Prediction of Reaction Barriers and Thermochemical Properties with Explicitly Correlated Coupled-Cluster Methods: A Basis Set Assessment

Jinmei Zhang and Edward F. Valeev*

Department of Chemistry, Virginia Tech, Blacksburg, Virginia 24061, United States

S *Supporting Information*

ABSTRACT: We assessed the performance of our perturbative explicitly correlated coupled-cluster method, $CCSD(T)_{\overline{F12}}$, for accurate prediction of chemical reactivity. The reference data included reaction barrier heights, electronic reaction energies, atomization energies, and enthalpies of formation from the following sources: (1) the DBH24/08 database of 22 reaction barriers (Truhlar et al.), (2) the HJO12 set of isogyric reaction energies (Helgaker et al.), and (3) a HEAT set of atomization energies and heats of formation (Stanton et al.). We performed two types of analyses targeting the two distinct uses of explicitly correlated $CCSD(T)$ models: as a replacement for basis-set-extrapolated $CCSD(T)$ in highly accurate composite methods like HEAT and as a distinct model chemistry for standalone applications. Hence, we analyzed in detail (1) the *basis set error* of each component of the $CCSD(T)_{\overline{F12}}$ contribution to the chemical energy difference in question and (2) the *total error* of the $CCSD(T)_{\overline{F12}}$ model chemistry relative to the benchmark values. Two basis set families were utilized in the calculations: the standard aug-cc-p(C)VXZ-F12 (X = D, T, Q) basis sets for the conventional correlation methods and the cc-p(C)VXZ-F12 (X = D, T, Q) basis sets of Peterson and co-workers that are specifically designed for explicitly correlated methods. Our conclusion is that the performance of the two families for CCSD correlation contributions (which are the only components affected by the explicitly correlated terms in our formation) are nearly identical with triple- and quadruple-ζ quality basis sets, with some differences at the double-ζ level. Chemical accuracy ($\sim$4.18 kJ/mol) for reaction barrier heights, electronic reaction energies, atomization energies, and enthalpies of formation is attained on average with the aug-cc-pVDZ, aug-cc-pVTZ, cc-pCVTZ-F12/aug-cc-pCVTZ, and cc-pCVDZ-F12 basis sets, respectively, at the $CCSD(T)_{\overline{F12}}$ level of theory. The corresponding mean unsigned errors are 1.72 kJ/mol, 1.5 kJ/mol, $\sim$2 kJ/mol, and 2.17 kJ/mol, and the corresponding maximum unsigned errors are 4.44 kJ/mol, 3.6 kJ/mol, $\sim$5 kJ/mol, and 5.75 kJ/mol.

1. INTRODUCTION

Quantitative computational low-temperature thermochemistry and kinetics are challenging because they require prediction of electronic energy differences (reaction energies, atomization energies, enthalpies of formation, and reaction barriers) with *chemical accuracy*, usually defined as errors not exceeding 1 kcal/mol (4.18 kJ/mol) or even as 1 kJ/mol. Despite the impressive recent progress of density functional theory (DFT) methodology, its applicability to computational thermochemistry and kinetics is still limited. For example, a comprehensive testing of DFT and wave function methods by Zheng and co-wrkers[1] against the DBH24/08 benchmark database of reaction barriers[2,3] revealed that the performance of the best empirically tailored hybrid DFT model chemistries approaches the 1 kcal/mol mean unsigned error, whereas the majority of hybrid DFT functionals result in average errors of several kilocalories per mole. The high-end many-body wave function methods appear to be the only way at the moment to reliably reduce the errors below the chemical accuracy mark; e.g., the "gold standard" coupled-cluster singles, doubles, and perturbative triples- [CCSD(T)] method with a relatively large aug-cc-pCV(T +d)Z basis, while orders of magnitude more expensive than the DFT counterparts, was in error by only $\sim$0.5 kcal/mol.

We should note that the electronic components of the molecular energies are the principal—but not the only— stumbling block here. The most accurate first-principles thermochemical models[4] inform us that prediction of relative energies to sub-1-kcal/mol accuracy reliably requires accounting for the effects of (1) anharmonicity, (2) special relativity (kinematic, Darwin, and spin−orbit terms), and (3) adiabatic coupling of electronic and nuclear motions. Yet the biggest remaining challenge is accurate computation of relative "bare" electronic energies: it requires (4) high-end wave function models and (5) a reduction of the basis set error (typically, via the basis set extrapolation). This paper is concerned with the last issue, the basis set error, its role in prediction of the chemical energy differences, and to what extent it is alleviated by the use of relatively novel explicitly correlated wave function methods.

The most troublesome origin of the basis set error of many-body wave functions is the sole use of Slater determinants as the building blocks. Standard wave function methods, such as CCSD(T) or multireference configuration interaction (MRCI), predict electron correlation energies with basis sets errors that decrease in atoms as $O[(L_{max} + 1)^{-3}]$ for a basis set saturated to angular momentum L_{max}.[5] The basis set problem is due to the orbital product nature of Slater determinants; the Coulomb hole[6] that appears at short interelectronic distances cannot be

efficiently described by orbital products alone.[7] Explicitly correlated methods overcome this problem by expressing the Coulomb hole in terms of the interelectronic distances (r_{ij}). For atoms, the explicitly correlated methods of the F12 type (pioneered by Kutzelnigg[8] and commonly known as "R12 methods") have a basis set error of $O[(L_{max} + 1)^{-7}]$. The need to compute expensive three- and four-electron integrals, characteristic of all explicitly correlated methods, is avoided in F12 methods by the resolution of the identity;[9−12] only two-electron, albeit nonstandard, integrals are needed in F12 methods.

Rapid improvement of the F12 technology over the past decade has made it sufficiently robust for use by nonspecialists. Only some key developments can be mentioned here: the development of robust RI approximations[11,12] and their utilization to compute many-electron integrals;[13] the development of efficient nonlinear correlation factors;[14−16] the introduction of the full coupled-cluster (CC) F12 method;[17−20] the development of practical approximations to CC-F12, both iterative and noniterative;[21−27] extensions to high-order excitations[19,28] and multireference models;[29−33] relativistic extensions of F12 methods;[34,35] and the development of basis sets specifically tailored for F12 methods.[36−38] The modern F12 methods are efficient and competitive with extrapolation: to achieve the same basis set error, the approximate CC-F12 methods require a basis set of two cardinal numbers lower than the comparable standard CC computation.

Yet despite the tremendous progress, the F12 technology is not yet sufficiently robust for black-box applications across the periodic table. One of the significant challenges, among many, is the requirement for particular orbital basis set types (OBS). Specifically, for robust performance in the F12 context, the OBS needs to include diffuse atomic orbitals; this requirement makes difficult the F12 method applications to large systems as well as to some electronically excited states. The only recommended options for F12 computations remain the cc-pVXZ-F12 series of F12-optimized basis sets of Peterson et al.[36−38] and the standard aug-cc-pVXZ series of Dunning et al.[39−41] In this work, our goal is to systematically evaluate these two series of basis sets, with an eye toward future improvement along the lines of the recent work of some of us.[42]

In this paper, we document the extension of the perturbative coupled-cluster F12 method, $CCSD(T)_{\overline{F12}}$, developed in our group for the treatment of high-spin open-shell molecules and implemented in the open-source freely available Massively Parallel Quantum Chemistry (MPQC) package.[43] We assessed its performance for accurate studies of chemical reactivity by performing benchmark calculations of reaction barrier heights and thermochemical properties including electronic reaction energies, atomization energies, and heats of formation. The performance of the $CCSD(T)_{\overline{F12}}$ method for reaction barrier heights is assessed by benchmarking against the DBH24/08 database (Table 1). The performance for electronic reaction energies is gauged for the HJO12 set of 12 isogyric reactions[44] (Table 2), which range from nearly thermoneutral to highly exothermic and involve breaking and forming an assortment of chemical bonds. The performance for computation of the atomization energies and enthalpies of formation of the molecules is measured for the standard high-accuracy extrapolated ab initio thermochemistry (HEAT) set.[45] The HEAT test set includes 31 atoms and molecules for which both experimentally derived and theoretical enthalpies of formation (at 0 K) are available in better than a 1 kJ/mol agreement.

Table 1. The Reactions Including Forward and Reverse Directions in the DBH24/08 Database[1−3,a]

reaction no.	reaction
	heavy-atom transfer
1	$H + N_2O \leftrightarrow OH + N_2$
2	$H + ClH \leftrightarrow HCl + H$
3	$CH_3 + FCl \leftrightarrow CH_3F + Cl$
	nucleophilic substitution
4	$Cl^- \cdots CH_3Cl \leftrightarrow ClCH_3 \cdots Cl^-$
5	$F^- \cdots CH_3Cl \leftrightarrow FCH_3 \cdots Cl^-$
6	$OH^- + CH_3F \leftrightarrow HOCH_3 + F^-$
	unimolecular and association
7	$H + N_2 \leftrightarrow HN_2$
8	$H + C_2H_4 \leftrightarrow CH_3 + CH_2$
9	$HCN \leftrightarrow HNC$
	hydrogen transfer
10	$OH + CH_4 \leftrightarrow CH_3 + H_2O$
11	$H + OH \leftrightarrow O + H_2$
12	$H + H_2S \leftrightarrow H_2 + HS$

[a]There are 22 unique reaction barriers (the forward and reverse reactions in reactions 2 and 4 are the same).

