

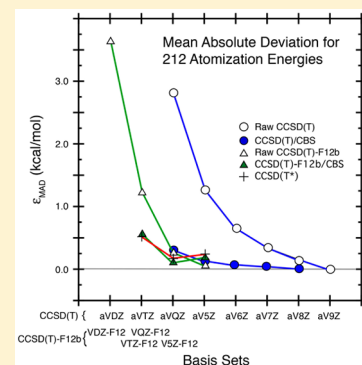
Statistical Electronic Structure Calibration Study of the CCSD(T*)-F12b Method for Atomization Energies

David Feller*

Department of Chemistry, Washington State University, Pullman, Washington 99164-4630, United States

S Supporting Information

ABSTRACT: In the explicitly correlated CCSD(T)-F12b coupled cluster method only the singles and doubles component of the energy benefits from inclusion of terms involving the interelectronic distance. Consequently, only that component exhibits accelerated convergence with respect to the 1-particle basis set. The smaller perturbative triples component converges at the same rate as the corresponding piece in standard CCSD(T). With the alternative CCSD(T*)-F12b method the triples correlation energy is scaled up by the ratio of explicitly correlated to standard second-order perturbation theory correlation energies in an attempt to better approximate the basis set limit. An extensive and diverse 212 molecule collection of reference total atomization energies, developed with large basis sets (up to aug-cc-pV9Z in some cases) and standard CCSD(T), was used to calibrate the performance of CCSD(T*). Scaling of the (T) energy led to improved results relative to raw F12b values but only provided a statistical advantage over previously proposed complete basis set extrapolation techniques for the smallest basis sets. With larger sets, scaling (T) produced noticeably poorer results, sometimes by a factor of 2. In agreement with earlier studies, basis set extrapolated CCSD(T)-F12b was found to exhibit a systematic bias toward overestimating reference atomization energies with an error that increases with the magnitude of the valence correlation energy.



I. INTRODUCTION

Standard, correlated electronic structure methods, such as configuration interaction and coupled cluster theory, are plagued by painfully slow convergence to the 1-particle complete basis set (CBS) limit. Expanding the solution of the many-electron Schrödinger equation in a finite, (typically) atom-centered collection of Gaussian functions inevitably leads to basis set truncation errors that are difficult to quantify *a priori* and may be sufficiently large so as to render the results of questionable usefulness. In addition, calculations with very extended basis sets require long run times (possibly days to weeks) or they may exceed the available computer resources (central processing unit cycles, memory or nonvolatile storage) and simply be intractable.

The complementary n -particle error, measured with respect to the exact solution of the Schrödinger equation, has been referred to as the “intrinsic error” of the method by Woon and Dunning.¹ Full configuration interaction (FCI) wave functions provide one method for determining the n -particle limit for any specified 1-particle basis set, but the computational cost of FCI is so exorbitant that as a practical matter the technique is severely limited. For many purposes, coupled cluster theory through iterative single and double excitations plus a quasiperturbative treatment of “connected” triple excitations, a method labeled CCSD(T),^{2–5} provides a good alternative to FCI because of its reduced (but by no means trivial) computational requirements. High accuracy studies that rely on CCSD(T) as a *starting point* for the n -particle expansion encounter a wide range of 1-particle basis set requirements that

vary with the type of molecular properties being sought. In the most challenging cases, large basis sets and lengthy computer runs remain a necessity for achieving the desired accuracy, at least compared to the requirements for more approximate methods such as density functional theory. Hopes of simultaneously extending electronic structure methods to larger molecules while continuing to achieve high accuracy has spurred the development of approaches for dealing with the 1-particle problem. One such method for accomplishing this objective will be the focus of the current work.

Shortly after the introduction of the correlation consistent basis sets of Dunning, Peterson, and co-workers^{6–20} researchers began exploiting the uniform convergence properties of the new basis sets in an effort to accelerate convergence and lessen the need for very expensive calculations. A variety of simple extrapolation formulas were proposed for estimating the CBS limit.^{21–26} The correlation consistent basis set sequence is conventionally denoted cc-pVnZ, where the basis set index n assumes values of D(2), T(3), Q(4), 5, ..., 9 (or aug-cc-pVnZ for the diffuse function augmented sequence). Throughout this paper we will abbreviate the basis set labels to aVnZ, where it is understood that for second row elements the corresponding

Special Issue: 100 Years of Combustion Kinetics at Argonne: A Festschrift for Lawrence B. Harding, Joe V. Michael, and Albert F. Wagner

Received: January 16, 2015

Revised: March 1, 2015

aug-cc-pV(n+d)Z sets, which include important tight d functions, were used.

When judging the degree of 1-particle convergence associated with any particular basis set or multibasis set extrapolation technique, one might assume that either the total energy or the correlation component of the energy would represent the natural quantity to be monitored. However, the total energy of an isolated chemical system is rarely of chemical interest. A wide range of atomic and molecular properties provide alternative choices for use as a metric. As noted above, because of the variable rate at which different properties converge, some are more sensitive probes of the 1-particle expansion than others. Thermochemical properties, including the total atomization energy, ΣD_e , or heat of formation, are well-known to be very slowly convergent. Consequently, they pose one of the most demanding challenges confronting electronic structure methods. For this reason they will be used to gauge the 1-particle convergence in the present investigation.

Statistical analyses of atomization energy errors, defined as the difference between $aVnZ$ values and a collection of the best available reference data, indicated that raw standard CCSD(T) errors, i.e., those obtained without help from CBS extrapolations, decrease by approximately a factor of 2 for every increment in n .²⁷ In our 2013 study the error metrics covered 186 molecules and included: (1) root-mean-square (RMS) and (2) mean absolute deviation (MAD), mean signed deviation (MSD) and the maximum absolute deviation. The analysis relied on data contained in the Computational Results Database (CRDB).²⁸ The current version of the database holds information on 464 chemical systems spanning the first five periods of the periodic table and stored in more than 128 000 entries. Here an “entry” might correspond to an optimized structure or a set of normal-mode frequencies. In this study all calculations were performed using the frozen core (FC) approximation, in which the inner shell electrons are excluded from the correlation treatment. The CRDB also contains core/valence results. The reference set consisted of very large basis set CCSD(T)(FC)/CBS extrapolated values obtained with an expression involving the inverse of the maximum angular momentum present in the basis set (I_{\max}):

$$E(I_{\max}) = E_{\text{CBS}} + A/(I_{\max} + 1/2)^4 \quad (1)$$

which was proposed by Martin²³ who built upon earlier work by Schwartz,^{29,30} Carroll,³¹ Hill,³² and Kutzelnigg and Morgan.³³ This formula produces results that are numerically similar to those obtained from a five formula average, but it is conceptually simpler. In the context of the present study, use of the five formula average for defining the reference atomization energies would have decreased their values by a relatively small 0.01–0.12 kcal/mol, depending on the quality of the basis set and the size of the molecule.

As with the previous study,²⁷ the majority of the current reference ΣD_e values will be based on CCSD(T)(FC) performed with very large basis sets and subsequently extrapolated to the CBS limit. Calculations at this level of theory represent just one component, albeit the largest, in a typical composite theoretical approach to thermochemistry. Experience has shown that close agreement with *experiment* can require as many as 12 components.^{34–37} Some degree of cancellation of error among the various components is likely, but the extent varies widely from one system to another. If high uniform accuracy across many different types of bonding is

desired, the component errors must be carefully monitored. Specifically, for thermochemical studies hoping to achieve an *overall* target accuracy on the order of 0.2–0.4 kcal/mol the error in the CCSD(T) component should not exceed 0.1–0.2 kcal/mol unless cancellation of error is overtly relied on for reaching the target accuracy. Note that this is much tighter than so-called “chemical accuracy” at ± 1 kcal/mol. Methods such as the Feller–Peterson–Dixon (FPD) procedure,^{34–37} Martin’s W4 method³⁸ and HEAT345-(Q)³⁹ have all been reported capable of predicting atomization energies in agreement with accurate experimental/Active Thermochemical Table (ATcT)^{40,41} values to better than 0.1 kcal/mol for small-to-medium size molecules. For our purposes, a medium size system would be something on the order of benzene or hexane. Larger systems can be tackled, but the uncertainties will rapidly grow. Consequently, although errors of several tenths of a kcal/mol may be insignificant for many investigations, in the present study they are judged to be important.

Previous studies indicated that convergence to within ± 1 kcal/mol in the raw ΣD_e values required basis sets of at least $aV7Z$ or $aV8Z$ quality, which is only practical for very small systems. Several extrapolation formulas were found to be capable of a level of accuracy equivalent to increasing the basis set index from n to $n + 2$.⁴² By optimizing extrapolation parameters strictly for the purpose of reproducing atomization energies, it proved possible to achieve even better results (equivalent to $n + 3$), offering enormous savings in computational effort.⁴³ In general, basis sets of at least VQZ (or $aVQZ$) quality were required for achieving ± 1 kcal/mol accuracy when used with extrapolation formulas. Although Truhlar and co-workers have advocated extrapolations with the smaller VDZ/VTZ basis set combination,^{44,45} several authors have criticized this choice as being unreliable.^{42,46} For example, in the case of n -propane and n -butane we found VDZ/VTZ errors in the atomization energies exceeding 11 and 14 kcal/mol, respectively. Truhlar’s recommendation was based on a very small collection of molecules.

An alternative approach to the use of *ad hoc* extrapolation formulas relies on the introduction of nonlinear terms in r_{ij} , the interelectronic distance into the theoretical treatment. These so-called “explicitly correlated” methods include the recently developed CCSD(T)-F12a/b/c techniques^{47–52} as implemented in the MOLPRO suite of programs.⁵³ In a pair of studies attempting to calibrate F12x comparison was made with standard CCSD(T)(FC).^{27,54} Calculations were performed with the cc-pVnZ-F12 orbital basis sets of Peterson et al., which were specifically designed for F12a/b.^{55,56} Following the abbreviation convention adopted for the original correlation consistent basis sets, these F12 basis sets will be referred to in the text as VnZ-F12. Among the conclusions reached in the earlier studies was that F12a was found to frequently overestimate both the correlation energy and ΣD_e values, even with the relatively small VTZ-F12 basis set and without extrapolation. Werner and co-workers have rationalized this in terms of a missing term in the Lagrangian.^{47,50} Due to fortuitous cancellation of error, F12a with VDZ-F12 or VTZ-F12 basis sets often produced better agreement with the reference values than the corresponding F12b results. Measured across all basis sets, the CCSD(T)-F12b method more closely matched standard CCSD(T), while displaying a considerably smaller tendency to overestimate the CCSD(T)/CBS limit.

