extreme case. We therefore, do not recommend the application of the G4 procedure for sulfur oxides and highly fluorinated or chlorinated systems. We note that on removing the above twelve systems from the training set the RMSD for the G4 procedure is reduced from  $0.95$  to  $0.71 \text{ kcal mol}^{-1}$ .

As expected, the computationally more economical G4(MP2) procedure results in a larger RMSD of  $1.29 \text{ kcal mol}^{-1}$  for the entire W4-17-nonMR database (and a MAD of  $0.85 \text{ kcal mol}^{-1}$ ). In addition, deviations larger than  $1 \text{ kcal mol}^{-1}$  are obtained for 48 of the 183 systems. For example, large underestimations (given in parenthesis) are obtained for (i) boron containing molecules: borole ( $1.13$ ),  $BH_3$  ( $2.03$ ),  $BHF_2$  ( $2.40$ ),  $BF_3$  ( $2.63$ ), and  $B_2H_6$  ( $4.46$ ); (ii) chlorine oxides:  $HClO_3$  ( $2.03$ ) and  $HClO_4$  ( $2.95$ ); and (iii) *n*-alkanes: propane ( $1.04$ ), butane ( $1.19$ ), and pentane ( $1.21 \text{ kcal mol}^{-1}$ ). Apart from that, large overestimations are

obtained for the perchlorinated compounds, namely:  $CHCl_3$  ( $1.89$ ),  $CCl_2$  ( $2.45$ ),  $C_2Cl_2$  ( $2.93$ ),  $CCl_4$  ( $3.44$ ),  $C_2Cl_4$  ( $4.91$ ), and  $C_2Cl_6$  ( $7.53 \text{ kcal mol}^{-1}$ ). Table S5 of the Supporting Information lists all the deviations larger than  $2 \text{ kcal mol}^{-1}$  (in absolute value) in the W4-17-nonMR database for the CCSD(T) composite methods.

We note that evaluating the performance of the G4-type procedures against  $TAE_0$ , rather than  $TAE_{er}$ , values has a very small effect on the overall error statistics. For example, the RMSD over the 183 atomization energies in the W4-17-nonMR dataset changes from  $0.95$  to  $0.93 \text{ kcal mol}^{-1}$  for G4 theory and from  $1.29$  to  $1.30 \text{ kcal mol}^{-1}$  for G4(MP2) theory.

The G4(MP2)-6X and ROG4(MP2)-6X procedures have the same computational cost as G4(MP2) but use additional empirical parameters. These procedures show deteriorated performance for the W4-17-nonMR dataset relative to G4(MP2), with an RMSD of  $1.65 \text{ kcal mol}^{-1}$ . In addition, for over 75 systems, deviations larger than  $1 \text{ kcal mol}^{-1}$  are obtained. The G4(MP2)-6X and ROG4(MP2)-6X procedures show similar performance to G4(MP2) for the boron containing molecules. However, the performance for the Cl-containing molecules is further deteriorated; namely, deviations between  $3$ – $10 \text{ kcal mol}^{-1}$  are obtained for systems such as  $CCl_2$ ,  $CCl_2O$ ,  $CF_2Cl_2$ ,  $CCl_3H$ ,  $CCl_4$ ,  $C_2Cl_2$ ,  $C_2F_2Cl_2$ ,  $C_2Cl_4$ , and  $C_2Cl_6$ .

In contrast to the G4 and G4(MP2) procedures, which show reasonably good performance, the G3 and G3(MP2) procedures result in poor performance with RMSDs of  $1.88$  and  $2.22 \text{ kcal mol}^{-1}$ , respectively. Similar observations can be made for the G3B3 and G3(MP2)B3 procedures. These results suggest that the G4-type procedures should be favored over the earlier the G3-type procedures for the calculation of TAEs and heats of formation.

The CBS-type procedures perform rather poorly; in particular, we obtain the following RMSDs for the W4-17-nonMR

**Table 4.** Performance of a representative set of CCSD(T)-based composite thermochemical procedures for the 183 atomization reactions involving first- and second-row elements in the W4-17-nonMR dataset (error statistics are given in kcal mol<sup>-1</sup>).<sup>[a]</sup>

Method	RMSD	MAD	MSD	# of LDs	LND	LPD
G4	0.95	0.68	-0.06	36	-2.52 (HClO <sub>4</sub> )	5.73 (C <sub>2</sub> Cl <sub>6</sub> )
G4(MP2)	1.29	0.85	0.16	48	-4.46 (B <sub>2</sub> H <sub>6</sub> )	7.53 (C <sub>2</sub> Cl <sub>6</sub> )
G4(MP2)-6X	1.65	1.11	0.90	77	-3.52 (B <sub>2</sub> H <sub>6</sub> )	10.24 (C <sub>2</sub> Cl <sub>6</sub> )
ROG4(MP2)-6X	1.65	1.14	0.90	78	-3.19 (B <sub>2</sub> H <sub>6</sub> )	9.86 (C <sub>2</sub> Cl <sub>6</sub> )
G3 <sup>[b]</sup>	1.88	1.32	-0.68	86	-7.52 (SF <sub>6</sub> )	8.65 (C <sub>2</sub> Cl <sub>6</sub> )
G3B3	1.94	1.22	-0.72	76	-10.85 (SF <sub>6</sub> )	5.39 (C <sub>2</sub> Cl <sub>6</sub> )
G3(MP2) <sup>[b]</sup>	2.22	1.47	-0.71	97	-10.56 (SF <sub>6</sub> )	8.93 (C <sub>2</sub> Cl <sub>6</sub> )
G3(MP2)B3	2.22	1.33	-0.48	84	-12.94 (HClO <sub>4</sub> )	7.45 (C <sub>2</sub> Cl <sub>6</sub> )
CBS-4M	4.23	2.89	0.33	130	-16.30 (SF <sub>6</sub> )	22.74 (C <sub>2</sub> Cl <sub>6</sub> )
CBS-QB3	2.35	1.53	0.74	101	-4.41 (SF <sub>6</sub> )	15.88 (C <sub>2</sub> Cl <sub>6</sub> )
ROCBS-QB3	2.04	1.34	0.32	88	-6.10 (SF <sub>6</sub> )	13.52 (C <sub>2</sub> Cl <sub>6</sub> )
ccCA-PS3	0.92	0.63	0.25	36	-1.59 (BN)	5.81 (C <sub>2</sub> Cl <sub>6</sub> )
W1	0.74	0.57	0.23	30	-1.95 (HClO <sub>4</sub> )	2.25 (C <sub>2</sub> F <sub>6</sub> )
W1-F12	0.72	0.51	-0.45	21	-2.60 (S <sub>2</sub> O)	0.81 (C <sub>6</sub> H <sub>6</sub> )
W1-F12 <sup>[c]</sup>	0.75	0.51	-0.46	22	-2.87 (P <sub>4</sub> )	0.88 (C <sub>6</sub> H <sub>6</sub> )
W1X-1	0.67	0.45	-0.19	22	-2.30 (SO <sub>3</sub> )	3.49 (C <sub>2</sub> Cl <sub>6</sub> )
W1X-2	0.78	0.54	-0.15	32	-2.53 (HClO <sub>4</sub> )	3.23 (C <sub>2</sub> Cl <sub>6</sub> )
W2	0.57	0.37	-0.12	10	-2.52 (HO <sub>3</sub> )	2.71 (C <sub>2</sub> Cl <sub>6</sub> )
W2-F12	0.55	0.38	-0.07	9	-2.07 (HO <sub>3</sub> )	3.30 (C <sub>2</sub> Cl <sub>6</sub> )
W2X	0.63	0.45	-0.01	18	-2.23 (HO <sub>3</sub> )	3.25 (C <sub>2</sub> Cl <sub>6</sub> )

[a] RMSD = root-mean-square deviation, MAD = mean-absolute deviation, MSD = mean-signed deviation, # of LDs = number of deviations exceeding 1 kcal mol<sup>-1</sup> (see also Table S5 of the Supporting Information), LND = largest negative deviation, LPD = largest positive deviation. [b] cis-HO<sub>3</sub> and trans-HO<sub>3</sub> were removed due to problems with the geometry optimizations. [c] A new version of W1-F12 theory developed in the present work, in which the CCSD component of the CV correction is calculated using explicitly correlated F12 techniques (see text).

database: 4.23 (CBS-4M), 2.35 (CBS-QB3), and 2.04 (ROCBS-QB3) kcal mol<sup>-1</sup>. The poor performance of the CBS methods for atomization energies is also demonstrated by the large numbers of errors exceeding 1 kcal mol<sup>-1</sup>, specifically: 130 (CBS-4M), 101 (CBS-QB3), and 88 (ROCBS-QB3) kcal mol<sup>-1</sup>. It is noteworthy that the popular CBS-QB3 method severely overestimates the TAEs of highly chlorinated/fluorinated species. For example, deviations of 5–16 kcal mol<sup>-1</sup> are obtained for AlCl<sub>3</sub>, CF<sub>2</sub>Cl<sub>2</sub>, CHCl<sub>3</sub>, CCl<sub>4</sub>, C<sub>2</sub>F<sub>6</sub>, C<sub>2</sub>F<sub>2</sub>Cl<sub>2</sub>, C<sub>2</sub>Cl<sub>4</sub>, and C<sub>2</sub>Cl<sub>6</sub> (see Table S5 of the Supporting Information).

