AIP The Journal of Chemical Physics

An expanded calibration study of the explicitly correlated CCSD(T)-F12b method using large basis set standard CCSD(T) atomization energies

David Feller and Kirk A. Peterson

Citation: J. Chem. Phys. 139, 084110 (2013); doi: 10.1063/1.4819125
View online: http://dx.doi.org/10.1063/1.4819125
View Table of Contents: http://jcp.aip.org/resource/1/JCPSA6/v139/i8
Published by the AIP Publishing LLC.

THE JOURNAL OF CHEMICAL PHYSICS **139**, 084110 (2013)

An expanded calibration study of the explicitly correlated CCSD(T)-F12b method using large basis set standard CCSD(T) atomization energies

David Feller[a] and Kirk A. Peterson
Department of Chemistry, Washington State University, Pullman, Washington 99164-4630, USA

(Received 30 June 2013; accepted 8 August 2013; published online 28 August 2013)

The effectiveness of the recently developed, explicitly correlated coupled cluster method CCSD(T)-F12b is examined in terms of its ability to reproduce atomization energies derived from complete basis set extrapolations of standard CCSD(T). Most of the standard method findings were obtained with aug-cc-pV7Z or aug-cc-pV8Z basis sets. For a few homonuclear diatomic molecules it was possible to push the basis set to the aug-cc-pV9Z level. F12b calculations were performed with the cc-pV*n*Z-F12 (n = D, T, Q) basis set sequence and were also extrapolated to the basis set limit using a Schwenke-style, parameterized formula. A systematic bias was observed in the F12b method with the (VTZ-F12/VQZ-F12) basis set combination. This bias resulted in the underestimation of reference values associated with small molecules (valence correlation energies <0.5 E_h) and an even larger overestimation of atomization energies for bigger systems. Consequently, caution should be exercised in the use of F12b for high accuracy studies. Root mean square and mean absolute deviation error metrics for this basis set combination were comparable to complete basis set values obtained with standard CCSD(T) and the aug-cc-pVDZ through aug-cc-pVQZ basis set sequence. However, the mean signed deviation was an order of magnitude larger. Problems partially due to basis set superposition error were identified with second row compounds which resulted in a weak performance for the smaller VDZ-F12/VTZ-F12 combination of basis sets. © *2013 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4819125]

I. INTRODUCTION

The majority of high accuracy electronic structure calculations on "well behaved" chemical systems, i.e., those not suffering from excessive multi-configurational character, make use of the single reference coupled cluster technique involving single and double excitations, combined with a non-iterative, quasiperturbative estimate of the effect of triple excitations, CCSD(T).[1–4] This method recovers a large enough fraction of the electron correlation energy that it can be applied to an extensive range of molecules, including many whose wave functions exhibit moderate multiconfigurational character, such as C_2, O_3, and FOO.[5,6] At the current state of the art, theoretical methods are capable of reliably predicting an assortment of molecular properties. While some properties require only modest computational effort, such as the electric dipole moment, others are more challenging. Among the latter are the class of thermochemical properties that includes heats of formation or atomization energies.[7] Numerous studies have shown that the principal source of error in CCSD(T) calculations arises from the incompleteness in the 1-particle basis set.[8,9] Consequently, for studies attempting to achieve high accuracy the treatment of basis set issues is critical. Unfortunately, the steep increase in computational cost accompanying improvements in the basis set and the lack of formal *a priori* error bars makes a brute force assault on the problem very difficult. To put matters into perspective, high accuracy investigations normally require levels of theory beyond frozen core (FC) CCSD(T) in which the inner shell electrons are excluded from the recovery of correlation effects. These include consideration of core/valence correlation, scalar relativistic, and higher order correlation effects, such as exemplified in the FPD,[7,9,10] Gaussian-*n*,[11–13] W*n*,[14,15] HEATx,[16,17] ccCA,[18–20] FPA,[21–23] and Petersson-style CBS[24–26] approaches. The calculations in this study are limited to a treatment of valence-only correlation.

In an effort to reduce the impact of the basis set incompleteness problem a variety of simple extrapolation formulas that exploit the smooth 1-particle convergence behavior characteristic of the correlation consistent basis sets of Dunning and co-workers have been proposed.[27–41] Extrapolation formulas provide a practical route to estimating the complete basis set (CBS) limit while avoiding the need to perform extremely expensive or possibly intractable calculations. Correlation consistent basis sets are available for elements from the 1st to 6th periods and include levels as high as 10ζ, making them ideally suited for a wide range of high accuracy investigations. When the correlation consistent basis sets are augmented with extra diffuse functions they are conventionally denoted aug-cc-pV*n*Z (first and third row) and aug-cc-pV(*n*+d)Z (second row). For the sake of brevity we will use the notation aV*n*Z to mean a combination of aug-cc-pV*n*Z for first and third row elements and aug-cc-pV(*n*+d)Z for second row elements throughout this work.

In previous work we compared 7 extrapolation formulas against a collection of 141 atomization energies obtained with large basis set, standard CCSD(T).[42] Even the poorest performing scheme was found to produced results comparable to

---

[a] E-mail: dfeller@owt.com

or better than raw CCSD(T) results with the next higher level basis set, i.e., CBS(aV*n*Z) ≈ raw[aV(*n*+1)Z]. The *most* effective schemes produced results comparable to raw [aV(*n*+2)Z] values. While no single formula proved superior across all basis sets and classes of molecules, a five formula average and an expression involving the inverse of the maximum angular momentum present in the basis set ($\ell_{\max}$):

$$E(\ell_{\max}) = E_{\mathrm{CBS}} + A/(\ell_{max} + 1/2)^4, \tag{1}$$

which was proposed by Martin[43] were statistically superior. Martin's formula built upon the earlier works of Schwartz,[44,45] Carroll,[46] Hill,[47] and Kutzelnigg and Morgan.[48] A follow-up study used an expanded 183-member reference set to explore improved extrapolation formulas, one of which achieved another increase in the level of basis set, i.e., CBS(aV*n*Z) ≈ raw[aV(*n*+3)Z].[49]

An alternate route to reducing the basis set truncation error is provided by explicitly correlated methods, such as the recently developed CCSD(T)-F12a/b techniques.[50–52] Through the inclusion of nonlinear terms in the interelectronic distance, r_{ij}, CCSD(T)-F12 methods achieve dramatically improved convergence with respect to the basis set. Note that in the current implementation of the F12 method, only the CCSD step is explicitly correlated. Thus, the triple excitation piece of the energy converges in essentially the same manner as in standard CCSD(T). We compared F12a and F12b predictions of structures, harmonic frequencies, and atomization energies against large basis set standard CCSD(T) for C_2 and 15 small hydrocarbons.[53] F12b, which differs from F12a by the inclusion of an additional energy correction which approximately doubles the magnitude of the coupling between the conventional and explicitly correlated pieces of the calculation,[51,54,55] was found to produce results closer to standard CCSD(T). F12a frequently overestimated the true basis set limit for the correlation energy, especially with quadruple zeta basis sets. Raw CCSD(T)-F12b/cc-pV*n*Z-F12 results were comparable in accuracy to CCSD(T)/aug-cc-V(*n*+2)Z values, but were obtained at a fraction of the computational cost. In addition to the raw CCSD(T)-F12b results, we also studied the performance of extrapolated F12b energies using the formula suggested by Hill *et al.*[56] In that procedure the self-consistent field (SCF) limit was approximated by the SCF + complementary auxiliary basis set (CABS) singles correction[50] determined with the largest basis set (cc-pVTZ-F12 for the DT extrapolation, cc-pVQZ-F12 for the TQ extrapolation). A slight tendency to overshoot the CCSD(T)/CBS(aVQ56) estimates was noted in larger systems, but the uncertainty in the reference values was sufficiently large that it placed in question the significance of this observation.