Even with the relatively large VQZ-F12 basis set, raw F12b results often do not meet the accuracy requirements for some

Table 1. List of Molecules and Corresponding Largest Basis Sets^a

H ₂ (¹ Σ _g ⁺) aV9Z	OFO (² B ₂) aV8Z	CS ₂ (¹ Σ _g ⁺) aV8Z	FCN (¹ Σ ⁺) aV7Z	CF ₃ (² A ₁) aV6Z	LiCl (¹ Σ ⁺) aV6Z	LiF (¹ Σ ⁺) aV6Z	C ₃ H ₄ allene aV7Z
CH (² Π) aV9Z	CCl (² Π) aV8Z	S ₂ (³ Σ _g ⁻) aV8Z	FNC (¹ Σ ⁺) aV7Z	CF ₄ (¹ A ₁) aV6Z	LiH (¹ Σ ⁺) aV6Z	LiN (³ Σ ⁻) aV6Z	LiO (² Π) aV6Z
C ₂ (¹ Σ _g ⁺) aV9Z	CINO (¹ A') aV8Z	Cl ₂ O (¹ A ₁) aV7Z	FS (² Π) aV8Z	CH ₃ OH aV6Z	HNCO (¹ A') aV7Z	HCNO (¹ A') aV7Z	HOCN (¹ A') aV7Z
O ₂ (³ Σ _g ⁻) aV9Z	PS (² Π) aV8Z	SO (³ Σ ⁻) aV8Z	ClF (¹ Σ ⁺) aV8Z	CH ₃ F (¹ A ₁) aV7Z	H ₂ CS (¹ A ₁) aV7Z	AlP (³ Σ ⁻) aV8Z	H ₂ S ₂ (¹ A') aV7Z
H ₂ O (¹ A ₁) aV9Z	B ₂ (³ Σ _g ⁻) aV8Z	SO ₂ (¹ A ₁) aV7Z	HF (¹ Σ ⁺) aV9Z	NS (² Π) aV8Z	ZnP (⁴ Σ ⁻) aV6Z	ZnO (¹ Σ ⁺) aV6Z	CH ₃ NH ₂ (¹ A') aV6Z
CH ₂ (³ B ₁) aV8Z	BH ₃ (¹ A ₁) aV7Z	SO ₃ (¹ A ₁) aV7Z	HS (² Π) aV8Z	C ₂ H ₆ (¹ A _g) aV7Z	N ₂ H ₂ trans aV7Z	N ₂ H ₂ cis aV7Z	N ₂ H ₂ iso aV7Z
CH ₃ (¹ A ₁) aV9Z	CH ₃ (² A ₂ ^{''}) aV8Z	P ₂ (¹ Σ _g ⁺) aV8Z	Br ₂ (¹ Σ _g ⁺) aV7Z	N ₂ H ₄ (¹ A) aV7Z	NH ₂ OH (¹ A') aV7Z	SiF ₄ (¹ A ₁) aV7Z	NCO (² Π) aV7Z
CO (¹ Σ ⁺) aV8Z	CH ₄ (¹ A ₁) aV7Z	PN (¹ Σ ⁺) aV8Z	HBr (¹ Σ ⁺) aV7Z	C ₂ HF (¹ Σ) aV7Z	As ₂ (¹ Σ _g ⁺) aV6Z	AsH (³ Σ ⁻) aV6Z	AsH ₂ (² B ₁) aV6Z
C ₂ H ₂ (¹ Σ _g ⁺) aV8Z	NH (³ Σ ⁻) aV9Z	PO (² Π) aV7Z	CBr (² Π) aV6Z	AsH ₃ (¹ A ₁) aV6Z	AsF (³ Σ ⁻) aV6Z	AsF ₂ (² B ₁) aV6Z	AsF ₃ (¹ A ₁) aV6Z
N ₂ (¹ Σ _g ⁺) aV9Z	NH ₂ (² B ₁) aV8Z	PH (³ Σ ⁻) aV7Z	BrO (² Π) aV6Z	AsN (¹ Σ _g ⁺) aV6Z	PF ₂ (² B ₁) aV7Z	CH ₃ CHO acetaldehyde aV6Z	HONO trans aV7Z
CN (² Σ ⁺) aV8Z	NH ₃ (¹ A ₁) aV8Z	PH ₂ (² B ₁) aV7Z	HO ₂ (² A ^{''}) aV8Z	HONO cis aV7Z	CHF ₃ (¹ A ₁) aV6Z	N ₂ O ₂ linear aV7Z	N ₂ O ₂ nonlinear aV7Z
CO ₂ (¹ Σ _g ⁺) aV8Z	C ₂ H (² Σ) aV8Z	PH ₃ (¹ A ₁) aV7Z	AlN (³ Π) aV7Z	N ₂ O ₃ (¹ A') aV6Z	N ₂ O ₄ (¹ A _g) aV6Z	C ₂ H ₂ O ₂ glyoxal aV6Z	HCOH trans aV7Z
NO (² Π) aV8Z	HCO (² A')	CCL ₂ (¹ A ₁) aV7Z	Si ₂ (¹ Σ _g ⁺) aV8Z	C ₂ H ₄ O oxir aV6Z	CH ₃ NH aV6Z	CH ₂ O ₂ aV7Z	HOF (¹ A') aV7Z
HNO (¹ A') aV8Z	H ₂ CO (¹ A ₁) aV8Z	CCl ₃ (² A ₁) aV6Z	SiO (¹ Σ ⁺) aV8Z	CH ₂ F ₂ (¹ A ₁) aV6Z	C ₂ N ₂ (¹ Σ _g ⁺) aV7Z	NF ₂ (² B ₁) aV7Z	C ₂ H ₃ vinyl aV7Z
F ₂ (¹ Σ _g ⁺) aV9Z	CF ₂ O (¹ A ₁) aV7Z	BP (³ Π) aV7Z	AlH (¹ Σ ⁺) aV6Z	C ₃ H ₄ propyne aV6Z	C ₄ H ₁₀ butane aV6Z	C ₅ H ₁₂ pentane F12	C ₄ H ₆ <i>trans</i> -1,3- butadiene aV6Z
CF (² Π) aV8Z	CF (⁴ Σ ⁻) aV8Z	PF (³ Σ) aV8Z	BN (³ Π) aV7Z	C ₆ H ₁₄ hexane F12	C ₂ H ₅ ethyl aV6Z	C ₄ H ₄ O furan F12	C ₃ H ₆ propene aV6Z
CF ₂ (¹ A ₁) aV8Z	NO ₂ (² A ₁) aV8Z	BCl (¹ Σ ⁺) aV7Z	H ₂ O ₂ (¹ A') aV7Z	C ₄ H ₅ N pyrrole aV6Z	C ₆ H ₆ benzene F12	CCl ₄ (¹ A ₁) aV6Z	C ₂ Cl ₄ (¹ A _g) aV6Z
CF ₂ (³ B ₁) aV8Z	O ₃ (¹ A ₁) aV8Z	H ₂ S (¹ A ₁) aV8Z	SiH (² Π) aV8Z	BS (² Σ ⁺) aV7Z	BO (² Σ ⁺) aV8Z	C ₂ H ₂ F ₂ 1,1-difluo aV6Z	CHFO (¹ A') aV7Z
FO (² Π) aV8Z	F ₂ O (¹ A ₁) aV7Z	Al ₂ (³ Π _u) aV7Z	SiH ₂ (¹ A ₁) aV8Z	C ₃ H ₅ allyl aV7Z	H ₂ O dimer aV7Z	HF dimer aV6Z	AsO (² Π) aV6Z
NF (³ Σ ⁻) aV8Z	FOO (² A ^{''}) aV7Z	ClO ₂ (² B ₁) aV7Z	SiH ₄ (¹ A ₁) aV7Z	BrF (¹ Σ ⁺) aV7Z	CH ₂ Cl (² B ₁) aV7Z	C ₃ (¹ Σ _g ⁺) aV7Z	C ₃ H ₈ <i>n</i> -propane aV6Z
OH (² Π) aV8Z	SiF (² Π) aV8Z	SiF ₂ (¹ A) aV8Z	AlF (¹ Σ ⁺) aV8Z	C ₂ H ₂ O oxirene aV7Z	HCOH cis aV7Z	C ₃ H ₂ O ketene aV6Z	ClH ₂ N (¹ A') aV7Z
C ₂ H ₄ (¹ A _g) aV8Z	FOOF (¹ A) aV7Z	C ₃ H ₄ cyclopropene aV7Z	C ₂ H ₃ F ₂ <i>trans</i> -1,2- aV6Z	HCP (¹ Σ ⁺) aV7Z	C ₂ H ₃ F stag. aV6Z	C ₂ H ₅ N aziridine aV6Z	C ₃ O ₂ aV6Z
Cl ₂ (¹ Σ _g ⁺) aV8Z	HCN (¹ Σ ⁺) aV8Z	C ₂ F ₂ (¹ Σ _g ⁺) aV7Z	C ₂ H ₂ F ₂ <i>cis</i> -1,2- aV6Z	HNO ₂ (¹ A ₁) aV6Z	H ₂ N ₂ O nitro aV6Z	NH ₂ F (¹ A') aV7Z	C ₂ H ₅ N aziridine aV6Z
HCl (¹ Σ ⁺) aV8Z	HNC (¹ Σ ⁺) aV8Z	N ₂ O (¹ Σ ⁺) aV7Z	SiS (¹ Σ ⁺) aV8Z	H ₂ C=C vinyl aV7Z	C ₂ H ₃ N aV7Z	NF ₃ (¹ A ₁) aV6Z	NH ₃ O (¹ A ₁) aV7Z
C ₃ H ₆ cyclop aV6Z	CH ₂ Cl ₂ (¹ A ₁) aV7Z	HNO ₃ (¹ A') aV7Z	C ₂ F ₄ (¹ A _g) aV6Z	PF ₃ (¹ A ₁) aV7Z	C ₂ H ₆ O aV6Z		
CICN (¹ Σ ⁺) aV8Z	H ₂ SiO (¹ A ₁) aV7Z	BrCl (¹ Σ ⁺) aV6Z	I ₂ (¹ Σ _g ⁺) aV6Z				
HOCl (¹ A') aV7Z	CH ₃ Cl (¹ A ₁) aV6Z	Li ₂ (¹ Σ _g ⁺) aV6Z	AlF ₃ (¹ A ₁) aV6Z				

^aFor second row elements (Na–Cl) aVDZ = aug-cc-pV(D+d)Z, aVTZ = aug-cc-pV(T+d)Z, etc. The entries labeled F12 correspond to CCSD(T)-F12b/CBS(Q5) estimates.

purposes. Consequently, Hill et al. proposed a parametrized basis set extrapolation formula designed for explicitly correlated coupled cluster theory and the VnZ -F12, $n = D, T, Q$, basis sets.⁵⁷ The procedure of Hill et al. does not involve an extrapolation of the self-consistent field (SCF) component of the energy, relying instead on the SCF + complementary auxiliary basis set (CABS) energy in the largest basis set. A small, systematic bias was observed in the CCSD(T)-F12b(CBS VTZ-F12/VQZ-F12) extrapolated atomization energies.⁵⁴ For small molecules the reference values were underestimated. For larger molecules, with valence correlation energies $>0.5E_h$, they were overestimated by 0.2–0.6 kcal/mol.

