AIP | The Journal of Chemical Physics

What are the most efficient basis set strategies for correlated wave function calculations of reaction energies and barrier heights?

Ewa Papajak and Donald G. Truhlar

Citation: J. Chem. Phys. **137**, 064110 (2012); doi: 10.1063/1.4738980
View online: http://dx.doi.org/10.1063/1.4738980
View Table of Contents: http://jcp.aip.org/resource/1/JCPSA6/v137/i6
Published by the AIP Publishing LLC.

Additional information on J. Chem. Phys.

Journal Homepage: http://jcp.aip.org/
Journal Information: http://jcp.aip.org/about/about_the_journal
Top downloads: http://jcp.aip.org/features/most_downloaded
Information for Authors: http://jcp.aip.org/authors

THE JOURNAL OF CHEMICAL PHYSICS **137**, 064110 (2012)

What are the most efficient basis set strategies for correlated wave function calculations of reaction energies and barrier heights?

Ewa Papajak and Donald G. Truhlar[a)]
Department of Chemistry and Minnesota Supercomputing Institute, University of Minnesota, Minneapolis, Minnesota 55455-0431, USA

(Received 12 May 2012; accepted 10 July 2012; published online 10 August 2012)

As electronic structure methods are being used to obtain quantitatively accurate reaction energies and barrier heights for increasingly larger systems, the choice of an efficient basis set is becoming more critical. The optimum strategy for achieving basis set convergence can depend on the way that electron correlation is treated and can take advantage of flexibility in the order in which basis functions are added. Here we study several approaches for estimating accurate reaction energies and barrier heights from post-Hartree–Fock electronic structure calculations. First and second, we evaluate methods of estimating the basis set limit of second order Møller-Plesset perturbation theory and of coupled cluster theory with single and double excitations and a quasiperturbative treatment of connected triple excitations by using explicitly correlated basis functions (in the F12a implementation) along with valence, polarization, and diffuse one-electron basis functions. Third, we test the scheme of adding a higher-order correction to MP2 results (sometimes called MP2/CBS + ΔCCSD(T)). Finally, we evaluate the basis set requirements of these methods in light of comparisons to Weizmann-3.2, Weizmann-4, and CCSDT(2)$_Q$/CBS+CV+R results. © *2012 American Institute of Physics*. [http://dx.doi.org/10.1063/1.4738980]

I. INTRODUCTION

Accurate evaluation of energies of reaction and chemical reaction barrier heights is important for calculating the thermodynamics and kinetics of chemical processes. Coupled cluster calculations of these quantities based on increasingly larger basis sets are widely used in the literature, where they are used both for comparison with the experimental results and as benchmark values for evaluation or parameterization of less computationally demanding methods. In particular, the coupled cluster method with singles, doubles, and quasiperturbative connected triples, CCSD(T),[1] is often said to be "the gold standard of quantum chemistry" and is often accurate to 1 kcal/mol or better when a complete enough basis set is used. However it is also computationally demanding for moderate-sized and large systems. For this reason it has become popular to *estimate* the complete-basis-set (CBS) limit of CCSD(T) calculations rather than calculate it directly.

There are several methods in common use for obtaining CCSD(T)/CBS limits. The first method, which will be called the straight method, involves carrying out CCSD(T) results with larger and larger basis sets until the results converge. The efficiency of this approach can be greatly improved by the use of basis functions that depend explicitly on the interelectronic distances (R12, F12, F12a, and F12b methods).[2–21]

The second method, which will be called extrapolation, is similar to the straight method, but one does not actually carry out a calculation with a nearly complete basis set; instead one extrapolates to the limit from two or three calculations with smaller but systematically increasing basis sets. In the light of recent developments and the impressive efficiency of the explicitly correlated methods, extrapolation without explicit correlation is becoming of only historical interest, since similar accuracy can be achieved with R12 (Refs. 22–24) and F12a (Refs. 14, 19, and 25) or F12b (Refs. 14, 19, and 25) methods in a single step.[26]

The third method will be called the dual-level method. In it one first writes the CCSD(T)/CBS energy as

$$E^{\mathrm{CBS}}_{\mathrm{CCSD(T)}} = E^{\mathrm{CBS}}_{\mathrm{MP2}} + \Delta E^{\mathrm{CBS}}, \tag{1}$$

where $E^{\mathrm{CBS}}_{\mathrm{MP2}}$ is the CBS limit of the energy calculated by second-order Møller–Plesset perturbation theory (MP2),[27] and

$$\Delta E^{\mathrm{CBS}}_{\mathrm{CCSD(T)}} = E^{\mathrm{CBS}}_{\mathrm{CCSD(T)}} - E^{\mathrm{CBS}}_{\mathrm{MP2}} \tag{2}$$

is the coupled cluster correction. Then we recognize, that, as has been pointed out and explained in various contexts,[28–31,61] E_{MP2} converges much more slowly with respect to basis set size than does ΔE. The first step in using Eq. (1) is to estimate $E^{\mathrm{CBS}}_{\mathrm{MP2}}$. The second step is to estimate the CCSD(T) correction by calculating it with larger and larger basis set until ΔE converges. One can use smaller basis sets for step 2 than for step 1.

$$\Delta E^{\mathrm{CBS}}_{\mathrm{CCSD(T)}} \approx E^{\mathrm{FBS}}_{\mathrm{CCSD(T)}} - E^{\mathrm{FBS}}_{\mathrm{MP2}}, \tag{3}$$

where FBS denotes the finite basis set used to calculate ΔE. One could also use extrapolation for one or both terms of Eq. (1).

Approaches similar to Eqs. (1)–(3) have been used in a variety of additive schemes of greater complexity, where multiple correction terms are involved. The Gaussian-*n* approach by Pople, Curtiss, Raghavachari, and co-workers[32–35] was

---

[a)] Author to whom correspondence should be addressed. Electronic mail: truhlar@umn.edu.

the first popular example of such a scheme. Hobza and co-workers used a similar strategy, but without following the precise scheme of G-*n* methods[28,36] Focal point analysis of Allen and co-workers[31,37] and multicoefficient correlation methods (MCCMs) (Refs. 39–42) by our group are based on a similar idea and may also be considered to be elaborations of the older SEC (Ref. 43) and SAC (Ref. 44) schemes. Employing sequences of basis sets and extending them until convergence is attained may also be viewed as special cases of the focal point method.[45,46] The CBS methods[47,48] of Petersson and co-workers involve more detailed analyses of the asymptotic convergence of the correlation energy but may be considered to be the foundational papers supporting all the future developments. The "interference" term of the CBS methods may also be used to provide a correction to Eq. (3).[49]