The goal of the present work is to extend the calibration of the F12b method beyond C_2 and the hydrocarbons included in the 2010 investigation by using a much larger collection of reference atomization energies. As with the hydrocarbon study, both raw F12b/V*n*Z-F12 (*n* = D, T, Q) and F12b/CBS extrapolated values will be examined. Particular attention will be paid to the accuracy of the reference set in light of the presumed high accuracy of the explicitly correlated method. The current reference set began with the 183-member collection of molecules,[49] but increased the number of systems treated with 7ζ- and 8ζ-level basis sets. Additional measures that were taken to further improve the reference set accuracy, will be described. The statistical analyses reported here will be performed with information stored in the Computational Results Database (CRDB).[57] The database currently holds approximately 121 000 experimental and theoretical entries covering 432 neutral, anion, and cation atomic and molecular species in their ground and low-lying excited states. While the emphasis in populating the database was on high accuracy results, a wide range of methods is covered.

The focus in this study will be on energy *differences* rather than total energies. Obtaining total energies that are sufficiently well converged with respect to the 1-particle basis set to serve as accurate CBS reference values has proven to be extremely difficult. A simple illustration can be found in Table I, where full configuration interaction (FCI) total energies for the H_2 molecule are shown as a function of the basis set. Essentially exact values are available from the works of Wolniewicz[58] and Cencek and Kutzelnigg.[59] Due to the well-known problems of reproducing the electron-electron Coulomb cusp with orbital expansion products, even the aV9Z raw energy remains in error by 3 × 10^{-5} E_h. CBS extrapolation with either the $1/(\ell_{\max} + 1/2)^4$ or $1/\ell_{\max}{}^3$

TABLE I. H_2 ($^1\Sigma_g{}^+$) full configuration interaction energies.

				$1/(\ell_{\max} + 1/2)^4$		$1/\ell_{\max}{}^3$	
Basis	Composition	No. of Functions	Raw E(FCI)	E(FCI)/CBS	Error[a]	E(FCI)/CBS	Error[a]
aV5Z	[6s,5p,4d,3f,2g]	160	−1.1742522	−1.174565	− 0.000089	−1.174655	− 0.000179
aV6Z	[7s,6p,5d,4f,3g,2h]	254	−1.1743600	− 1.174473	0.000003	−1.174508	− 0.000032
aV7Z	[8s,7p,6d,5f,4g,3h,2i]	378	−1.1744109	− 1.174477	− 0.000001	−1.174497	− 0.000021
aV8Z	[9s,8p,7d,6f,5g,4h,3i,2k]	536	−1.1744306	− 1.174470	0.000006	−1.174481	− 0.000005
aV9Z	[10s,9p,8d,7f,6g,5h,4i,3k,2l]	732	−1.1744464	− 1.174469	0.000007	−1.174476	0.000000
CK[b]			−1.1744757				
W[c]			−1.1744757				

[a] Error is measured with respect to the Cencek and Kutzelnigg value.[59] The complete basis set (CBS) extrapolated values are based on the maximum angular momentum ($\ell_{\max}$) found in the basis set. Energies were evaluated at the optimal bond lengths; aV5Z = 0.74160 Å, aV6Z = 0.74150 Å, aV7Z = 0.74147 Å, aV8Z = 0.74145 Å, and aV9Z = 0.74144 Å.

[b] Cencek and Kutzelnigg[59] at 1.40 bohrs using a basis set of 1200 explicitly correlated functions.

[c] Wolniewicz.[58]

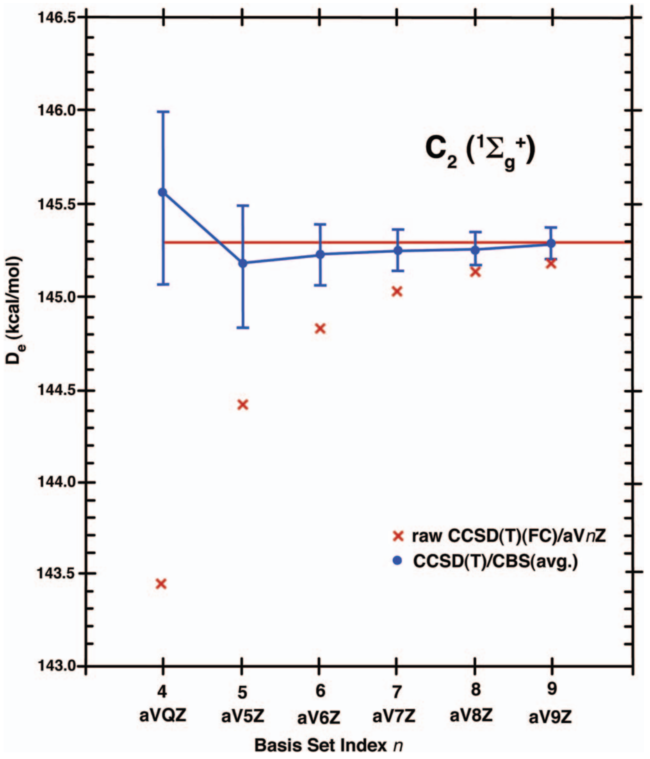

FIG. 1. Frozen core CCSD(T) dissociation energies for C_2 as a function of the basis set index. The complete basis set estimates used a multi-formula average. Error bars were based on the spread among the CBS values.

formulas[43,60] significantly reduces the error. Replicating these results for larger systems is currently not possible.

II. APPROACH

Frozen core CCSD(T) molecular energies were obtained at the respective aV*n*Z (n = D, T, Q, …, 8) optimal geometries using MOLPRO version 2012.1.[61] The largest CCSD(T) calculation with MOLPRO involved slightly more than 1800 basis functions. In a few instances it was possible to perform CCSD(T)/aV9Z calculations with new basis sets developed for this study (see supplementary material).[84] The correlation consistent basis sets are well known for producing smoothly convergent properties as a function of the basis set index n. An illustration of this characteristic in the present context is shown in Figure 1, where the raw aV*n*Z dissociation energies (D_e) of C_2 ($^1\Sigma_g{}^+$) are plotted against the basis set index. Figure 1 also shows the behavior of the CBS estimates based on an average of multiple extrapolation formulas. The extrapolation formula in Eq. (1) produces similar results.

MOLPRO currently lacks support for basis functions beyond $\ell_{\max} = 6$ (i functions). Consequently, most of the aV7Z and aV8Z optimal structures were obtained without k and l functions. For a few homonuclear diatomics it proved possible to retain k and l functions in the optimizations. Calculations with the full aV7Z and aV8Z basis sets were performed with the Dalton 2.0[62] and Gaussian 09[63] applications. Open shell calculations were based on the R/UCCSD(T) method, which begins with restricted open-shell Hartree-Fock (ROHF) orbitals, but allows a small amount of spin contamination in the solution of the CCSD equations.[64,65] Atomic symmetry in the orbitals was imposed in calculations on the atomic asymptotes. The current version of Dalton only supports closed shell systems and Gaussian 09 only handles unrestricted UCCSD(T) open shell calculations. Consequently, we were forced to estimate the aV7Z and aV8Z basis sets R/UCCSD(T) energies. This was accomplished by assuming that the difference between UCCSD(T) and R/UCCSD(T) energies for basis sets containing k and l functions could be determined with sufficient accuracy using basis sets containing no higher than i-functions. An example is shown in Table II for the carbon atom where the UCCSD(T)-R/UCCSD(T) difference in the right hand column has apparently converged to 10^{-6} E_h or better at the aV6Z level.

Due to the time-consuming nature of aV7Z and aV8Z calculations, in most cases the effects of k and l functions were estimated by a short extrapolation, as discussed in Ref. 8. For the sake of completeness, we briefly summarized it here. In the case of the aV7Z basis set, a series of three calculations were performed at the optimal geometry using the aV7Z(no h,i,k), aV7Z(no i,k) and aV7Z(no k) basis sets. The corresponding Hartree-Fock (HF) energies are extrapolated using a 3-parameter exponential function. Typically, the HF/aV7Z(no k) energies are within a few tenths of a microhartree of the full

TABLE II. Carbon atom (^3P) CCSD(T)(FC) energies (E_h).