The ccCA composite methods are computationally more expensive than the *Gn* and CBS methods but are devoid of empirical scaling factors. With the exception of the G4 procedure, ccCA-PS3 significantly outperforms the *Gn* and CBS methods with an overall RMSD of 0.92 kcal mol<sup>-1</sup> and a MAD of 0.63 kcal mol<sup>-1</sup>. Deviations larger than 1 kcal mol<sup>-1</sup> are still obtained for 36 species, where particularly large underestimations (of 3–6 kcal mol<sup>-1</sup>) are obtained for the second-row species: CCl<sub>4</sub>, C<sub>2</sub>Cl<sub>4</sub>, and C<sub>2</sub>Cl<sub>6</sub> (see Table S5 of the Supporting Information). These statistics are considerably better than those reported in Ref. [1]: in the present work, the core-valence contribution for second-row elements was evaluated using a purpose-designed core-valence basis set reported by the Wilson group in Ref. [101].

**Wn methods.** The *Gn*, CBS, and ccCA-PS3 methods discussed in the previous section are computationally cost-effective, as they use basis-set extrapolations and/or additivity schemes at the MP2 level. Let us now move to the performance of the computationally more demanding and more robust *Wn* methods, which use basis set extrapolation at the HF, CCSD, and CCSD(T) levels. We note that the W1X-*n* procedures are an

exception as they still calculate the core-valence and scalar relativistic corrections at the MP2 level.

The original W1 method outperforms all the considered *Gn*, CBS, and ccCA-PS3 methods and attains an overall RMSD of 0.74 kcal mol<sup>-1</sup> and a MAD of 0.57 kcal mol<sup>-1</sup>. Nevertheless, deviations larger than 1 kcal mol<sup>-1</sup> are obtained for as many as 30 systems. Nearly half of these systems are highly fluorinated and chlorinated species with deviations (all overestimations) ranging from 1.09 (CF<sub>2</sub>Cl<sub>2</sub>) to 2.25 (C<sub>2</sub>F<sub>6</sub>) kcal mol<sup>-1</sup>. Relatively large overestimations are also obtained for medium-sized organic compounds, such as: 1.14 (pyrrole), 1.25 (pentane), and 1.50 (benzene) kcal mol<sup>-1</sup>.

The W1-F12 method uses explicitly correlated techniques in the CCSD calculations and has a significantly reduced computational cost relative to the original W1 method in terms of CPU time, disk space, and RAM.<sup>[60]</sup> Thus, W1-F12 has been applied to large hydrocarbons (e.g., sumanene,<sup>[100]</sup> corannulene,<sup>[102]</sup> and even dodecahedrane<sup>[4]</sup>) as well as to systems of biological relevance (e.g., guanine<sup>[60]</sup> and arginine).<sup>[103]</sup> Overall, W1-F12 shows similar performance to W1 and results in an overall RMSD of 0.72 kcal mol<sup>-1</sup> and a MAD of 0.51 kcal mol<sup>-1</sup>. We note that in contrast to W1 theory, which tends to overestimate the atomization energies, W1-F12 tends to underestimate them as indicated from a mean-signed deviation (MSD) of -0.45 kcal mol<sup>-1</sup>. The number of deviations larger than 1 kcal mol<sup>-1</sup> is reduced from 30 for W1 theory to 21 for W1-F12 theory. Large underestimations ranging between 2.0–2.6 kcal mol<sup>-1</sup> are obtained for second-row systems such as: P<sub>2</sub>, P<sub>4</sub>, SO<sub>3</sub>, S<sub>2</sub>O, HClO<sub>4</sub>, and AlCl<sub>3</sub>. This is consistent with our previous observation that W1-F12 shows better performance for species containing only first-row elements.<sup>[60]</sup> For the subset of 120 first-row systems in the W4-17-nonMR

**Table 5.** Error statistics (in kcal mol<sup>-1</sup>) for the CV corrections in the W4-17-nonMR dataset.

Method	RMSD	MAD	MSD	LND	LPD
W1	0.24	0.18	-0.13	-0.78 (C <sub>6</sub> H <sub>6</sub> )	0.80 (SF <sub>6</sub> )
W1-F12 <sup>[a]</sup>	0.16	0.10	-0.07	-0.98 (AlF <sub>3</sub> )	0.41 (SF <sub>6</sub> )
W1-F12 <sup>[b]</sup>	0.10	0.08	-0.08	-0.38 (PF <sub>5</sub> )	0.23 (OCS)
W1X- <i>n</i> <sup>[c]</sup>	0.95	0.74	-0.72	-2.72 (C <sub>6</sub> H <sub>6</sub> )	0.79 (C <sub>2</sub> Cl <sub>6</sub> )
W2-F12	0.15	0.09	0.00	-0.45 (Si <sub>2</sub> H <sub>6</sub> )	0.82 (C <sub>2</sub> Cl <sub>6</sub> )

The reference CV corrections are taken from W4 (or higher) theory.<sup>[d]</sup> [a] From the original W1-F12 theory in which the CV correction is calculated using conventional CCSD(T) calculations (see Ref. 60). [b] From the new version of W1-F12 theory developed here, in which the CV correction is calculated using CCSD-F12 and CCSD(T) calculations (see text). [c] The same CV and scalar relativistic corrections are used in W1X-1 and W1X-2. RMSD = root-mean-square deviation, MAD = mean-absolute deviation, MSD = mean-signed deviation, # of LDs = number of deviations exceeding 1 kcal mol<sup>-1</sup>, LND = largest negative deviation, LPD = largest positive deviation.

database, W1-F12 attains an RMSD of 0.46 kcal mol<sup>-1</sup>, a MAD of 0.33 kcal mol<sup>-1</sup>, and the largest deviation of +0.81 kcal mol<sup>-1</sup> is obtained for benzene. For this subset, W1-F12 significantly outperforms the original W1 theory (*vide infra*).

The W1X procedure comes in two flavors (W1X-1 and W1X-2). These methods reduce the computational cost of W1-F12 by calculating the core-valence and scalar relativistic corrections at the MP2 level.<sup>[62]</sup> In the W1-F12 and W1X-1 procedures, the valence (T) correction is obtained from conventional CCSD(T) calculations in conjunction with the aug'-cc-pV(*n*+*d*)Z basis sets (*n* = D, T).<sup>[60,61]</sup> The W1X-2 method is a simplified version of W1X-1, in which the (T) correction is obtained in conjunction with the cc-pV*n*Z-F12 basis sets (*n* = D, T). This simplification reduces the number of single-point energy calculations, and leads to a further reduction in the computational cost. Interestingly, the W1X-1 method slightly outperforms the W1-F12 method, with an RMSD of 0.67 kcal mol<sup>-1</sup>, a MAD of 0.45 kcal mol<sup>-1</sup>, and 22 deviations larger than 1 kcal mol<sup>-1</sup>. As expected, W1X-2 results in deteriorated performance compared to W1X-1. In particular, it attains an RMSD of 0.78 kcal mol<sup>-1</sup>, a MAD of 0.54 kcal mol<sup>-1</sup>, and 32 deviations larger than 1 kcal mol<sup>-1</sup>.

Table 5 lists the error statistics for the CV correction used in W1, W1-F12, and W1X-*n* theories over the W4-17-nonMR dataset. The CV corrections from these theories tend to systematically underestimate the CV correction from W4 (or higher) theory. However, while W1 and W1-F12 result in MSDs of -0.13 and -0.07 kcal mol<sup>-1</sup>, respectively, the CV correction in W1X-*n* leads to a very large MSD of -0.77 kcal mol<sup>-1</sup> and an overall RMSD of 0.95 kcal mol<sup>-1</sup>. It should be noted that underestimations larger than 1 kcal mol<sup>-1</sup> are obtained for 52 species, and that particularly large deviations (larger than 2 kcal mol<sup>-1</sup>) are obtained for benzene, cyclopentadiene, pentane, borole, silole, AlF<sub>3</sub>, and SiF<sub>4</sub>. Overall, the W1X-*n* procedures benefit from a systematic error cancelations between the CV component (which tends to underestimate the CCSD(T)/CBS values) and the valence CCSD(T) component (which tends to overestimate the CCSD(T)/CBS values).

In an attempt to develop an improved core-valence correlation correction for W1-F12 theory, we experimented with core-valence basis sets developed especially for this purpose, cc-pCV*n*Z-F12 (*n* = D, T, Q).<sup>[104]</sup> As recommended there, we used different geminal exponents for valence and core-valence pairs: details of MOLPRO's implementation are given in Ref. [105]. For evaluation purposes, we evaluated the core-valence corrections for the entire W4-15 set<sup>[22]</sup> at both the CCSD(T)/aug-cc-pCV{5,6}Z and CCSD(T)/aug-cc-pwCV{Q,5}Z levels. The former uses basis sets with larger angular momentum, the latter basis sets specifically weighted for core-valence correlation. As can be seen in Table 6, the RMSD between both sets of inner-shell corrections is just 0.008 kcal mol<sup>-1</sup>, which means that either set should be suitable for calibration purposes. We have hence decided to list RMSDs with respect to both datasets below.