In employing the straight- and dual-level methods, one usually employs a sequence of correlation-consistent (cc) basis sets (although MCCMs have been shown to be often more efficient with other basis sets[40]). Correlation consistent basis sets by Dunning and co-workers (including cc-pV*n*Z,[50–54] and cc-pV(*n*+d)Z,[55]) have the advantage that they have been designed to converge to the complete basis set limit systematically, thereby in principle allowing extrapolation to the limit from much smaller basis sets. Originally, they have been augmented with large and gradually increasing sets of diffuse functions to accompany increasing degrees of decontraction of the valence space *n* (aug-cc-pV*n*Z (Refs. 50–54, 56, and 57) and aug-cc-pV(*n*+d)Z (Ref. 55)). It has been shown previously that in cases where diffuse functions are needed, partial augmentation of Dunning's basis sets is a more efficient approach than the originally suggested fully augmented approach.[58–63] It also leads to fewer SCF convergence problems. Partial augmentation of cc-pV*n*Z basis sets leads to jul-cc-pV*n*Z, jun-cc-pV*n*Z, may-cc-pV*n*Z, …, etc. basis sets, where "jul-" indicates a full "aug" set of diffuse functions on heavy atoms, but no augmentation on hydrogen atoms, and the "jun", "may", "apr", … etc. basis sets (so called seasonal basis sets) are the same as "jul" minus successive subshells of higher-angular-momentum diffuse functions on heavy atoms.[61,63]

Rather than actually employing a sequence of basis sets, it is also very common to use only one basis set that, based on experience, is expected to be nearly complete. Clearly this is more convenient. Based on experience accumulated to date, many workers choose the aug-cc-pVTZ, cc-pVQZ, or aug-cc-pVQZ basis sets or extrapolation schemes involving these sets without further testing. *However, to provide guidance it is useful to have available systematic studies of which basis sets are large enough to yield results within given tolerances of the CBS limit. This is one key goal of the present study.*

Great efficiency in approaching the complete basis set limit can be attained by using explicitly correlated basis function, as in the R12, F12, F12a, and F12b schemes. The combination of F12 explicitly correlated basis functions with, e.g., jun-cc one-electron basis functions is expected to be a powerful approach to reach the CBS limit of MP2, and the analogous F12a scheme is known to be a powerful approach to reach the CCSD(T)/CBS limit.

At this point, we insert a note on nomenclature. The basis sets considered here do not include core correlation or core polarization functions, thus they converge to what might best be called the valence CBS limit. It is very common in practical applications in the literature (almost universal) to call this the CBS limit, so we will use the shorter "CBS limit", but the reader should keep this in mind. Core correlation contributions can be estimated,[64,65] but that is beyond our present scope. Although core correlation and core polarization effects will be included in our best estimates (see below), our goal here is to test the widely employed methods that do not include core correlation or core polarization. (Note that core polarization is included at the Hartree–Fock level; when we refer to core polarization in this paragraph we mean the beyond-Hartree–Fock part.)

In the present article we therefore address the following questions:

(1) *What is the most efficient one-electron basis set from among the cc, aug-cc, and seasonal sequences for using the F12 scheme to reach the MP2-F12/CBS limit within various tolerances by the straight approach?*

In order to answer this question, we use MP2-F12/CBS results as our benchmark values. It is understood that, even with an infinite basis set, MP2 theory results carry large errors and so the purpose of this first query is to establish which basis sets efficiently approximate MP2-F12/CBS barrier height and reaction energy values, not how to calculate the most accurate values of these properties. As illustrated by Eqs. (1)–(3), MP2-F12/CBS results near to the CBS limit can then be used together with higher-order corrections in order to obtain accurate data.

(2) *What is the most efficient one-electron basis set from among the cc, aug-cc, and seasonal sequences for using the F12a scheme to reach the CCSD(T)/CBS limit within various tolerances by the straight approach?*
(3) *What are the most efficient one-electron basis set pairs to employ in MP2-F12 and CCSD(T)-F12a calculations to for dual-level calculations employing Eq. (3)?*

In order to address questions (2) and (3), we use CCSD(T)/CBS results as our benchmark values. We should keep in mind though that CCSD(T) is not an exact theory. It is widely appreciated that decreasing a source of error in a computational method reaches a point of diminishing returns when that source of error is reduced to the point where it is no longer dominant. For example, if CCSD(T)/CBS has a typical error of 0.4 kcal/mol, one expects to get more accurate results with a basis set with a typical incompleteness error of 0.5 kcal/mol than with a basis set with a typical incompleteness error of 1.0 kcal/mol. However, increasing the basis set to reduce basis set errors from 0.3 kcal/mol to 0.15 kcal/mol is hardly warranted since the error can only decrease below ~0.4 kcal/mol by cancellation of errors. This brings us to the fourth question:

(4) *At what basis set size do CCSD(T) and dual-level calculations stop becoming more accurate on average?*

To answer question (4) we compare CCSD(T) and dual-level calculations with various sizes of basis sets to the most accurate data available for the quantities studied here – Weizmann-4 (W4) (Ref. 66), Weizmann-3.2 (W3.2), and CCSDT(2)$_Q$/CBS+CV+R calculations (where CV denotes core-valence correlation and R denotes relativistic corrections) – rather than to CCSD(T)/CBS calculations, which are relevant to questions (2) and (3). We note that W4 calculations have 95% confidence interval of 0.16 kcal/mol for thermochemical data.[66] Thus it is seldom called for to get closer than 0.16 kcal/mol to the CBS limit. Note also that W4 corresponds approximately to CCSDTQ5/CBSD+CV+R+DBOC, where DBOC denotes the diagonal Born-Oppenheimer correction, and W3.2 corresponds approximately to CCSDT(Q)/CBS+CV+R+DBOC.