					Difference		
Basis	UHF[a]	UCCSD(T)[a]	ROHF[b]	R/UCCSD(T)[b]	HF	Correlation	Total
aVDZ	– 37.687763	– 37.764866	– 37.683070	– 37.764798	0.004693	−0.004625	0.000068
aVTZ	– 37.691811	– 37.781825	– 37.686777	– 37.781725	0.005034	−0.004934	0.000100
aVQZ	– 37.693352	– 37.786875	– 37.688256	– 37.786770	0.005096	−0.004991	0.000105
aV5Z	– 37.693690	– 37.788380	– 37.688573	– 37.788267	0.005117	−0.005004	0.000113
aV6Z	– 37.693732	– 37.788869	– 37.688612	– 37.788755	0.005120	−0.005006	0.000114
aV7Z	– 37.693738	– 37.789092	– 37.688618	– 37.788978 (est.)	0.005120	−0.005006[c]	0.000114 (est.)
aV8Z	– 37.693739	– 37.789186	– 37.688619	– 37.789072 (est.)	0.005120	−0.005006[c]	0.000114 (est.)
aV9Z	– 37.693739	– 37.789256	– 37.688619	– 37.789120 (est.)	0.005120	−0.005006[c]	0.000114 (est.)

[a] No symmetry equivalencing of the atomic orbitals.
[b] Symmetry equivalencing imposed on the atomic orbitals.
[c] Based on the correlation energy differences without k (aV7Z), kl (aV8Z) or klm (aV9Z) functions.

TABLE III. Comparison of atomization energies (kcal/mol) obtained from explicit *k*- and *k*,*l*-function frozen core calculations with values obtained from short extrapolations.[a]

	aug-cc-pV7Z			aug-cc-pV8Z		
Molecule	Full basis	Extrap.	Error	Full basis	Extrap.	Error
CH ($^2\Pi$)	83.85	83.84	− 0.01	83.88	83.87	− 0.01
CH_2 ($^3B_1$)	189.77	189.77	0.00	139.82	139.80	− 0.02
CH_2 ($^1A_1$)	180.56	180.56	0.00	180.59	180.60	0.01
C_2 ($^1\Sigma_g{}^+$)	145.03	145.03	0.00	145.13	145.11	− 0.02
CN ($^2\Sigma^+$)	178.89	178.89	0.00	179.03	179.02	− 0.01
C_2H_2 ($^1\Sigma_g{}^+$)	402.48	402.47	− 0.01			
CO ($^1\Sigma^+$)	258.29	258.29	0.00	258.39	258.41	0.02
CO_2 ($^1\Sigma_g{}^+$)	387.83	387.84	0.01	387.87	387.83	− 0.04
CF_2 ($^1A_1$)	257.85	257.85	0.00			
CS ($^1\Sigma^+$)	170.63	170.63	0.00	170.76	170.75	− 0.01
B_2 ($^3\Sigma_g{}^-$)	65.10	65.10	0.00	65.14	65.14	0.00
BH ($^1\Sigma^+$)	84.69	84.69	0.00	84.71	84.70	− 0.01
N_2 ($^1\Sigma_g{}^+$)	226.69	226.71	0.02	226.86	226.87	0.01
NH ($^3\Sigma^-$)	82.74	82.73	− 0.01	82.77	82.75	− 0.02
NH_2 ($^3B_1$)	181.89	181.89	0.00	181.96	181.93	− 0.03
NO ($^2\Pi$)	151.43	151.44	0.01	151.58	151.57	− 0.01
O_2 ($^3\Sigma_g{}^-$)	119.67	119.67	0.00	119.78	119.76	− 0.02
OH ($^2\Pi$)	106.90	106.90	0.00	106.95	106.93	− 0.02
F_2 ($^1\Sigma_g{}^+$)	38.36	38.36	0.00	38.42	38.41	− 0.01
HF ($^1\Sigma^+$)	141.43	141.42	− 0.01	141.48	141.44	− 0.04
Al_2 ($^3\Pi_u$)	33.26	33.27	0.01	33.28	33.29	0.00
AlH ($^1\Sigma^+$)	73.68	73.68	0.00	73.69	73.69	0.00
AlF ($^1\Sigma^+$)	163.30	163.31	0.01	163.39	163.37	− 0.02
Si_2 ($^3\Sigma_g{}^-$)	75.78	75.77	− 0.01	75.85	75.85	0.00
SiH ($^2\Pi$)	73.75	73.75	0.00	73.77	73.76	− 0.01
SiO ($^1\Sigma^+$)	191.47	191.48	0.01	191.62	191.60	− 0.02
SiS ($^1\Sigma^+$)	146.68	146.69	0.01	146.83	146.82	− 0.01
P_2 ($^1\Sigma_g{}^+$)	115.45	115.44	− 0.01	115.61	115.61	0.00
PH ($^3\Sigma^-$)	73.23	73.22	− 0.01	73.25	73.24	− 0.01
PN ($^1\Sigma^+$)	144.74	144.75	0.01	144.91	144.91	0.00
S_2 ($^3\Sigma_g{}^-$)	103.20	103.20	0.00	103.35	103.34	− 0.01
SO ($^3\Sigma^-$)	125.50	125.50	0.00	125.64	125.62	− 0.02
SH ($^2\Pi$)	87.49	87.49	0.00	87.53	87.52	− 0.01
Cl_2 ($^1\Sigma_g{}^+$)	59.59	59.60	0.01	59.68	59.68	0.00
HCl ($^1\Sigma^+$)	107.43	107.43	0.00	107.46	107.45	− 0.01
ClF ($^1\Sigma^+$)	62.55	62.55	0.00	62.63	62.63	0.00
H_2O ($^1A_1$)	232.43	232.43	0.00	232.49	232.47	− 0.02
MAD			<0.01			0.01
Max. abs.			0.02			0.04

[a] The aV7Z extrapolations were based on an unweighted average of the exponential and $\ell_{max}{}^{-3}$ estimates of the missing *k*-function contribution to the correlation energy. The aV8Z extrapolation for the missing k,l-functions used a 2/3:1/3 weighted average (exponential: $\ell_{max}{}^{-3}$).

aV7Z value, so the extrapolation contributes very little. In the next step the correlation energies are extrapolated using an exponential and a $1/\ell_{max}{}^3$ formula to account for the residual contribution from the missing *k*-functions. Experience has shown that the average of the two formulas yields a good approximation to energy *differences* in the full aV7Z basis set. For the aV8Z basis set a 2/3(exp) to 1/3($\ell_{max}{}^3$) weighting is used. Additional tests of the accuracy of this approximation were conducted for this study and the results are shown in Table III. For aV7Z the level of agreement is excellent, with a mean absolute deviation (MAD) of <0.01 kcal/mol and a maximum deviation of 0.02 kcal/mol. With the aV8Z basis set, where the extrapolation is slightly "longer," the agreement is only moderately worse. Most of the molecules in Table III are small. The magnitude of the error could increase as the size of the system and the magnitude of the correlation energy associated with missing *k*- and *l*-functions increases, but it appears likely that the contributions from these very high angular momentum functions can be adequately estimated for our current purpose.

Using energies obtained from the short extrapolation in subsequent CBS extrapolations increases the error somewhat, relative to the raw basis set results, as might be expected. For example, in the case of the diatomic molecule PN ($^1\Sigma^+$) the short aV7Z extrapolation predicts a D_e value that is 0.01 kcal/mol larger than the full, *k*-function inclusive calculation. After CBS extrapolation, the difference increases to 0.03 kcal/mol, an error that remains within the uncertainty (±0.18 kcal/mol) we normally associate with the extrapolated value. In this case the CBS limit was estimated as the average of 4 formulas, including Eq. (1) and the following:

$$E(n) = E_{CBS} + Ae^{-bn}, \tag{2}$$

$$E(n) = E_{CBS} + Ae^{-(n-1)} + Be^{-(n-1)^2}, \tag{3}$$

$$E(\ell_{max}) = E_{CBS} + A/\ell_{max}{}^3. \tag{4}$$

The 3-point exponential formula represented by Eq. (2) was proposed in 1992[66] and the mixed Gaussian/exponential several years later.[67] Eq. (4) was proposed by Helgaker *et al.* in 1997.[60] In the past we have often adopted the spread among the various formulas as a conservative, albeit crude estimate of the uncertainty in the CBS extrapolation. Equations (1)–(4) are free of embedded coefficients derived from empirical fits to defined data sets.