Concerning conventional approaches, it is clear that awCV{T,Q}Z, as used in W4 theory, works very well, with an RMSD of just 0.02 kcal mol<sup>-1</sup> (either reference). The complete individual contributions can be found in the Supporting Information to the present article.

One can achieve similar accuracies at the CCSD(Ts)-F12b/cc-pCV{T,Q}Z-F12 level, where Ts stands for uniform scaling of the (T) contribution with a scaling factor given in Ref. [106] (Ecorr[MP2-F12]/E[MP2] scaling<sup>[107]</sup> of the triples is not suitable for the present purpose because of its lack of size extensivity). However, this is a somewhat hollow victory, as these calculations are considerably more expensive than CCSD(T)/awCV{T,Q}Z. Clearly, a computational protocol in which the lion's share of CPU time would have to be spent on the relatively small core-valence contribution is poorly balanced.

The original W1-F12 and W2-F12 protocols have core-valence contributions with RMSD of 0.12 and 0.09 kcal mol<sup>-1</sup>, respectively. Can we improve on that at manageable computational cost? CCSD-F12b/cc-pCV{D,T}Z-F12, combined with a (Ts) from the larger basis set, achieves an RMSD of just 0.06 kcal mol<sup>-1</sup>. Still, the largest calculation required, namely CCSD(T)-F12b/cc-pCVTZ-F12, could represent an obstacle for larger systems. Using the smaller of the two basis sets for (Ts) instead increases RMSD to an unacceptable 0.15 kcal mol<sup>-1</sup>. If we, however, obtain (T) from a CCSD(T)/cc-pwCVTZ(no f) calculation as in the original W1-F12 paper,<sup>[60]</sup> and combine that with CCSD-F12b/cc-pCV{D,T}Z-F12, we find RMSD = 0.073 kcal mol<sup>-1</sup> relative to aCV{5,6}Z, and 0.077 kcal mol<sup>-1</sup> relative to the awCV{Q,5}Z reference data. We deem this to be an acceptable compromise between accuracy and computational cost, and hence select this as the revised core-valence level for W1-F12 theory.

The W2-type methods attempt to approximate the all-electron, relativistic CCSD(T) energy closer to the CBS limit compared to the W1-type methods. This, of course, comes with a significant increase in the computational cost. For example, whereas W1-type methods have been applied to systems with ~20 non-hydrogen atoms (*vide supra*), the largest systems W2-type methods have been applied to include up to ~10 non-hydrogen atoms (e.g., cubane<sup>[4]</sup> and adenine).<sup>[60]</sup> The original W2 method results in a respectable RMSD of 0.57 kcal mol<sup>-1</sup> and a MAD of 0.37 kcal mol<sup>-1</sup> for the W4-17-

**Table 6.** RMSDs (in kcal mol<sup>-1</sup>) relative to two different reference levels for the W4-15 dataset.

	Conventional CCSD(T)						
	ACV{T,Q}Z	ACV{Q,5}Z	awCVTZ	awCVQZ	awCV5Z	awCV{T,Q}Z	awCV{Q,5}Z
RMSD <sup>[a]</sup>	0.053	0.020	0.161	0.061	0.030	0.022	0.008
RMSD <sup>[b]</sup>	0.058	0.016	0.163	0.064	0.033	0.021	0.000
	CCSD(T)-F12b			CCSD(Ts)-F12b			
	CVDZ-F12	CVTZ-F12	CVQZ-F12	CVDZ-F12	CVTZ-F12	CVQZ-F12	CV{T,Q}Z-F12
RMSD <sup>[a]</sup>	0.275	0.091	0.034	0.267	0.081	0.030	0.023
RMSD <sup>[b]</sup>	0.287	0.097	0.039	0.276	0.087	0.035	0.028
	CCSD-F12b/CV{D,T}Z-F12			CCSD <sup>[d]</sup>			
	(Ts)/CVDZ-F12	(Ts)/CVTZ-F12	[c]	[c]	[c]		
			New W <sub>n</sub> -F12	Old W1-F12	Old W2-F12		
RMSD <sup>[a]</sup>	0.152	0.059	0.073	0.122	0.085		
RMSD <sup>[b]</sup>	0.154	0.062	0.077	0.123	0.091		

[a] Relative to CCSD(T)/aug-cc-pCV{5,6}Z reference data. [b] Relative to CCSD(T)/aug-cc-pwCV{Q,5}Z reference data. [c] Conventional CCSD(T)/pwCVTZ(no f). [d] CCSD contribution CCSD/pwCVTZ [W1-F12] or CCSD/awCVTZ [W2-F12] scaled by 1.1; (T) contribution see footnote c. See Table S6 of the Supporting Information for complete data.

nonMR database. We note that only for three challenging systems (*trans*-HO<sub>3</sub>, *cis*-HO<sub>3</sub>, and C<sub>2</sub>Cl<sub>6</sub>) deviations larger than 2 kcal mol<sup>-1</sup> are obtained. These deviations are dominated by large post-CCSD(T) contributions of -2.31, -1.85, and -1.71 kcal mol<sup>-1</sup>, respectively.

W2-F12 theory is the computationally economical, explicitly correlated cognate of W2, which uses explicitly correlated techniques in the valence CCSD(T) steps and smaller basis sets in the CCSD(T) CV calculations. For example, a W2 calculation for C<sub>2</sub>Cl<sub>6</sub> requires a total of 555 GB of scratch disk and 667 CPU-hours to run to completion, while a W2-F12 calculation requires only 347 GB of scratch disk and 102 CPU-hours. W2-F12 produces results in practically the same error statistics as W2 (Table 4).

The W2X procedure further reduces the computational cost of W2-F12 by (i) using the aug-cc-pV<sub>n</sub>Z basis sets in the valence CCSD(T) extrapolations, rather than the larger cc-pV<sub>n</sub>Z-F12 basis sets (which were specifically optimized for explicitly correlated calculations), and (ii) using a cost-effective MP2-based additivity scheme in the CV extrapolations. Overall, W2X shows a slight deterioration in performance compared to W2 and W2-F12. For example, W2X attains an RMSD of 0.63 kcal mol<sup>-1</sup> for the W4-17-nonMR dataset, with 18 deviations larger than 1 kcal mol<sup>-1</sup> (compared with 9–10 such deviations for W2 and W2-F12). In particular, large deviations, which range between 1–3 kcal mol<sup>-1</sup>, are obtained for halogenated systems such as: PF<sub>5</sub>, SF<sub>6</sub>, CF<sub>4</sub>, CF<sub>2</sub>Cl<sub>2</sub>, CCl<sub>4</sub>, C<sub>2</sub>F<sub>2</sub>Cl<sub>2</sub>, C<sub>2</sub>Cl<sub>4</sub>, C<sub>2</sub>F<sub>6</sub>, C<sub>2</sub>Cl<sub>6</sub>.

**Table 7.** Performance of a representative set of CCSD(T) composite thermochemical procedures for the 120 atomization reactions involving only first-row elements in the W4-17-nonMR dataset (error statistics are given in kcal mol<sup>-1</sup>).<sup>[a]</sup>

Method	RMSD	MAD	MSD	# of LDs	LND	LPD
G4	0.64	0.49	0.04	12	-1.74 (BF <sub>3</sub> )	2.49 (N <sub>2</sub> O <sub>4</sub> )
G4(MP2)	0.87	0.62	-0.07	22	-4.46 (B <sub>2</sub> H <sub>6</sub> )	1.74 (CO)
G4(MP2)-6X	0.98	0.75	0.51	38	-3.52 (B <sub>2</sub> H <sub>6</sub> )	2.97 (C <sub>2</sub> F <sub>6</sub> )
ROG4(MP2)-6X	1.00	0.77	0.53	36	-3.19 (B <sub>2</sub> H <sub>6</sub> )	3.30 (C <sub>2</sub> F <sub>6</sub> )
G3 <sup>[b]</sup>	1.27	1.00	-0.61	48	-3.36 (pentane)	3.16 (C <sub>2</sub> F <sub>6</sub> )
G3B3	1.06	0.83	-0.38	38	-2.83 (B <sub>2</sub> H <sub>6</sub> )	2.04 (NCCN)
G3(MP2) <sup>[b]</sup>	1.42	1.09	-0.77	55	-6.05 (B <sub>2</sub> H <sub>6</sub> )	1.76 (NH)
G3(MP2)B3	1.19	0.90	-0.36	44	-5.76 (B <sub>2</sub> H <sub>6</sub> )	2.06 (NH)
CBS-4M	3.02	2.23	-0.23	79	-8.61 (HO <sub>2</sub> )	9.52 (benzene)
CBS-QB3	1.58	1.21	0.24	61	-3.70 (borole)	6.30 (C <sub>2</sub> F <sub>6</sub> )
ROCBS-QB3	1.43	1.08	-0.06	53	-4.19 (borole)	4.87 (C <sub>2</sub> F <sub>6</sub> )
CBS-APNO	1.18	0.86	0.50	36	-3.86 (HO <sub>3</sub> )	4.81 (C <sub>2</sub> F <sub>6</sub> )
ccCA-PS3	0.64	0.52	0.13	14	-1.59 (BN)	1.42 (C <sub>2</sub> F <sub>6</sub> )
W1	0.63	0.48	0.28	12	-1.82 (HO <sub>3</sub> )	2.25 (C <sub>2</sub> F <sub>6</sub> )
W1-F12	0.46	0.33	-0.26	4	-2.20 (HO <sub>3</sub> )	0.81 (benzene)
W1-F12 <sup>[c]</sup>	0.47	0.33	-0.26	5	-2.29 (HO <sub>3</sub> )	0.88 (benzene)
W1X-1	0.39	0.30	-0.10	2	-1.39 (B <sub>2</sub> H <sub>6</sub> )	1.12 (C <sub>2</sub> F <sub>6</sub> )
W1X-2	0.54	0.38	-0.03	7	-1.14 (B <sub>2</sub> H <sub>6</sub> )	2.89 (C <sub>2</sub> F <sub>6</sub> )
W2	0.52	0.33	-0.19	5	-2.52 (HO <sub>3</sub> )	0.99 (benzene)
W2-F12	0.45	0.33	-0.19	3	-2.07 (HO <sub>3</sub> )	1.02 (C <sub>2</sub> F <sub>6</sub> )
W2X	0.53	0.38	-0.05	4	-2.23 (HO <sub>3</sub> )	1.51 (benzene)