II. COMPUTATIONAL DETAILS

In this article, we evaluate methods based on their predictions of the forward and reverse barrier heights and the energy of reaction for the following 16 reactions:

- *Radical hydrogen transfer:*

$$\mathrm{HO_2\cdot + CH_3OH \rightarrow H_2O_2 + \cdot CH_2OH} \tag{R1}$$

$$\mathrm{HO_2\cdot + CH_3OH \rightarrow H_2O_2 + CH_3O\cdot} \tag{R2}$$

$$\mathrm{\cdot\ CH_3 + CH_3OH \rightarrow CH_4 + \cdot CH_2OH} \tag{R3}$$

$$\mathrm{\cdot\ CH_3 + CH_3OH \rightarrow CH_4 + CH_3O\ \cdot} \tag{R4}$$

$$\mathrm{\cdot\ OH + CH_4 \rightarrow \cdot CH_3 + H_2O} \tag{R5}$$

$$\mathrm{H\cdot + OH \rightarrow O(^3P) + H_2} \tag{R6}$$

$$\mathrm{H\cdot + H_2S \rightarrow H_2 + HS} \tag{R7}$$

- *Heavy atom transfer:*

$$\mathrm{H\cdot + N_2O \rightarrow \cdot OH + N_2} \tag{R8}$$

$$\mathrm{H\cdot + HCl \rightarrow H\cdot + HCl} \tag{R9}$$

$$\mathrm{\cdot\ CH_3 + FCl \rightarrow CH_3F + Cl\cdot} \tag{R10}$$

- *Nucleophilic substitution:*

$$\mathrm{Cl^- + CH_3Cl \rightarrow Cl^- + CH_3Cl} \tag{R11}$$

$$\mathrm{F^- + CH_3Cl \rightarrow Cl^- + CH_3F} \tag{R12}$$

$$\mathrm{OH^- + CH_3F \rightarrow F^- + CH_3OH} \tag{R13}$$

- *Unimolecular and association reactions:*

$$\mathrm{H\cdot + N_2 \rightarrow HN_2\cdot} \tag{R14}$$

$$\mathrm{H\cdot + C_2H_4 \rightarrow \cdot CH_2CH_3} \tag{R15}$$

$$\mathrm{HCN \rightarrow HNC} \tag{R16}$$

Reactions (R5)–(R16) constitute DBH24/08 database that has been developed as a statistically representative set of reactions for barrier height calculations and has been used before for evaluation computational methods for calculation of barrier heights.[67] To that set we add hydrogen abstraction reactions from methanol (reactions (R1)–(R4)) since they can be representative of larger and more complicated organic systems.[68]

Our goal is to draw general conclusions about basis set strategy, and we believe that the best way to do this is to average the results of basis set convergence tests over multiple data. The full set of 48 data is called the thermochemical kinetics 48 (TK48) database; it contains the energy of reaction and the forward and reverse barrier heights of the sixteen reactions listed above.

In order to test sensitivity of the basis set errors to the diffuse basis sets, TK48 is divided into ATK9 and NTK39. ATK9 includes barriers and reaction energies of the reactions that involve anions (R11)–(R13), which are expected to be more sensitive to the saturation of the diffuse space. NTK39 consists of analogous data for the neutral reactions (R1)–(R10) and (R14)–(R16), which should be less sensitive to augmentation.

III. METHODS

All calculations in the present tests are single-point energy calculations at pre-optimized geometries; we consider both both open- and closed-shell species.

All the MP2 results presented here are based on the MP2-F12 procedure,[17] and all the CCSD(T) results presented here were obtained with the F12a (Ref. 20) procedure. These procedures augment the MP2 calculation and the CCSD part of the CCSD(T) calculation by explicitly correlated terms. These explicitly correlated terms do not make a direct contribution to the (T) correction of the CCSD(T) method, but the (T) correction changes due to changes in the CCSD amplitudes.

Barrier height and reaction energy calculations were performed using the MOLPRO 2009 (Ref. 69) program. In order to reduce cost of the calculations, we used standard density fitting basis sets. In particular, all the MP2-F12 and CCSD(T)-F12a calculations employing aug-cc-pV(n+d)Z calculations were carried out using the aug-cc-pV*n*Z/JKFIT and aug-cc-pV*n*Z/MP2FIT density-fitting basis sets, and all the calculations involving cc-pV(*n*+d)Z), seasonal (*mon*-cc-pV(*n*+d)Z), and cc-pV*n*Z-F12 basis sets employ cc-pV*n*Z/JKFIT and cc-pV*n*Z/MP2FIT as the density fitting basis sets. All calculations use 3C(FIX) ansatz with orbital invariant amplitudes and were carried out with full non-linear fit of the geminal expansion.

For the stationary points of the reactions R5 through R16, we used QCISD/MG3 geometries from previous studies.[70,71] Geometries for the methanol reactions (R1–R4) were M06-2X/MG3S calculations taken from Ref. 68. In all cases we include spin-orbit contributions for all open-shell systems by the same procedure as used previously.[72] For the single-point calculations of the open-shell systems we used the restricted formalism for all levels; in particular, we used RHF, RMP2-F12,[73] and RRCCSD(T)-F12a in this study.

IV. COST ESTIMATES

One cannot adequately judge the relative success of various computational methods without considering cost, which we express here in units of CPU time. Unfortunately, precise comparisons of CPU times are impossible, because such times depend on the software, the computer, the load on that computer, the way that the method is coded, the number of processors used, and many other variables. Nevertheless the trends in computer time are meaningful if one does not interpret them too finely; for example, it is usually meaningful if two computer times (on the same computer with the same number of processors) vary by more than a factor of two on the same computer.

In order to quantify the relative cost of the methods that we test in this study, we have calculated single-point energies of the same molecule, on the same machine (Chinook at the Pacific Northwest National Laboratory), using the same number (8) of processors, same amount of memory (437 MW), and same software (MOLPRO 2009) for all the combinations of levels and basis sets. For this purpose phosphinomethanol was chosen, because it has a reasonable ratio of first period:second period:third period atoms (6:2:1), and it is affordable even for very complete calculations. Unfortunately, it is rather small so the times for expensive methods have not reached their asymptotic scaling regime.

We report CPU time of these single-point calculations as relative to the most inexpensive (HF/cc-pV(D+d)Z) calculation in Table I.

TABLE I. The cost of HF, MP2-F12, and CCSD(T)-F12a single point calculation (CPU time) of CH2(OH)PH2 molecule normalized to the smallest calculation (HF/cc-pVDZ).

	HF[a]	MP2-F12	CCSD(T)-F12a
aug-cc-pV(Q+d)Z	1090	4300	68900
jul-cc-pV(Q+d)Z	491	2250	34900
jun-cc-pV(Q+d)Z	386	2010	27000
may-cc-pV(Q+d)Z	278	1840	21500
apr-cc-pV(Q+d)Z	237	1740	18600
cc-pV(Q+d)Z	205	1660	16000
aug-cc-pV(T+d)Z	90	1070	7740
jul-cc-pV(T+d)Z	45	532	3840
jun-cc-pV(T+d)Z	27	469	2710
may-cc-pV(T+d)Z	21	430	2110
cc-pV(T+d)Z	16	400	1710
aug-cc-pV(D+d)Z	4.8	322	775
jul-cc-pV(D+d)Z	2.7	180	396
jun-cc-pV(D+d)Z	1.5	160	285
cc-pV(D+d)Z	1.0	146	223

[a] An average of two HF runs: for the MP2-F12 and CCSD(T)-F12a calculation.