Basis sets as large as the ones used here can cause convergence difficulties in the HF and CCSD iterations or irregularities in the convergence of the atomization energies due to near linear dependency. The smallest eigenvalues of the overlap matrix were observed to reach 10^{-7} to 10^{-8}. To deal with this problem, the 2-electron integral accuracy in MOLPRO was increased from the default 10^{-12} to 10^{-15} (gthresh,twoint = 1.0d-15). In Gaussian 09 the integral accuracy was set to 10^{-13} with the Int = (Acc2E = 13) command. Two other commands were used with Gaussian. The first (nobasistransform) instructed the program to disable the automatic transformation of generally contracted basis sets into segmented contractions. While this transformation can be applied in such a way as to span exactly the same function space as the original basis set, we found that with the aV7Z and aV8Z sets the function spaces were being slightly truncated, leading to energies that were somewhat too high. The small differences in energy were immediately obvious due to the very smooth convergence pattern found with the correlation consistent basis sets. The second command tightened the threshold for the elimination of functions associated with small eigenvalues of the overlap matrix. The default in Gaussian 09 (10^{-6}) was reset to 10^{-7} with IOp(3/59 = 7). Using this strategy, no obvious symptoms of linear dependency were observed.

The current 186-member reference set was based on the slightly smaller collection of molecules used in our earlier investigation of improved basis set extrapolation formulas.[49] It represents a compromise between the desire to cover a wide portion of the Periodic Table and the practical need to limit the size of the systems so that CCSD(T) with very large basis sets remained possible with our current hardware and software. While any finite reference set runs the risk of potentially biasing the conclusions of the study, it is hoped that the present set is adequate for providing statistically significant measures of the performance of CCSD(T)-F12b. The complete list of molecules is provided in the supplementary material[84] along with the corresponding largest basis sets. In all but four cases (C_4H_{10} n-butane, C_5H_{12} n-pentane, C_6H_6 benzene, and C_6H_{14} n-hexane) estimates of the CBS atomization energies were based on at least aV6Z calculations. For this group of four molecules, aV6Z calculations proved to be intractable. Consequently, the reference values were based on CBS(VTQZ-F12) extrapolations. These values were only used to judge the quality of non-extrapolated and CBS(VDTZ-F12) results. With the exception of I_2, the reference atomization energies did not rely on the use pseudopotentials. Because optimal F12 basis sets are currently not available for transition metals, ZnP, ZnO were not included in the F12 comparisons, although they were part of the earlier study.[49]

Another complication arises in the case of third row molecules. The only available F12 basis sets are the main group sets recently developed by Peterson and Hill,[68] which are intended for use with the small core, relativistic pseudopotentials of Metz *et al.*[69] and Peterson *et al.*[38] In the case of Br_2 and As_2 we adjusted our reference all-electron results with scalar relativistic, second order Douglas-Kroll-Hess (DKH) CCSD(T)(FC) corrections[70,71] using DKH recontracted basis sets.[72] These corrections amounted to −0.72 kcal/mol (As_2) and −0.54 kcal/mol (Br_2) and are intended to make the numbers comparable.

Trends in statistical error metrics, such as the mean absolute deviation, the root mean square (RMS), the mean signed deviation (MSD), and the maximum negative error (Max) errors, for standard CCSD(T) atomization energies are shown in Table IV. CBS reference values were based on Eq. (1). All metrics exhibit smooth, albeit slow, convergence, with each increment in the basis set index n reducing the error by approximately a factor of two. The MAD and MSD values differ only in sign because all of the finite basis set results are smaller than the corresponding CBS values. Similar behavior was reported in our previous study[49] with slightly smaller basis sets and a multi-formula CBS average. This suggests that the findings are likely stable with respect to further moderate increases in the number of reference systems and improvement in the basis set quality. Without the use of CBS extrapolation, basis sets of aV7Z to aV8Z quality would be needed in order to achieve guaranteed accuracy within ±1 kcal/mol, the so-called "chemical accuracy."

TABLE IV. Raw standard CCSD(T)(FC) atomization energies errors (kcal/mol).[a]

Basis set	No. of items	ε_{RMS}	ε_{MAD}	Max. Pos.	Max. Neg.
aug-cc-pVTZ	186	9.38	7.87		−27.86
aug-cc-pVQZ	186	3.26	2.77		−9.16
aug-cc-pV5Z	183	1.42	1.22		−3.95
aug-cc-pV6Z	181	0.72	0.62		−2.13
aug-cc-pV7Z	140	0.39	0.34		−0.93
aug-cc-pV8Z	85	0.19	0.16		−0.56

[a] The errors are measured with respect to the best available $1/(\ell_{max} + 1/2)^4$ estimates of the complete basis set limits. Use of the multi-formula average would have resulted in very similar statistics. MAD = mean absolute deviation. RMS = root mean square deviation. MSD = mean signed deviation. Max. = maximum error (negative sign indicates the raw value is smaller than the best estimate of the CBS limit).

All CCSD(T)-F12b calculations were performed with MOLPRO using the cc-pVnZ-F12, (n = D, T, Q) orbital basis sets designed for this method by Peterson *et al.*[73] The diagonal, fixed amplitude *Ansatz*, in which the amplitudes of the explicitly correlated configurations are determined by the wave function cusp conditions[74] was used. CCSD(T)-F12b also requires auxiliary basis sets, where the cc-pVnZ-F12/OptRI sets[75] were used for the CABS resolution of the identity (RI) step.[76] Density fitting of the Fock and exchange matrices was accomplished with the cc-pVnZ/JKFIT basis sets of Weigend[77] while other 2-electron integrals were density fit using aVnZ/MP2Fit sets.[78,79] VDZ-F12 orbital basis set calculations used VTZ/JKFIT fitting basis sets and aVTZ MP2Fit basis sets. This differs from the default in MOLPRO which uses a DZ-quality set. The larger set was chosen for the sake of consistency with our earlier work.[53] The geminal Slater exponent (β) values were taken from the recommendations of Hill *et al.*,[56] namely, 0.9 a_0^{-1} (VDZ-F12), 1.0 a_0^{-1} (VTZ-F12), and 1.0 a_0^{-1} (VQZ-F12). With the exception of the choices of the orbital basis set, the preceding items have no counterpart in standard CCSD(T) theory but they can significantly affect the results. For example, if the VDZ-F12 β is increased from 0.9 to 1.4, a value that has been used in some earlier studies,[73] the atomization energy for C_3H_4 (cyclopropane) decreases by more than 0.7 kcal/mol. The method is orbital invariant, size consistent, and free of geminal basis set superposition error. In the current implementation, only the CCSD step involves explicit correlation. As was the case with the standard CCSD(T) calculations, all F12b atomization energies were evaluated at the optimal F12b geometries.

Complete basis set extrapolations for the F12 method relied on a parameterized two-point, Schwenke-style[80] formula proposed by Hill *et al.*[56] and has the form

$$E^{C}{}_{CBS} = (E_{n+1} - E_n)\, F^{C}{}_{n+1} + E_n, \qquad (5)$$

where the coefficients ($F^{C}{}_{n+1}$) are specific for C = CCSD and (T) components of the CCSD(T) energy and for different basis set pairs, e.g., VDTZ-F12, VTQZ-F12. The Hartree-Fock energy piece of the CCSD(T) energy was not explicitly extrapolated. Instead, the Hartree-Fock CBS limit was estimated by the Hartree-Fock+CABS singles correction in the largest basis set. The coefficients were optimized to reproduce reference energies for the Ne and Ar atoms and 12 small first and second row molecules. Note that atomization energies were *not* considered in defining the extrapolation coefficients.

Based on past experience, we expect CBS(aV56Z) to be the smallest basis set combination capable of achieving the level of accuracy (equivalent to raw aV8Z or better) required for our purpose. The majority of molecules in the reference set were treated with aV7Z or aV8Z basis sets. After CBS extrapolation, these results should be comparable in quality to raw aV9Z or aV10Z atomization energies.

When compiling statistical data for the standard CCSD(T) method, we restricted comparisons to reference values derived from basis sets at least one level higher in quality than the largest basis set appearing in the combination being benchmarked. For example, reference values used to judge the accuracy of a CBS(aVQ56Z) atomization energy would be restricted to either CBS(aV567Z) or CBS(aV678Z) values, assuming the latter was available.