Table 2. The HJO12 Set of Isogyric Reactions[44] **Involving 20 Small Molecules**

reaction no.	reaction
1	$CO + H_2 \rightarrow CH_2O$
2	$N_2 + 3H_2 \rightarrow 2NH_3$
3	$C_2H_2 + H_2 \rightarrow C_2H_4$
4	$CO_2 + 4H_2 \rightarrow CH_4 + 2H_2O$
5	$CH_2O + 2H_2 \rightarrow CH_4 + H_2O$
6	$CO + 3H_2 \rightarrow CH_4 + H_2O$
7	$HCN + 3H_2 \rightarrow CH_4 + NH_3$
8	$HNO + 2H_2 \rightarrow NH_3 + H_2O$
9	$C_2H_2 + 3H_2 \rightarrow CH_4 + NH_3$
10	$CH_2 + H_2 \rightarrow CH_4$
11	$F_2 + H_2 \rightarrow 2HF$
12	$2CH_2 \rightarrow C_2H_4$

We performed two types of analyses targeting the two distinct uses of explicitly correlated CCSD(T) models: as a replacement for basis-set-extrapolated CCSD(T) in highly accurate composite methods like HEAT and as a distinct model chemistry for standalone applications. Hence, we analyzed in detail (1) the *basis set error* of each component of the $CCSD(T)_{\overline{F12}}$ contribution to the chemical energy difference in question and (2) the *total error* of the $CCSD(T)_{\overline{F12}}$ model chemistry relative to the benchmark values.

We begin with a brief description of the theoretical approach and computational details in section 2. Results of computations and their discussion are given in section 3. We summarize our findings and discuss their relevance to the future applications of F12 methods in section 4.

2. COMPUTATIONAL METHODS

The original diagonal orbital-invariant formulation of the perturbative coupled-cluster F12 method, $CCSD(T)_{\overline{F12}}$, was reported for closed-shell species in ref 46. An open-shell variant of the $CCSD(T)_{\overline{F12}}$ method has been reported by our group before[24] based on the nondiagonal ansatz with optimized geminal coefficients. In this work, we implemented the open-shell method using the diagonal orbital-invariant (SP) ansatz of Ten-no;[12] this development is similar to work by others[26,47] in the context of the iterative CC-F12 methods. Thus, we only report the essential details of the open-shell variant of the method; the full programmable equations as implemented in the MPQC package are reported in the Supporting Information.

2.1. The $CCSD(T)_{\overline{F12}}$ Formalism. The CCSD wave function captures the effects of two-electron correlations and the resulting orbital relaxation in a finite basis:

$$\Psi_{\text{CCSD}} \equiv \exp(\hat{T})|\Psi_0\rangle \tag{1}$$

$$\hat{T} \equiv \hat{T}_1 + \hat{T}_2 \tag{2}$$

$$\hat{T}_1 \equiv t_a^i \tilde{a}_i^a \tag{3}$$

$$\hat{T}_2 \equiv \frac{1}{(2!)^2} t_{ab}^{ij} \tilde{a}_{ij}^{ab} \tag{4}$$

where Ψ_0 is the Hartree–Fock reference function (the notation is explained in the Appendix). In the $CCSD(T)_{\overline{F12}}$ method, the CCSD wave function is extended perturbatively to account for the effect of three-electron correlations in the finite basis (via standard triple excitations), residual orbital relaxation effects (via standard singlet excitations into complementary auxiliary basis set (CABS)), as well as for short-distance two-electron correlations (via explicitly correlated geminals).[24] The corresponding $CCSD(T)_{\overline{F12}}$ energy has four contributions:

$$E_{CCSD(T)_{\overline{F12}}} = E_{CCSD} + E_{(T)} + E_{(2)_S} + E_{(2)\overline{F12}'} \tag{5}$$

where the first two terms add up to the standard CCSD(T) energy, the next contribution is the usual "CABS singles" correction[22] (we will not discuss its evaluation here; more details can be found in the Supporting Information for ref 48), and the last contribution is due to the explicitly correlated F12 terms. In the diagonal orbital-invariant approach to CCSD-$(T)_{\overline{F12}}$,[46] the $E_{(2)\overline{F12}}$ contribution is evaluated directly from the usual Hylleraas functional, without any parameter optimization:

$$E_{(2)\overline{F12}} = \langle 1|\hat{H}^{(0)}|1\rangle + \langle 0|\hat{H}^{(1)}|1\rangle + \langle 1|\hat{H}^{(1)}|0\rangle \tag{6}$$

The zeroth- and first-order Hamiltonians are defined from the similarity-transformed CCSD Hamiltonian $\bar{H} \equiv \exp(-\hat{T})\hat{H}\exp(\hat{T})$ by matrix (Löwdin) partitioning[49] so that $\hat{H}^{(0)}$ in eq 6 becomes the standard Fock operator, $\hat{F}$, and $\hat{H}^{(1)}$ becomes $\bar{H}$. The left- and right-hand zeroth-order wave functions are the standard left- and right-hand CCSD "wave functions":

$$|0\rangle \equiv |\Psi_0\rangle \tag{7}$$

$$\langle 0| \equiv \langle \Psi_0|(1 + \hat{\Lambda}) \tag{8}$$

where Ψ_0 is the Hartree–Fock reference function, and Λ represents the undetermined multipliers of the CCSD Lagrangian.[44] Since $\hat{\Lambda}$ is not necessary for computing the CCSD energy (it is, however, necessary for computing

molecular properties), by analogy with the CCSD(T) method we invoke approximation $\hat{\Lambda} \approx \hat{T}^{\dagger}$.[50]

The first-order wave function

$$|1\rangle \equiv \frac{1}{2!} \sum_{ij} |\Gamma_{ij}^{ij}\rangle \tag{9}$$

is composed of geminal functions, defined as

$$|\Gamma_{ij}^{ij}\rangle = \frac{1}{2!} \bar{R}_{\alpha\beta}^{ij} \tilde{a}_{ij}^{\alpha\beta}|\Psi_0\rangle \tag{10}$$

where $\bar{R}_{\alpha\beta}^{ij}$ is the antisymmetrized matrix element of the geminal correlation factor $f(r_{12})$:

$$R_{\alpha\beta}^{ij} = \langle \alpha\beta|\hat{Q}_{12}f(r_{12})|ij\rangle \tag{11}$$

Projector $\hat{Q}_{12}$ ensures that the geminal functions in eq 10 are strongly orthogonal (orthogonal to the $|\Psi_0\rangle$ and to the single excitations) as well as orthogonal to the standard double excitations:[11]

$$\hat{Q}_{12} = (1 - \hat{O}_1)(1 - \hat{O}_2) - \hat{V}_1\hat{V}_2 \tag{12}$$

where $\hat{O}$ and $\hat{V}$ are the projectors on the occupied and virtual orbitals.

The geminal correlation factor is designed so that functions $|\Gamma_{ij}^{ij}\rangle$ describe the short-range two-electron correlation according to the first-order cusp condition for the pair function $|\psi_{ij}\rangle$:

$$r_{12} \to 0: |\psi_{ij}^{(1)}\rangle = C_l r_{12}|\psi_{ij}^{(0)}\rangle \tag{13}$$

where C_l is the cusp coefficient that depends on permutational symmetry of $\psi_{ij}^{(0)}$. In practice, it is sufficient to consider only two cases: for singlet pairs $C_0 = 1/2$ and for triplet pairs $C_1 = 1/4$ (more exotic cases, such as the unnatural parity singlet for which $C_2 = 1/6$, can be realized in heavy atoms[51]). The appropriate selection of the proper coefficient can be attained using a spatial coordinate permutation operator[52] or, equivalently, with spin projectors:

$$\hat{P}_0 = |\alpha\beta\rangle_0 \langle\alpha\beta|_0 \tag{14}$$

$$\hat{P}_1 = |\beta\beta\rangle\langle\beta\beta| + |\alpha\beta\rangle_1\langle\alpha\beta|_1 + |\alpha\alpha\rangle\langle\alpha\alpha| \tag{15}$$

with

$$|\alpha\beta\rangle_0 = \frac{1}{\sqrt{2}}\{\alpha(1)\beta(2) - \beta(1)\alpha(2)\}$$

$$|\alpha\beta\rangle_1 = \frac{1}{\sqrt{2}}\{\alpha(1)\beta(2) + \beta(1)\alpha(2)\}$$

Thus, the spin-adapted geminal correlation factor is written as

$$f(r_{12}) \equiv (C_0\hat{P}_0 + C_1\hat{P}_1)\gamma(r_{12}) \tag{16}$$

In this work, $\gamma(r_{12})$ was the standard exponential correlation factor of Ten-no[14] expanded as a linear combination of $N_g = 6$ Gaussian geminals[15]

$$\gamma(r_{12}) \equiv -l_c \exp(-r_{12}/l_c) \approx \sum_i^{N_g} c_i \exp(-\alpha_i r_{12}^2) \tag{17}$$

The correlation length scale $l_c \sim 1a_0$ was set to a value that depends empirically on the orbital basis set.

