

## Sources of error in electronic structure calculations on small chemical systems

David Feller<sup>a)</sup> and Kirk A. Peterson<sup>b)</sup>

*Department of Chemistry, Washington State University, Pullman, Washington 99164-4630*

T. Daniel Crawford<sup>c)</sup>

*Department of Chemistry, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061*

(Received 12 September 2005; accepted 20 October 2005; published online 2 February 2006)

The sources of error in electronic structure calculations arising from the truncation of the one-particle and  $n$ -particle expansions are examined with very large correlation consistent basis sets, in some cases up through valence 10- $\zeta$  quality, and coupled cluster methods, up through connected quadruple excitations. A limited number of full configuration interaction corrections are also considered. For cases where full configuration interaction calculations were unavailable or prohibitively expensive, a continued fraction approximation was used. In addition, errors arising from core/valence and relativistic corrections are also probed for a number of small chemical systems. The accuracies of several formulas for estimating total energies and atomization energies in the complete basis set limit are compared in light of the present large basis set findings. In agreement with previous work, the CCSD(T) method is found to provide results that are closer to the CCSDTQ and full configuration-interaction results than the less approximate CCSDT method. © 2006 American Institute of Physics. [DOI: 10.1063/1.2137323]

### I. INTRODUCTION

The scope and predictive capability of *ab initio* electronic structure approaches to solving the multielectron Schrödinger equation have undergone an enormous change since their introduction over 50 years ago. Initially, molecular-orbital wave functions expanded in terms of analytical basis sets met with very limited success. For example, early theoretical attempts to calculate the dissociation energy ( $D_e$ ) of  $N_2$  were in error by more than 200 kcal/mol.<sup>1</sup> The inclusion of electron correlation corrections via the configuration interaction (CI) technique reduced the error by only 50 kcal/mol.<sup>2</sup> Over time, improvements in the underlying mathematical models, the computational algorithms, computer software, and hardware yielded improvements in accuracy. Nonetheless, the lack of a practical approach to error estimation and reduction often led to widespread difficulty distinguishing between merely a fortuitous agreement between theory and experiment and genuine methodological progress. This situation was exacerbated by several undesirable characteristics of the standard approach, especially the steep scaling in computational cost with the size of the chemical system and the slowly convergent nature of the so-called one-particle and  $n$ -particle expansions. Twenty years after the first calculations on  $N_2$ , the error in  $D_e$  remained  $\sim 60$  kcal/mol,<sup>3</sup> relative to an experimental value of 228.4 kcal/mol.<sup>4,5</sup> The possibility of achieving “chemical accuracy,”  $\sim 1$  kcal/mol (0.24 kJ/mol), via a computational

procedure that did not depend upon large fortuitous cancellation of error seemed remote.

Fortunately, the present situation, at least in the area of small molecules, has improved to the point where high accuracy is achievable in most cases. For the purposes of this discussion, “small molecules” will be loosely interpreted to mean chemical systems consisting of fewer than 10 first or second row elements, combined with 20 or fewer hydrogens. This definition is clearly subjective and is likely to grow over time as capabilities grow. Although the scope of the current study does not include molecules at the upper end of this range, the topics to be discussed are nevertheless pertinent insofar as such chemical systems are amenable to the same theoretical techniques either at present or in the not-too-distant future.

Despite their power, *ab initio* electronic structure methods continue to suffer from a number of widely recognized weaknesses. As already mentioned, the methods exhibit steep scaling in computational cost with respect to the size of the chemical system and level of sophistication of the treatment. They also exhibit slow convergence and there exist no practical means to assign *a priori* error bars to a set of computed properties. The goal of this work is to help address the latter problem. We will probe several of the potential sources of error in the calculations of thermochemical and spectroscopic properties, in an effort to better understand the evolving limitations inherent to our approach. It is hoped that this knowledge can guide future studies seeking to achieve high accuracy. The chemical systems to be discussed were selected because of their role in past studies and because of their involvement in ongoing projects related to ground-state and low-lying excited-state potential-energy surfaces.