[a] RMSD = root-mean-square deviation, MAD = mean-absolute deviation, MSD = mean-signed deviation, # of LDs = number of deviations exceeding 1 kcal mol<sup>-1</sup>, LND = largest negative deviation, LPD = largest positive deviation. [b] *cis*-HO<sub>3</sub> and *trans*-HO<sub>3</sub> were removed due to problems with the geometry optimizations. [c] A new version of W1-F12 theory developed in the present work, in which the CCSD component of the CV correction is calculated using explicitly correlated F12 techniques.



**Table 8.** Performance of a representative set of CCSD(T) composite thermochemical procedures for the 17 atomization reactions in the W4-17-MR dataset (error statistics are given in kcal mol<sup>-1</sup>).<sup>[a]</sup>

Method	RMSD	MAD	MSD	# of LDs	LND	LPD
G4	1.71	1.28	-0.67	9	-4.19 (ClOO)	2.94 (C <sub>2</sub> )
G4(MP2)	2.29	1.84	-0.24	10	-4.53 (ClOO)	4.65 (C <sub>2</sub> )
G4(MP2)-6X	1.60	1.23	0.86	9	-3.13 (ClOO)	3.89 (C <sub>2</sub> )
ROG4(MP2)-6X	1.23	0.91	0.71	6	-0.90 (FO <sub>2</sub> )	3.58 (C <sub>2</sub> )
G3 <sup>[b]</sup>	4.26	2.59	-2.20	8	-12.68 (FO <sub>2</sub> )	2.71 (C <sub>2</sub> )
G3B3	3.52	2.62	-2.20	11	-7.54 (ClO <sub>3</sub> )	2.87 (C <sub>2</sub> )
G3(MP2) <sup>[b]</sup>	4.96	3.32	-2.47	12	-13.86 (FO <sub>2</sub> )	4.49 (C <sub>2</sub> )
G3(MP2)B3	4.38	3.23	-2.10	15	-10.58 (ClF <sub>5</sub> )	4.70 (C <sub>2</sub> )
CBS-QB3	2.70	2.05	-0.12	10	-6.59 (ClF <sub>5</sub> )	3.00 (F <sub>2</sub> O <sub>2</sub> )
ROCBS-QB3	2.55	1.66	-0.61	10	-8.00 (ClF <sub>5</sub> )	2.10 (F <sub>2</sub> O <sub>2</sub> )
ccCA-PS3	1.20	0.98	-0.26	9	-2.37 (B <sub>2</sub> )	1.50 (Cl <sub>2</sub> O)
W1	1.76	1.40	-1.36	9	-3.54 (S <sub>4</sub> )	0.27 (BN)
W1-F12	2.13	1.74	-1.68	10	-5.08 (S <sub>4</sub> )	0.50 (BN)
W1-F12 <sup>[c]</sup>	2.13	1.74	-1.68	10	-5.08 (S <sub>4</sub> )	0.50 (BN)
W1X-1	1.15	0.87	-0.69	7	-2.51 (S <sub>4</sub> )	0.70 (BN)
W1X-2	1.16	0.97	-0.55	7	-2.36 (S <sub>4</sub> )	0.82 (ClF <sub>5</sub> )
W2	1.85	1.54	-1.51	11	-3.63 (ClOO)	0.20 (BN)
W2-F12	1.39	1.05	-0.99	7	-2.90 (ClOO)	0.45 (BN)
W2X	1.65	1.34	-1.17	7	-3.19 (ClOO)	0.85 (ClF <sub>5</sub> )

[a] RMSD = root-mean-square deviation, MAD = mean-absolute deviation, MSD = mean-signed deviation, # of LDs = number of deviations exceeding 1 kcal mol<sup>-1</sup>, LND = largest negative deviation, LPD = largest positive deviation. [b] FO<sub>2</sub> was removed due to problems with the geometry optimizations. [c] A new version of W1-F12 theory developed in the present work, in which the CCSD component of the CV correction is calculated using explicitly correlated F12 techniques.

### Performance of CCSD(T) composite methods for first-row systems

Table 7 gives the error statistics for the CCSD(T) composite procedures for the 120 first row atomization reactions in the W4-17-nonMR database. An inspection of the error statistics in Tables 4 and 7 reveals that the performance of the *Gn*, CBS, ccCA-PS3, W1-F12, and W1X-*n* methods for the subset of first-row systems is significantly better than that for the entire W4-17-nonMR database. This is demonstrated by significant reductions in both RMSD values and the number of deviations exceeding 1 kcal mol<sup>-1</sup>.

The G4-type procedures attain RMSDs of 0.64 (G4), 0.87 (G4(MP2)), 0.98 (G4(MP2)-6X), and 1.00 (ROG4(MP2)-6X) kcal mol<sup>-1</sup> for the subset of first-row systems. For comparison, for the entire W4-17-nonMR database, the RMSDs for these procedures are higher by amounts ranging from 0.31 (G4) to 0.67 (G4(MP2)-6X) kcal mol<sup>-1</sup>. Notably, the G4 and G4(MP2) procedures result in only 12 and 22 deviations larger than 1 kcal mol<sup>-1</sup>, respectively. Even the G3-type procedures attain respectable RMSDs for the subset of 120 first-row systems; these range between 1.06 (G3B3) to 1.42 (G3(MP2)) kcal mol<sup>-1</sup>. For comparison, for the entire W4-17-nonMR database, the RMSDs for these procedures approach or exceed the 2 kcal mol<sup>-1</sup> mark.

The CBS-QB3 and ROCBS-QB3 procedures show relatively poor performance for the first-row systems in the W4-17-nonMR database. In particular, they attain RMSDs of 1.58 (CBS-QB3) and 1.43 (ROCBS-QB3) kcal mol<sup>-1</sup> with over 50 deviations exceeding 1 kcal mol<sup>-1</sup>. The CBS-APNO method, conversely, attains an RMSD of 1.18 kcal mol<sup>-1</sup> for the 120 first-row systems.

The RMSD for the ccCA-PS3 method over the first-row systems is lower than that for the entire W4-17-nonMR database

by 0.28 kcal mol<sup>-1</sup>. The ccCA-PS3 method results in excellent performance with an RMSD of 0.64 kcal mol<sup>-1</sup> and 14 deviations larger than 1 kcal mol<sup>-1</sup> (Table 7). We note that these error statistics are similar to those obtained for G4 and W1 theories, while computational cost increases in the order: G4 < ccCA-PS3 < W1.

Similarly, the RMSDs for the explicitly correlated W1 methods for the first-row systems are lower than those for the entire W4-17-nonMR database by 0.26 (W1-F12), 0.28 (W1X-1), and 0.24 (W1X-2) kcal mol<sup>-1</sup>. Among these, the W1X-1 procedure puts in the best performance with an RMSD of merely 0.39 kcal mol<sup>-1</sup> and only two deviations larger than 1 kcal mol<sup>-1</sup> (for B<sub>2</sub>H<sub>6</sub> and C<sub>2</sub>F<sub>6</sub>). W1-F12 theory attains a slightly higher RMSD of 0.46 kcal mol<sup>-1</sup>, together with four deviations larger than 1 kcal mol<sup>-1</sup>, all of which are characterized by considerable multi-reference effects (namely, *cis/trans*-HO<sub>3</sub>, NO<sub>2</sub>, and N<sub>2</sub>O<sub>4</sub>).