The latter behavior was reminiscent of what was observed with F12a, although in that case the magnitude of the bias was much more pronounced. At its current state of development, CCSD(T)-F12b appears to require the assistance of extrapolation formulas to reach the highest levels of accuracy in thermochemical properties. Numerical examples will be discussed below.

For noncovalent interactions, Marchetti and Werner proposed scaling the triples contribution (T) by the ratio of the second order Møller–Plesset perturbation theory (MP2)⁵⁸ correlation energies with the F12 approximation and the value without explicit correlation, i.e., $E_{\text{corr}}(\text{MP2-F12})/E_{\text{corr}}(\text{MP})$.^{59,60}

This suggestion is motivated by the lack of explicit correlation in the (T) component of CCSD(T)-F12b causing it to converge at essentially the same rate as the triples correction in standard CCSD(T). The modification with scaled triples, known as CCSD(T*), can be viewed as an alternate estimate of the CBS limit. Because correlation energy recovery occurs approximately an order of magnitude faster for triple excitations than for singles and doubles,⁶¹ combining an explicit treatment of $E_{\text{CCSD}}(\text{corr})$ with a more approximate treatment of the triples appeared reasonable. The goal of the current work is to document the strengths and weaknesses of CCSD(T*) in regard to atomization energies to better understand its limitations as a practical, predictive tool.

II. APPROACH

The complete list of 212 molecules considered in this study is provided in Table 1 along with the corresponding largest basis sets used in the CBS extrapolations required to obtain the reference values. This list is a superset of the 105 closed shell, neutral, H, C, N, O, and F-containing molecules selected by Bakowies and several other investigators.⁶² All of the molecular energies were obtained at their respective optimal CCSD(T)-(FC)/aVnZ or CCSD(T)-F12b(FC)/VnZ-F12 geometries, which helps minimize the arbitrariness associated with using geometries obtained from experiment or from some lower level of theory. It also matches the way the CBS extrapolation step is used in the Feller–Peterson–Dixon (FPD) procedure.^{34–37} Thus, the reference CBS atomization energies correspond to near-CBS limit structures. The only exceptions occurred when a single point calculation required several days or more to complete, as was the case with calculations involving more than 1600 basis functions. In these situations the optimal structure was estimated by an exponential extrapolation of the internal coordinates from the next three smaller basis sets.

Note that other authors have chosen different approaches for defining CBS benchmark quality reference values. For example, Bakowies focused on second-order Møller–Plesset perturbation theory (MP2(FC)) and CCSD(FC) correlation energies based on very large, uncontracted R12A and R12B energies or on a weighted average of R12A and R12B for a small collection of 12 molecules using R12 values reported by Fliegl et al.^{62,63} In a smaller collection of 7 molecules, Bakowies also considered CBS reference values for atomization energies but the uncontracted basis set used in the R12 calculations lacked i functions. The author concluded that the small size of this reference set was “likely neither representative nor accurate enough” to be used for developing an effective CBS extrapolation procedure. Hill et al.⁵⁷ computed large basis set CCSD-F12b reference correlation energies for 14 molecules along with a smaller number of (T) reference values to develop an extrapolation technique for F12b. Köhn and D. P. Tew selected a subset of 30 molecules taken from the 105 molecule collection of Bakowies to define CCSD correlation energies using an uncontracted version of the aug-cc-pV6Z basis set.⁶⁴

The diagonal, fixed amplitude *ansatz*, in which the amplitudes of the explicitly correlated configurations are determined by the wave function cusp conditions⁶⁵ was chosen for all CCSD(T)-F12b calculations. The method also requires auxiliary basis sets. Whenever possible, the cc-pVnZ-F12/OptRI sets⁶⁶ were used for the CABS resolution of the identity (RI) step.⁶⁷ The lack of OptRI auxiliary basis sets for the VSZ-F12 orbital sets⁶⁶ caused occasional problems. For example, in some cases geometry optimizations could only be converged to

$\sim 0.001 E_h/\text{bohr}$, which should not affect the atomization energies by more than 0.001 kcal/mol. Such problems could sometimes be avoided by using a smaller step size when numerical gradients are computed. MOLPRO does not currently provide analytical gradients for CCSD(T)-F12b. A recently released patch to the MOLPRO software has reduced this problem.

Density fitting of the Fock and exchange matrices was accomplished with the cc-pVnZ/JKFIT basis sets of Weigend⁶⁸ whereas other 2-electron integrals were density fit using aVnZ/MP2Fit sets.^{69,70} 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.⁵⁴ Geminal exponents of 0.9 (VDZ-F12), 1.0 (VTZ-F12 and VQZ-F12) and 1.2 (VSZ-F12) were applied. For hydrogen the rev2 variant of the VSZ-F12 basis set was chosen because it contains a larger number of correlating functions than the original set, [6s5p4d3f2g] vs [6s5p3d2f1g]. With the exception of the orbital basis set, the preceding items have no counterpart in standard CCSD(T) theory but making different choices for parameters such as the geminal exponents can cause nontrivial changes in the results.

Most CCSD(T) calculations were performed with MOLPRO 2012.1.⁵³ The largest of these involved 2026 basis functions. To avoid basis set linear dependency problems in some cases, the accuracy of the 2-electron integrals was increased from the default 10^{-12} to 10^{-15} with the MOLPRO command `gthresh,twoint=1.0d-15`. This represents the limit of what can be achieved with 64 bit arithmetic. With this threshold no obvious signs of linear dependency, such as unusually large numbers of SCF or CCSD iterations or deviations from the typically smooth convergence of the energy with respect to the basis set index n , were observed. The smallest eigenvalue of the overlap matrix encountered in this work was 10^{-8} , a value that is similar in magnitude to what was successfully handled in a recent study of Rydberg excited states in ethylene.⁷¹ 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.^{72,73} Full atomic symmetry in the orbitals was imposed in calculations on the isolated atoms. Because the current version of MOLPRO does not support basis functions beyond $l_{\text{max}} = 6$ (i functions) the aV7Z through aV9Z calculations, which involved k , l , and m functions, were performed with Dalton 2.0⁷⁴ or Gaussian 09.⁷⁵ In Gaussian 09 the integral accuracy was increased to 10^{-13} with the `Int=(Acc2E=13)` command. Three 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 because we found that with the largest basis sets the function spaces were being slightly truncated, leading to energies that were too high. Though the differences in energy were small, the deviation from the otherwise smooth convergence pattern rendered what might have been an overlooked source of error immediately obvious. The second command tightened the threshold for the elimination of functions associated with small eigenvalues of the overlap matrix. The G09 default (10^{-6}) was reduced to 10^{-7} with `IOp(3/59=7)`. Finally, the in-core SCF and CCSD algorithms were disabled with the commands `SCF(NoInCore)` and `CCSD(T,NoInCore)` because a problem was found with calculations involving large memory when high

Table 2. Differences (kcal/mol) between CBS(aV56Z), CBS(aV67Z), CBS(aV78Z), and CBS(aV89Z) Reference Atomization Energies^a

molecule	CBS(aV56Z)	Δ	CBS(aV67Z)	Δ	CBS(aV78Z)	Δ
H ₂ (¹ Σ_g^+)	109.478	-0.003	109.486	0.005	109.476	-0.005
BH (¹ Σ^+)	84.717	-0.015	84.743	0.011	84.748	0.016
B ₂ (³ Σ_g^-)	65.181	-0.018	65.179	-0.012	65.201	0.002
CH (² Π)	83.898	-0.030	83.932	0.004	83.941	0.013
CH ₂ (³ B ₁)	189.855	-0.037	189.884	0.020	189.892	0.028
CH ₂ (¹ A ₁)	180.664	-0.038	180.718	0.016	180.722	0.020
C ₂ (¹ Σ_g^+)	145.248	-0.049	145.267	-0.030	145.289	-0.008
N ₂ (¹ Σ_g^+)	227.103	-0.144	227.152	-0.095	227.171	-0.076
O ₂ (³ Σ_g^-)	119.906	-0.060	119.915	-0.051	119.951	-0.015
F ₂ (¹ Σ_g^+)	38.465	-0.011	38.476	0.000	38.524	0.048
HF (¹ Σ^+)	141.483	-0.062	141.540	-0.005	141.549	0.004
OH (² Π)	106.967	-0.051	107.021	0.003	107.022	0.004
H ₂ O (¹ A ₁)	232.554	0.081	232.646	-0.038	232.635	0.000
NH (³ Σ^-)	82.903	0.036	82.859	-0.008	82.868	0.001
NH ₂ (² B ₁)	182.030	-0.110	182.115	-0.025	182.066	-0.074
ϵ_{MAD}		0.050		0.022		0.021
ϵ_{MSD}		-0.028		-0.012		-0.003
ϵ_{Max}		-0.144		-0.095		-0.076

^aAll CBS estimates obtained with the $1/(l_{\text{max}} + 1/2)^4$ formula. The CBS(aV89Z) values are H₂ = 109.481, BH = 84.732, B₂ = 65.199, CH = 83.928, CH₂ (³B₁) = 189.864, CH₂ (¹A₁) = 180.702, C₂ = 145.297, N₂ = 227.247, O₂ = 119.966, F₂ = 38.476, HF = 141.545, OH = 107.018, H₂O = 232.635, NH = 82.867, and NH₂ = 182.140 kcal/mol.

angular momentum functions were present in the basis set. Multiple comparisons between G09 and Dalton on closed shell species confirmed that CCSD(T) total energies were in agreement to a microhartree or better, which was sufficient for our purposes.