With the new form of geminal functions, the second-order F12 correction (eq 6) for the first-order wave function ansatz eq 9 has the following form:
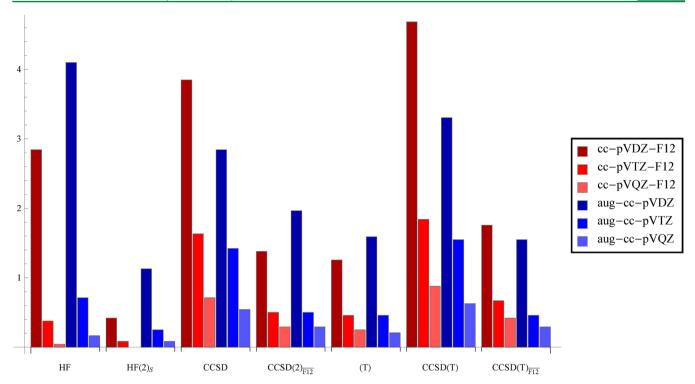
Figure 1. Average unsigned basis set errors (kJ/mol) of various electronic contributions to the reaction barriers. "HF" and "HF(2)$_S$" refer to the Hartree–Fock energies without and with inclusions of the CABS singles correction (eq 5). Other labels refer to the correlation energy contributions only ("(T)" refers to the $E_{(T)}$ energy only). Definitions of the CBS limits for each component are provided in the text.

$$E_{(2)\overline{F12}} = \sum_{i<j} \varepsilon_{ij}^{(2)}$$

$$\varepsilon_{ij}^{(2)} = 2\tilde{V}_{ij}^{ij} + \tilde{B}_{ij}^{ij} \tag{18}$$

where $\tilde{V}_{ij}^{ij}$ is the geminal matrix of the first-order Hamiltonian and has the following form:

$$\tilde{V}_{ij}^{ij} \equiv V_{ij}^{ij} + \frac{1}{2}(V_{ab}^{ij} + C_{ab}^{ij})t_{ij}^{ab} + V_{ia}^{ij}t_j^a + V_{aj}^{ij}t_i^a \tag{19}$$

$$V_{pq}^{ij} \equiv \frac{1}{2}\bar{R}_{\alpha\beta}^{ij}\bar{g}_{pq}^{\alpha\beta} \tag{20}$$

$$C_{ab}^{ij} \equiv F_a^\alpha \bar{R}_{\alpha b}^{ij} + F_b^\alpha \bar{R}_{a\alpha}^{ij} \tag{21}$$

$\tilde{B}_{ij}^{ij}$ is the geminal matrix of the zero-order Hamiltonian. If we assume canonical Hartree–Fock orbitals, it is expressed as

$$\tilde{B}_{ij}^{ij} \equiv B_{ij}^{ij} - (F_i^i + F_j^j)X_{ij}^{ij} \tag{22}$$

$$B_{ij}^{ij} \equiv \bar{R}_{\alpha\beta}^{ij}F_\gamma^\beta \bar{R}_{ij}^{\alpha\gamma} \tag{23}$$

$$X_{ij}^{ij} \equiv \frac{1}{2}\bar{R}_{\alpha\beta}^{ij}\bar{R}_{ij}^{\alpha\beta} \tag{24}$$

The infinite-range sums in the matrices V, C, and X were approximated by finite sums over the complementary auxiliary basis set (CABS),[11] while B used approximation C of Kedzuch et al.[13] as well as the CABS approach (complete details have been described by one of us elsewhere[48]).

Because the (T) correction to CCSD(T) is most commonly formulated in the spin-unrestricted framework, we assumed that the spatial parts of spin–orbitals are spin-dependent. After integrating out the spin degrees of freedom, we arrived at the final expression for eq 18:

$$E_{(2)\overline{F12}} = \sum_{I<J} \varepsilon_{IJ}^{(2)} + \sum_{I,\bar{J}} \varepsilon_{I\bar{J}}^{(2)} + \sum_{\bar{I}<\bar{J}} \varepsilon_{\bar{I}\bar{J}}^{(2)} \tag{25}$$

$$\varepsilon_{IJ}^{(2)} = 2C_1\tilde{V}_{IJ}^{IJ} + C_1^2\tilde{B}_{IJ}^{IJ} \tag{26}$$

$$\varepsilon_{I\bar{J}}^{(2)} = (C_0 + C_1)\tilde{V}_{I\bar{J}}^{I\bar{J}} + (C_0 - C_1)\tilde{V}_{I\bar{J}}^{\bar{J}I} + \frac{(C_0 + C_1)^2}{4}\tilde{B}_{I\bar{J}}^{I\bar{J}}$$
$$+ \frac{C_0^2 - C_1^2}{4}\tilde{B}_{I\bar{J}}^{\bar{J}I} + \frac{C_0^2 - C_1^2}{4}\tilde{B}_{\bar{J}I}^{I\bar{J}} + \frac{(C_0 - C_1)^2}{4}\tilde{B}_{\bar{J}I}^{\bar{J}I}$$

$$\varepsilon_{\bar{I}\bar{J}}^{(2)} = 2C_1\tilde{V}_{\bar{I}\bar{J}}^{\bar{I}\bar{J}} + C_1^2\tilde{B}_{\bar{I}\bar{J}}^{\bar{I}\bar{J}} \tag{27}$$

The detailed expressions for the spin-free matrices $\tilde{V}$ and $\tilde{B}$ are presented in the Supporting Information.

2.2. Computational Details. The (U)QCISD/MG3 molecular geometries for the DBH24/08 database were obtained from ref 2. Two types of orbital basis sets were utilized for the reaction barrier calculation:

- The cc-pVXZ-F12 family of basis sets (X = D, T, Q) designed by Peterson and co-workers[36] to be used specifically with explicitly correlated methods. The complementary auxiliary basis set (CABS) was constructed from the matching cc-pVXZ-F12/OptRI basis sets[37] using the CABS+ approach.[11]
- The standard aug-cc-pVXZ family of basis sets (X = D, T, Q) that was designed by Dunning et al.[39−41] for computations with standard correlated methods (for S and Cl atoms, the corresponding basis sets augmented with tight d functions, aug-cc-pV(X+d)Z,[53] were used). The complementary auxiliary basis set (CABS) was constructed using the matching aug-cc-pVXZ/OptRI basis sets[54] using the CABS+ approach.[11]
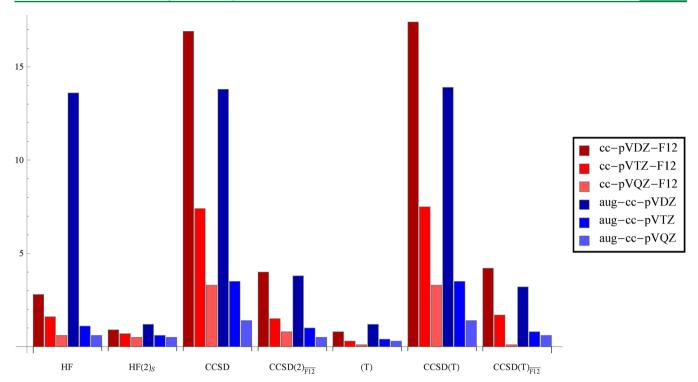
Figure 2. Average unsigned basis set errors (kJ/mol) of various electronic contributions to the reaction energies. "HF" and "HF(2)$_S$" refer to the Hartree−Fock energies without and with inclusions of the CABS singles correction (eq 5). Other labels refer to the correlation energy contributions only ("(T)" refers to the $E_{(T)}$ energy only). Definitions of the CBS limits for each component are provided in the text.

Evaluation of eq 25 utilized the robust density fitting[55] in a cc-pV(X+1)Z-RI basis set for the cc-pVXZ-F12 calculation and an aug-cc-pVXZ-RI for the aug-cc-pVXZ calculation, where X is the cardinal number of the corresponding orbital basis; the conventional CC wave functions and energies were evaluated without density fitting. The recommended correlation length scales[36] were used; correlation factors were expanded in terms of six Gaussian geminals.[56]

Although there are only 22 unique barriers in the DBH24/08 database, to be consistent with Zheng et al.[1] in our analysis we also considered all 24 barriers statistically independent. The recommended spin−orbit energies[1] were used to correct the energies of Cl, O, OH, and HS radicals. Other relativistic effects were neglected. In the calculations of reaction barriers here, only the valence electron correlations were considered (core correlation contributions were estimated to be below 0.1 kcal/mol, as described below).

Experimental geometries were used for the 20 closed-shell molecules involved in the HJO12 set of reactions.[44] The computational procedure was identical to that used for the DBH20/08 set. The core-correlation corrections were obtained from ref 44.