Finally, we note that the W1, W2, W2-F12, and W2X methods attain a balanced performance for the first and second-row systems. The RMSDs over the subset of first-row systems are lower than those for the entire W4-17-nonMR database by 0.1 kcal mol<sup>-1</sup> (or less). For the first-row systems, W1 theory attains an overall RMSD of 0.63 kcal mol<sup>-1</sup>, W2 and W2X show similar performance with RMSDs of ~0.52 kcal mol<sup>-1</sup>, and W2-F12 results in an RMSD of 0.45 kcal mol<sup>-1</sup>. Somewhat disappointingly, these RMSDs represent little or no improvement over the corresponding W1-type theories (Table 7).

### Performance of CCSD(T) composite methods for multireference species

So far we have examined the performance of CCSD(T) composite protocols for the 183 atomization energies in the W4-17-nonMR dataset. Let us now examine the performance of

**Table 9.** Performance of a representative set of DHDFT and MP $n$ -based procedures in conjunction with the aug'-cc-pV(5+d)Z basis set for the 183 atomization reactions involving first- and second-row elements in the W4-17-nonMR dataset (error statistics are given in kcal mol<sup>-1</sup>).<sup>[a]</sup>

Method	RMSD	MAD	MSD	# of LDs	LND	LPD
MP2	14.90	11.34	9.42	121	-15.15 (CN)	48.18 (N <sub>2</sub> O <sub>4</sub> )
SCS-MP2	8.14	6.45	5.59	98	-18.80 (CN)	30.92 (N <sub>2</sub> O <sub>4</sub> )
MP2.5	4.84	3.50	-1.36	41	-22.60 (HO <sub>3</sub> )	11.24 (C <sub>2</sub> Cl <sub>6</sub> )
MP3	14.41	12.14	-12.14	154	-54.59 (N <sub>2</sub> O <sub>4</sub> )	0.48 (AlCl <sub>3</sub> )
SCS-MP3	3.64	2.31	0.24	18	-22.26 (CN)	9.30 (NCCN)
$\omega$ B97X-2 (LP) <sup>[b]</sup>	2.88	2.05	1.12	15	-12.66 (HO <sub>3</sub> )	12.59 (P <sub>4</sub> )
$\omega$ B97X-2 (TQZ) <sup>[b]</sup>	2.62	2.11	-1.83	8	-12.51 (HO <sub>3</sub> )	5.76 (B <sub>2</sub> H <sub>6</sub> )
B2-PLYP	3.43	2.46	-0.70	24	-15.41 (C <sub>2</sub> Cl <sub>6</sub> )	11.97 (N <sub>2</sub> O <sub>4</sub> )
B2-PLYP-D3BJ	2.89	2.01	0.62	14	-9.52 (SF <sub>6</sub> )	13.69 (N <sub>2</sub> O <sub>4</sub> )
B2GP-PLYP	3.00	2.30	-1.90	12	-12.94 (HO <sub>3</sub> )	4.40 (HN <sub>3</sub> )
B2GP-PLYP-D3BJ	2.32	1.68	-0.83	7	-12.63 (HO <sub>3</sub> )	5.60 (P <sub>4</sub> )
B2K-PLYP	3.36	2.69	-2.57	18	-14.52 (HO <sub>3</sub> )	1.93 (BH <sub>3</sub> )
B2K-PLYP-D3BJ	2.97	2.27	-2.06	14	-14.36 (HO <sub>3</sub> )	3.25 (benzene)
B2T-PLYP	3.45	2.68	-2.24	27	-13.15 (SF <sub>6</sub> )	5.17 (HN <sub>3</sub> )
DSD-BLYP	2.80	2.01	1.25	11	-11.80 (HO <sub>3</sub> )	11.93 (N <sub>2</sub> O <sub>4</sub> )
DSD-PBEP86	2.67	1.93	0.86	12	-5.03 (SiF <sub>4</sub> )	13.68 (N <sub>2</sub> O <sub>4</sub> )
DSD-PBEP86-D3BJ	3.71	2.71	1.88	27	-12.22 (HO <sub>3</sub> )	14.99 (N <sub>2</sub> O <sub>4</sub> )
PWPB95	2.83	1.99	0.97	14	-8.83 (SiF <sub>4</sub> )	13.43 (P <sub>4</sub> )
PWPB95-D3BJ	3.29	2.34	1.60	19	-8.21 (SiF <sub>4</sub> )	17.52 (P <sub>4</sub> )

[a] RMSD = root-mean-square deviation, MAD = mean-absolute deviation, MSD = mean-signed deviation, # of LDs = number of deviations exceeding 5 kcal mol<sup>-1</sup> (in absolute value), LND = largest negative deviation, LPD = largest positive deviation. [b] Due to convergence issues with some of the systems, these calculations were carried out in conjunction with the Def2-QZVPP basis set.<sup>[109]</sup>

these methods for the subset of 17 highly multireference systems in the W4-17-MR subset; the relevant results are given in Table 8. The performance of the G3-type procedures significantly deteriorates compared to their performance for the non-multireference systems. In particular, they result in RMSDs ranging between 3.52 (G3B3) and 4.96 (G3(MP2)) kcal mol<sup>-1</sup>. The CBS-QB3 and ROCBS-QB3 methods do not perform as poorly, yet they still result in relatively large RMSDs (2.70 and 2.55 kcal mol<sup>-1</sup>, respectively). The G4-type methods perform much better than the G3-type and CBS-type procedures, with RMSDs ranging between 1.23 (ROG4(MP2)-6X) and 2.29 (G4(MP2)) kcal mol<sup>-1</sup>. Interestingly, the performance the G4(MP2)-6X and ROG4(MP2)-6X procedures is actually better than that for the W4-17-nonMR database. The ccCA-PS3 method provides similar performance to the ROG4(MP2)-6X method with an RMSD of 1.20 kcal mol<sup>-1</sup>. With the exception of W1-F12, which attains an RMSD of 2.13 kcal mol<sup>-1</sup>, the W $n$ -type procedures show reasonably good performance with RMSDs ranging between 1.15 (W1X-1) and 1.85 (W2) kcal mol<sup>-1</sup>.

### Performance of DHDFT methods for non-multireference species

In recent years, DHDFT methods have been found to obtain atomization energies with accuracies that are approaching the threshold of "chemical accuracy," that is, with RMSDs between 1–2 kcal mol<sup>-1</sup>.<sup>[1,70–75,108]</sup> Here, we will evaluate the performance of a representative set of DHDFT methods for the 183 atomization energies in the W4-17-nonMR database. We will also consider the performance of a number of standard and modified MP2 methods, as they have a similar computational cost to the DHDFT procedures. All of these relevant results are

presented in Table 9. The performance of higher-order MP $n$ -based methods for atomization energies and other thermochemical properties have been discussed in detail elsewhere.<sup>[110]</sup> As expected for low-order many-body perturbation theory, MP2 systematically overestimates the TAEs and results in an RMSD of 14.90 kcal mol<sup>-1</sup> with two thirds of the deviations exceeding 5 kcal mol<sup>-1</sup>. Scaling the same-spin and opposite-spin components of the MP2 correlation energy significantly improves the performance; nevertheless, spin-component-scaled (SCS) MP2 still results in an unacceptably large RMSD of 8.14 kcal mol<sup>-1</sup>. In contrast, MP3 systematically underestimates the TAEs and results in a similar RMSD to the MP2 method. Thus, MP2.5, which is an average of MP2 and MP3, provides significantly better performance with an RMSD of 4.84 kcal mol<sup>-1</sup> for the W4-17-nonMR dataset. The SCS-MP3 procedure fares better with an RMSD of 3.64 and 18 deviations larger than 5 kcal mol<sup>-1</sup>.

A quick look at Table 9 reveals that all the DHDFT methods offer better performance than the the MP $n$ -based procedures. Nevertheless, a number of DHDFT methods still offer relatively poor performance with RMSDs larger than 3.0 kcal mol<sup>-1</sup>, namely: B2GP-PLYP, B2-PLYP, B2K-PLYP, and B2T-PLYP. These methods tend to systematically underestimate the TAEs, and therefore including an empirical dispersion correction, which systematically increases the TAEs, improves their performance. For example, we obtain the following RMSDs for the dispersion corrected functionals: 2.32 (B2GP-PLYP-D3BJ), 2.89 (B2-PLYP-D3BJ), and 2.97 (B2K-PLYP-D3BJ) kcal mol<sup>-1</sup>. Particularly large dispersion corrections are obtained for the larger/heavier systems in the W4-17-nonMR database. For example, the D3BJ correction for the B2-PLYP functional increases the TAEs by: 4.0 (butane), 4.2 (cyclopentadiene and dithiotane), 4.5 (thiophene), 4.7 (CCl<sub>4</sub>), 5.3 (silole), 5.4 (pentane), 5.5 (C<sub>2</sub>Cl<sub>4</sub>), and 10.5 (C<sub>2</sub>Cl<sub>6</sub>)

kcal mol<sup>-1</sup>. It has been previously noted that dispersion corrections can have a significant effect on the performance of DHDFT methods for atomization energies of alkanes with up to eight carbons.<sup>[5]</sup>

The spin-component scaled DHDFT methods result in RMSDs of 2.83 (PWPB95), 2.80 (DSD-BLYP), and 2.67 (DSD-PBEP86) kcal mol<sup>-1</sup>. In contrast to conventional double hybrids, the DSD (dispersion-corrected, spin-component scaled, double hybrids<sup>[72]</sup>) are biased toward overestimating the TAEs. Therefore, their performance deteriorates on inclusion of the D3BJ dispersion correction. The long-range corrected spin-component scaled DHDFT methods of the Berkeley group show similar performance, with RMSDs of 2.88 ( $\omega$ B97X-2(LP)) and 2.62 ( $\omega$ B97X-2(TQZ)) kcal mol<sup>-1</sup>.