In cases involving the aV7Z or aV8Z basis sets where the number of functions exceeded 1500, the contributions of k and l functions were estimated by a short extrapolation, as discussed elsewhere.^{27,61} Comparison with explicit k and l calculations indicate that this approximation should be accurate to 0.01–0.02 kcal/mol when atomization energies are computed. A new aV9Z boron basis set was created for this study. Because Dalton and Gaussian do not support the R/UCCSD(T) method for open shell systems, R/U energies with aV7Z through aV9Z basis sets were estimated by assuming that the differences between UCCSD(T) and R/UCCSD(T) energies were essentially converged at the aV6Z level. Though it was impossible to directly test this assumption, the sequence of UCCSD(T) – R/UCCSD(T) differences as a function of basis set size suggests that the estimated R/U energies should be accurate to $\sim 10^{-6} E_h$.

Several improvements were made to the current atomization energy reference set compared to the older 2013 reference set.²⁷ First, the number of molecules was increased, including some noncovalent species. Second, whenever possible, the highest basis set level was increased. Assessing the degree of convergence in the reference set molecule-by-molecule is difficult because the present calculations utilized the largest feasible basis sets allowed by our present hardware and software. Normally, the basis set index n serves as a reasonable qualitative indicator of the expected accuracy. In spite of that, with large basis sets an equally important factor is the number of functions per valence electron pair. Consider for example the two isoelectronic molecules F₂ and C₂H₆. With the aV6Z basis set F₂ has 54 functions per valence electron pair, whereas C₂H₆ has 163. For F₂ aV6Z is capable of recovering 99.0% of the valence CCSD(T)/CBS correlation energy, whereas for C₂H₆ it

was closer to 99.8%. This difference produced better 1-particle convergence in the atomization energy for C₂H₆. Further discussion of this topic was provided in our earlier investigation.²⁷ Another perspective on the accuracy of the reference set can be found by comparing the CBS(aV56Z) values, the lowest level considered in this study, with the admittedly limited number of CBS(aV89Z) values. As seen in Table 2, across all available comparisons we find mean absolute deviations of 0.05 (aV56Z), 0.02 (aV67Z), and 0.02 (aV78Z) kcal/mol. For this small set of molecules the CBS(aV56Z) extrapolation systematically undershoots the CBS(aV89Z) reference values ($\epsilon_{\text{MSD}} = -0.03$ kcal/mol), whereas the more expensive CBS(aV67Z) and CBS(aV78Z) extrapolations display little systematic bias.

Table 3 contains the error statistics for standard CCSD(T) with and without CBS extrapolation for the current reference set. Values are little changed compared to the findings in the 2013 study.²⁷ The raw results across all basis sets (aVDZ to aV9Z) consistently decline by a factor of 2 in the error metrics for every increment in the basis set index n . CBS extrapolation is effective in further decreasing the error, leading to results equivalent to $n + 2$ or $n + 3$, where n is the largest basis set used in the extrapolation. After considering the data in Tables 2 and 3, as well as other approaches to gauging the reference set error discussed in our earlier study, we propose ϵ_{MAD} (mean absolute deviations relative to the reference values) of ± 0.06 CBS(aV56Z), ± 0.04 CBS(aV67Z) and ± 0.02 kcal/mol CBS(aV78Z), which we believe to be conservative. If the CBS(aV89Z) extrapolation were to follow the trend established by the smaller sets, an uncertainty on the order of ± 0.01 kcal/mol would be expected. Individual uncertainties could obviously be much larger. Data for calibrating the CBS(aV89Z) values currently do not exist due to the need for aV10Z energies. Among the molecules included in Table 3, triply bonded N₂ with its comparatively short 1.0991 Å bond length displays a D_e that is particularly difficult to converge.

Table 3. Raw and CBS Extrapolated Standard CCSD(T)(FC) Atomization Energies Errors (kcal/mol)^a

basis set	no. of items	ϵ_{MAD}	ϵ_{RMS}	ϵ_{MSD}	ϵ_{max}
aug-cc-pVDZ	212	25.48	30.35	-25.48	-101.53
aug-cc-pVTZ	212	8.07	9.47	-8.07	-27.90
aug-cc-pVQZ	212	2.84	3.28	-2.84	-9.16
aug-cc-pV5Z	209	1.28	1.47	-1.287	-3.95
aug-cc-pV6Z	209	0.66	0.76	-0.66	-2.13
aug-cc-pV7Z	157	0.35	0.40	-0.35	-0.93
aug-cc-pV8Z	85	0.16	0.19	-0.16	-0.56
aug-cc-pV9Z	15	0.06	0.08	-0.06	-0.12
CBS(aVTQZ)	212	0.33	0.42	0.22	1.62
CBS(aVQ5Z)	212	0.13	0.20	-0.09	-0.92
CBS(aV56Z)	156	0.06	0.09	-0.02	0.61
CBS(aV67Z)	85	0.05	0.07	0.03	0.25
CBS(aV78Z)	15	0.02	0.03	0.00	-0.07

^aThe errors are measured with respect the best available CBS($1/[I_{\text{max}} + 1/2]^4$) estimates of the basis set limits. The CBS(xy) estimates were also based on the $1/[I_{\text{max}} + 1/2]^4$ formula. 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).

Note that although some of results appearing in the text and tables will be reported to 10^{-3} kcal/mol, this is not meant to imply that any of the reference values are accurate to 10^{-3} . As the above discussion implies, although a few values might fortuitously be converged to better than 0.01 kcal/mol, for most of the molecules in the reference set it is not possible to achieve such accuracy. Occasional reporting of results to a thousandth of a kcal/mol simply reflects the need to track differences to that level.

In previous studies of the effectiveness of CBS extrapolation, when used to predict atomization energies, we stressed that no single formula provided superior statistical results across all correlation consistent basis sets and types of molecules.^{16,42,43} These studies included a detailed examination of six formulas, with several other expressions receiving a more cursory examination. The current improved reference set led to the same conclusion. Although the $1/[I_{\text{max}} + 1/2]^4$ extrapolation formula is a good compromise for standard CCSD(T), for particular small collections of molecules or specific basis set

combinations, it is possible that other choices may be more accurate. For example, in the case of the first row molecules in Table 2 the CBS(aV56Z) mean signed deviation indicates that this level of theory systematically underestimates the CBS-(aV89Z) reference values. Slightly reducing the exponent in the denominator from 4 to 3.7 would bring the numbers into better agreement with the reference values, but it is unclear whether this is a limited effect or if it has some more general significance.

For several of the largest systems (e.g., C_6H_6) CCSD(T)/aV6Z calculations proved impossible with the standard method, either because of the amount of memory or disk storage or because of problems with linear dependence in the basis set. In such cases we generated reference values from explicitly correlated CCSD(T)-F12b calculations performed with the cc-pVQZ-F12 and cc-pV5Z-F12 basis sets.

As with standard CCSD(T), F12b energies exhibit uniform convergence and are, therefore, amenable to CBS extrapolation. With F12 we use the parametrized formula of Hill et al.⁵⁷ for basis sets VDZ-F12 through VQZ-F12. In the study introducing the V5Z-F12 basis sets for hydrogen and B-Ne, Peterson et al.⁵⁶ noted that no extrapolation of the CCSD-F12b correlation energy was necessary to reach an accuracy comparable to CBS(aV56Z) with standard CCSD for their TAE28 test set consisting of small closed-shell molecules. The (T) component of the energy, which does not benefit from explicit correlation, was extrapolated to the CBS limit using a $1/I_{\text{max}}^3$ formula. In this study the impact of extrapolating the CCSD-F12b component of the energy was also examined by considering three different strategies, all of which were combined with a $1/I_{\text{max}}^3$ extrapolation of the (T) component. These included (1) no extrapolation of E(CCSD-F12b), (2) a $1/I_{\text{max}}^5$ extrapolation of E(CCSD-F12b), and (3) a $1/I_{\text{max}}^7$ extrapolation of E(CCSD-F12b). Although the asymptotic convergence of F12 is expected to go as $1/I_{\text{max}}^{7,76}$, the work of Hill et al. indicated that at the $S\zeta$ basis set level it is actually better described by $1/I_{\text{max}}^{4,9}$ (see Table 10 in Hill et al.⁵⁷). Following the precedent of the earlier F12 extrapolation procedure of Hill et al., the SCF + CABS energy was used as an estimate of the Hartree-Fock limit contribution to the atomization energy.

Table 4. Raw and CBS Extrapolated CCSD(T)-F12b(FC) Atomization Energies Errors (kcal/mol)^a

basis set	ϵ_{MAD}	ϵ_{RMS}	ϵ_{MSD}	ϵ_{max}	approximate CCSD(T) equivalent
cc-pVDZ-F12	3.68	4.47	-3.65	-14.94	<aVQZ
cc-pVTZ-F12	1.25	1.69	-1.25	-14.70	aV5Z
cc-pVQZ-F12	0.29	0.34	-0.29	-1.26	aV7Z or CBS(aVTQZ)
cc-pV5Z-F12(rev 2) ^b	0.06	0.07	-0.03	-0.21	aV9Z or CBS(aV67Z)
CBS(VDTZ-F12)	0.54	0.79	-0.48	-3.51	<CBS(aVTQZ)
CBS(VTQZ-F12)	0.14	0.18	0.11	0.65	CBS(aVQ5Z)
CBS(VQ5Z-F12)-(T) ^c	0.12	0.15	0.10	-0.60	CBS(aVQ5Z)
CBS(VQ5Z-F12)-both ^d	0.16	0.19	0.16	0.56	CBS(aVQ5Z)

^aThe errors are measured with respect to the best available CBS($1/[I_{\text{max}} + 1/2]^4$) estimates of the basis set limits. The CBS(VDTZ-F12) and CBS(VTQZ-F12) estimates were based on the F12b extrapolation formula proposed by Hill et al.⁵⁷ 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). The number of comparisons = 209 for cc-pVnZ-F12, $n = \text{D, T, Q}$ and for CBS(VDTZ-F12) and CBS(VTQZ-F12). Finally, 121 for cc-pV5Z(rev2)-F12. ^bThe cc-pV5Z-F12 basis sets are only available for H, B - Ne. ^cCBS estimate included a $1/I_{\text{max}}^3$ extrapolation of the (T) energy component only. ^dCBS estimate included both a $1/I_{\text{max}}^5$ extrapolation of the CCSD-F12b energy component and a $1/I_{\text{max}}^3$ extrapolation of the (T) component. Using a $1/I_{\text{max}}^5$ formula for the CCSD component would have reduced ϵ_{MSD} by roughly 0.01 kcal/mol.