The CCSD(T)/cc-pVQZ method geometries from ref 45 were used for the molecules in the HEAT test set. Two core correlation basis sets were utilized for the calculations of atomization energies and enthalpies of formation. The first sets are the cc-pCVXZ-F12 (X = D, T, Q) basis sets, which are also designed by Peterson et al.[57] for the explicitly correlated methods, and the second sets are the standard aug-cc-pCVXZ (X = D, T, Q) basis sets,[39,41] which are used for the standard correlation methods. The CABS was constructed with the cc-pCVXZ-F12/OptRI basis sets[57] using the CABS+ approach.[11] The aug-cc-pwCV(X+1)Z-RI basis set was used for the robust density fitting[55] of evaluating eq 25. The anharmonic zero-point

energies calculated within the HEAT model were used, while other relativistic corrections were neglected.

All explicitly correlated calculations were performed using the trunk version of the Massively Parallel Quantum Chemistry (MPQC) package[43] (it can be obtained for free under the GNU General Public License (GPL) at http://www.mpqc.org/). The CC wave functions and energies were computed with the PSI3 suite.[58]

3. DISCUSSION OF RESULTS

3.1. Basis Set Errors of Reaction Barriers. The mean unsigned basis set errors of the standard and explicitly correlated CCSD(T) energies (and their selected components) for the DBH24/08 reaction barrier set are presented in Figure 1. The basis set errors (BSE) were defined with respect to the complete basis set (CBS) limits defined as follows: (1) the HF(2)$_S$/cc-pVQZ-F12 energies were used as the CBS Hartree−Fock limit; (2) the CBS limits for the CCSD and (T) correlation energy components were obtained using the standard X^{-3} extrapolation formula[59] from the aug-cc-pVTZ and aug-cc-pVQZ values.

First, we note that the basis set errors of the Hartree−Fock contributions to reaction barriers cannot be neglected (∼ 4 kJ/mol with the aug-cc-pVDZ basis set). Although the HF BSEs are smaller with the cc-pVXZ-F12 basis sets than with the aug-cc-pVXZ counterparts (due to the extra s and p functions in the former), at the double-ζ level the use of CABS singles correction is mandatory. Even with the correction, the aug-cc-pVDZ HF energy is still in error by ∼1 kJ/mol.

The basis set errors of CCSD energies are large: ∼3−4 kJ/mol with the double-ζ basis set, and ∼1.5 kJ/mol with a triple-ζ basis set. Explicitly correlated CCSD energies have greatly reduced basis set errors: for example, the mean unsigned BSE of CCSD(2)$_{\overline{F12}}$/cc-pVDZ-F12 is 1.38 kJ/mol. Note that when
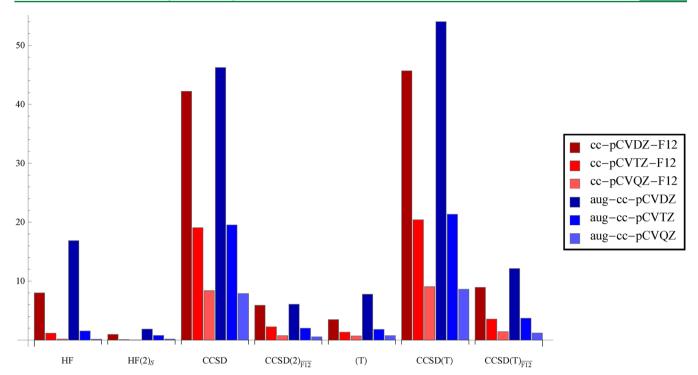
Figure 3. Average unsigned basis set errors (kJ/mol) of various electronic contributions to the atomization energies. "HF" and "HF(2)$_S$" refer to the Hartree–Fock energies without and with inclusions of the CABS singles correction (eq 5). Other labels refer to the correlation energy contributions only ("(T)" refers to the $E_{(T)}$ energy only). Definitions of the CBS limits for each component are provided in the text.

used with non-F12 methods the F12-optimized cc-pVXZ-F12 basis sets result in larger basis set errors than the standard aug-cc-pVXZ basis sets, as expected. In combination with F12 methods only the cc-pVDZ-F12 basis set is clearly preferred over the aug-cc-pVDZ counterpart; at the triple- and quadruple-ζ levels the two basis set families result in essentially identical BSEs, on the order of 0.2 kJ/mol or below. Unfortunately, it is not possible to judge the relative performance of the two families any further because of the limited accuracy of the CBS limits. We estimate the current CBS limit for CCSD energy obtained by extrapolation to be only accurate to 0.1 kcal/mol (~ 0.4 kJ/mol). This figure is in line with the mean unsigned difference of 0.25 kJ/mol between the cc-pVQZ-F12 and aug-cc-pVQZ CCSD(2)$_{\overline{F12}}$ energies. It is unlikely that the CBS limits can be deduced more accurately without performing time-consuming quintuple-ζ computations; however, since the 0.1 kcal/mol accuracy is quite satisfactory, we do not consider it worthwhile at the moment.

Finally, the CCSD(T) BSE is greatly reduced also by the use of explicit correlation. Only a double-ζ basis set is sufficient with the CCSD(T)$_{\overline{F12}}$ method for ~1.6 kJ/mol BSE; triple-ζ basis sets further reduce basis set errors to ~0.5 kJ/mol. The conventional CCSD(T) method required basis sets one cardinal number greater (triple and quadruple-ζ, respectively) to achieve the same precision. Even with the explicitly correlated CCSD method, the remaining basis set errors are comparable to the basis set errors of the (T) contribution to the CCSD(T) energy. Therefore it is, in our opinion, acceptable to treat the (T) energy without the use of explicitly correlated three-body functions, unless extremely high accuracy is needed.

We observed a peculiar behavior of basis set errors of CCSD(2)$_{\overline{F12}}$ and (T) contributions: (1) With the cc-pVXZ-F12 basis sets, the two errors add up constructively; thus the mean

unsigned error of CCSD(T)$_{\overline{F12}}$ is greater than that of CCSD(2)$_{\overline{F12}}$, e.g., for cc-pVDZ-F12: 1.76 kJ/mol vs 1.38 kJ/mol. (2) With the aug-cc-pVXZ basis sets, the errors cancel; thus, the mean unsigned error of CCSD(T)$_{\overline{F12}}$ is smaller than that of CCSD(2)$_{\overline{F12}}$, e.g., for cc-pVDZ-F12: 1.55 kJ/mol vs 1.97 kJ/mol. It is not clear whether this behavior is systematic; it will be important in the discussion of total errors in section 3.5.

3.2. Basis Set Errors of Reaction Energies. Figure 2 shows the basis set errors of the HF energies and various correlation contributions to the electronic reaction energies for the HJO12 set[44] of isogyric reactions. The CBS limits for the basis set errors were defined as follows: (1) The CBS Hartree–Fock values were the reaction energies computed with the aug-cc-pV6Z basis set. (2) The valence (T) CBS limit was obtained by the basis set extrapolation with the Schwenke method[60] using aug-cc-pVQZ and aug-cc-pV5Z basis sets, (3) The CBS limit for the CCSD(T) valence correlation energies was obtained by subtracting from the CBS CCSD(T) energies[44] (obtained by the X^{-3} extrapolation using the cc-pCV5Z and cc-pCV6Z basis sets) the cc-pV6Z HF energies and the core-correlation energies given in ref 44. (4) The CCSD CBS limit was obtained by subtracting the Schwenke extrapolated CBS limit of the valence (T) energy from the CCSD(T) CBS limit.

A very large HF basis set error results from the use of the aug-cc-pVDZ basis set (13.6 kJ/mol). Fortunately, the CABS singles correction reduces the basis error to 1.2 kJ/mol. Although the HF BSE to reaction energies with the cc-pVDZ-F12 basis set is much smaller (2.8 kJ/mol), the CABS singles correction is still needed to reduce the error around 1 kJ/mol. At the triple-ζ level, however, the HF BSE with the F12 basis set is slightly larger than that with the non-F12 counterpart, but the HF BSEs with the two basis sets become identical at the quadruple-ζ level.
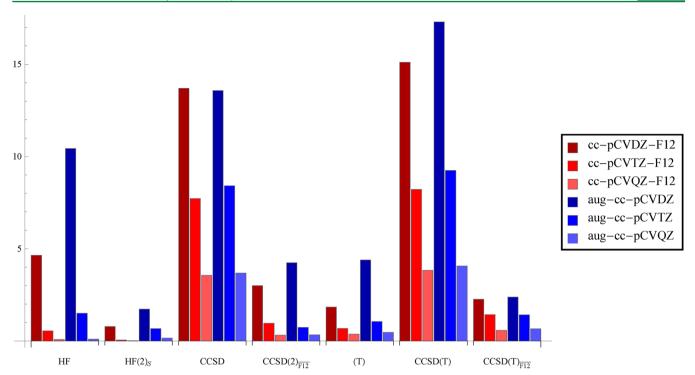
Figure 4. Average unsigned basis set errors (kJ/mol) of various electronic contributions to the enthalpies of formation. "HF" and "HF(2)$_S$" refer to the Hartree–Fock energies without and with inclusions of the CABS singles correction (eq 5). Other labels refer to the correlation energy contributions only ("(T)" refers to the $E_{(T)}$ energy only). Definitions of the CBS limits for each component are provided in the text.