We conclude by noting that B2GP-PLYP-D3BJ is the best performing DHDFT method, which results in an overall RMSD of 2.32 kcal mol<sup>-1</sup> and only eight individual deviations larger than 5 kcal mol<sup>-1</sup>.

## Conclusions

We have used the first-principles Weizmann-4 (W4) computational thermochemistry protocol to generate the W4-17 dataset of 200 TAEs with 3 $\sigma$  CIs of 1 kJ mol<sup>-1</sup>. The W4-17 dataset includes first- and second-row molecules and radicals with up to eight non-hydrogen atoms; these cover a broad spectrum of bonding situations and multireference character, and as such are an excellent benchmark for the parameterization and validation of high-level *ab initio* methods. The W4-17 dataset includes two subsets: (i) a subset of 183 systems characterized by dynamical or mild-to-moderate nondynamical correlation effects (denoted W4-17-nonMR) and (ii) a subset of 17 systems with severe multireference effects (W4-17-MR). We have used these subsets to evaluate the performance of a wide range of CCSD(T)-based composite procedures and DHDFT methods.

With regard to the performance of the CCSD(T) composite procedures for the W4-17-nonMR database of 183 atomization energies (with %TAE([T]) values < 10%), we have reached the following conclusions:

- A range of CCSD(T)-based composite methods result in RMSDs smaller than 1 kcal mol<sup>-1</sup>. In particular, the *W<sub>n</sub>*, *W<sub>n</sub>-F12*, *W<sub>n</sub>X*, ccCA-PS3, and G4 procedures result in RMSDs ranging between 0.55 and 0.95 kcal mol<sup>-1</sup>.
- Conversely, the RMSDs for the G3-based, *G<sub>n</sub>(MP2)* and CBS methods exceed the 1 kcal mol<sup>-1</sup> mark.
- Of the *G<sub>n</sub>*-type theories, G4 gives the best overall performance with an RMSD of 0.95 kcal mol<sup>-1</sup>. However, very large deviations, ranging between 1–5 kcal mol<sup>-1</sup>, are obtained for highly fluorinated and chlorinated systems (such as SF<sub>6</sub>, PF<sub>5</sub>, PF<sub>3</sub>, BF<sub>3</sub>, CCl<sub>4</sub>, C<sub>2</sub>Cl<sub>2</sub>, C<sub>2</sub>Cl<sub>4</sub>, and C<sub>2</sub>Cl<sub>6</sub>). On removal of these systems from the training set, the RMSD for G4 is reduced to merely 0.71 kcal mol<sup>-1</sup>. The G3 and G3B3 procedures have a similar computational cost to G4, but show inferior performance in terms of both the RMSD (i.e., 1.88 and 1.94 kcal mol<sup>-1</sup>, respectively), and the number of errors exceeding 1 kcal mol<sup>-1</sup>

(i.e., 36 for G4 vs. over 75 for the G3-type procedures). If a choice has to be made between these methods, we recommend using the more recent G4 procedure.

- Of the *G<sub>n</sub>(MP2)*-type procedures, G4(MP2) puts in the best overall performance with an RMSD of 1.29 kcal mol<sup>-1</sup> and 48 errors exceeding 1 kcal mol<sup>-1</sup>. The G4(MP2)-6X, ROG4(MP2)-6X, G3(MP2), and G3(MP2)B3 procedures perform less well with RMSDs ranging between 1.65 and 2.22 kcal mol<sup>-1</sup> and over 75 errors exceeding 1 kcal mol<sup>-1</sup>.
- The CBS-type procedures (CBS-QB3, ROCBS-QB3, and CBS-4M) result in RMSDs between 2.04–4.23 kcal mol<sup>-1</sup> for the W4-17-nonMR database. Therefore, we would not recommend these methods for the calculation of atomization energies.
- Let us move on to the computationally more expensive ccCA-PS3 method. This method offers similar performance to the G4 method and results in overall RMSD of 0.92 kcal mol<sup>-1</sup>.
- The original W1 theory results in a lower RMSD of 0.74 kcal mol<sup>-1</sup> and 30 deviations larger than 1 kcal mol<sup>-1</sup>. The W1-F12 and W1X-*n* methods give comparable performance to W1 theory at a significantly reduced computational cost; namely, they result in RMSDs between 0.67 (W1X-1) and 0.78 (W1X-2) kcal mol<sup>-1</sup>.
- The computationally more expensive W2-type methods show somewhat better performance with RMSDs of 0.55 (W2-F12), 0.57 (W2), and 0.63 (W2X) kcal mol<sup>-1</sup>, and 9 (W2-F12), 10 (W2), and 18 (W2X) deviations larger than 1 kcal mol<sup>-1</sup>.
- Finally, we note that the performance of the *G<sub>n</sub>*, CBS, ccCA-PS3, W1-F12, and W1X-*n* methods for the subset of 120 first-row systems in the W4-17-nonMR database is significantly better than that for the entire W4-17-nonMR database. This is demonstrated by significant reductions in both RMSD values and the number of deviations exceeding 1 kcal mol<sup>-1</sup>.


The W4-17-nonMR database represents a challenging test for DHDFT methods. With regard to the performance of these procedures, we draw the following conclusions:

- The first-generation DHDFT methods result in relatively large RMSDs of 3.00 (B2GP-PLYP), 3.36 (B2K-PLYP), 3.43 (B2-PLYP), and 3.45 (B2T-PLYP) kcal mol<sup>-1</sup>. For comparison, all the MP2 and MP3-based procedures result in larger RMSDs ranging between 3.64 (SCS-MP3) and 14.90 (MP2) kcal mol<sup>-1</sup>.
- Dispersion corrections significantly improve the performance of the first-generation DHDFT methods. In particular, we obtain the following RMSDs 2.32 (B2GP-PLYP-D3BJ), 2.89 (B2-PLYP-D3BJ), and 2.97 (B2K-PLYP-D3BJ) kcal mol<sup>-1</sup>.
- The DSD double hybrids result in RMSDs of 2.67 (DSD-PBEP86), 2.80 (DSD-BLYP), and 2.83 (PWPB95) kcal mol<sup>-1</sup>. The long-range corrected DHDFT methods show similar performance with RMSDs of 2.62 ( $\omega$ B97X-2(TQZ)) and 2.88 ( $\omega$ B97X-2(LP)) kcal mol<sup>-1</sup>.

In summary, the results presented here will help computational chemists and experimentalists to select appropriate composite *ab initio* procedures for the calculation of TAEs (or their cognates, molecular heats of formation). We make recommendations based on the performance of these methods for the entire W4-17-nonMR database and for selected subsets. In addition, we believe that the W4-17 database will be a useful resource for critical evaluation of future composite *ab initio* methods and DHDFT functionals. We hope that the results presented here will inspire the development of next-generation methods with improved performance for TAEs.

**Keywords:** composite *ab initio* methods · coupled cluster theory · Weizmann-4 theory · G4 theory · CCSD(T)-based methods

How to cite this article: A. Karton, N. Sylvetsky, J. M. L. Martin. *J. Comput. Chem.* **2017**, DOI: 10.1002/jcc.24854

 Additional Supporting Information may be found in the online version of this article.