The triples scale factor used to compute CCSD(T*) energies for closed shell molecules was taken as the ratio of two density fitted correlation energies. In terms of MOLPRO's output syntax the ratio is expressed as [DF-MP2-F12/3C(FIX)/DF-MP2]/[DF-MP2]. This definition differs slightly from the one used by MOLPRO when the scale_trip option is invoked. The latter uses the ratio of the standard MP2 correlation energy, $E_{\text{corr}}(\text{MP})$, adjusted by the F12 C3(FIX) correlation correction, to $E_{\text{corr}}(\text{MP})$. In practice, the two scale factors are numerically very close but the first definition avoids the need to carry out an additional standard MP2 calculation. Tests on a few closed shell molecules showed that the different ratios yielded CCSD(T*) energies that differed by $1 \mu E_h$ or less. Given the semi-quantitative nature of the scaling and the use of different scaling factors for the same molecule in different situations, e.g., H₂O as an isolated molecule vs H₂O as a fragment of a water cluster, the approximation used here was judged to be acceptable. Atomic scale factors were taken from the parent molecule to ensure size-consistency in the results. Scale factors for open shell molecules were based on [RMP2-F12/3C(FIX)]/[RMP2], again in terms of the MOLPRO output file syntax. Although scaling factors varied from system to system, values for first row molecules typically ranged from 1.02 (VSZ-F12) to 1.1 (VTZ-F12).

III. RESULTS AND DISCUSSION

A. CCSD(T)-F12b with and without Extrapolation.

Error statistics for explicitly correlated CCSD(T)-F12b with and without CBS extrapolation are shown in Table 4. Each row in the table contains an approximate standard CCSD(T) equivalent. Raw VnZ-F12, $n = D, T$ results correspond roughly to CCSD(T)/aV($n+2$)Z; i.e., the inclusion of explicit correlation achieves two basis set indices better convergence in the 1-particle expansion for these two basis sets. The VQZ-F12 basis set picks up three basis set indices and the raw VSZ-F12 set corresponds roughly to aV9Z, although the number of aV9Z values is quite limited. Each increment in the basis set index, n , reduces the F12b errors by a factor of ~ 4 . Given the method's n^3N^4 scaling in the computational cost, where n and N are the number of occupied and unoccupied molecular orbitals, this represents an enormous savings in computational cost.

The remarkable ability of CCSD(T)-F12b/VnZ-F12, $n = Q, S$, to reproduce the expensive reference atomization energies is achieved, in part, by the method's tendency to overestimate ΣD_e . This observation was initially reported in our 2013 study.²⁷ Following CBS extrapolation, the CBS(VQ5-F12) error metrics actually increase relative to the raw VSZ-F12 results. Figure 1 shows the F12b/CBS errors as a function of the absolute value of the valence correlation energy. Although there is considerable scatter, a linear least-squares fit of the data predicts errors on the order of 0.2 kcal/mol for molecules with valence correlation energies $\geq 0.8 E_h$ (e.g., CH₂F₂ or NF₂). The data in Table 4 also indicate that for this reference set extrapolated F12b atomization energies are unable to achieve a level of accuracy better than CCSD(T)/CBS(aVTQ5Z).

The small basis set CBS(VDTZ-F12) extrapolated results, which rely on the formula of Hill et al.,⁵⁷ were found to be somewhat less accurate than the CBS(aVTQZ) results with standard CCSD(T). Thus, it may not be accurate enough for some studies, although its computational cost is low. CBS-(VTQZ-F12) yields error metrics in close agreement with the much more expensive CBS(VQ5Z-F12) method, our best

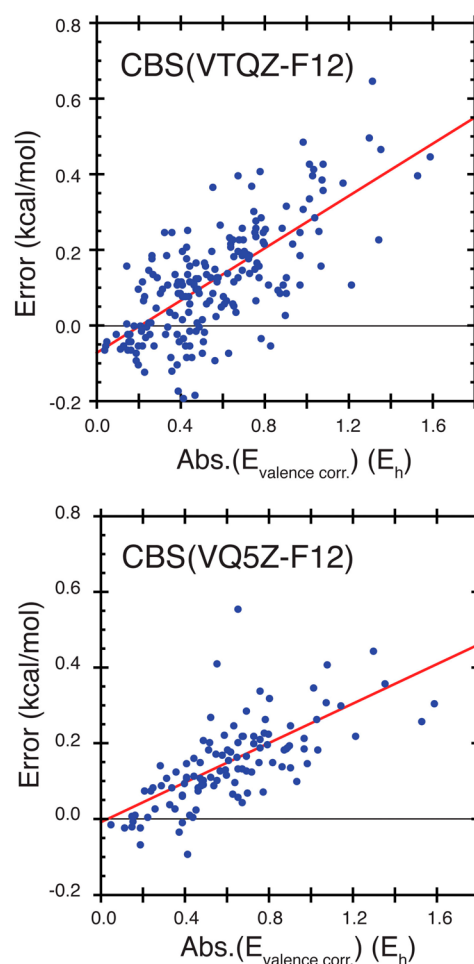


Figure 1. Error in extrapolated CCSD(T)-F12b/CBS(VTQZ-F12) and CBS(VQ5Z-F12) atomization energies as a function of the valence correlation energy. The reference values were based on the $1/(l_{\text{max}} + 1/2)^4$ formula combined with standard CCSD(T). The red line represents a linear least-squares fit.

estimate of the F12b basis set limit. When comparing the results in Table 4, one should emphasize that VSZ basis sets are not currently available for second row elements. If we limit the CBS(VTQZ-F12) analysis to only those molecules covered by the VSZ-F12 basis set, the error metrics increase only slightly, e.g., $\epsilon_{\text{MAD}} = 0.16$ (limited comparison) vs 0.14 kcal/mol (all molecules).

In light of the issues just discussed, CCSD(T)-F12b should be used with caution in studies seeking the very highest levels of accuracy for small-to-medium size molecules unless a deliberate decision is made to avoid extrapolation to the CBS limit, instead relying on fortuitous cancellation of error. For specific classes of molecules it may prove possible to develop an empirically based adjustment to compensate for the systematic behavior shown in Figure 1.

Changing the details of the CBS extrapolation procedure can produce minor changes in the VSZ-F12 atomization energies, as illustrated in Table 5 for six selected molecules. As with standard CCSD(T), extrapolation universally increases the F12b atomization energy. However, unlike CCSD(T) there are several instances (C₂H₂O and CF₄) where the raw value already slightly exceed the reference value, i.e., prior to extrapolation. After extrapolation, all but one of the F12b values is larger than the corresponding reference value. Despite the size of the VSZ-

Table 5. Selected CCSD(T)-F12b/V5Z-F12 and CCSD(T)-F12b/CBS(Q5) Atomization Energies (kcal/mol) with Respect to the Details of the Extrapolation^a

molecule	raw V5Z-F12	CCSD-F12(no extr.) (T) used $1/I_{\max}^3$	CCSD-F12($1/I_{\max}^5$) (T) used $1/I_{\max}^3$	CCSD-F12($1/I_{\max}^7$) (T) used $1/I_{\max}^3$	ref value
C ₂ (¹ Σ _g ⁺)	145.121	145.228	145.207	145.217	145.297
C ₃ H ₈ <i>n</i> -propane	1004.097	1004.283	1004.376	1004.333	1004.171
CO ₂ (¹ Σ _g ⁺)	388.169	388.297	388.314	388.307	388.243
C ₂ H ₂ O ketene	530.360	530.536	530.603	530.572	530.354
CF ₄ (¹ A ₁)	478.044	478.241	478.292	478.268	477.926
N ₂ O ₄ (¹ A _g)	468.643	469.004	469.109	469.062	468.800

^aBased on CCSD(T)-F12b energy components.**Table 6. Examples of CCSD(T), CCSD(T)-F12b, and (T*) Correlation Energies^a**

CCSD(T)(FC)			CCSD(T)-F12b(FC)				
basis	$E_{\text{corr}}(\text{CCSD})$	$E_{\text{corr}}(\text{T})$	basis	$E_{\text{corr}}(\text{CCSD})$	$E_{\text{corr}}(\text{T})$	scaled $E_{\text{corr}}(\text{T})$	scale factors
C ₂ (¹ Σ _g ⁺)							
aVDZ	-0.317946	-0.028957	VDZ-F12	-0.363209	-0.031661	-0.035553	1.122943
aVTZ	-0.350702	-0.034030	VTZ-F12	-0.366136	-0.034283	-0.036017	1.050553
aVQZ	-0.361062	-0.035313	VQZ-F12	-0.367017	-0.035294	-0.036186	1.025273
aV5Z	-0.364447	-0.035776	V5Z-F12	-0.367315	-0.035558	-0.036209	1.018311
aV6Z	-0.365772	-0.035940					
aV7Z	-0.366419	-0.036016					
aV8Z	-0.366729	-0.036053					
aV9Z	-0.366941	-0.036076					
CBS ^b	-0.367319	-0.036132	CBS(Q5) ^c	-0.367337	-0.036072		
N ₂ (¹ Σ _g ⁺)							
aVDZ	-0.325032	-0.014200	VDZ-F12	-0.398026	-0.016675	-0.019714	1.182277
aVTZ	-0.378073	-0.019410	VTZ-F12	-0.404972	-0.019616	-0.021090	1.075142
aVQZ	-0.395736	-0.020539	VQZ-F12	-0.406873	-0.020558	-0.021334	1.037766
aV5Z	-0.401950	-0.020985	V5Z-F12	-0.407408	-0.020862	-0.021388	1.025195
aV6Z	-0.404433	-0.021143					
aV7Z	-0.405632	-0.021217					
aV8Z	-0.406315	-0.021258					
aV9Z	-0.406727	-0.021278					
CBS ^b	-0.407462	-0.021325	CBS(Q5) ^c	-0.407669	-0.021180		
Cl ₂ (¹ Σ _g ⁺)							
aV(D+d)Z	-0.329955	-0.010732	VDZ-F12	-0.443807	-0.011155	-0.015471	1.386926
aV(T+d)Z	-0.421271	-0.019652	VTZ-F12	-0.464294	-0.019219	-0.022103	1.150035
aV(Q+d)Z	-0.452183	-0.022192	VQZ-F12	-0.470810	-0.021706	-0.023406	1.078338
aV(S+d)Z	-0.462635	-0.023164	aV(S+d)Z	-0.472094	-0.022917	-0.023864	1.041310
aV(6+d)Z	-0.467640	-0.023585					
aV(7+d)Z	-0.469627	-0.023756					
aV(8+d)Z	-0.470700	-0.023840					
CBS ^b	-0.472350	-0.024011	CBS(TQ) ^d	-0.473178	-0.023620		

^aAll energies were obtained at the optimal bond length for each basis set and level of theory. The CCSD-F12b correlation energy was measured with respect to the SCF + CABS energy. Geminal exponent values were 0.9 (VDZ-F12), 1.0 (VTZ-F12), 1.0 (VQZ), 1.2 (V5Z-F12), and 1.5 (aV5Z). ^bCBS estimate for $E_{\text{corr}}(\text{CCSD})$ with conventional CCSD(T) is based on a $1/(I_{\max} + 1/2)^2$ extrapolation. The estimate for (T) was based on $1/I_{\max}^3$. ^cCBS(Q5) estimates for the explicitly correlated method are based on a $1/I_{\max}^5$ extrapolation of the CCSD correlation energy and $1/I_{\max}^3$ extrapolation of (T). ^dCBS(TQ) estimate for the explicitly correlated method are based on the coefficients recommended by J. G. Hill et al. (*J. Chem. Phys.* **2009**, *131*, 194105).