V. RESULTS AND DISCUSSION

Tables II–V summarize mean unsigned basis-set errors of all the single- and dual-level methods for barrier heights and reaction energies using various basis sets. First, we examine convergence of MP2-F12 with respect to the basis set in order to address ***question 1***. Table II lists mean unsigned errors for the barrier heights and reaction energies in TK48. As benchmark values we used MP2-F12/jul-cc-pV(5+d)Z results. Since the sensitivity of the neutral and anionic reactions to diffuse functions is different, we list separately the errors on the reactions containing anions (in ATK9 column) and neutral reactants only (NTK39 column). The table shows that triple-ζ basis sets provide results already converged to within $\sim$0.1 kcal/mol to the quintuple-ζ values. For very highly converged MP2/CBS estimates, may-cc-pV(Q+d)Z can be employed; however jun-cc-pV(T+d)Z already converges all three columns to better than the 0.16 kcal/mol limit mentioned earlier. Notice that may-cc-pV(T+d)Z is about a factor of two more accurate than the only slightly less expensive (see MP2-F12 column of Table I) aug-cc-pV(D+d)Z basis set.

In order to address ***question 2***, in Table III we list deviations of the CCSD(T) theory with various basis sets from the near-CBS values (CCSD(T)-F12a/apr-cc-pV(5+d)Z). As in MP2, triple- and quadruple-ζ results are within 0.2 and 0.1 kcal/mol of the quintuple zeta results. Beyond the jun-cc-pV(T+d)Z basis set, improvement is very slow compared to the rapid increase in the costs (which can be seen in the last column of Table I). For example, consecutive basis sets, despite nearly doubling the cost, improve the average errors by hundredths of kcal/mol or not at all. The jul-cc-pV(T+d)Z basis set achieve better than 0.16 kcal/mol for all three columns of Table III with a cost lower than aug-cc-pV(T+d)Z by a

TABLE II. Mean unsigned errors (in kcal/mol) of MP2 barrier heights and reaction energies using MP2-F12/jul-cc-pV(5+d)Z values as a benchmark

	TK48	ATK9	NTK39
apr-cc-pV(5+d)Z	0.01	0.02	0.01
aug-cc-pV(Q+d)Z	0.03	0.03	0.03
jul-cc-pV(Q+d)Z	0.04	0.05	0.04
jun-cc-pV(Q+d)Z	0.03	0.04	0.03
may-cc-pV(Q+d)Z	0.04	0.04	0.04
apr-cc-pV(Q+d)Z	0.07	0.14	0.05
cc-pV(Q+d)Z	0.69	2.74	0.22
aug-cc-pV(T+d)Z	0.07	0.09	0.07
jul-cc-pV(T+d)Z	0.12	0.11	0.13
jun-cc-pV(T+d)Z	0.13	0.15	0.12
may-cc-pV(T+d)Z	0.18	0.26	0.16
cc-pV(T+d)Z	1.19	3.92	0.56
aug-cc-pV(D+d)Z	0.32	0.36	0.31
jul-cc-pV(D+d)Z	0.46	0.37	0.48
jun-cc-pV(D+d)Z	0.70	1.30	0.56
cc-pV(D+d)Z	2.45	6.25	1.58

TABLE III. Mean unsigned errors (in kcal/mol) on CCSD(T)-F12a barrier heights and reaction energies as compared to the CCSD(T)-F12a/apr-cc-pV(5+d)Z results.

	TK48	ATK9	NTK39
aug-cc-pV(Q+d)Z	0.05	0.05	0.05
jul-cc-pV(Q+d)Z	0.05	0.07	0.05
jun-cc-pV(Q+d)Z	0.04	0.03	0.04
may-cc-pV(Q+d)Z	0.04	0.02	0.04
apr-cc-pV(Q+d)Z	0.09	0.18	0.07
cc-pV(Q+d)Z	0.63	2.46	0.21
aug-cc-pV(T+d)Z	0.14	0.20	0.12
jul-cc-pV(T+d)Z	0.15	0.15	0.15
jun-cc-pV(T+d)Z	0.14	0.18	0.14
may-cc-pV(T+d)Z	0.20	0.24	0.19
cc-pV(T+d)Z	1.09	3.50	0.54
aug-cc-pV(D+d)Z	0.32	0.45	0.29
jul-cc-pV(D+d)Z	0.45	0.42	0.46
jun-cc-pV(D+d)Z	0.68	1.37	0.52
cc-pV(D+d)Z	2.23	5.79	1.40

factor of 2.0, and jun-cc-pV(T+d)Z does about as well with a cost lowering of a factor of 2.9. The may-cc-pV(T+d)Z basis set raises the basis set errors to only 0.19–0.24 kcal/mol with a cost lowering (again relative to aug-) of a factor of 3.7, whereas the error is very large when diffuse basis functions are totally omitted. These comparisons provide a dramatic illustration of the efficiency of seasonal basis sets.

In order to answer ***question 3***, we need to balance errors introduced by the two terms of Eq. (1). We denote the basis used for the first term as X and that for the second step as Y. When choosing X, the demand on the accuracy should depend on the uncertainty introduced by the level of theory (and the basis set used) at which higher order corrections are going to be calculated. However, as mentioned above, triple-ζ basis sets provide results already converged to within ~0.1 kcal/mol of CBS values, which, as we will show below, is more than an order of magnitude lower than the error of the MP2/CBS result itself. For the very accurate dual-level schemes, may-cc-pV(Q+d)Z would be recommended, but as discussed in conjunction with question 1, that will usually be overkill. To find the best strategy we will explicitly consider combinations of X and Y.

TABLE IV. Mean unsigned errors (in kcal/mol) of ΔCCSD(T)-F12a corrections to barrier heights and reaction energies as compared to the CCSD(T)-F12a/apr-cc-pV(5+d)Z values of the corrections.

	TK48	TK9	TK39
aug-cc-pV(Q+d)Z	0.03	0.03	0.03
jul-cc-pV(Q+d)Z	0.03	0.02	0.03
jun-cc-pV(Q+d)Z	0.03	0.03	0.03
may-cc-pV(Q+d)Z	0.03	0.02	0.03
apr-cc-pV(Q+d)Z	0.05	0.06	0.04
cc-pV(Q+d)Z	0.11	0.28	0.07
aug-cc-pV(T+d)Z	0.11	0.15	0.10
jul-cc-pV(T+d)Z	0.10	0.12	0.09
jun-cc-pV(T+d)Z	0.09	0.10	0.09
may-cc-pV(T+d)Z	0.11	0.10	0.12
cc-pV(T+d)Z	0.22	0.47	0.16
aug-cc-pV(D+d)Z	0.23	0.22	0.23
jul-cc-pV(D+d)Z	0.25	0.22	0.25
jun-cc-pV(D+d)Z	0.28	0.24	0.28
cc-pV(D+d)Z	0.44	0.68	0.38

TABLE V. Mean unsigned errors[a] on data in NK48 database calculated by the dual-level method using Eq. (3).