- [1] A. Karton, S. Daon, J. M. L. Martin, *Chem. Phys. Lett.* **2011**, *510*, 165.
- [2] S. E. Wheeler, K. N. Houk, P. V. R. Schleyer, W. D. Allen, *J. Am. Chem. Soc.* **2009**, *131*, 2547.
- [3] S. E. Wheeler, *WIREs Comput. Mol. Sci.* **2012**, *2*, 204.
- [4] A. Karton, P. R. Schreiner, J. M. L. Martin, *J. Comput. Chem.* **2016**, *37*, 49.
- [5] A. Karton, D. Gruzman, J. M. L. Martin, *J. Phys. Chem. A* **2009**, *113*, 8434.
- [6] R. O. Ramabhadran, K. Raghavachari, *J. Chem. Theory Comput.* **2011**, *7*, 2094.
- [7] R. O. Ramabhadran, K. Raghavachari, *J. Phys. Chem. A* **2012**, *116*, 7531.
- [8] R. O. Ramabhadran, K. Raghavachari, *Acc. Chem. Res.* **2014**, *47*, 3596.
- [9] E. I. Izgorodina, M. L. Coote, L. Radom, *J. Phys. Chem. A* **2005**, *109*, 7558.
- [10] M. D. Wodrich, C. Corminboeuf, S. E. Wheeler, *J. Phys. Chem. A* **2012**, *116*, 3436.
- [11] R. J. O'Reilly, A. Karton, L. Radom, *Int. J. Quantum Chem.* **2012**, *112*, 1862.
- [12] M. D. Wodrich, C. Corminboeuf, P. V. R. Schleyer, *Org. Lett.* **2006**, *8*, 3631.
- [13] M. D. Wodrich, C. Corminboeuf, P. R. Schreiner, A. A. Fokin, P. V. R. Schleyer, *Org. Lett.* **2006**, *9*, 1851.
- [14] S. Grimme, *Angew. Chem. Int. Ed.* **2006**, *45*, 4460.
- [15] P. R. Schreiner, *Angew. Chem. Int. Ed.* **2007**, *46*, 4217.
- [16] S. Grimme, M. Steinmetz, M. Korth, *J. Org. Chem.* **2007**, *72*, 2118.
- [17] S. Grimme, *Org. Lett.* **2010**, *12*, 4670.
- [18] A. Karton, J. M. L. Martin, *Mol. Phys.* **2012**, *110*, 2477.
- [19] L.-J. Yu, A. Karton, *Chem. Phys.* **2014**, *441*, 166.
- [20] A. Karton, *J. Comput. Chem.* **2017**, *38*, 370.
- [21] A. Karton, E. Rabinovich, J. M. L. Martin, B. Ruscic, *J. Chem. Phys.* **2006**, *125*, 144108.
- [22] N. Sylvetsky, K. A. Peterson, A. Karton, J. M. L. Martin, *J. Chem. Phys.* **2016**, *144*, 214101.
- [23] A. Karton, P. R. Taylor, J. M. L. Martin, *J. Chem. Phys.* **2007**, *127*, 064104.
- [24] A. Karton, *WIREs Comput. Mol. Sci.* **2016**, *6*, 292.
- [25] K. A. Peterson, D. Feller, D. A. Dixon, *Theor. Chem. Acc.* **2012**, *131*, 1079.
- [26] J. T. Margraf, D. Ranasinghe, R. J. Bartlett, *Phys. Chem. Chem. Phys.* **2017**, *19*, 9798.
- [27] N. Mardirossian, M. Head-Gordon, *J. Chem. Phys.* **2016**, *144*, 214110.
- [28] R. Peverati, D. G. Truhlar, *Philos. Trans. R. Soc. A* **2014**, *372*, 20120476.
- [29] A. J. Cohen, P. Mori-Sánchez, W. Yang, *Chem. Rev.* **2012**, *112*, 289.
- [30] L. Goerigk, S. Grimme, *Phys. Chem. Chem. Phys.* **2011**, *13*, 6670.
- [31] L. A. Curtiss, P. C. Redfern, K. Raghavachari, *WIREs Comput. Mol. Sci.* **2011**, *1*, 810.
- [32] N. J. DeYonker, T. R. Cundari, A. K. Wilson, In *Advances in the Theory of Atomic and Molecular Systems (Progress in Theoretical Chemistry and Physics, Vol. 19)*; P. Piecuch, J. Maruani, G. Delgado-Barrio, S. Wilson, Eds.; Springer Netherlands: Dordrecht, Netherlands, **2009**; pp. 197–224.
- [33] D. Feller, K. A. Peterson, D. A. Dixon, *J. Chem. Phys.* **2008**, *129*, 204105.
- [34] T. Helgaker, W. Klopper, D. P. Tew, *Mol. Phys.* **2008**, *106*, 2107.
- [35] W. M. F. Fabian, *Monatsch. Chem.* **2008**, *139*, 309.
- [36] J. M. L. Martin, *Annu. Rep. Comput. Chem.* **2005**, *1*, 31.
- [37] T. Helgaker, W. Klopper, K. L. Bak, A. Halkier, P. Jørgensen, J. Olsen, In *Quantum-Mechanical Prediction of Thermochemical Data, Vol. 22*; J. Cioslowski, Ed.; Kluwer: Dordrecht, **2001**; pp. 1–30.
- [38] J. M. L. Martin, S. Parthiban, In *Quantum-Mechanical Prediction of Thermochemical Data, Vol. 22*, J. Cioslowski, Ed.; Kluwer: Dordrecht, **2001**; pp. 31–65.
- [39] P. R. Spackman, D. Jayatilaka, A. Karton, *J. Chem. Phys.* **2016**, *145*, 104101.
- [40] A. Karton, J. M. L. Martin, *Theor. Chem. Acc.* **2006**, *115*, 330.
- [41] T. H. Dunning, *J. Chem. Phys.* **1989**, *90*, 1007.
- [42] R. A. Kendall, T. H. Dunning, R. J. Harrison, *J. Chem. Phys.* **1992**, *96*, 6796.
- [43] T. H. Dunning, K. A. Peterson, A. K. Wilson, *J. Chem. Phys.* **2001**, *114*, 9244.
- [44] K. A. Peterson, T. H. Dunning, *J. Chem. Phys.* **2002**, *117*, 10548.
- [45] M. Douglas, N. M. Kroll, *Ann. Phys.* **1974**, *82*, 89.
- [46] B. A. Heß, *Phys. Rev. A* **1986**, *33*, 3742.
- [47] W. A. de Jong, R. J. Harrison, D. A. Dixon, *J. Chem. Phys.* **2001**, *114*, 48.
- [48] Available at: <https://www.nist.gov/pml/basic-atomic-spectroscopic-data-handbook> (National Institute of Standards and Technology: Gaithersburg, MD) (accessed on March 30, **2017**).
- [49] B. Ruscic, R. E. Pinzon, M. L. Morton, N. K. Srinivasan, M.-C. Su, J. W. Sutherland, J. V. Michael, *J. Phys. Chem. A* **2006**, *110*, 6592.
- [50] B. Ruscic, R. E. Pinzon, G. von Laszewski, D. Kodeboyina, A. Burcat, D. Leahy, D. Montoya, A. F. Wagner, *J. Phys. Conf. Ser.* **2005**, *16*, 561.
- [51] B. Ruscic, *Encyclopedia of Science and Technology (2005 Yearbook of Science and Technology)*; McGraw-Hill: New York, **2004**; pp. 3–7.
- [52] B. Ruscic, R. E. Pinzon, M. L. Morton, G. von Laszewski, S. J. Bittner, S. G. Nijsure, K. A. Amin, M. Minkoff, A. F. Wagner, *J. Phys. Chem. A* **2004**, *108*, 9979.
- [53] L. A. Curtiss, P. C. Redfern, K. Raghavachari, *J. Chem. Phys.* **2007**, *126*, 084108.
- [54] L. A. Curtiss, P. C. Redfern, K. Raghavachari, *J. Chem. Phys.* **2007**, *127*, 124105.
- [55] B. Chan, J. Deng, L. Radom, *J. Chem. Theory Comput.* **2011**, *7*, 112.
- [56] B. Chan, A. Karton, K. Raghavachari, L. Radom, *J. Phys. Chem. A* **2016**, *120*, 9299.
- [57] J. W. Ochterski, G. A. Petersson, J. A. Montgomery, Jr., *J. Chem. Phys.* **1996**, *104*, 2598.
- [58] J. A. Montgomery, Jr., M. J. Frisch, J. W. Ochterski, G. A. Petersson, *J. Chem. Phys.* **1999**, *110*, 2822.
- [59] J. M. L. Martin, G. de Oliveira, *J. Chem. Phys.* **1999**, *111*, 1843.
- [60] A. Karton, J. M. L. Martin, *J. Chem. Phys.* **2012**, *136*, 124114.
- [61] B. Chan, L. Radom, *J. Chem. Theory Comput.* **2012**, *8*, 4259.
- [62] B. Chan, L. Radom, *J. Chem. Theory Comput.* **2015**, *11*, 2109.
- [63] B. Chan, *Pure Appl. Chem.* (in press). DOI: 10.1515/pac-2016-1116
- [64] N. J. DeYonker, T. R. Cundari, A. K. Wilson, *J. Chem. Phys.* **2006**, *124*, 114104.
- [65] N. J. DeYonker, T. Grimes, S. Yockel, A. Dinescu, B. Mintz, T. R. Cundari, A. K. Wilson, *J. Chem. Phys.* **2006**, *125*, 104111.
- [66] N. J. DeYonker, B. R. Wilson, A. W. Pierpont, T. R. Cundari, A. K. Wilson, *Mol. Phys.* **2009**, *107*, 1107.
- [67] A. Mahler, A. K. Wilson, *J. Chem. Theory Comput.* **2013**, *9*, 1402.
- [68] M. K. Kesharwani, B. Brauer, J. M. L. Martin, *J. Phys. Chem. A* **2015**, *119*, 1701.
- [69] S. Grimme, *J. Chem. Phys.* **2006**, *124*, 034108.
- [70] A. Karton, A. Tarnopolsky, J. F. Lamere, G. C. Schatz, J. M. L. Martin, *J. Phys. Chem. A* **2008**, *112*, 12868.
- [71] A. Tarnopolsky, A. Karton, R. Sertchook, D. Vuzman, J. M. L. Martin, *J. Phys. Chem. A* **2008**, *112*, 3.