F12 basis set, failure to treat the CCSD basis set truncation error is seen to lead to differences of 0.02–0.11 kcal/mol in ΣD_e . Smaller differences (0.01–0.04 kcal/mol) are found between the $1/I_{\max}^5$ and $1/I_{\max}^7$ formulas used for extrapolating the CCSD-F12b correlation energy. The latter choice produces smaller atomization energies, as expected. For reasons discussed previously, the $1/I_{\max}^5$ formula seems more appropriate for this quality basis set.

B. CCSD(T*)-F12b. Although absolute accuracy in the total energy is not a prerequisite for achieving accuracy in atomization energies, a brief discussion of the breakdown in the convergence pattern for CCSD, CCSD-F12b, (T), and (T*) energy components for three diatomic examples (C₂, N₂, and Cl₂) will be presented. As shown in Table 6, the pace of convergence with the standard method is painfully slow along the correlation consistent sequence of basis sets. Even with the aV9Z basis set the CCSD correlation energy remains 0.0007 E_h (worst case) higher than the CBS limit. For each of these cases

F12b appears to slightly overestimate the CCSD/CBS value, as is often the case among the molecules we have studied. Counterbalancing this is a tendency for the F12b triples component, to be somewhat underestimated. As expected, scaling (T) substantially improves the triples component, but the resulting values are a bit too large. Conclusions drawn from the data in Table 6 are, in general, valid across all of the systems treated with the aV9Z basis set.

Prior to performing the statistical analysis it was difficult to predict the overall impact of substituting (T*) for (T) as the perturbative triples component due to the following competing effects: (1) basis set truncation error in the CCSD energy, (2) the propensity of F12b to overestimate E_{CCSD} , and (3) results from early calculations that indicated (T*) might yield a triples contribution that was frequently larger than the true value. Subsequent to the analysis, we find that (T*) significantly improves atomization energies relative to the raw F12b values, at least for the smaller VDZ-F12 and VTZ-F12 basis sets. Yet when (T*) results are compared to other methods for improving upon the raw results, such as those obtained from the CBS extrapolation formula of Hill et al.,⁵⁷ the findings are mixed. With the small VDZD-F12b/VTZ-F12 basis set combination, the CCSD(T*) method yields error metrics in Table 7 that are slightly smaller than the CBS values in Table 4.

Table 7. CCSD(T*)-F12b(FC) Atomization Energies Errors (kcal/mol)^a

basis set	ϵ_{MAD}	ϵ_{RMS}	ϵ_{MSD}	ϵ_{max}
cc-VTZ-F12	0.41	0.51	-0.29	-1.58
cc-VQZ-F12	0.22	0.33	0.15	1.69
cc-VSZ-F12	0.25	0.30	0.24	0.86

^aThe errors are measured with respect to the best available CBS($[I_{\text{max}} + 1/2]^4$) estimates of the basis set limits. All calculations assumed the frozen core approximation.

Both approaches display a tendency to underestimate the reference values, -0.48 (CBS) vs -0.29 kcal/mol (CCSD(T*)), although for somewhat different reasons. Despite the lack of a correction for basis set truncation in the CCSD component, CCSD(T*) benefits from a partial cancellation of factors (1) and (3) noted above. With the larger basis set combinations, CCSD(T*) fails to achieve the accuracy found with the formula of Hill et al.⁵⁷ Specifically, CCSD(T*) with VTZ-F12/VQZ-F12 basis sets display better balance between factors (1) and (3), causing ϵ_{MSD} to change sign and decrease to 0.15 kcal/mol. But all four error metrics are significantly larger than the CBS(VTQZ-F12) counterparts. Ultimately, with the VQZ-F12/VSZ-F12 combination the CCSD(T*) error arising from item (1) becomes relatively insignificant and factors (2) and (3) combine to yield atomization energies that clearly overestimate the reference values ($\epsilon_{\text{MSD}} = 0.24$ kcal/mol) and again exceed the corresponding CBS error metrics. In addition to the numerical data in Table 7, error metrics as a function of basis set are also displayed graphically in Figure 2.

Dimerization energies for selected hydrogen-bonded dimers; $(\text{H}_2\text{O})_2$, $(\text{NH}_3)_2$, and $(\text{HF})_2$ are shown in Table 8. In all cases, the MP2 binding energies lie very close to the corresponding CCSD(T) values. Note that in noncovalently bonded systems with a significant dispersion component, such as the π - π stacked interaction in the indole-benzene complex (Figure 3), the CCSD(T) correction to the MP2 binding energy can exceed 1 kcal/mol.⁷⁷ Unfortunately, systems of that size are too

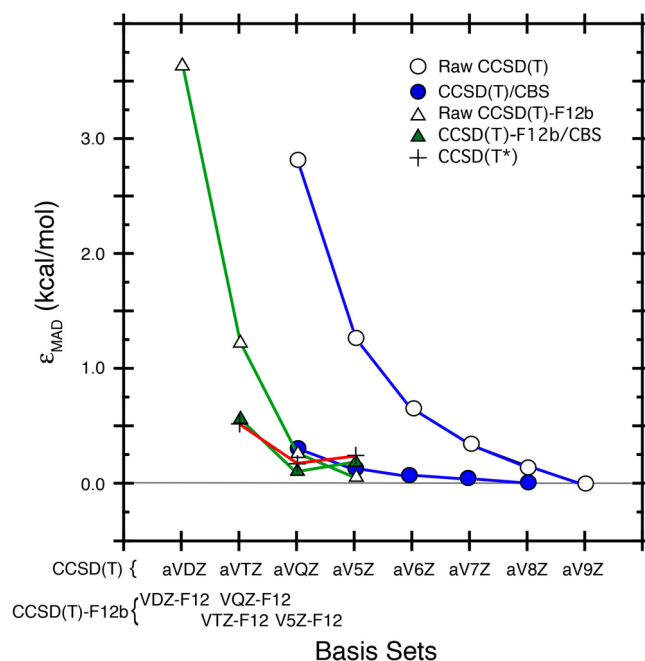


Figure 2. Mean absolute deviation for atomization energies with respect to the reference values obtained at various levels of theory.

large for use with the quality of basis sets chosen for this study and were not included in the reference set. For the complexes in Table 8 it is well-known that the standard MP2 and CCSD(T) approach the binding energy basis set limit from above, whereas the counterpoise-corrected values approach from below. CCSD(T)-F12b also approaches the dimerization basis set limit from above for the water and ammonia dimers, but convergence is somewhat erratic for the hydrogen fluoride dimer. With the small VTZ-F12 basis set both F12b/CBS and CCSD(T*) worsen agreement with the CCSD(T)/CBS limit because they increase a raw binding value that is already too large. With larger F12 basis sets, extrapolation or use of (T) scaling makes little difference. The CCSD(T)-F12b basis set limit differs from the standard CCSD(T) limit by 0.01–0.03 kcal/mol, and CCSD(T*)-F12b provides no apparent advantage over F12b CBS extrapolation.

As a side note, the water dimer binding energy has been the subject of dozens of studies. The current MP2(FC) CBS limit (4.93 kcal/mol) compares favorably with the 4.95 ± 0.05 kcal/mol²¹ estimate first reported in 1992 and a 4.9 ± 0.1 kcal/mol⁷⁸ value reported in 1996, despite the present use of much larger basis sets and the inclusion of geometry optimization. The earlier studies also explored the impact of higher order correlation and core/valence correlation on the binding energy. Basis set superposition error declines smoothly along the correlation consistent sequence from 0.24 kcal/mol with aVDZ to 0.07 kcal/mol with aV7Z.

IV. CONCLUSION

By exploiting partial cancellation of errors associated with underestimating $E_{\text{CCSD}}(\text{corr})$ and overestimating $E_{\text{T}}(\text{corr})$, CCSD(T*)-F12b does well for small basis sets but could not match the performance of a CBS extrapolation technique⁵⁷ designed to estimate the basis set limits for both components of the energy with VQZ-F12 and VSZ-F12(rev 2) basis sets. Both (T*) and the CBS(F12) methods effectively reduce the error with respect to the reference set when compared to raw F12b

Table 8. Selected Noncovalent Electronic Binding Energies (kcal/mol)^a

basis	standard methods		basis	explicitly correlated methods		
	MP2	CCSD(T)		CCSD(T)-F12b	CBS(VnmZ)	CCSD(T*)-F12b
H ₂ O Dimer						
aVDZ	5.263	5.296	VDZ-F12	5.018		
aVTZ	5.181	5.217	VTZ-F12	5.017	5.040	5.042
aVQZ	5.093	5.101	VQZ-F12	4.999	4.998	5.011
aVSZ	5.034	5.034	VSZ-F12	4.988	4.988	4.996
aV6Z	5.007	5.009				
aV7Z	4.975	4.987				
CBS(aV67Z)	4.934	4.960				
NH ₃ Dimer						
aVDZ	3.624	3.587	VDZ-F12	3.165		
aVTZ	3.286	3.269	VTZ-F12	3.162	3.177	3.184
aVQZ	3.220	3.190	VQZ-F12	3.146	3.160	3.157
aVSZ	3.182	3.155	VSZ-F12	3.142	3.144	3.148
aV6Z	3.169	3.143				
CBS(aV56Z)	3.153	3.130				
HF Dimer						
aVDZ	4.686	4.813	VDZ-F12	4.466		
aVTZ	4.708	4.829	VTZ-F12	4.584	4.641	4.603
aVQZ	4.633	4.723	VQZ-F12	4.577	4.582	4.586
aVSZ	4.580	4.652	VSZ-F12	4.575	4.580	4.581
aV6Z	4.532	4.607				
CBS(aV56Z)	4.483	4.560				

^aCBS extrapolations with standard MP2 and CCSD(T) were based in the $1/[l_{\max} + 1/2]^4$ formula.