X	jun-Q	may-Q	apr-Q	aug-T	jul-T	jun-T	may-T	T	aug-D
Y									
jun-Q	…	…	…	…	…	…	…	…	…
may-Q	0.04	…	…	…	…	…	…	…	…
apr-Q	0.06	0.06	…	…	…	…	…	…	…
Q	0.11	0.11	0.10	…	…	…	…	…	…
aug-T	0.11	0.11	0.13	…	…	…	…	…	…
jul-T	0.10	0.10	0.12	0.12	…	…	…	…	…
jun-T	0.09	0.09	0.12	0.11	0.13	…	…	…	…
may-T	0.11	0.11	0.13	0.11	0.13	0.14	…	…	…
T	0.21	0.21	0.20	0.20	0.20	0.22	0.25	…	…
aug-D	0.22	0.23	0.25	0.24	0.24	0.25	0.29	1.18	…
jul-D	0.24	0.25	0.26	0.25	0.26	0.28	0.30	1.17	0.34
jun-D	0.27	0.27	0.28	0.28	0.27	0.27	0.29	1.16	0.31
D	0.42	0.43	0.41	0.42	0.41	0.41	0.40	1.01	0.49

[a] Mean unsigned errors (kcal/mol) as compared to the CCSD(T)/apr-5 values.

The suitable choice of balanced pairs of basis sets X and Y can be made using Tables II and IV. By comparing these tables pair X = may-cc-pV(Q+d)Z, Y = jun-cc-pV(T+d)Z is well balanced for very high accuracy (0.10 kcal/mol) since with the basis set combination the MP2-F12 and ΔCCSD(T) terms would be introducing comparable cost and error. For the somewhat more realistic goal of better than 0.16 kcal/mol, Tables II and IV show that Y = may-cc-pV(T+d)Z is a good candidate to match with X = jun-cc-pV(T+d)Z, and Y = jul-cc-pV(D+d)Z or even jun-cc-pV(D+d)Z appear to be good candidates to match with X = may-cc-pV(T+d)Z.

A more direct evaluation of the combinations may be based on Table V which lists basis set errors of results obtained using Eq. (3) as determined by comparing to the near-CCSD(T) limit. For example, in order to evaluate the combination of X = jun-T with Y = jun-D (where we have now introduced an obvious shorthand notation), consider that CCSD(T)/jun-D by itself has an error of 0.68 kcal/mol (Table III). However, when we combine it with additional MP2-F12/jun-T calculation, the error goes down to 0.27 kcal/mol (Table V). In this case, the dual method is twice as expensive as the CCSD(T) component calculation. If we wanted to use only CCSD(T) (straight approach), similar improvement could only be achieved if the basis set was increased to may-T, which is roughly eight times more expensive than the CCSD(T)/jun-D calculation.

To put the orders of magnitude of errors discussed here in perspective and address ***question 4***, which provides the ultimate criterion for basis set selection, one should consider the error of the CCSD(T)/CBS and dual-level approaches. Therefore, we have compared the CCSD(T) results to the best available estimates. For the reactions R6–R16 we used Weizmann 3.2 and 4 results of Martin and co-workers as the best

TABLE VI. Mean unsigned errors (in kcal/mol) on MP2-F12 and CCSD(T)-F12a barrier heights and reaction energies as compared to the best available estimates: W3.2, W4, and CCSDT(2)$_Q$/CBS.

	MP2-F12			CCSD(T)-F12a		
	TK48	ATK9	NTK39	TK48	ATK9	NTK39
apr-5	2.85	0.89	3.30	0.44	0.11	0.52
aug-Q	2.84	0.87	3.30	0.43	0.16	0.49
jul-Q	2.84	0.84	3.31	0.44	0.18	0.50
jun-Q	2.85	0.87	3.30	0.44	0.14	0.51
may-Q	2.85	0.88	3.30	0.45	0.13	0.53
apr-Q	2.83	0.76	3.31	0.50	0.29	0.55
Q	3.12	2.32	3.30	0.89	2.57	0.50
aug-T	2.83	0.87	3.28	0.41	0.18	0.46
jul-T	2.84	0.81	3.30	0.41	0.14	0.48
jun-T	2.86	0.92	3.31	0.44	0.12	0.51
may-T	2.88	0.98	3.32	0.50	0.18	0.58
T	3.36	3.56	3.31	1.14	3.62	0.57
aug-D	2.88	0.98	3.32	0.45	0.39	0.47
jul-D	2.92	0.94	3.38	0.52	0.38	0.56
jun-D	3.23	2.15	3.48	0.79	1.26	0.68
D	3.99	6.03	3.52	2.19	5.90	1.33

TABLE VII. Mean unsigned errors[a] on data in TK48 database calculated by the dual-level method using Eq. (3).

X	jun-Q	may-Q	apr-Q	aug-T	jul-T	jun-T	may-T	T	aug-D
Y:									
jun-Q	…	…	…	…	…	…	…	…	…
may-Q	0.45	…	…	…	…	…	…	…	…
apr-Q	0.47	0.47	…	…	…	…	…	…	…
Q	0.49	0.50	0.49	…	…	…	…	…	…
aug-T	0.43	0.43	0.45	…	…	…	…	…	…
jul-T	0.44	0.45	0.47	0.42	…	…	…	…	…
jun-T	0.46	0.46	0.48	0.43	0.43	…	…	…	…
may-T	0.50	0.50	0.52	0.47	0.46	0.47	…	…	…
T	0.57	0.58	0.57	0.55	0.53	0.55	0.57	…	…
aug-D	0.49	0.49	0.51	0.47	0.46	0.47	0.51	1.21	…
jul-D	0.52	0.53	0.54	0.51	0.50	0.53	0.56	1.20	0.53
jun-D	0.54	0.54	0.55	0.53	0.50	0.52	0.54	1.18	0.52
D	0.70	0.70	0.69	0.69	0.66	0.67	0.67	1.12	0.69

[a] Mean unsigned errors (kcal/mol) as compared to the Weizmann and CCSDT(2)$_Q$/CBS limit values.

available estimates.[74] Since our calculations do include spin-orbit corrections, but do not include DBOC corrections, the analogous Weizmann values were used from Ref. 74. In the case of reactions R1–R4, CCSDT(2)$_Q$/CBS (Ref. 75) benchmarks[68] were used. These CBS values were obtained[68] by extrapolation from cc-pVTZ and cc-pVQZ basis sets and have been shown to agree well with experiment. They also include spin-orbit and relativistic corrections. The errors of MP2-F12 and CCSD(T)-F12a are summarized in Table VI. We see that on average CCSD(T)/5Z results differ from these benchmarks by 0.44 kcal/mol, which is much higher than the hundredths of kcal/mol in Table III. This means that in the search for the most accurate results, beyond the CCSD(T) with triple-ζ basis sets, additional computational resources should be invested in higher-order correlation and relativistic effects rather than increasing basis set.