- [72] S. Kozuch, D. Gruzman, J. M. L. Martin, *J. Phys. Chem. C* **2010**, *114*, 20801.
- [73] S. Kozuch, J. M. L. Martin, *Phys. Chem. Chem. Phys.* **2011**, *13*, 20104.
- [74] S. Kozuch, J. M. L. J. Martin, *Comput. Chem.* **2013**, *34*, 2327.
- [75] L. Goerigk, S. Grimme, *J. Chem. Theory Comput.* **2010**, *6*, 107.
- [76] J.-D. Chai, M. Head-Gordon, *J. Chem. Phys.* **2009**, *131*, 174105.
- [77] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, et al. *Gaussian09*; Gaussian, Inc.: Wallingford, CT, **2009**.
- [78] Y. Shao, Z. Gan, E. Epifanovsky, A. T. B. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P. R. Horn, L. D. Jacobson, I. Kaliman, R. Z. Khaliullin, T. Kus, A. Landau, J. Liu, E. I. Proynov, Y. M. Rhee, R. M. Richard, M. A. Rohrdanz, R. P. Steele, E. J. Sundstrom, H. L. Woodcock III, P. M. Zimmerman, D. Zuev, B. Albrecht, E. Alguire, B. Austin, G. J. O. Beran, Y. A. Bernard, E. Berquist, K. Brandhorst, K. B. Bravaya, S. T. Brown, D. Casanova, C.-M. Chang, Y. Chen, S. H. Chien, K. D. Closser, D. L. Crittenden, M. Diedenhofen, R. A. DiStasio Jr., H. Do, A. D. Dutoi, R. G. Edgar, S. Fatehi, L. Fusti-Molnar, A. Ghysels, A. Golubeva-Zadorozhnaya, J. Gomes, M. W. D. Hanson-Heine, P. H. P. Harbach, A. W. Hauser, E. G. Hohenstein, Z. C. Holden, T.-C. Jagau, H. Ji, B. Kaduk, K. Khistyayev, J. Kim, J. Kim, R. A. King, P. Klunzinger, D. Kosenkov, T. Kowalczyk, C. M. Krauter, K. U. Lao, A. D. Laurent, K. V. Lawler, S. V. Levchenko, C. Y. Lin, F. Liu, E. Livshits, R. C. Lochan, A. Luenser, P. Manohar, S. F. Manzer, S.-P. Mao, N. Mardirossian, A. V. Marenich, S. A. Maurer, N. J. Mayhall, E. Neuscammann, C. M. Oana, R. Olivares-Amaya, D. P. O'Neill, J. A. Parkhill, T. M. Perrine, R. Peverati, A. Prociuk, D. R. Rehn, E. Rosta, N. J. Russ, S. M. Sharada, S. Sharma, D. W. Small, A. Sodt, T. Stein, D. Stück, Y.-C. Su, A. J. W. Thom, T. Tsuchimochi, V. Vanovschi, L. Vogt, O. Vydrov, T. Wang, M.A. Watson, J. Wenzel, A. White, C. F. Williams, J. Yang, S. Yeganeh, S. R. Yost, Z.-Q. You, I. Ying Zhang, X. Zhang, Y. Zhao, B. R. Brooks, G. K. L. Chan, D. M. Chipman, C. J. Cramer, W. A. Goddard III, M. S. Gordon, W. J. Hehre, A. Klamt, H. F. Schaefer III, M. W. Schmidt, C. D. Sherrill, D. G. Truhlar, A. Warshel, X. Xu, A. Aspuru-Guzik, R. Baer, A. T. Bell, N. A. Besley, J.-D. Chai, A. Dreuw, B. D. Dunietz, T. R. Furlani, S. R. Gwaltney, C.-P. Hsu, Y. Jung, J. Kong, D. S. Lambrecht, W. Liang, C. Ochsenfeld, V. A. Rassolov, L. V. Slipchenko, J. E. Subotnik, T. Van Voorhis, J. M. Herbert, A. I. Krylov, P. M. W. Gill & M. Head-Gordon. *Mol. Phys.* **2015**, *113*, 184.
- [79] S. Grimme, S. Ehrlich, L. Goerigk, *J. Comput. Chem.* **2011**, *32*, 1456.
- [80] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, *132*, 154104.
- [81] L. Goerigk, S. Grimme, *J. Chem. Theory Comput.* **2011**, *7*, 291.
- [82] A. D. Becke, E. R. Johnson, *J. Chem. Phys.* **2005**, *123*, 154101.
- [83] S. Grimme, A. Hansen, J. G. Brandenburg, C. Bannwarth, *Chem. Rev.* **2016**, *116*, 5105.
- [84] A. Karton, L. Goerigk, *J. Comp. Chem.* **2015**, *36*, 622.
- [85] S. Grimme, *J. Chem. Phys.* **2003**, *118*, 9095.
- [86] M. Pitonak, P. Neogrady, J. Cerny, S. Grimme, P. Hobza, *Chem. Phys. Chem.* **2009**, *10*, 282.
- [87] S. Grimme, *J. Comput. Chem.* **2003**, *24*, 1529.
- [88] T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut, T. B. Adler, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, P. Celani, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut, K. R. Shamasundar, T. B. Adler, R. D. Amos, A. Bernhardsson, A. Berning, D.L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, E. Goll, C. Hampel, A. Hesselmann, G. Hetzer, T. Hrenar, G. Jansen, C. Köppl, Y. Liu, A. W. Lloyd, R. A. Mata, A. J. May, S. J. McNicholas, W. Meyer, M. E. Mura, A. Nicklass, D. P. O'Neill, P. Palmieri, D. Peng, K. Pflüger, R. Pitzer, M. Reiher, T. Shiozaki, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, M. Wang, Molpro, version 2012.1, a package of ab initio programs 2012, <http://www.molpro.net>.
- [89] H. J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schutz, *WIREs Comput. Mol. Sci.* **2012**, *2*, 242.
- [90] M. Kállay, Z. Rolik, J. Csontos, P. Nagy, G. Samu, D. Mester, I. Ladjánszki, L. Szegedy, B. Ladóczki, K. Petrov, M. Farkas, B. Hégyely. MRCC: A Quantum Chemical Program Suite. See also: <http://www.mrcc.hu>.
- [91] Z. Rolik, L. Szegedy, I. Ladjánszki, B. Ladóczki, M. Kállay, *J. Chem. Phys.* **2013**, *139*, 094105.
- [92] T. J. Lee, P. R. Taylor, *Int. J. Quantum Chem.* **1989**, *36*, 199.
- [93] C. L. Janssen, I. M. B. Nielsen, *Chem. Phys. Lett.* **1998**, *290*, 423.
- [94] U. R. Fogueri, S. Kozuch, A. Karton, J. M. L. Martin, *Theor. Chem. Acc.* **2013**, *132*, 1291.
- [95] A. Karton, J. M. L. Martin, *J. Chem. Phys.* **2010**, *133*, 144102.
- [96] A. Karton, S. Parthiban, J. M. L. Martin, *J. Phys. Chem. A* **2009**, *113*, 4802.
- [97] J. D. Watts, R. J. Bartlett, *Int. J. Quantum Chem.* **1994**, *52*, 195.
- [98] J. D. Watts, M. Urban, R. J. Bartlett, *Theor. Chim. Acta* **1995**, *90*, 341.
- [99] J. Zheng, Y. Zhao, D. G. Truhlar, *J. Phys. Chem. A* **2007**, *111*, 4632.
- [100] W. Wan, A. Karton, *Chem. Phys. Lett.* **2016**, *643*, 34.
- [101] S. M. Yockel, A. K. Wilson, *Theor. Chem. Acc.* **2008**, *120*, 119.
- [102] A. Karton, B. Chan, K. Raghavachari, L. Radom, *J. Phys. Chem. A* **2013**, *117*, 1834.
- [103] A. Karton, L.-J. Yu, M. K. Kesharwani, J. M. L. Martin, *Theor. Chem. Acc.* **2014**, *133*, 1483.
- [104] J. G. Hill, S. Mazumder, J. M. L. Martin, *J. Chem. Phys.* **2010**, *132*, 054108.
- [105] H. J. Werner, G. Knizia, F. R. Manby, *Mol. Phys.* **2011**, *109*, 407.
- [106] K. A. Peterson, M. K. Kesharwani, J. M. L. Martin, *Mol. Phys.* **2015**, *113*, 1551.
- [107] O. Marchetti, H. J. Werner, *Mol. Phys.* **2008**, *10*, 3400.
- [108] L. Goerigk, S. Grimme, *WIREs Comput. Mol. Sci.* **2014**, *4*, 576.
- [109] F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297.
- [110] L.-J. Yu, W. Wan, A. Karton, *Chem. Phys.* **2016**, *480*, 23.

Received: 2 April 2017  
Revised: 10 May 2017  
Accepted: 18 May 2017  
Published online on 00 Month 2017