The basis set errors of CCSD energies are large, especially with the double-ζ basis sets (over 13 kJ/mol). Explicitly correlated CCSD energies have much smaller basis set errors (~ -4 kJ/mol for X = D and ~ 1 kJ/mol for X = T, Q). Compared to the CCSD(2)$_{\overline{F12}}$ BSE, the (T) BSEs are noticeably smaller, especially at the double-ζ level. Unexpectedly, we note that the cc-pVXZ-F12 (X = D, T) basis sets give slightly larger CCSD(2)$_{\overline{F12}}$ BSEs than the corresponding aug-cc-pVXZ basis sets for all of the cases but have slightly smaller basis set errors for the (T) contributions than the aug-cc-pVXZ basis sets.

The basis set errors of the standard CCSD(T) energies are very large and slowly convergent, but the use of explicit correlation significantly reduces the basis set errors. The triple-ζ CCSD(T)$_{\overline{F12}}$ calculations have BSEs twice smaller than the quadruple-ζ CCSD(T) calculations. With the quadruple basis sets, the basis set errors decrease to under 1 kJ/mol. Overall, the BSEs from the two basis set families do not differ significantly. While the aug-cc-pVXZ basis sets yield a smaller BSE than the cc-pVXZ-F12 basis sets when X = D and T, the aug-cc-pVQZ basis set gave larger errors than the cc-pVQZ-F12 basis set, which has a BSE close to zero.

3.3. Basis Set Errors of Atomization Energies. The basis set errors of the various electronic contributions to the atomization energies of the HEAT test set are shown in Figure 3. The HF/aug-cc-pCV5Z energies were defined as the CBS HF limit; the CBS limits for the CCSD and (T) correlation energy components were obtained using the X^{-3} extrapolation formula[59] with the aug-cc-pCVQZ and aug-cc-pCV5Z values.

As expected, the F12 basis sets give smaller HF basis set errors than the non-F12 basis sets. The basis set errors of the double-ζ HF contributions are large (>8 kJ/mol), and the use of CABS singles correction is necessary to reduce the errors around 1 kJ/mol. At the triple- and quadruple-ζ levels,

however, the HF BSE become quite small; the use of CABS singles correction can be optional.

The CCSD basis set errors are very large with the small basis sets (>40 kJ/mol and >19 kJ/mol for X = D and T, respectively). Even with the large quadruple-ζ basis sets, the BSEs are still around 8 kJ/mol. Unexpectedly, at the double- and triple-ζ levels, the F12 basis sets yield slightly smaller CCSD BSEs than the standard non-F12-optimized aug-cc-pCVXZ basis sets. The CCSD BSE are significantly reduced by the use of the explicit correlation. In fact, the double-ζ CCSD(2)$_{\overline{F12}}$ BSEs are noticeably smaller than the quadruple-ζ CCSD BSEs. With the quadruple-ζ basis sets, the BSE of the CCSD(2)$_{\overline{F12}}$ correlation energies are reduced to below 1 kJ/mol. The basis sets yield similar CCSD(2)$_{\overline{F12}}$ BSEs.

The (T) BSEs are slightly smaller than the explicitly correlated CCSD BSEs in most cases. The only exception is the (T) BSE with the aug-cc-pCVDZ basis set. As a result, while the BSE of the CCSD(T)$_{\overline{F12}}$ correlation energies with the aug-cc-pCVDZ basis set (12.12 kJ/mol) is noticeably larger than the BSE of the CCSD(T) correlation energies with the aug-cc-pCVQZ basis set (8.63 kJ/mol), the CCSD(T)$_{\overline{F12}}$ BSE with the cc-pCVDZ-F12 basis set (8.93 kJ/mol) is smaller than the CCSD(T) BSE with the cc-pCVQZ-F12 basis set (9.06 kJ/mol), which follows the same pattern for the CCSD(2)$_{\overline{F12}}$ BSE. With the quadruple-ζ basis sets, the basis set errors of the explicitly correlated CCSD(T) correlation energies reduce to ~ 1 kJ/mol. Different from the calculations in the previous sections, we find that the aug-cc-pCVXZ calculations yield larger basis set errors than the cc-pCVXZ-F12 basis set when X = D and T but give slightly smaller errors than the corresponding F12 basis set when X = Q.

3.4. Basis Set Errors of Enthalpies of Formation. Figure 4 presents the basis set errors of the HF energies and various

contributions to the enthalpies of formation. The CBS limits were defined in the same way as the previous section.

Again, the HF basis set errors with the double-ζ basis set are large, and the CABS singles correction reduces the errors to around 1−2 kJ/mol. As we increase the size of the basis set, the use of CABS singles correction becomes optional since the HF BSEs become small.

The basis set errors of the CCSD correlation energies range from ∼13 kJ/mol to ∼3 kJ/mol. The range is reduced to ∼4 kJ/mol to ∼0.3 kJ/mol with the explicit correlation. The two basis set families generate close basis set errors, and neither of them shows consistently better performance than the other. Overall, the CCSD(T)$_{\overline{F12}}$ BSEs are close to the (T) BSEs.

Again, the explicit correlation significantly improves the basis convergence of the CCSD(T) correlation energies. The CCSD(T)$_{\overline{F12}}$ BSEs with the double-ζ basis sets are almost twice smaller than the CCSD(T) BSEs with the quadruple-ζ basis sets. With the large quadruple-ζ basis sets, the basis set errors of explicitly correlated CCSD(T) energies reduce to around 0.6 kJ/mol. These values are close to the corresponding overall errors of the HEAT model (0.24 kJ/mol). The BSEs from the calculations with the two basis sets are very close, although the F12 basis sets give slightly better results for the standard and explicitly correlated CCSD(T) correlation energies than the non-F12 basis sets.

3.5. Overall Performance of CCSD(T)$_{\overline{F12}}$ for Reaction Barriers. To access the overall performance of the CCSD-(T)$_{\overline{F12}}$ model chemistry for computing reaction barriers, we computed the mean unsigned errors (MUEs) and maximum unsigned errors (MaxUEs) relative to the benchmark electronic barriers in the DBH24/08 database; the results are presented in Table 3.

Table 3. The Mean Unsigned Errors and Maximum Unsigned Errors (kJ/mol) of DBH24/08 Reaction Barriers Computed with CCSD and CCSD(T) Methods Using the cc-pVXZ-F12 and aug-cc-pVXZ Basis Set Families[a]

method	MUE			MaxUE		
	X = D	X = T	X = Q	X = D	X = T	X = Q
cc-pVXZ-F12						
CCSD	11.89	9.84	9.77	64.46	44.34	41.53
CCSD(2)$_{\overline{F12}}$	9.40	9.63	9.83	40.68	39.77	40.03
CCSD(T)	6.20	2.27	1.30	30.34	7.86	4.30
CCSD(T)$_{\overline{F12}}$	1.93	1.16	1.18	6.56	3.28	2.80
aug-cc-pVXZ						
CCSD	7.59	9.14	9.63	47.62	43.06	41.20
CCSD(2)$_{\overline{F12}}$	8.04	9.36	9.76	32.85	38.52	38.95
CCSD(T)	5.21	2.37	1.24	16.06	6.99	4.12
CCSD(T)$_{\overline{F12}}$	1.72	1.05	1.08	4.44[b]	2.45	2.54[c]

[a]Unless noted specifically, the maximum errors were observed for the reverse barrier of reaction 1 (see Table 1). [b]Reverse reaction 6. [c]Reverse reaction 3.

As expected, the barriers computed with the CCSD method have large mean errors: the majority of the resulting MUEs are larger than 9 kJ/mol, and the MaxUE's are around or more than 41 kJ/mol. The inclusions of the F12 correction into the CCSD method reduces the MUEs only slightly. This should not be surprising: the F12 correction only reduces the basis set error and does not affect the method error.

We noticed that the MaxUE's are all from the reverse reaction barrier of reaction 1 (OH + N_2 → H + N_2O), which

has a very high reaction barrier (345.05 kJ/mol). To evaluate the MaxUE influence on MUE, we also computed the MUEs of different methods without the MaxUE values (see the Supporting Information). However, the new MUEs do not show a significant decrease; the errors are still on the same level as those including the MaxUEs.