Similarly, when we evaluate the dual-level results using the errors with respect to the best available estimates in Table VII, we see that the improvement achieved by going from X = jul-T and Y = jun-D (within 0.27 kcal/mol of CCSD(T)/CBS) to X = may-Q and Y = jun-T (0.09 kcal/mol) is very insignificant in comparison with the total errors.

In fact, there seems to be no statistically significant advantage in going beyond X : Y = jun-T : may-T or X : Y = jun-T : aug-D, each of which has a mean total error (see Table VII) of only 0.47 kcal/mol. In fact X : Y = aug-D : jun-D has a mean basis set error only slightly larger at 0.52 kcal/mol.

In order to compare the efficiency of the single and dual methods on a quality vs. cost basis, Tables VIII–X list selected methods (MP2, CCSD(T), and dual-level) in order of increasing cost. Table VIII is based on the full database (TK48), and Tables IX and X correspond respectively to the subsets of reactions involving anions (ATK9) and those involving only neutral species (NK39). To avoid clutter with basis sets that do not offer significant improvement, the tables list only methods that offer improvement, with respect to less expensive methods, both in deviations relative to the CCSD(T)/CBS limit (second last column) and in deviations with respect to the best estimate.[76] In other words, some rows were deleted so that *both* of the final two columns are monotonically decreasing functions as the cost goes up—the idea being that only if both criteria improve is it worth increasing the cost. Methods that do not appear in at least one of Tables VIII–X can hardly be recommended in the final analysis.

The most astonishing conclusion concerning the errors with respect to the best estimate (these errors are listed in the last columns of Tables VIII–X) is that beyond CCSD(T)-F12a with aug-D basis set the errors on the barrier heights

TABLE VIII. Mean unsigned errors on 48 data (TK48) in kcal/mol with respect to coupled cluster limit benchmarks and the best available estimates in the order of the increasing cost.

Method[a]						
MP2-F12		CCSD(T)-F12a			CCSD(T)/	BE
Diffuse	ζ	Diffuse	ζ	Cost	CBS Error[b]	Error[c]
…	D	…	…	38	3.84	3.99
jun	D	…	…	42	3.07	3.23
jul	D	…	…	47	2.77	2.92
…	…	…	D	58	2.23	2.19
…	…	jun	D	74	0.68	0.79
…	…	jul	D	104	0.45	0.52
		aug	D	203	0.32	0.45
…	…	jun	T	711	0.14	0.44
jul	T	jun	T	860	0.13	0.43
aug	T	jul	T	1309	0.12	0.42

[a] For single-level calculations only one basis set is listed; for dual-level calculations using Eq. (3), first two columns (MP2-F12 basis set) list X bases, while third and fourth (CCSD(T)-F12a basis set) list Y bases used. In all cases basis sets with tight d functions on third row atoms (+d) were used.

[b] Error with respect to the CCSD(T)-F12a/apr-cc-pV(5+d)Z values.

[c] Error with respect to the Best Estimate: W3.2 and W4 values from Ref. 74 for DBH24 data and CCSDT(2)$_Q$/CBS (T,Q-extrapolation) values[68] for the methanol reactions.

TABLE IX. Mean unsigned errors on ATK9 data in kcal/mol with respect to coupled cluster limit and the best available estimates in the order of the increasing cost.

Method[a]						
MP2-F12		CCSD(T)-F12a				
Diffuse	ζ	Diffuse	ζ	Cost	CCSD(T)/ CBS Error[b]	BE Error[c]
…	D	…	…	38	5.92	6.03
jun	D	…	…	42	2.26	2.15
jul	D	…	…	47	1.05	0.94
…	…	jul	D	104	0.42	0.38
may	T	jun	D	191	0.35	0.30
jun	T	jun	D	203	0.26	0.21
…	…	may	T	553	0.24	0.18
jun	T	may	T	682	0.16	0.12
may	Q	may	T	1103	0.08	0.10

[a] For single-level calculations only one basis set is listed; for dual-level calculations using Eq. (3), first two columns (MP2-F12 basis set) list X bases, while third and fourth (CCSD(T)-F12a basis set) list Y bases used. In all cases basis sets with tight d functions on third row atoms (+d) were used.

[b] Error with respect to the CCSD(T)-F12/apr-cc-pV(5+d)Z values.

[c] Error with respect to the best estimate: W3.2 and W4 values from Ref. 74 for DBH24 data and CCSDT(2)$_Q$/CBS (T,Q-extrapolation) values[68] for the methanol reactions.

and reaction energies are dominated by the limitations of the method, not the basis set. For example increasing the basis set from aug-D to may-Q decreases the CCSD(T)/CBS error of 0.32 kcal/mol to 0.04 kcal/mol, but the BE error remains exactly the same (0.45 kcal/mol), so this method is not included in Table VIII.

It is beyond the scope of this article to evaluate and compare all the available basis sets for the present types of calculations. However, We note that some basis sets, called cc-pV*n*Z-F12 basis sets,[77] have been optimized specifically for the explicitly correlated methods, and therefore we have tested them. We note that, despite their name, these basis sets include some diffuse functions. Our application of these basis sets to the present database and subdatabases showed that they do not provide an exception to the conclusions we draw from the calculations above with seasonal basis sets. For example, in the case of the TK48 database, the BE error for CCSD(T)-F12a calculations drops from 0.46 kcal/mol to 0.45 kcal/mol, when we increase the basis set from cc-pVTZ-F12 to cc-pVQZ-F12, and no combination of cc-pV*n*Z-F12 basis sets for X and Y offers a smaller error than 0.45 kcal/mol in the dual-level approach. Full sets of errors for the tests of cc-pV*n*Z-F12 basis sets are provided in the supplementary material.

TABLE X. Mean unsigned errors on NTK39 data in kcal/mol with respect to coupled cluster limit and Weizmann results in order of increasing cost.