III. RESULTS AND DISCUSSION

While total energies were not the primary focus of this study, the availability of aV8Z basis set results allows us to briefly comment on the accuracy of the F12b energies. In developing the F12/CBS extrapolation formulas, Hill *et al.*[56] created a collection of 14 reference CCSD(T) energies (Ne, Ar, and 12 small molecules) using explicitly correlated CCSD energies obtained with a large, contracted (spdfgh) basis set. The perturbative triples (T) energies, which do not involve explicit correlation, were obtained with the above basis sets augmented with multiple *i*-type functions. They conservatively estimated the uncertainty in the resulting CCSD energies to be 0.2–0.3 mE_h, with the largest uncertainties associated with second row containing molecules. The current best standard method energies were based on: (1) a $1/\ell_{max}{}^3$ extrapolation of the singlet pair energies and the (T) energies and (2) a $1/\ell_{max}{}^5$ extrapolation of the triplet pair energies.[81] Comparing the F12b/CBS and best standard method energies, we find a level of agreement that is generally consistent with the uncertainty estimate of Hill *et al.*[56] For the size of the systems in the reference set of Hill *et al.*[56] there is a slight tendency to overestimate the CCSD component and to underestimate the much smaller (T) component. The largest difference in total CCSD(T) energies between the Hill reference set and the best current estimates is ~0.6 mE_h. This should be viewed as good agreement in light of the sensitivity of the F12 energy to the details of the calculation, such as the value of the geminal Slater exponent. F12b/CBS(VTQZ-F12) energies are very close to the current reference values, as might be expected due to the use of the latter in the training set for the former. For larger systems, where only F12b/CBS(VTQZ-F12) extrapolated values are available, the magnitude of the CCSD-F12b energy tends more strongly to exceed the best estimate from standard CCSD, while the (T) energy component from the F12b calculations continue to underestimate the corresponding CCSD(T) value.

The accuracy of the current reference set of atomization energies is critical to the strength of the conclusions drawn in this study. If the uncertainty in the reference set is not significantly smaller than the uncertainty in the F12b values, the error statistics are cast in doubt. Unfortunately, estimating the uncertainties in extrapolations based on aV7Z and aV8Z basis sets is difficult. Stability in the CBS values can be viewed as *suggestive* of accuracy, but stability alone is insufficient to prove the case. A typical example is shown in Table V where various raw and extrapolated CCSD(T) atomization energies for OCS ($^1\Sigma^+$) are listed in the left-hand columns. The CBS estimates for the standard method are shown to vary by ±0.06 kcal/mol. The spread among CBS(aV678Z) values from Eqs. (1)–(4) is 0.13 kcal/mol, a measure we have used in the past with the multi-formula average, but which experience shows is often overly conservative for large basis sets. The difference between the raw aV(8+*d*)Z value and the CBS(aV678Z) value, which can be viewed as the "length" of the extrapolation is 0.28 kcal/mol. Considering all of the above, we crudely estimate an uncertainty on the order of 0.05 kcal/mol for the CBS(aV678Z) value. The limited number of aV9Z examples lends support to the belief that CBS(aV567Z) and CBS(aV678Z) values are stable to well within 0.1 kcal/mol. In our previous study[49] we estimated the average uncertainty for the reference set to be ~0.04–0.05 kcal/mol. In light of the many increases in basis set size for the current version of the reference set, an *average* uncertainty of ±0.05 kcal/mol seems reasonable.

TABLE V. OCS ($^1\Sigma$) Atomization Energies (kcal/mol).

	CCSD(T)				CCSD(T)-F12b		
Basis	Raw ΣD_e	CBS[a] ΣD_e	Δ[b]	Basis	Raw ΣD_e	CBS[a] ΣD_e	Δ[b]
aV(6+d)Z	333.12	333.97	0.85	VDZ-F12	331.06		
aV(7+d)Z	333.47	334.03	0.56	VTZ-F12	332.79	332.95	0.16
aV(8+d)Z	333.69	333.97	0.28	VQZ-F12	333.75	334.20	0.45
				aV(Q+d)Z	333.91		
				aV(5+d)Z	334.11	334.25	0.14

[a] Based on the $1/(\ell_{max} + 1/2)^4$ formula.
[b] Difference between the raw and extrapolated values.

One of the most prominent features of the CCSD(T)/aV*n*Z sequence of calculations is the uniform way in which property values converge with respect to the 1-particle basis set expansion. We know of no instances where the aV*n*Z sequence of atomization energies fails to display monotonic convergence to the CBS limit. Similar behavior was observed with CCSD(T)-F12b/V*n*Z-F12 calculations for compounds composed of hydrogen and first row elements. However, some molecules involving main group second row elements (Al–Cl) were found to deviate from this pattern. For example, in the case of Cl_2 ($^1\Sigma_g{}^+$) we find D_e values of 59.8 (VDZ-F12), 59.0 (VTZ-F12), and 59.5 (VQZ-F12) kcal/mol. For CCl_4 the corresponding atomization energies were 315.5, 313.7, and 314.8 kcal/mol. The convergence patterns for other properties, such as r_e and ω_e, also failed to display monotonic convergence. Even when second row dissociation energies varied strictly monotonically, such as in the case of SiS ($^1\Sigma^+$), the convergence pattern lacked uniformity.

In part this is likely due to the impact of basis set superposition error (BSSE) with the cc-pVDZ-F12 basis set and the greater sensitivity to the choice of β with this small basis set. While BSSE is present in both standard and explicitly correlated calculations, its relative importance is greater in the latter. If we adopt the Boys and Bernardi[82] counterpoise correction (CP) as our definition of BSSE, we find CP corrections of −1.3 (VDZ-F12), −0.5 (VTZ-F12), and −0.1 (VQZ-F12)

TABLE VI. Raw and CBS extrapolated CCSD(T)-F12b atomization energies errors (kcal/mol).[a]

Basis set	$\varepsilon_{\mathrm{RMS}}$	$\varepsilon_{\mathrm{MAD}}$	$\varepsilon_{\mathrm{MSD}}$	Max. Pos.	Max. Neg.
cc-pVDZ-F12	4.40	3.56	− 3.56		− 14.94
cc-pVTZ-F12	2.05	1.31	− 1.31		− 14.25
cc-pVQZ-F12	0.34	0.28	− 0.28		− 1.26
CBS(VDTZ-F12)	0.86	0.57	− 0.51	0.62	− 3.51
CBS(VTQZ-F12)	0.19	0.15	0.12	0.65	− 0.19

[a] Errors are measured with respect to the reference $1/(\ell_{\max} + 1/2)^4$ estimates of the CCSD(T) complete basis set limits. Use of the multi-formula average would have resulted in very similar statistics. MAD = mean absolute deviation. RMS = root mean square deviation. MSD = mean signed deviation. Max. = maximum error (negative sign indicates the raw value is smaller than the best estimate of the CBS limit). Min. = minimum error.

kcal/mol. These lead to CP-corrected $D_e(Cl_2)$ values of 58.5, 58.5, and 59.3 kcal/mol. Thus, even with a CP correction the dissociation energies still fail to display the typical correlation consistent behavior of the standard method. This phenomenon has consequences for the CBS extrapolation with the smaller basis sets. The CBS(VDTZ-F12) D_e = 58.0 kcal/mol, which is *smaller* than either the raw or the CP-corrected VDZ-F12 values. Fortunately, the VTQZ-F12 pair of basis sets does not suffer from this problem and yields reasonable CBS values. By way of comparison, CP corrections for Cl_2 with the standard method are: −3.4 (VDZ), −1.6 (VTZ), and −0.6 (VQZ) kcal/mol. We find the magnitude of the correction drops by about a factor of 2 for every increment in the basis set index *n*. Thus, by the time aV6Z–aV8Z basis sets are used the CP corrections for the raw binding energies have fallen into the 0.2–0.05 kcal/mol range. BSSE disappears in the limit of a complete basis set.