The inclusion of the perturbative triple correction (T) is crucial to obtaining accurate reaction barriers, as becomes readily apparent from comparing the CCSD and CCSD(T) barriers. Yet a triple-ζ basis is needed to reduce the CCSD(T) MUEs to around 2 kJ/mol; the corresponding triple-ζ MaxUE's are still around or above 7 kJ/mol. Only with the large quadruple-ζ basis sets do the maximum errors of the CCSD(T) barriers reduce to ∼1 kJ/mol. However, for the largest DBH24/08 test cases, each increase of the cardinal number increases the computational cost of the CCSD(T) computation by an order of magnitude. Hence, with the transition from aug-cc-VDZ to aug-cc-pVQZ the mean and maximum CCSD(T) errors reduce by a factor of 4, whereas the cost increases by a factor of $O(100)$. The slow basis set convergence is clearly a severe problem for computing the reaction barriers in the DBH24/08 set and must be dealt with to unleash the inherent high accuracy of the CCSD(T) method.

Fortunately, the basis set errors of CCSD(T) are reduced dramatically by switching to the CCSD(T)$_{\overline{F12}}$ method. The use of F12 terms reduces the error by an equivalent of 1 to 2 cardinal numbers (i.e., a double-ζ basis result with F12 is in between the accuracy of conventional triple- and quadruple-ζ energies). Only a double-ζ basis set is sufficient to reduce the mean basis set errors to below 2 kJ/mol. The mean errors reduce further to close to 1 kJ/mol with a triple-ζ basis; use of a quadruple-ζ basis results in no additional error decrease. The maximum errors are also reduced dramatically compared to the standard CCSD(T) method, below 4 kJ/mol with triple-ζ basis sets. The mean and maximum errors obtained with the cc-pVXZ-F12 basis sets of Peterson and co-workers are greater than those obtained with the aug-cc-pVXZ counterparts, even though the former family was designed specifically for the use with F12 methods and have been found to outperform the standard basis sets.[61] The double-ζ results are especially surprising as the cc-pVDZ-F12 basis set is considerably larger than the aug-cc-pVDZ counterpart (the cc-pVDZ-F12 basis set has one more s and two more p functions). On the other hand, while the cc-pV{T,Q}Z-F12 basis sets are still larger than the corresponding aug-cc-pV{T,Q}Z basis sets for non-hydrogen atoms, they are actually smaller for hydrogen atoms. For the hydrogen atom, the cc-pVTZ-F12 basis set has one less p function than the aug-cc-pVTZ basis set, and the cc-pVQZ-F12 basis set has one less p and f functions compared to the aug-cc-pVQZ basis set.

To further investigate the effects of diffuse functions on the hydrogen atom, we carried the aug-cc-pVXZ calculations with the cc-pVXZ basis sets for hydrogen atoms. We find that the mean and maximum errors of the calculations (shown in Table 4) become larger, especially for the calculations at the double-ζ level. However, the mean errors of the triple- and quadruple-ζ level calculations are very close to those of the corresponding cc-pV{T,Q}Z-F12 calculations. This indicates that the extra diffuse functions are not the reason for the better performance of the aug-cc-pVXZ versus the cc-pVXZ-F12 basis. The discussion of basis set errors in section 3.1 suggests that the smaller overall errors obtained with the aug-cc-pVXZ basis sets are perhaps due to a fortuitous cancellation of the CCSD(2)$_{\overline{F12}}$

Table 4. The CCSD(T) and CCSD(T)$_{\overline{F12}}$ Mean and Maximum Unsigned Errors (kJ/mol) of DBH24/08 Reaction Barriers Using the aug-cc-pVXZ Basis Sets for Non-Hydrogen Atoms and the cc-pVXZ Basis Sets for Hydrogen Atom[a]

	MUE			MaxUE		
method	X = D	X = T	X = Q	X = D	X = T	X = Q
CCSD(T)	6.67	2.76	1.37	17.90	7.40	4.26
CCSD(T)$_{\overline{F12}}$	2.31	1.20	1.12	5.09[1]	3.54[2]	2.84[2]

[a]Unless noted specifically, the maximum errors were observed for the reverse barrier of reaction 1 (see Table 1). [b]Forward reaction 12. [c]Reverse reaction 12.

and (T) energies. This at the very least suggests that the cc-pVXZ-F12 basis sets may need a more thorough benchmarking for a variety of applications before the use of standard non-F12-optimized basis sets is eliminated.

In addition, we performed the cc-pVXZ-F12 and aug-cc-pVXZ-F12 (X = D, T) calculations with different geminal exponents (0.8, 1.0, and 1.2). The results are listed in Table 5.

Table 5. The CCSD(T)$_{\overline{F12}}$ Mean and Maximum Unsigned Errors (kJ/mol) of DBH24/08 Reaction Barriers with the cc-pVXZ-F12 and aug-cc-pVXZ Basis Set Families Using Different Geminal Exponents (β).[1a]

	MUE			MaxUE		
method	β = 0.8	β = 1.0	β = 1.2	β = 0.8	β = 1.0	β = 1.2
cc-pVDZ-F12	2.08	1.91	2.07	8.14	5.67	5.22
cc-pVTZ-F12	1.10	1.16	1.22	3.63	3.28	3.24
aug-cc-pVDZ	1.65	1.65	1.81	4.17	4.27[c]	4.36[c]
aug-cc-pVTZ	1.08	1.05	1.05	2.59	2.37[d]	2.45

[a]Unless noted specifically, the maximum errors were observed for the reverse barrier of reaction 1 (see Table 1). [b]The optimal geminal exponents for the aug-cc-pVDZ, aug-cc-pVTZ, cc-pVDZ-F12, and cc-pVTZ-F12 basis sets are 1.1, 1.2, 0.9, and 1.0, respectively.[36] [c]Reverse reaction 6. [d]Reverse reaction 12.

As expected, the calculations at the triple-ζ level using different geminal exponents give very similar results, which is also the case for the aug-cc-pVDZ calculations. For the cc-pVDZ-F12 calculations, while the MUEs of the calculations with different geminal exponents are still close, the discrepancy between the corresponding MaxUEs is much larger. We also find that the calculations with optimal geminal exponents did not necessarily give the best results. This suggests that the choice of geminal exponent is not crucial to the quality of the calculations as long as the geminal exponent is chosen in a reasonable range.

In the calculations above, we only included the valence electron correlations. To address the influence of the core and core−valence electron correlations, we computed the MUEs of CCSD(T)/aug-cc-pCVXZ (X = D, T) with and without a frozen core. The differences between them with both basis sets are around 0.4 kJ/mol, which means the influence of the core and core−valence electron correlations is not significant. Thus, it is reasonable to just consider the valence electron correlations in the calculations, and the resultant errors will be negligible.

3.6. Overall Performance of CCSD(T)$_{\overline{F12}}$ for Reaction Energies. The overall errors of the CCSD(T)$_{\overline{F12}}$ model chemistry for the HJO12 reaction energy set[44] are listed in Table 6. The errors are relative to the experimentally derived values from ref 44, which are obtained by subtracting the

Table 6. The Mean Unsigned Errors and Maximum Unsigned Errors (kJ/mol) of the Electronic Reaction Energies for the HJO12 Isogyric Reaction Set[44a]

	MUE			MaxUE		
method	X = D	X = T	X = Q	X = D	X = T	X = Q
cc-pVXZ-F12						
CCSD	11.5	6.4	6.8	37.5	24.5	19.7
CCSD(2)$_{\overline{F12}}$	6.0	7.5	8.3	15.3	16.2	16.1
CCSD(T)	15.3	8.8	3.7	32.8[b]	17.5[b]	8.4[b]
CCSD(T)$_{\overline{F12}}$	4.7	2.3	1.6	13.7[b]	6.4[b]	4.3[b]
aug-cc-pVXZ						
CCSD	15.4	9.7	9.5	45.9	27.3	20.1
CCSD(2)$_{\overline{F12}}$	6.8	8.3	9.0	14.6	16.6	16.4[c]
CCSD(T)	10.6	3.7	1.8	35.1	13.3	5.4
CCSD(T)$_{\overline{F12}}$	3.7	1.5	1.2	7.5[b]	3.6[b]	3.1[c]

[a]Unless noted specifically, the maximum errors were observed for reaction 12 (see Table 2). [b]Reaction 2. [c]Reaction 11.

vibrational and scalar relativistic energy corrections from the experimental reaction energies. Note that we employed *valence* conventional and explicitly correlated CCSD and CCSD(T) methods; hence the core correlation contributions from ref 44 were included additively.

As expected, the inclusion of the F12 correction into the CCSD(T) method significantly improves the accuracy of the reaction energies. Similarly to the reaction barriers, the introduction of F12 terms reduces the basis set requirements by between 1 and 2 cardinal numbers; e.g., we find that the errors of the triple-ζ CCSD(T)$_{\overline{F12}}$ calculations are smaller than those of the quadruple-ζ CCSD(T) calculations. With the quadruple-ζ basis sets, the CCSD(T)$_{\overline{F12}}$ calculations have MUEs of about 1 kJ/mol, and MaxUEs around 3−4 kJ/mol. Remarkably, the aug-cc-pVXZ basis sets are found to yield smaller errors than the cc-pVXZ-F12 basis sets. This is especially pronounced for the maximum errors. In fact, the aug-cc-pVTZ basis set gives a slightly smaller MUE and MaxUE than the cc-pVQZ-F12 basis set.