Method[a]						
MP2-F12		CCSD(T)-F12a				
Diffuse	ζ	Diffuse	ζ	Cost	CCSD(T)/ CBS Error[b]	BE Error[c]
…	D	…	…	38	3.36	3.52
jun	D	…	…	42	3.26	3.48
jul	D	…	…	47	3.17	3.38
…	…	…	D	58	1.40	1.33
…	…	jun	D	74	0.52	0.68
…	…	jul	D	104	0.46	0.56
aug	D	jun	D	159	0.29	0.55
jun	T	aug	D	332	0.24	0.53
jun	T	may	T	682	0.14	0.51
…	…	aug	T	2036	0.12	0.46

[a] For single-level calculations only one basis set is listed; for dual-level calculations using Eq. (3), first two columns (MP2-F12 basis set) list X bases, while third and fourth (CCSD(T)-F12a basis set) list Y bases used. In all cases basis sets with tight d functions on third row atoms (+d) were used.

[b] Error with respect to the CCSD(T)-F12a/apr-cc-pV(5+d)Z values.

[c] cError with respect to the best estimate: W3.2 and W4 values from Ref. 74 for DBH24 data and CCSDT(2)$_Q$/CBS (T,Q-extrapolation) values[68] for the methanol reactions.

VI. CONCLUSIONS

- MP2 results for barrier height and reaction energy should only be used with higher-order correlation corrections. MP2-F12 converges much faster with respect to the basis set than do calculations without explicitly correlated functions. Augmentation is necessary to get good results, but this can be done efficiently by using partial augmentation (seasonal basis sets) without needing full augmentation (aug-).
- For the dual level approach we found two reasonably efficient basis set combinations $X : Y = \text{may-Q} : \text{jun-T}$ and $X : Y = \text{jun-T} : \text{jun-D}$. If we consider reliability vs. cost these combinations of MP2-F12 and CCSD(T)-F12a with two different basis sets perform better than plain CCSD(T), but the difference in error at a similar cost is not large. CCSD(T)-F12a/jul-cc-pVDZ is already adequate for an accuracy of ~0.6 kcal/mol, and using this basis for CCSD(T) correction with a much large basis for the MP2 component brings the error down only to ~0.5 kcal/mol.
- CCSD(T)-F12a convergence with respect to the best estimates is nonexistent beyond partially augmented triple zeta basis sets, because the error (~0.4 kcal/mol) is dominated by other effects. Therefore including higher-order electron correlation (e.g., full triples and perturbative quadruples), core correlation, and relativistic effects are recommended before proceeding to quadruple-ζ basis sets at the CCSD(T) level.
- In order to achieve chemical accuracy (< 1 kcal/mol) it is important that all contributions to the error be converged to that level with respect to increasing the basis set.[78] However, the present results show that although estimating CBS limit values is interesting from the theoretical point of view, it often does not reduce errors with respect to the best estimates. Once the basis set is large enough to make the difference of a CCSD(T) calculation from CCSD(T)/CBS smaller than the difference of CCSD(T)/CBS from the accurate result, there is little to gain by increasing the basis set further.

ACKNOWLEDGMENTS

The authors are grateful to Gillian Shaw for preliminary calculations and to Antonio Varandas and Jingjing Zheng for stimulating discussions. This work was supported in part by the U. S. Department of Energy, Office of Basic Energy Sciences (Grant No. DE-FG02-86ER13579). This research was performed in part using the Molecular Science Computing Facility (MSCF) in the William R. Wiley Environmental Molecular Sciences Laboratory, a national scientific user facility sponsored by the U.S. Department of Energy's Office of Biological and Environmental Research and located at the Pacific Northwest National Laboratory, operated for the Department of Energy by Battelle.