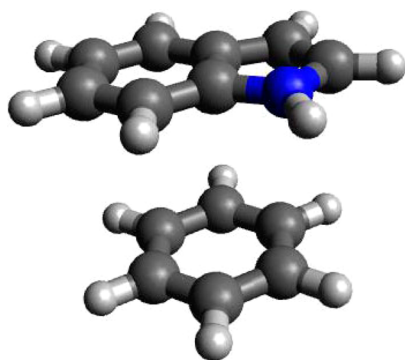


Figure 3. Indole-benzene π - π stacked bonding complex.

values. With the large VSZ-F12(rev 2) basis set the error in $E_{\text{CCSD}}(\text{corr})$ diminished to the point where it no longer provided an effective counterbalance to the (T*) error, leading to larger ϵ_{MAD} and ϵ_{MSD} error metrics than were found with smaller basis sets. In the case of noncovalent binding energies characterized by a minor dispersion component, such as hydrogen bond dimers, very little difference was observed between CCSD(T) and CCSD(T*)-F12b. Consequently, CCSD(T*)-F12b was found to provide no overall advantage compared to CBS extrapolation techniques. Peterson et al.⁵⁶ also noted that (T*) tended to overshoot the true basis set limit and proposed using basis set-specific scaling factors chosen to minimize the RMS deviation solely for the (T) components in a small collection of 28 atomization energies. Those authors found this choice of scale factor reduced the RMS deviation from 0.092 to 0.027 kcal/mol for the cc-pVSZ(rev 2) basis set. We have not examined this possibility here due to time constraints, nor have we considered the possibility of basing the scale factors on the ratio of CCSD correlation energies, which tends to yield smaller values than MP2 energies.

The current statistical analysis focused on total atomization energies and relied on an extensive, diverse reference set developed with large basis set, standard CCSD(T) results. Ancillary results on CCSD(T)-F12b strengthen the conclusions reached in an earlier study⁵⁴ that reported a systematic bias in CBS atomization energies, relative to standard CCSD(T), that grew with the size of the correlation energy. This characteristic suggests that the method should be used with caution in studies aiming for the very highest levels of accuracy. Neither CBS extrapolated CCSD(T)-F12b nor CCSD(T*)-F12b was capable of achieving a level of accuracy on par with standard CCSD(T)/CBS(aV56Z), although the latter entails much more expensive calculations.

Though a statistical perspective is well suited for highlighting the overall strengths and weaknesses of various levels of theory, the oftentimes unexpected differences among superficially similar molecules is one of the reasons chemistry provides such a fascinating field of study. Inferences drawn from the present study for molecules outside the reference set should be viewed critically. The predictive power of the statistics depends on both the extent to which the target molecule's electronic structure is well represented by the reference set, as well as the accuracy of the reference set. No finite reference set is capable of adequately modeling the enormous diversity of molecules composed of elements from throughout the Periodic Table. Nor is any reference set ever completely accurate. In time, the size, diversity and accuracy of reference sets such as the one used in this study is expected to improve, along with their utility. For example, if one were interested in an alcohol or a saturated hydrocarbon, it might be possible to limit the statistical analysis to only include alcohols or hydrocarbons, assuming that the reference set contained a sufficient number of similar molecules to make the statistics meaningful. The approach followed here is quite general and can, in principle,

be applied to a wide assortment of properties and levels of theory.

■ ASSOCIATED CONTENT

● Supporting Information

The newly developed boron aug-cc-pV9Z basis set. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: dfeller@owt.com.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Professors Kirk Peterson and Jan Martin for access to their cc-pV5Z-F12 basis sets and auxiliary functions prior to publication.

■ REFERENCES

- (1) Woon, D. E.; Dunning, T. H., Jr. Benchmark calculations with correlated molecular wave functions. I. Multireference configuration interaction calculations for second row diatomic hydrides. *J. Chem. Phys.* **1993**, *99*, 1914–1928.
- (2) Purvis, G. D., III; Bartlett, R. J. A full coupled-cluster singles and doubles model: The inclusion of disconnected triples. *J. Chem. Phys.* **1982**, *76*, 1910–1918.
- (3) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.* **1989**, *157*, 479–483.
- (4) Watts, J. D.; Gauss, J.; Bartlett, R. J. Coupled-cluster methods with non iterative triple excitations for restricted open-shell Hartree-Fock and other general single determinant reference functions. Energies and analytical gradients. *J. Chem. Phys.* **1993**, *98*, 8718–8733.
- (5) Bartlett, R. J.; Musiał, M. Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.* **2007**, *79*, 291–352.
- (6) Dunning, T. H., Jr. Gaussian basis sets for use in correlated calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- (7) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. Electron affinities of the first row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* **1992**, *96*, 6796–6806.
- (8) Woon, D. E.; Dunning, T. H., Jr. Gaussian basis sets for use in correlated calculations. IV. Calculation of static electric response properties. *J. Chem. Phys.* **1994**, *100*, 2975–2988.
- (9) Woon, D. E.; Dunning, T. H., Jr. Gaussian basis sets for use in correlated calculations. V. Core-valence basis sets for boron through neon. *J. Chem. Phys.* **1995**, *103*, 4572–4585.
- (10) Peterson, K. A.; Dunning, T. H., Jr. Accurate correlation consistent basis sets for molecular core-valence correlation effects. The second row atoms Al–Ar, and the first row atoms B–Ne revisited. *J. Chem. Phys.* **2002**, *117*, 10548–10560.
- (11) Koput, J.; Peterson, K. A. Ab Initio Potential Energy Surface and Vibrational-Rotational Energy Levels of $X^2\Sigma^+$ CaOH. *J. Phys. Chem. A* **2002**, *106*, 9595–9599.
- (12) Woon, D. E.; Dunning, T. H., Jr. Gaussian basis sets for use in correlated calculations. III. The atoms aluminum through argon. *J. Chem. Phys.* **1993**, *98*, 1358–1371.
- (13) Wilson, A. K.; Woon, D. E.; Peterson, K. A.; Dunning, T. H., Jr. Gaussian basis sets for use in correlated molecular calculations. IX. The atoms gallium through krypton. *J. Chem. Phys.* **1999**, *110*, 7667–7676.
- (14) Balabanov, N. B.; Peterson, K. A. Systematically convergent basis sets for transition metals. I. All-electron correlation consistent basis sets for the 3d elements Sc – Zn. *J. Chem. Phys.* **2005**, *123*, 064107–064122.
- (15) Wilson, A. K.; van Mourik, T.; Dunning, T. H., Jr. Gaussian basis sets for use in correlated calculations. VI. Sextuple zeta correlation consistent basis sets for boron through neon. *J. Mol. Struct.: THEOCHEM* **1996**, *388*, 339–349.
- (16) Feller, D.; Peterson, K. A. A re-examination of atomization energies for the Gaussian-2 set of molecules. *J. Chem. Phys.* **1999**, *110*, 8384–8396.
- (17) Peterson, K. A.; Figgen, D.; Goll, E.; Stoll, H.; Dolg, M. Systematically convergent basis sets with relativistic pseudopotentials. II. Correlation consistent basis sets for the post-d group 16 - 18 elements. *J. Chem. Phys.* **2003**, *119*, 11113–11123.
- (18) Peterson, K. A. Systematically convergent basis sets with relativistic pseudopotentials. I. Correlation consistent basis sets for the post-d group 13 - 15 elements. *J. Chem. Phys.* **2003**, *119*, 11099–11112.
- (19) Peterson, K. A.; Puzzarini, C. Systematically convergent basis sets for transition metals. II. Pseudopotential-based correlation consistent basis sets for the group 11 (Cu, Ag, Au) and 12 (Zn, Cd, Hg) elements. *Theor. Chem. Acc.* **2005**, *114*, 283–296.
- (20) Barnes, E. C.; Petersson, G. A.; Feller, D.; Peterson, K. A. The CCSD(T) complete basis set limit for Ne revisited. *J. Chem. Phys.* **2008**, *129*, 194115–1–194115–6.
- (21) Feller, D. Application of systematic sequences of wave functions to the water dimer. *J. Chem. Phys.* **1992**, *96*, 6104–6114.
- (22) Peterson, K. A.; Woon, D. E.; Dunning, T. H., Jr. Benchmark calculations with correlated molecular wave functions. IV The classical barrier height of the $H + H_2 \rightarrow H_2 + H$ reaction. *J. Chem. Phys.* **1994**, *100*, 7410–7415.
- (23) Martin, J. M. L. Ab initio total atomization energies of small molecules - towards the basis set limit. *Chem. Phys. Lett.* **1996**, *259*, 669–678.
- (24) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. Basis Set Convergence of Correlated Calculations on Water. *J. Chem. Phys.* **1997**, *106*, 9639–9646.
- (25) Klopper, W.; Bak, K. L.; Jørgensen, P.; Olsen, J.; Helgaker, T. Highly accurate calculations of molecular electronic structure. *J. Phys. B* **1999**, *32*, R103–R130.
- (26) Schwenke, D. W. The extrapolation of one-electron basis sets in electronic structure calculations: How it should work and how it can be made to work. *J. Chem. Phys.* **2005**, *122*, 014107–1–014107–7.
- (27) Feller, D.; Peterson, K. A. An expanded calibration study of the explicitly correlated CCSD(T)-F12b method using large basis set standard CCSD(T) atomization energies. *J. Chem. Phys.* **2013**, *139*, 084110–1–084110–10.
- (28) Feller, D. The Role of Databases in Support of Computational Chemistry Calculations. *J. Comput. Chem.* **1996**, *17*, 1571–1586.
- (29) Schwartz, C. Importance of angular correlations between atomic electrons. *Phys. Rev.* **1962**, *126*, 1015–1019.
- (30) Schwartz, C. In *Methods in Computational Physics*; Alder, B. J., Fernbach, S., Rotenberg, M., Eds.; Academic Press: New York, 1963; Vol. 2, p 262.
- (31) Carroll, D. P.; Silverstone, H. J.; Metzger, R. M. Piecewise polynomial configuration interaction natural orbital study of $1s^2$ helium. *J. Chem. Phys.* **1979**, *71*, 4142–4163.
- (32) Hill, R. N. Rates of convergence and error estimation formulas for the Rayleigh-Ritz variational method. *J. Chem. Phys.* **1985**, *83*, 1173–1196.
- (33) Kutzelnigg, W.; Morgan, J. D. Rates of convergence of partial-wave expansions of atomic correlation energies. *J. Chem. Phys.* **1992**, *96*, 4484–4508.
- (34) Feller, D.; Peterson, K. A.; Dixon, D. A. Further benchmarks of a composite, convergent, statistically-calibrated coupled cluster approach for thermochemical and spectroscopic studies. *Mol. Phys.* **2012**, *110*, 2381–2399.
- (35) Peterson, K. A.; Feller, D.; Dixon, D. A. Chemical accuracy in ab initio thermochemistry and spectroscopy: Current strategies and future challenges. *Theor. Chem. Acc.* **2012**, *131*, 1079–1099.
- (36) Dixon, D. A.; Feller, D.; Peterson, K. A. A Practical Guide to Reliable First Principles Computational Thermochemistry Predictions

Across the Periodic Table. In *Annual Report on Computational Chemistry*; Wheeler, R. A., Ed.; Elsevier: Amsterdam, 2012; Vol. 8.