Table VI summarizes the results of all CCSD(T)-F12b/V*n*Z-F12 calculations, both prior to and after CBS extrapolation. The raw F12b values display an impressive degree of accuracy, with the following rough equivalencies observed between explicitly correlated and standard methods: VDZ-F12 is slightly worse than aVQZ, VTZ-F12 ≈ aV5Z, and VQZ-F12 ≈ aV7Z. A thorough examination of the results reveals that this is due in part to a tendency for F12b to overestimate the reference values in the CBS limit, combined with the universal approach to the limit from below. Although F12b raw atomization energies were not found to actually exceed any reference value, the CBS(VTQZ-F12) entry in Table VI shows an overall positive mean signed deviation. Hidden among the overall statistics are subsets of the molecules where close agreement between CBS(VDTZ-F12) and CBS(VTQZ-F12) estimates was found, such as was the case with hydrocarbons. This is in agreement with our earlier findings.[53]

Figure 2, where the F12b errors are plotted as a function of the magnitude of the valence correlation energy, reveals a trend in which errors for small systems (correlation energies <0.5 E_h) are relatively evenly distributed about zero, but errors for larger systems systematically overshoot the reference values. On the right-hand side of the figure average errors approach 0.5 kcal/mol. To put that into perspective, high accuracy thermochemical studies frequently strive for agreement with well-established experimental values to within 0.1–0.2 kcal/mol. In addition, the size of the systems considered in this study was constrained to be small enough to allow geometry optimizations with at least aV6Z basis sets. Inclusion of larger molecules in the reference set would have exacerbated the mean or RMS differences between standard CCSD(T) and CCSD(T)-F12b. If we had based our reference energies on a multi-formula average, all of the data points in Figure 2 would have been shifted to larger values by 0.02–0.15 kcal/mol. As mentioned previously, the (T) component of the F12b energy is essentially obtained the standard way and the CBS(VTQZ-F12) estimate frequently underestimates the true (T)/CBS limit. This tends to help balance the overestimation of the CCSD energy. In the case of OCS ($^1\Sigma^+$), replacing the CBS(VTQZ-F12) (T) energy component with more accurate CBS(aV78Z) values from standard CCSD(T) estimates increases the atomization energy by a little more than 0.01 kcal/mol.

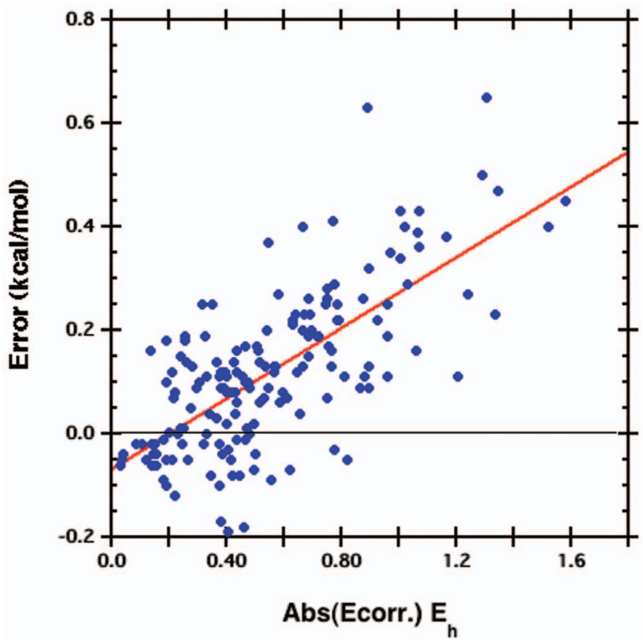

FIG. 2. The error in extrapolated CCSD(T)-F12b/CBS(VTQZ) atomization energies as a function of the valence correlation energy obtained from the largest basis set, standard CCSD(T) calculation. The reference values were based on the $1/(\ell_{\max} + 1/2)^4$ formula. If a multi-formula average were to have been used for the reference values, the data points would be shifted up by 0.02–0.15 kcal/mol. The red line represents a linear least squares fit.

In order to determine if the systematic overestimation of atomization energies was due to a problem with the F12 extrapolation procedure, explicitly correlated calculations were performed on OCS ($^1\Sigma$) with the aug-cc-pV*n*Z/aug-cc-pV(*n*+d)Z, *n* = Q,5 basis sets. Geminal Slater exponents of 1.4 a_0^{-1} (aVQZ) and 1.5 a_0^{-1} (aV5Z) were used, as recommended by Hill *et al.*[56] These were combined with VQZ/JKFIT and V5Z/JKFIT fitting basis sets, AVQZ and AV5Z MP2Fit basis sets and aVQZ+d and aV5Z+d OptRI auxiliary basis sets.[83] The raw and extrapolated results are shown in the right-hand portion of Table V, along with the corresponding V*n*Z-F12 values. Since the CCSD(T)-F12b/CBS(VTQZ-F12) and

CCSD(T)-F12b/CBS(aVQ5Z) atomization energies differ by only 0.05 kcal/mol and both are 0.23–0.28 kcal/mol larger than the reference CCSD(T)/CBS(aV678Z) value, we conclude that the difference between standard CCSD(T) and F12b results are likely inherent to the F12b method and not an artifact of the CBS extrapolation. If anything, this case suggests that the CBS(VTQZ-F12) estimates may be slightly smaller than the true F12b basis set limits.

It has been noted that CCSD(T)-F12a yields raw and CBS extrapolated correlation energies and atomization energies that are larger than the best standard method values, sometimes even with basis sets as small as VTZ-F12.[53] In this study we find evidence for similar, albeit much less severe, bias in the F12b method for which only the CBS limit values overshoot the reference values. With CCSD(T)-F12b this behavior required a detailed comparison with a high quality reference set. The overshooting of the apparent CCSD(T)/CBS limit appears to extend to bond lengths and harmonic frequencies, as well. In the case of N_2, where aV8Z and aV9Z results are available, the raw VQZ-F12 values already appear too short for the bond length and too large for the harmonic frequency.

The CBS(V*nm*Z-F12) mean signed deviations ($\varepsilon_{\mathrm{MSD}}$) in Table VI display a pronounced swing from underestimating the reference values by more than −0.5 kcal/mol to overestimating the same quantities by more than 0.1 kcal/mol. The limited data on F12b extrapolations with larger basis sets suggest that the VTQZ-F12 combination may still be ~0.1 kcal/mol smaller than the true F12b basis set limit for $\varepsilon_{\mathrm{MSD}}$.

Table VII contains the error statistics for CBS extrapolated atomization energies obtained from standard CCSD(T) calculations and a variable exponent Schwenke-style formula, described in our earlier report as[49]

$$E^{C}{}_{\mathrm{CBS}} = (E_n - E_{n+1})^{\alpha}\, D^{C}{}_{n+1} + (E_n - E_{n+1})^{\beta}\, F^{C}{}_{n+1} + E_n, \tag{6}$$

where C = SCF, CCSD, and (T) correlation components was used. The nine adjustable parameters were re-optimized for the current version of the reference set. While the reference values used for generating the results in Table VII were obtained from the $1/(\ell_{\max} + 1/2)^4$ formula (Eq. (1)), very similar results were found using the multi-formula average. Of course, in that case the fitting extrapolation parameters are necessarily slightly different.

Error statistics can be compared among Tables IV, VI, and VII. Extrapolating the VDTZ-F12 basis set combination decreases the RMS deviation relative to the raw VTZ-F12 value by roughly a factor of two. The CBS(VDTZ-F12) errors fall midway between the raw aV5Z and aV6Z values, but this method requires just a fraction of the computer time. Even greater improvements would have been possible were it not for the problems associated with second row elements, as already discussed. Overall, CBS(VDTZ-F12) is significantly biased towards underestimating the reference atomization energies. Comparable CBS(aVDTZ) standard method statistics are not available because the only available extrapolation formula is the one suggested by Schwenke[80] and in our experience such small basis sets do not yield results suitable for high accuracy studies.