3.7. Overall Performance of CCSD(T)$_{\overline{F12}}$ for Atomization Energies and Heats of Formation. The performance of the CCSD(T)$_{\overline{F12}}$ model chemistry for thermochemical computations was further tested against the HEAT test set. We computed the atomization energies and enthalpies of formations with the *all-electron* conventional and explicitly correlated CCSD and CCSD(T) methods. The results were augmented with the (anharmonic) zero-point energies taken from the HEAT reference database; all other corrections were neglected.

The mean and maximum unsigned errors of atomization energies are listed in Table 7. The atomization energies computed within the HEAT model were used as the reference. The inclusions of F12 terms again reduces the errors by an equivalent of 1.5 to 2 cardinal numbers, with better improvements observed for the cc-pCVXZ-F12 series. In almost all instances the cc-pCVXZ-F12 CCSD(T)$_{\overline{F12}}$ errors are smaller than the corresponding aug-cc-pCVXZ errors (the lone exception is the mean error of the cc-pCVTZ-F12 data).

Similar trends were observed for the heats of formation (see Table 8). The reference values are the enthalpies of formation computed within the HEAT model using the elemental reaction approach. It can be seen that the MaxUEs come from either the C/N atoms or molecules which contain them.

Table 7. The Mean Unsigned Errors and Maximum Unsigned Errors (kJ/mol) of Atomization Energies for HEAT Test Set[a]

method	MUE			MaxUE		
	X = D	X = T	X = Q	X = D	X = T	X = Q
cc-pCVXZ-F12						
CCSD	75.38	45.39	33.71	152.14	91.82	70.51
$CCSD(2)_{\overline{F12}}$	31.55	27.48	25.92	62.03	57.87	55.22
CCSD(T)	52.65	20.50	8.18	100.36	35.55	14.30[b]
$CCSD(T)_{\overline{F12}}$	8.82	2.62	0.95	15.88[c]	5.16[d]	2.87[d]
aug-cc-pCVXZ						
CCSD	88.30	46.22	33.09	167.51	98.72	70.84
$CCSD(2)_{\overline{F12}}$	29.91	26.28	25.47	55.54[c]	54.28	54.20
CCSD(T)	69.85	21.81	7.60	133.14[c]	43.58	15.36[b]
$CCSD(T)_{\overline{F12}}$	11.46	2.13	0.99	29.46[c]	5.84[d]	3.14

[a]Unless noted specifically, the maximal errors were from calculations of CO_2. [b]N_2. [c]C_2H_2. [d]CN.

Table 8. The Mean Unsigned Errors and Maximum Unsigned Errors (kJ/mol) of Enthalpies of Formation for HEAT Test Set[a]

method	MUE			MaxUE		
	X = D	X = T	X = Q	X = D	X = T	X = Q
cc-pCVXZ-F12						
CCSD	29.93	21.17	17.68	83.09	52.55	41.20[b]
$CCSD(2)_{\overline{F12}}$	13.74	15.26	15.19	35.66	34.08	32.66
CCSD(T)	17.77	7.86	3.11	52.74	19.84	7.80
$CCSD(T)_{\overline{F12}}$	2.17	0.89	0.90	5.75[c]	2.47[c]	2.46[d]
aug-cc-pCVXZ						
CCSD	34.24	23.31	18.26	98.05	57.99	41.71
$CCSD(2)_{\overline{F12}}$	11.15	14.90	15.18	29.87	32.58	32.25
CCSD(T)	24.10	9.47	3.48	73.21	26.08	8.51
$CCSD(T)_{\overline{F12}}$	2.79	1.27	1.11	11.52[e]	3.07[f]	2.71[d]

[a]The molecule, for which the error is maximal, is the C atom unless noted specifically. H_2, N_2, O_2, F_2, and CO were not included in the analysis since the elemental reaction approach, which was used to calculate the enthalpies of formation, defines these molecules as the "standard states" for elements H, N, O, F, and C, respectively. [b]CCH. [c]N. [d]CN. [e]C_2H_2. [f]CO_2.

In all cases, the cc-pCVXZ-F12 $CCSD(T)_{\overline{F12}}$ errors are smaller than the corresponding aug-cc-pCVXZ errors.

4. SUMMARY AND CONCLUSIONS

In this work, we evaluated the performance of the perturbative explicitly correlated coupled-cluster method $CCSD(T)_{\overline{F12}}$ in combination with the two basis set families commonly used for the explicitly correlated coupled-cluster calculations, namely, cc-pVXZ-F12 and aug-cc-pVXZ. The results of the reaction barriers, reaction energies, atomization energy, and enthalpy of formation show that the inclusion of the perturbative F12 correction significantly reduce the basis set error of the correlation energies. The effects of the F12 terms were least substantial for reaction barriers (DZ F12 errors were comparable to the TZ standard errors, i.e., an extra gain of one cardinal number) and the greatest for atomization energies and heats of formation (a gain of two cardinal numbers). Overall, the performance of the two families was similar at the TZ and QZ levels, with some differences observed at the double-ζ level.

We conclude that the aug-cc-pVDZ $CCSD(T)_{\overline{F12}}$ model chemistry is an excellent choice for computing electronic reaction barriers. With the aug-cc-pVDZ basis the mean unsigned error is only 1.72 kJ/mol (that surpasses all model chemistries considered by Zheng et al.[1]) and the maximum error of only 4.44 kJ/mol. The computational cost of the aug-cc-pVDZ $CCSD(T)_{\overline{F12}}$ computation is only a small factor greater than that of the corresponding standard aug-cc-pVDZ CCSD(T) energy computation and dramatically lower than that of the aug-cc-pVTZ CCSD(T) energy computation (hence, the F12 approach is clearly superior to the basis set extrapolation). While at least an order of magnitude more expensive than routine density functional model chemistries, the aug-cc-pVDZ $CCSD(T)_{\overline{F12}}$ predicts the barriers in the DBH24/08 database far more accurately, in the range of chemical accuracy.

For the electronic reaction energy calculations, the aug-cc-pVTZ $CCSD(T)_{\overline{F12}}$ model is shown to be the best choice. The mean unsigned error of the $CCSD(T)_{\overline{F12}}$ calculations with the aug-cc-pVTZ basis set is 1.5 kJ/mol, and its maximum unsigned error is 3.6 kJ/mol, which are below the chemical accuracy mark. While the computational cost of the aug-cc-pVTZ $CCSD(T)_{\overline{F12}}$ calculations is significantly less than that of the aug-cc-pVQZ CCSD(T) calculations, the aug-cc-pVTZ CCSD-$(T)_{\overline{F12}}$ model gives a more accurate prediction compared to the aug-cc-pVQZ CCSD(T) model.

The accurate predictions of atomization energies (in chemical accuracy range) require the $CCSD(T)_{\overline{F12}}$ method with the triple-ζ basis sets. The aug-cc-pCVTZ basis set gives slightly smaller mean unsigned errors than the cc-pCVTZ-F12 basis set (2.13 vs 2.62 kJ/mol) but yields a slightly larger maximum unsigned error compared to the cc-pCVTZ-F12 basis set (5.84 vs 5.16 kJ/mol). On the other hand, the more expensive quadruple-ζ CCSD(T) calculations yield errors about 3 times larger than the errors of the triple-ζ $CCSD(T)_{\overline{F12}}$ calculations. Increasing the basis set from triple-ζ to quadruple-ζ, the errors of the $CCSD(T)_{\overline{F12}}$ calculations decrease to under the chemical accuracy bar.

The cc-pCVDZ-F12 $CCSD(T)_{\overline{F12}}$ method is preferred for the calculations of enthalpy of formation. It has a mean unsigned error of 2.17 kJ/mol and a maximum unsigned error of 5.75 kJ/mol, which are in the range of the chemical accuracy. The similar accuracy requires the CCSD(T) calculations with the cc-pCVQZ-F12 basis set, which are far more expensive. The errors of the $CCSD(T)_{\overline{F12}}$ calculations for enthalpies of formation can be further reduced to below the chemical accuracy with the triple-ζ basis sets. By switching to the $CCSD(T)_{\overline{F12}}$ method, we reduce the computational cost for the calculations of the enthalpies of formation and improve the accuracy of the predictions at the same time.

■ APPENDIX

Following the convention, we used $i, j, ...; a, b, ...;$ and $p, q, ...$ for the occupied, virtual, and general orbitals in the Hartree−Fock (HF) basis, respectively; $\alpha, \beta, ...$ for the virtual orbitals in the complete basis; and $a', b', ...$ for the virtual orbitals in the complete basis that do not belong to the HF basis. In addition, we labeled spin orbitals with the lowercase indices ($i, j, \alpha, \beta...$), the spatial parts of alpha orbitals with uppercase indices ($I, J, A, B, ...$), and the spatial parts of beta orbitals with uppercase indices with bars ($\bar{I}, \bar{J}, \bar{A}, \bar{B}$, etc.).