(37) Feller, D.; Peterson, K. A.; Ruscic, B. Improved accuracy benchmarks of small molecules using correlation consistent basis sets. *Theor. Chem. Acc.* **2013**, *133*, 1407–1423.

(38) Karton, A.; Rabinovich, E.; Martin, J. M. L. W4 theory for computational thermochemistry: In pursuit of confident sub-kJ/mol predictions. Ruscic, B. *J. Chem. Phys.* **2006**, *125*, 144108–1–144108–17.

(39) Bomble, Y. J.; Vázquez, J.; Kállay, M.; Michauk, C.; Szalay, P. G.; Császár, A. G.; Gauss, J. High-accuracy extrapolated ab initio thermochemistry. II. Minor improvements to the protocol and a vital simplification. Stanton, J. F. *J. Chem. Phys.* **2006**, *125*, 064108–1–064108–8.

(40) Ruscic, B.; Pinzon, R. E.; Morton, M. L.; Laszewski, G. v.; Bittner, S.; Nijsure, S. G.; Amin, K. A.; Minkoff, M.; Wagner, A. F. Introduction to Active Thermochemical Tables: Several “Key” Enthalpies of Formation Revisited. *J. Phys. Chem. A* **2004**, *108*, 9979–9997.

(41) Ruscic, B.; Pinzon, R. E.; Laszewski, G. v.; Kodeboyina, D.; Burcat, A.; Leahy, D.; Montoya, D.; Wagner, A. F. Active Thermochemical Tables: Thermochemistry for the 21st Century. *J. Phys. Conf. Ser.* **2005**, *16*, S61–S70.

(42) Feller, D.; Peterson, K. A.; Hill, J. G. On the effectiveness of CCSD(T) complete basis set extrapolations for atomization energies. *J. Chem. Phys.* **2011**, *135*, 044102–1–044102–18.

(43) Feller, D. Benchmarks of improved complete basis set extrapolation schemes designed for standard CCSD(T) atomization energies. *J. Chem. Phys.* **2013**, *138*, 074103–1–074103–9.

(44) Truhlar, D. G. Basis-set extrapolation. *Chem. Phys. Lett.* **1998**, *294*, 45–48.

(45) Chuang, Y.; Truhlar, D. G. Geometry Optimization with an Infinite Basis Set. *J. Phys. Chem. A* **1999**, *103*, 651–652.

(46) Halkier, A.; Helgaker, T.; Klopper, W.; Jørgensen, P.; Császár, A. G. *Chem. Phys. Lett.* **1999**, *310*, 385.

(47) Adler, T. B.; Knizia, G.; Werner, H.-J. A simple and efficient CCSD(T)-F12 approximation. *J. Chem. Phys.* **2007**, *127*, 221106–1–221106–4.

(48) Werner, H.-J.; Adler, T. B.; Manby, F. R. General orbital invariant MP2-F12 theory. *J. Chem. Phys.* **2007**, *126*, 164102–1–164102–18.

(49) Knizia, G.; Werner, H. J. Explicitly correlated RMP2 for high-spin open-shell reference states. *J. Chem. Phys.* **2008**, *128*, 154103–1–154103–12.

(50) Knizia, G.; Adler, T. B.; Werner, H.-J. Simplified CCSD(T)-F12 methods: Theory and benchmarks. *J. Chem. Phys.* **2009**, *130*, 054104–1–054104–20.

(51) Rauhut, G.; Knizia, G.; Werner, H.-J. Accurate calculation of vibrational frequencies using explicitly correlated coupled-cluster theory. *J. Chem. Phys.* **2009**, *130*, 054105–1–054105–10.

(52) Hättig, C.; Tew, D. P.; Köhn, A. Accurate and efficient approximations to explicitly correlated coupled-cluster singles and doubles. *J. Chem. Phys.* **2010**, *132*, 231102–1–231102–4.

(53) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M.; Celani, P.; Korona, T.; Lindh, R.; Mitrushenkov, A.; Rauhut, G.; et al. *MOLPRO* Version 2012.1; Universität Stuttgart and Cardiff University: Stuttgart, Germany, and Cardiff, United Kingdom, 2012; <http://www.molpro.net>.

(54) Feller, D.; Peterson, K. A.; Hill, J. G. Calibration study of the CCSD(T)-F12a/b methods for C2 and small hydrocarbons. *J. Chem. Phys.* **2010**, *133*, 184102–1–184102–17.

(55) Peterson, K. A.; Adler, T. B.; Werner, H.-J. Systematically convergent basis sets for explicitly correlated wavefunctions. The atoms H, He, B–Ne, and Al–Ar. *J. Chem. Phys.* **2008**, *128*, 084102–1–084102–12.

(56) Peterson, K. A.; Kesharwani, M. K.; Martin, J. M. L. The cc-pVSZ-F12 basis set: reaching the basis set limit in explicitly correlated calculations. *J. M. L. Mol. Phys.* **2014**, DOI: 10.1080/00268976.2014.985755.

(57) Hill, J. G.; Peterson, K. A.; Knizia, G.; Werner, H.-J. Extrapolating MP2 and CCSD explicitly correlated correlation energies to the complete basis set limit with first and second order correlation consistent basis sets. *J. Chem. Phys.* **2009**, *131*, 194105–1–194105–13.

(58) Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618–622.

(59) Marchetti, O.; Werner, H.-J. Accurate calculations of intermolecular interaction energies using explicitly correlated wave functions. *Phys. Chem. Chem. Phys.* **2008**, *10*, 3400–3409.

(60) Marchetti, O.; Werner, H.-J. Accurate Calculations of Intermolecular Interaction Energies Using Explicitly Correlated Coupled Cluster Wave Functions and a Dispersion-Weighted MP2Method. *J. Phys. Chem. A* **2009**, *113*, 11580–11585.

(61) Feller, D.; Peterson, K. A.; Crawford, T. D. Sources of Error in Electronic Structure Calculations on Small Chemical Systems. *J. Chem. Phys.* **2006**, *124*, 054107–1–054107–17.

(62) Bakowies, D. Extrapolation of electron correlation energies to finite and complete basis set targets. *J. Chem. Phys.* **2007**, *127*, 084105–1–084105–23.

(63) Fliegl, H.; Klopper, W.; Hättig, C. Coupled-cluster theory with simplified linear- r_{12} corrections: The CCSD(R12) model. *J. Chem. Phys.* **2005**, *122*, 084107–1–084107–10.

(64) Köhn, A.; Tew, D. P. Towards the Hartree–Fock and coupled-cluster singles and doubles basis set limit: A study of various models that employ single excitations into a complementary auxiliary basis set. *J. Chem. Phys.* **2010**, *132*, 024101–1–024101–10.

(65) Ten-no, S. Initiation of explicitly correlated Slater-type geminal theory. *Chem. Phys. Lett.* **2004**, *398*, 56–61.

(66) Yousaf, K. E.; Peterson, K. A. Optimized auxiliary basis sets for explicitly correlated methods. *J. Chem. Phys.* **2008**, *129*, 184108–1–184108–7.

(67) Valeev, E. F. Improving on the resolution of the identity in linear R12 ab initio theories. *Chem. Phys. Lett.* **2004**, *395*, 190–195.

(68) Weigend, F. A fully direct RI-HF algorithm: Implementation, optimized auxiliary basis sets, demonstration of accuracy and efficiency. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4285–4291.

(69) Hättig, C. Optimization of auxiliary basis sets for RI-MP2 and RI-CC2 calculations: Core–valence and quintuple- ζ basis sets for H to Ar and QZVPP basis sets for Li to Kr. *Chem. Phys. Phys. Chem.* **2005**, *7*, 59–66.

(70) Weigend, F.; Kohn, A.; Hättig, C. Efficient use of the correlation consistent basis sets in resolution of the identity MP2 calculations. *J. Chem. Phys.* **2002**, *116*, 3175–3183.

(71) Feller, D.; Peterson, K. A.; Davidson, E. R. A systematic approach to vertically excited states of ethylene using configuration interaction and coupled cluster techniques. *J. Chem. Phys.* **2014**, *141*, 104302–1–104302–20.

(72) Knowles, P. J.; Hampel, C.; Werner, H.-J. Coupled cluster theory for high spin, open shell reference wave functions. *J. Chem. Phys.* **1993**, *99*, 5219–5227.

(73) Deegan, M. J. O.; Knowles, P. J. Perturbative corrections to account for triple excitations in closed and open shell coupled cluster theories. *Chem. Phys. Lett.* **1994**, *227*, 321–326.

(74) Angeli, C.; Bak, K. L.; Bakken, V.; Christiansen, O.; Cimiraglia, R.; Coriani, S.; Dahle, P.; Dalskov, E. K.; Enevoldsen, T.; Fernandez, B.; et al. *Dalton* 2.0, 2005.

(75) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; et al. *Gaussian 09*, revision D.01; Gaussian, Inc.: Wallingford, CT, 2009.

(76) Klopper, W.; Samson, C. C. M. Explicitly correlated second-order Møller–Plesset methods with auxiliary basis sets. *J. Chem. Phys.* **2002**, *116*, 6397–6410.

(77) Jurecka, P.; Sponer, J.; Cerny, J.; Hobza, P. Benchmark database of accurate (MP2 and CCSD(T) complete basis set limit) interaction energies of small model complexes, DNA base pairs, and amino acid pairs. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985–1993.

(78) Feyereisen, M. W.; Feller, D.; Dixon, D. A. Hydrogen bond energy of the water dimer. *J. Phys. Chem.* **1996**, *100*, 2993–2997.