TABLE VII. RMS atomization deviations (kcal/mol) for the variable exponent Schwenke-style extrapolation formula.[a]

Basis set combinations	$\varepsilon_{\mathrm{RMS}}$	$\varepsilon_{\mathrm{MAD}}$	$\varepsilon_{\mathrm{MSD}}$	Max. Pos.	Max. Neg.
CBS(aVTQZ)[b]	0.25	0.17	− 0.01	0.77	−1.30
CBS(aVQ5Z)[c]	0.10	0.07	0.01	0.28	−0.54
CBS(aV56Z)[d]	0.06	0.05	− 0.01	0.18	−0.18
CBS(aV67Z)[e]	0.07	0.05	− 0.01	0.21	−0.21

[a] Number of comparisons: 186 aVTQZ, 182 aVQ5Z, 138 aV56Z, 86 aV67Z. Errors are measured with respect to the reference $1/(\ell_{\max} + 1/2)^4$ estimates of the CCSD(T) complete basis set limits.

[b] Optimized aVTQZ parameters: SCF = 1.662627, D^{CCSD} = 1.823422, $D^{(\mathrm{T})}$ = 4.726121, F^{CCSD} = −0.731498, $F^{(\mathrm{T})}$ = 1.470283, α^{CCSD} = 1.397524, $\alpha^{(\mathrm{T})}$ = 1.199429, β^{CCSD} = 1.034092, $\beta^{(\mathrm{T})}$ = 1.199931. There was a single instance of an error exceeding 1 kcal/mol in magnitude.

[c] Optimized aVQ5Z parameters: SCF = 1.587158, D^{CCSD} = 3.119781, $D^{(\mathrm{T})}$ = 0.647576, F^{CCSD} = −2.396420, $F^{(\mathrm{T})}$ = −1.527814, α^{CCSD} = 1.558135, $\alpha^{(\mathrm{T})}$ = 0.997786, β^{CCSD} = 1.098693, $\beta^{(\mathrm{T})}$ = 0.719363.

[d] Optimized aV56Z parameters: SCF = 1.683266, D^{CCSD} = 2.726857, $D^{(\mathrm{T})}$ = 1.486013, F^{CCSD} = −5.561467, $F^{(\mathrm{T})}$ = −0.043761, α^{CCSD} = 1.747614, $\alpha^{(\mathrm{T})}$ = 1.490661, β^{CCSD} = 1.040605, $\beta^{(\mathrm{T})}$ = 1.045935.

[e] Optimized aV67Z parameters: SCF = 2.544380, D^{CCSD} = 4.838980, $D^{(\mathrm{T})}$ = 1.130880, F^{CCSD} = −7.405482, $F^{(\mathrm{T})}$ = −0.376773, α^{CCSD} = 1.415931, $\alpha^{(\mathrm{T})}$ = 2.111643, β^{CCSD} = 1.090784, $\beta^{(\mathrm{T})}$ = 2.658115.

The CBS(VTQZ-F12) method, which does not suffer from the same problems as CBS(VDTZ-F12) with second row compounds, produces results that are comparable to the standard CCSD(T)/CBS(aVTQZ) values when judged by RMS and MAD values. The maximum positive and negative errors are smaller. However, the mean signed deviation from the explicitly correlated approach (0.12 kcal/mol) is much larger than with the corresponding standard method (−0.01 kcal/mol). In systems where the valence correlation energy does not exceed 0.5 E_h, accuracy to better than ±0.2 kcal/mol is not required or for specific classes of molecules (e.g., hydrocarbons), this is an attractive alternative to standard methods. If CBS extrapolation isn't desired, the raw CCSD(T)-F12b/VQZ-F12 method provides a very cost effective alternative to achieving results comparable to much larger standard method calculations.

IV. CONCLUSION

Statistical analysis comparing complete basis set atomization energies derived from CCSD(T)-F12b and its counterpart standard CCSD(T) reveals a tendency for the explicitly correlated approach to overestimate the latter in cases where the valence correlation energy exceeds ~0.5 E_h. Limited data suggest that the use of basis sets larger than VQZ-F12 in the CBS extrapolation will exacerbate the problem. Extrapolations based on the cost effective VDZ-F12/VTZ-F12 basis set combination produced a RMS error of 0.9 kcal/mol, just within the margins of chemical accuracy (±1 kcal/mol), but with a relatively large maximum error of −3.5 kcal/mol. The more computationally expensive VTZ-F12/VQZ-F12 combination agreed much better with the reference values, achieving a RMS error of 0.2 kcal/mol, slightly smaller than the standard aVDTQZ value. However, the VTZ-F12/VQZ-F12 mean signed deviation was significantly larger than with the standard method. If the current reference energies were replaced by values based on a multi-formula average or if the reference set were more heavily weighted towards larger systems, the CBS(VTZ-F12/VQZ/F12) mean signed deviation would increase. Systematic errors observed with F12b serve as a cautionary note for use of this method in high accuracy studies. Due to the complexities inherent in the F12b method, it is difficult to identify the cause of the tendency to overestimate reference values obtained from standard CCSD(T). F12b is still a relatively new technique and improvements may be introduced to address this issue.

CBS reference sets are always subject to improvement, whether it be in terms of the largest basis set applied to a given molecule or the number and types of molecules in the collection. In acknowledgement of this reality, we feel that the error metrics reported here should be viewed as a circumscribed benchmark of the performance of CCSD(T)-F12b. However, the current accuracy is likely sufficient to insure the overall validity of the conclusions reached here.