[1] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, Chem. Phys. Lett. **157**, 479 (1989).
[2] W. Kutzelnigg, Theor Chim Acta **68**, 445 (1985).
[3] W. Klopper and W. Kutzelnigg, Chem. Phys. Lett. **134**, 17 (1987).
[4] W. Kutzelnigg and W. Klopper, J. Chem. Phys. **94**, 1985 (1991).
[5] W. Klopper and C. C. M. Samson, J. Chem. Phys. **116**, 6397 (2002).
[6] F. R. Manby, J. Chem. Phys. **119**, 4607 (2003).
[7] S. Ten-no and F. R. Manby, J. Chem. Phys. **119**, 5358 (2003).
[8] S. Ten-no, Chem. Phys. Lett. **398**, 56 (2004).
[9] W. Klopper J. Chem. Phys. **120**, 10890 (2004).
[10] S. Ten-no, J. Chem. Phys. **121**, 117 (2004).
[11] A. J. May and F. R. Manby, J. Chem. Phys. **121**, 4479 (2004).
[12] E. F. Valeev, Chem. Phys. Lett. **395**, 190 (2004).
[13] E. F. Valeev and C. L. Janssen, J. Chem. Phys. **121**, 1214 (2004).
[14] D. P. Tew and W. Klopper, J. Chem. Phys. **123**, 074101 (2005).
[15] E. F. Valeev, Chem. Phys. Lett. **418**, 333 (2006).
[16] F. R. Manby, H.-J. Werner, T. B. Adler, and A. J. May, J. Chem. Phys. **124**, 094103 (2006).
[17] H.-J. Werner, T. B. Adler, and F. R. Manby, J. Chem. Phys. **126**, 164102 (2007).
[18] J. Noga, S. Kedžuch, and J. Šimunek, J. Chem. Phys. **127**, 034106 (2007).
[19] T. B. Adler, G. Knizia, and H.-J. Werner, J. Chem. Phys. **127**, 221106 (2007).
[20] G. Knizia, T. B. Adler, and H.-J. Werner, J. Chem. Phys. **130**, 054104 (2009).
[21] T. D. Crawford, C. D. Sherrill, E. F. Valeev, J. T. Fermann, R. A. King, M. L. Leininger, S. T. Brown, C. L. Janssen, E. T. Seidl, J. P. Kenny, and W. D. Allen, J. Comput. Chem. **28**, 1610 (2007).
[22] H. Fliegl, W. Klopper, and C. Hättig, J. Chem. Phys. **122**, 084107 (2005).
[23] H. Fliegl, C. Hättig, and W. Klopper, Int. J. Quantum Chem. **106**, 2306 (2006).
[24] D. P. Tew and W. Klopper, J. Phys. Chem. A, **111**, 11242 (2007).
[25] S. Kedžuch, M. Milko, and J. Noga, Int. J. Quantum Chem. **105**, 929 (2005).
[26] J. Noga, P. Valiron, and W. Klopper, J. Chem. Phys. **115**, 2022 (2001).
[27] C. Møller and M. S. Plesset, Phys. Rev. **46**, 618 (1934).
[28] P. Hobza and J. Šponer, J. Am. Chem. Soc. **124**, 11802 (2002).
[29] P. Jurecka, J. Sponer, J. Cerny, and P. Hobza, Phys. Chem. Chem. Phys. **8**, 1985 (2006).
[30] G. A. Petersson and M. Braunstein, J. Chem. Phys. **83**, 5129 (1985).
[31] A. L. L. East and W. D. Allen, J. Chem. Phys. **99**, 4638 (1993).
[32] L. A. Curtiss, K. Raghavachari, G. W. Trucks, and J. A. Pople, J. Chem. Phys. **94**, 7221 (1991).
[33] L. A. Curtiss, K. Raghavachari, P. C. Redfern, V. Rassolov, and J. A. Pople, J. Chem. Phys. **109**, 7764 (1998).
[34] L. A. Curtiss, P. C. Redfern, K. Raghavachari, and J. A. Pople, J. Chem. Phys. **114**, 108 (2001).
[35] L. A. Curtiss and K. Raghavachari, ACS Symp. Ser. **677**, 176 (1998).
[36] P. Jurecka and P. Hobza, J. Am. Chem. Soc. **125**, 15608 (2002).
[37] H. Feng and W. D. Allen, J. Chem. Phys. **132**, 094304 (2010).
[38] A. G. Császár, W. D. Allen, and H. F. Schaefer, J. Chem. Phys. **108**, 9751 (1998).
[39] P. L. Fast, J. C. Corchado, M. L. Sanchez, and D. G. Truhlar, J. Phys. Chem. A **103**, 5129 (1999).
[40] B. J. Lynch and D. G. Truhlar, J. Phys. Chem. A **107**, 3898 (2003).
[41] B. J. Lynch, Y. Zhao, and D. G. Truhlar, J. Phys. Chem. A **109**, 1643 (2005).
[42] T.-H. Li, H.-R. Chen, and W.-P. Hu Chem. Phys. Lett. **412**, 430 (2005).
[43] F. B. Brown and D. G. Truhlar, Chem. Phys. Lett. **117**, 307 (1985).
[44] M. S. Gordon and D. G. Truhlar, J. Am. Chem. Soc. **108**, 5412 (1986).
[45] M. S. Schuurman, S. R. Muir, W. D. Allen, and H. F. Schaefer J. Chem. Phys. **120**, 11586 (2004).
[46] R. A. King, W. D. Allen, B. Ma, and H. F. Schaefer III, Faraday Discuss. **110**, 23 (1998).
[47] M. R. Nyden and G. A. Petersson, J. Chem. Phys. **75**, 1843 (1981).
[48] G. A. Petersson, ACS Symp. Ser. **677**, 237 (1998).
[49] K. D. Vogiatzis, E. C. Barnes, and W. Klopper, Chem. Phys. Lett. **503**, 157 (2011).
[50] D. E. Woon and T. H. Dunning, Jr., J. Chem. Phys. **100**, 2975 (1994).
[51] D. E. Woon, T. H. Dunning, Jr., J. Chem. Phys. **98**, 1358 (1993).
[52] J. Koput and K. A. Peterson, J. Phys. Chem. A **106**, 9595 (2002).
[53] N. B. Balabanov and K. A. Peterson, J. Chem. Phys. **123**, 064107 (2005).
[54] A. K. Wilson, D. E. Woon, K. A. Peterson, and T. H. Dunning, Jr., J. Chem. Phys. **110**, 7667 (1999).
[55] T. H. Dunning, Jr., K. A. Peterson, and A. K. Wilson, J. Chem. Phys. **114**, 9244 (2001).
[56] T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989).
[57] R. A. Kendall, T. H. Dunning, and R. J. Harrison, Jr., J. Chem. Phys. **96**, 6796 (1992).
[58] E. Papajak, H. R. Leverentz, J. Zheng, and D. G. Truhlar, J. Chem. Theory Comput. **5**, 1197 (2009); **5**, 3330 (2009).
[59] E. Papajak and D. G. Truhlar, J. Chem. Theory Comput. **6**, 597 (2010).
[60] J. Zheng, X. Xu, and D. G. Truhlar, Theor. Chem. Acc. **128**, 295 (2011).
[61] E. Papajak and D. G. Truhlar, J. Chem. Theory Comput. **7**, 10 (2011).
[62] J. R. Cheeseman and M. J. Frisch, J. Chem. Theory Comput. **7**, 3323 (2011).
[63] E. Papajak, J. Zheng, X. Xu, H. R. Leverentz, and D. G. Truhlar, J. Chem. Theory Comput. **7**, 3027 (2011).
[64] J. M. L. Martin, A. Sundermann, P. L. Fast, and D. G. Truhlar, J. Chem. Phys. **113**, 1348 (2000).
[65] K. A. Peterson and T. H. Dunning, Jr., J. Chem. Phys. **117**, 10548 (2002).
[66] A. Karton, E. Rabinovitch, J. M. L. Martin, and B. Ruscic, J. Chem. Phys. **125**, 144108 (2006).
[67] J. Zheng, Y. Zhao, and D. G. Truhlar, J. Chem. Theory Comput. **5**, 808 (2009).
[68] J. Alecu and D. G. Truhlar, J. Phys. Chem. A **115**, 2811 (2011).
[69] H.-J. Werner, P. J. Knowles, F. R. Manby, M. Schütz *et al.*, MOLPRO, version 2010.1, a package of *ab initio* programs, 2010, see http://www.molpro.net.
[70] Y. Zhao, N. González-García, and D. G. Truhlar, J. Phys. Chem. A **109**, 2012 (2005).
[71] B. J. Lynch, Y. Zhao, and D. G. Truhlar, J. Phys. Chem. A **107**, 1384 (2003).
[72] B. J. Lynch, Y. Zhao, and D. G. Truhlar, J. Phys. Chem. A **109**, 1643 (2005).
[73] P. J. Knowles, J. S. Andrews, R. D. Amos, N. C. Handy, and J. A. Pople, Chem. Phys. Lett. **186**, 130 (1991).
[74] A. Karton, A. Tarnopolsky, J.-F. Lamere, G. C. Schatz, and J. M. L. Martin, J. Phys. Chem. A **112**, 12868 (2008).
[75] S. Hirata, P.-D. Fan, A. A. Auer, M. Nooijen, and P. Piecuch, J. Chem. Phys. **121**, 12197 (2004).
[76] See supplementary material at http://dx.doi.org/10.1063/1.4738980 for the full versions of Tables VIII–X and for mean unsigned errors with cc-pVnZ-F12 basis sets, including all single- and dual-level methods tested in this study.
[77] K. A. Peterson, T. B. Adler, and H.-J. Werner, J. Chem. Phys. **128**, 084102 (2008).
[78] K. A. Peterson, D. Feller, and D. A. Dixon, Theor. Chem. Acc. **131**, 1079 (2012).