Tensor notation for second-quantized expressions has been described in detail elsewhere, e.g. ref 9. The relationship between tensor and Dirac notations is as follows:

$$O_p^q = \langle p|O(1)|q \rangle \tag{28}$$

$$G_{pq}^{rs} = \langle pq|G(1,2)|rs \rangle \tag{29}$$

Antisymmetrized matrix elements are denoted with a horizontal line above the operator symbol:

$$\bar{G}_{pq}^{rs} = G_{pq}^{rs} - G_{pq}^{sr} \tag{30}$$

Summation is implied over all pairs of indices that appear in the same term of an expression both in bra and ket (the Einstein summation convention), *unless* the summation is shown explicitly.

Matrix elements of one- and two-body operators featured in the R12/F12 theory are denoted as follows:

$$F_p^q = \langle p|\hat{F}|q \rangle \quad g_{pq}^{rs} = \langle pq|r_{12}^{-1}|rs \rangle$$

where $\hat{F}$ is the Fock operator.

■ ASSOCIATED CONTENT

ⓢ Supporting Information

Detailed programmable equations for the $\text{CCSD(T)}_{\overline{F12}}$ are provided. This information is available free of charge via the Internet at http://pubs.acs.org/.

■ AUTHOR INFORMATION

Corresponding Author
*E-mail: efv@vt.edu.

Notes
The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Zheng, J.; Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2009**, *5*, 808.

(2) Zheng, J.; Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput* **2007**, *3*, 569.

(3) Karton, A.; Tarnopolsky, A.; Lamere, J.-F.; Schatz, G. C.; Martin, J. M. L. *J. Phys. Chem. A* **2008**, *112*, 12868.

(4) Harding, M. E.; Vázquez, J.; Ruscic, B.; Wilson, A. K.; Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **2008**, *128*, 114111.

(5) Kutzelnigg, W.; Morgan, J. D., III. *J. Chem. Phys.* **1992**, *96*, 4484.

(6) Coulson, C. A.; Neilson, A. H. *Proc. Phys. Soc.* **1961**, *78*, 831.

(7) Pack, R. T.; Brown, W. *J. Chem. Phys.* **1966**, *45*, 556.

(8) Kutzelnigg, W. *Theor. Chim. Acta* **1985**, *68*, 445.

(9) Kutzelnigg, W.; Klopper, W. *J. Chem. Phys.* **1991**, *94*, 1985.

(10) Klopper, W.; Samson, C. C. M. *J. Chem. Phys.* **2002**, *116*, 6397.

(11) Valeev, E. F. *Chem. Phys. Lett.* **2004**, *395*, 190.

(12) Ten-no, S. *J. Chem. Phys.* **2004**, *121*, 117.

(13) Kedžuch, S.; Milko, M.; Noga, J. *Int. J. Quantum Chem.* **2005**, *105*, 929.

(14) Ten-No, S. *Chem. Phys. Lett.* **2004**, *398*, 56.

(15) May, A. J.; Valeev, E. F.; Polly, R.; Manby, F. *Phys. Chem. Chem. Phys.* **2005**, *7*, 2710.

(16) Valeev, E. F. *J. Chem. Phys.* **2006**, *125*, 244106.

(17) Noga, J.; Kutzelnigg, W.; Klopper, W. *Chem. Phys. Lett.* **1992**, *199*, 497.

(18) Shiozaki, T.; Kamiya, M.; Hirata, S.; Valeev, E. F. *Phys. Chem. Chem. Phys.* **2008**, *10*, 3358.

(19) Shiozaki, T.; Kamiya, M.; Hirata, S.; Valeev, E. F. *J. Chem. Phys.* **2009**, *130*, 054101.

(20) Köhn, A.; Richings, G. W.; Tew, D. P. *J. Chem. Phys.* **2008**, *129*, 201103.

(21) Fliegl, H.; Klopper, W.; Hättig, C. *J. Chem. Phys.* **2005**, *122*, 084107.

(22) Adler, T. B.; Knizia, G.; Werner, H.-J. *J. Chem. Phys.* **2007**, *127*, 221106.

(23) Valeev, E. F. *Phys. Chem. Chem. Phys.* **2008**, *10*, 106.

(24) Valeev, E. F.; Crawford, T. D. *J. Chem. Phys.* **2008**, *128*, 244113.

(25) Knizia, G.; Werner, H.-J. *J. Chem. Phys.* **2008**, *128*, 154103.

(26) Knizia, G.; Adler, T. B.; Werner, H.-J. *J. Chem. Phys.* **2009**, *130*, 054104.

(27) Hättig, C.; Tew, D. P.; Köhn, A. *J. Chem. Phys.* **2010**, *132*, 231102.

(28) Köhn, A. *J. Chem. Phys.* **2009**, *130*, 131101.

(29) Ten-no, S. *Chem. Phys. Lett.* **2007**, *447*, 175.

(30) Torheyden, M.; Valeev, E. F. *J. Chem. Phys.* **2009**, *131*, 171103.

(31) Shiozaki, T.; Werner, H.-J. *J. Chem. Phys.* **2010**, *133*, 141103.

(32) Kong, L.; Valeev, E. F. *J. Chem. Phys.* **2011**, *135*, 214105.

(33) Shiozaki, T.; Knizia, G.; Werner, H.-J. *J. Chem. Phys.* **2011**, *134*, 034113.

(34) Bischoff, F.; Klopper, W. *J. Chem. Phys.* **2010**, *132*, 094108.

(35) Bischoff, F.; Valeev, E. F.; Klopper, W.; Janssen, C. *J. Chem. Phys.* **2010**, *132*, 214104.

(36) Peterson, K. A.; Adler, T. B.; Werner, H.-J. *J. Chem. Phys.* **2008**, *128*, 084102.

(37) Yousaf, K. E.; Peterson, K. A. *J. Chem. Phys.* **2008**, *129*, 184108.

(38) Peterson, K. A.; Krause, C.; Stoll, H.; Hill, J. G.; Werner, H.-J. *Mol. Phys.* **2011**, *109*, 2607.

(39) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007.

(40) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.

(41) Woon, D. E.; Dunning, T. H. *J. Chem. Phys.* **1993**, *98*, 1358.

(42) Neese, F.; Valeev, E. F. *J. Chem. Theory Comput.* **2011**, *7*, 33.

(43) Janssen, C. L.; Nielsen, I. B.; Leininger, M. L.; Valeev, E. F.; Kenny, J. P.; Seidl, E. T. *The Massively Parallel Quantum Chemistry Program (MPQC)*, version 3.0 alpha; Sandia National Laboratories: Livermore, CA, 2008. http://www.mpqc.org (accessed July 2012).

(44) Helgaker, T.; Jørgensen, P.; Olsen, J. *Modern Electronic Structure Theory*, 1st ed.; Wiley: Chichester, U. K., 2000.

(45) Tajti, A.; Szalay, P.; Császár, A.; Kállay, M.; Gauss, J.; Valeev, E.; Flowers, B.; Vázquez, J.; Stanton, J. *J. Chem. Phys.* **2004**, *121*, 11599.

(46) Torheyden, M.; Valeev, E. F. *Phys. Chem. Chem. Phys.* **2008**, *10*, 3410.

(47) Bokhan, D.; Bernadotte, S.; Ten-no, S. *J. Chem. Phys.* **2009**, *131*, 084105.

(48) Huang, X.; Valeev, E. F.; Lee, T. J. *J. Chem. Phys.* **2010**, *133*, 244108.

(49) Löwdin, P.-O. *J. Math. Phys.* **1962**, *3*, 969.

(50) Stanton, J. F. *Chem. Phys. Lett.* **1997**, *281*, 130.

(51) Kutzelnigg, W.; Morgan, J. D. *J. Chem. Phys.* **1992**, *96*, 4484.

(52) Ten-no, S. *J. Chem. Phys.* **2004**, *121*, 117.

(53) Dunning, T. H.; Peterson, K. A.; Wilson, A. K. *J. Chem. Phys.* **2001**, *114*, 9244.

(54) Yousaf, K. E.; Peterson, K. A. *Chem. Phys. Lett.* **2009**, *476*, 303.

(55) Manby, F. R. *J. Chem. Phys.* **2003**, *119*, 4607.

(56) Tew, D. P.; Klopper, W. *J. Chem. Phys.* **2005**, *123*, 074101.

(57) Hill, J. G.; Mazumder, S.; Peterson, K. A. *J. Chem. Phys.* **2010**, *132*, 054108.

(58) Crawford, T. D.; Sherrill, C. D.; Valeev, E. F.; Fermann, J. T.; King, R. A.; Leininger, M. L.; Brown, S. T.; Janssen, C. L.; Seidl, E. T.; Kenny, J. P.; Allen, W. D. *J. Comput. Chem.* **2007**, *28*, 1610.

(59) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. *J. Chem. Phys.* **1997**, *106*, 9639.

(60) Schwenke, D. W. *J. Chem. Phys.* **2005**, *122*, 014107.

(61) Bischoff, F. A.; Wolfsegger, S.; Tew, D. P.; Klopper, W. *Mol. Phys.* **2009**, *107*, 963.