[1] G. D. Purvis III and R. J. Bartlett, J. Chem. Phys. **76**, 1910–1918 (1982).
[2] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, Chem. Phys. Lett. **157**, 479 (1989).
[3] J. D. Watts, J. Gauss, and R. J. Bartlett, J. Chem. Phys. **98**, 8718–8733 (1993).
[4] R. J. Bartlett and M. Musiał, Rev. Mod. Phys. **79**, 291–352 (2007).
[5] D. Feller and K. A. Peterson, J. Chem. Phys. **126**, 114105 (2007).
[6] D. Feller, K. A. Peterson, and D. A. Dixon, J. Chem. Phys. **129**, 204105-1–204105-32 (2008).
[7] K. A. Peterson, D. Feller, and D. A. Dixon, Theor. Chem. Acc. **131**, 1079 (2012).
[8] D. Feller, K. A. Peterson, and T. D. Crawford, J. Chem. Phys. **124**, 054107 (2006).
[9] D. A. Dixon, D. Feller, and K. A. Peterson, in *Annual Report in Computer Chemistry*, edited by R. A. Wheeler (Elsevier, 2012), Vol. 8.
[10] D. Feller, K. A. Peterson, and D. A. Dixon, Mol. Phys. **110**, 2381 (2012).
[11] J. A. Pople, M. Head-Gordon, D. J. Fox, K. Raghavachari, and L. A. Curtiss, J. Chem. Phys. **90**, 5622 (1989).
[12] L. A. Curtiss, K. Raghavachari, G. W. Trucks, and J. A. Pople, J. Chem. Phys. **94**, 7221 (1991).
[13] L. A. Curtiss, P. C. Redfern, and K. Raghavachari, J. Chem. Phys. **126**, 084108 (2007).
[14] J. M. L. Martin and G. d. Oliveira, J. Chem. Phys. **111**, 1843 (1999).
[15] A. Karton, E. Rabinovich, J. M. L. Martin, and B. Ruscic, J. Chem. Phys. **125**, 144108 (2006).
[16] 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).
[17] M. E. Harding, J. Vázquez, B. Ruscic, A. K. Wilson, J. Gauss, and J. F. Stanton, J. Chem. Phys. **128**, 114111 (2008).
[18] N. J. DeYonker, T. R. Cundari, and A. K. Wilson, J. Chem. Phys. **124**, 114104 (2006).
[19] N. J. DeYonker, T. Grimes, S. Yokel, A. Dinescu, B. Mintz, T. R. Cundari, and A. K. Wilson, J. Chem. Phys. **125**, 104111 (2006).
[20] N. J. DeYonker, B. R. Wilson, A. W. Pierpont, T. R. Cundari, and A. K. Wilson, Mol. Phys. **107**, 1107 (2009).
[21] A. L. L. East and W. D. Allen, J. Chem. Phys. **99**, 4638 (1993).
[22] A. L. L. East, W. D. Allen, and A. G. Császár, in *Structure and Conformations of Non-Rigid Molecules*, edited by J. Laane, B. v. d. Veken, and H. Overhammer (Kluwer, Dordrecht, 1993), pp. 343–373.
[23] A. G. Császár, W. D. Allen, and H. F. Schaefer III, J. Chem. Phys. **108**, 9751 (1998).
[24] G. A. Petersson, A. Bennett, T. G. Tensfeldt, M. A. Al-Laham, W. A. Shirley, and J. Mantzaris, J. Chem. Phys. **89**, 2193 (1988).
[25] G. A. Petersson, T. G. Tensfeldt, and J. A. Montgomery, Jr., J. Chem. Phys. **94**, 6091 (1991).
[26] J. W. Ochterski, G. A. Petersson and J. A. Montgomery, Jr., J. Chem. Phys. **104**, 2598 (1996).
[27] T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007–1023 (1989).
[28] R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, J. Chem. Phys. **96**, 6796–6806 (1992).
[29] D. E. Woon and T. H. Dunning, Jr., J. Chem. Phys. **100**, 2975–2988 (1994).
[30] D. E. Woon and T. H. Dunning, Jr., J. Chem. Phys. **103**, 4572–4585 (1995).
[31] K. A. Peterson and T. H. Dunning, Jr., J. Chem. Phys. **117**, 10548–10560 (2002).
[32] J. Koput and K. A. Peterson, J. Phys. Chem. A **106**, 9595–9599 (2002).
[33] D. E. Woon and T. H. Dunning, Jr., J. Chem. Phys. **98**, 1358–1371 (1993).
[34] A. K. Wilson, D. E. Woon, K. A. Peterson, and T. H. Dunning, Jr., J. Chem. Phys. **110**, 7667–7676 (1999).
[35] N. B. Balabanov and K. A. Peterson, J. Chem. Phys. **123**, 064107-1–064107-15 (2005).
[36] A. K. Wilson, T. van Mourik, and T. H. Dunning, Jr., J. Mol. Struct.: THEOCHEM **388**, 339 (1996).
[37] D. Feller and K. A. Peterson, J. Chem. Phys. **110**, 8384 (1999).
[38] K. A. Peterson, D. Figgen, E. Goll, H. Stoll, and M. Dolg, J. Chem. Phys. **119**, 11113 (2003).
[39] K. A. Peterson, J. Chem. Phys. **119**, 11099 (2003).
[40] K. A. Peterson and C. Puzzarini, Theor. Chem. Acc. **114**, 283 (2005).
[41] E. C. Barnes, G. A. Petersson, D. Feller, and K. A. Peterson, J. Chem. Phys. **129**, 194115 (2008).
[42] D. Feller, K. A. Peterson, and J. G. Hill, J. Chem. Phys. **135**, 044102 (2011).
[43] J. M. L. Martin, Chem. Phys. Lett. **259**, 669 (1996).
[44] C. Schwartz, Phys. Rev. **126**, 1015 (1962).
[45] C. Schwartz, in *Methods in Computational Physics*, edited by B. J. Alder, S. Fernbach, and M. Rotenberg (Academic Press, New York, 1963), Vol. 2, pp. 262–265.
[46] D. P. Carroll, H. J. Silverstone, and R. M. Metzger, J. Chem. Phys. **71**, 4142 (1979).
[47] R. N. Hill, J. Chem. Phys. **83**, 1173 (1985).
[48] W. Kutzelnigg and J. D. Morgan, J. Chem. Phys. **96**, 4484 (1992).
[49] D. Feller, J. Chem. Phys. **138**, 074103 (2013).
[50] T. B. Adler, G. Knizia, and H.-J. Werner, J. Chem. Phys. **127**, 221106 (2007).
[51] G. Knizia, T. B. Adler, and H.-J. Werner, J. Chem. Phys. **130**, 054104 (2009).
[52] G. Rauhut, G. Knizia, and H.-J. Werner, J. Chem. Phys. **130**, 054105 (2009).
[53] D. Feller, K. A. Peterson, and J. G. Hill, J. Chem. Phys. **133**, 184102 (2010).
[54] H.-J. Werner, T. B. Adler, and F. R. Manby, J. Chem. Phys. **126**, 164102 (2007).
[55] G. Knizia and H.-J. Werner, J. Chem. Phys. **128**, 154103 (2008).
[56] J. G. Hill, K. A. Peterson, G. Knizia, and H.-J. Werner, J. Chem. Phys. **131**, 194105 (2009).
[57] D. Feller, J. Comput. Chem. **17**, 1571–1586 (1996).
[58] L. Wolniewicz, J. Chem. Phys. **103**, 1792 (1995).
[59] W. Cencek and W. Kutzelnigg, J. Chem. Phys. **105**, 5878 (1996).
[60] T. Helgaker, W. Klopper, H. Koch, and J. Noga, J. Chem. Phys. **106**, 9639 (1997).
[61] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz *et al.*, MOLPRO version 2012.1, a package of *ab initio* programs, 2012, see http://www.molpro.net.
[62] C. Angeli, K. L. Bak, V. Bakken, O. Christiansen, R. Cimiraglia, S. Coriani, P. Dahle, E. K. Dalskov, T. Enevoldsen, B. Fernandez, C. Hättig, K. Hald, A. Halkier, H. Heiberg, T. Helgaker, H. Hettema, H. J. A. Jensen, D. Jonsson, P. Jørgensen, S. Kirpekar, W. Klopper, R. Kobayashi, H. Koch, A. Ligabue, O. B. Lutnaes, K. V. Mikkelsen, P. Norman, J. Olsen, M. J. Packer, T. B. Pedersen, E. R. Z. Rinkevicius, T. A. Ruden, K. Ruud, P. Sałek, A. S. d. Meras, T. Saue, S. P. A. Sauer, B. Schimmelpfennig, K. O. Sylvester-Hvid, P. R. Taylor, O. Vahtras, D. J. Wilson, and H. Ågren, Dalton 2.0, 2005.
[63] M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, Gaussian 09, Gaussian, Inc., Wallingford, CT, 2013.
[64] P. J. Knowles, C. Hampel, and H. J. Werner, J. Chem. Phys. **99**, 5219 (1993).
[65] M. J. O. Deegan and P. J. Knowles, Chem. Phys. Lett. **227**, 321 (1994).
[66] D. Feller, J. Chem. Phys. **96**, 6104 (1992).
[67] K. A. Peterson, D. E. Woon, and T. H. Dunning, Jr., J. Chem. Phys. **100**, 7410 (1994).

[68] K. A. Peterson and J. G. Hill, (unpublished).

[69] B. Metz, H. Stoll, and M. Dolg, J. Chem. Phys. **113**, 2563 (2000).

[70] M. Douglas and N. M. Kroll, Ann. Phys. (N.Y.) **82**, 89 (1974).

[71] G. Jansen and B. A. Hess, Phys. Rev. A **39**, 6016–6017 (1989).

[72] W. A. de Jong, R. J. Harrison, and D. A. Dixon, J. Chem. Phys. **114**, 48–53 (2001).

[73] K. A. Peterson, T. B. Adler, and H.-J. Werner, J. Chem. Phys. **128**, 084102 (2008).

[74] S. Ten-no, Chem. Phys. Lett. **398**, 56 (2004).

[75] K. E. Yousaf and K. A. Peterson, J. Chem. Phys. **129**, 184108 (2008).

[76] E. F. Valeev, Chem. Phys. Lett. **395**, 190 (2004).

[77] F. Weigend, Phys. Chem. Chem. Phys. **4**, 4285 (2002).

[78] C. Hättig, Chem. Phys. Phys. Chem. **7**, 59 (2005).

[79] F. Weigend, A. Köhn, and C. Hättig, J. Chem. Phys. **116**, 3175 (2002).

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

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

[82] S. F. Boys and F. Bernardi, Mol. Phys. **19**, 553 (1970).

[83] K. E. Yousaf and K. A. Peterson, Chem. Phys. Lett. **476**, 303 (2009).

[84] See supplementary material at http://dx.doi.org/10.1063/1.4819125 for the list of the reference molecules and the largest associated basis sets, as well as the newly developed basis sets and information on the CBS estimates for each molecule.