<sup>a)</sup>Electronic mail: dfeller@owt.com

<sup>b)</sup>Author to whom correspondence should be addressed. Electronic mail: kipeters@wsu.edu

<sup>c)</sup>Electronic mail: crawdad@vt.edu

## II. APPROACH

In previous work, we statistically analyzed the errors in the theoretical predictions of a variety of molecular properties, including geometries, harmonic frequencies, atomization energies, electron affinities, ionization potentials, etc. Our studies encompassed many different combinations of basis sets and correlation treatments. The original test set consisted of the Gaussian-2 (G2) collection of molecules.<sup>6-8</sup> This analysis was facilitated by information contained in the Environmental Molecular Sciences Computational Results Database (CRDB), which currently contains approximately 78 000 entries.<sup>9</sup> The insights gained from the analysis were helpful in developing a nonparametrized, multistep, composite theoretical approach that has been applied to more than 300 chemical systems.<sup>10-14</sup> The largest molecule studied with this approach is C<sub>8</sub>H<sub>18</sub>. It required a coupled cluster singles and doubles with perturbative triples [CCSD(T)] calculation with nearly 1500 basis functions.<sup>15</sup> Recent developments have extended the method to compounds containing first row transition metals.<sup>16</sup> In general, this composite approach appears to be capable of achieving a target accuracy of approximately 1 kcal/mol for thermochemical properties when used with a high-quality correlation recovery method, such as coupled cluster or configuration interaction wave functions, combined with basis sets of quadruple zeta quality or better.

Ideally, all electronic structure calculations would be performed with an effectively complete basis set at the full configuration interaction (FCI) level of theory using a relativistic Hamiltonian. Practical considerations preclude such a brute force attack on the problem and this situation is unlikely to change in the foreseeable future. Out of necessity, a more subtle strategy has been adopted by researchers seeking to achieve high accuracy results. This alternative approach proceeds by individually identifying and addressing all major sources of error. Smaller corrections are implicitly assumed to be additive. The final, best estimate for the thermodynamic or spectroscopic properties of interest is typically achieved by a combination of direct calculation and extrapolation. Although extrapolation would seem to introduce an element of uncertainty into the process, experience has shown that, when applied with care, it can reliably reduce the overall error. For matters of convenience, we have organized the following discussion along the lines of the steps in our composite approach, beginning with the one-particle and *n*-particle expansions since they have received a widespread recognition as the two largest sources of error in electronic structure calculations.

In order to predict thermochemical properties, such as heats of formation at 298 K, it is necessary to account for the vibrational zero-point energy (ZPE). For this purpose, accurate theoretical, anharmonic ZPEs based on high-quality potential surfaces are optimal, but such values are seldom reported for systems with more than five or six atoms. Depending upon the availability of experimental fundamentals, we have frequently estimated the anharmonic ZPE by performing a 1:1 averaging of theoretical harmonic frequencies and experimental fundamentals, as suggested by Grev *et al.*<sup>17</sup> In a previous study we compared this approach

against accurate anharmonic ZPEs taken from the literature for 31 polyatomic molecules. We have since expanded the data set to 36 molecules. It is relatively easy to predict anharmonic ZPEs for diatomic molecules with an accuracy of 0.1 kcal/mol or better. Using CCSD(T) harmonic frequencies, the root-mean-square (rms) errors ranged from 0.22 kcal/mol, with a double zeta (DZ) quality basis set, to 0.13 kcal/mol, with a quadruple zeta (QZ) quality basis set.<sup>11</sup> We also tested a 3:1 weighting, which should perform better on purely formal grounds when very accurate harmonic frequencies are available. For the same collection of small molecules, the 3:1 weighting produced only slight improvements in the errors. Unfortunately, CCSD(T) frequencies are prohibitively expensive for many systems, especially with the aug-cc-pVQZ basis set. Further discussion of the error arising from inaccuracies in the ZPE is beyond the scope of the present work.

## III. ONE-PARTICLE BASIS SET CONSIDERATIONS

### A. Basis set strategies

The quantum chemistry literature contains references to well over 300 families of one-particle basis sets, where the term “family” is meant to imply a collection of similar quality basis sets supporting multiple elements. In addition, a large number of *ad hoc* basis sets were developed to meet the needs of various specialized research projects and have not been adopted by the general computational chemistry community. Despite the multitude of basis set families, the design and calibration of new sets remains a very active area of research, reflecting both the importance of the one-particle expansion and the difficulty in constructing compact sets that are simultaneously effective at reproducing a multitude of chemical properties. Several reviews cover the topic.<sup>18</sup>

In 1987 Almlöf and Taylor published a seminal paper outlining a procedure that, in principle, allowed one to construct a systematic sequence of basis sets for approaching the complete basis set (CBS) limit.<sup>19</sup> It combined large Gaussian primitive expansions with atomic natural orbital (ANO) contractions. Natural orbital occupation numbers were used as the selection criterion for inclusion of functions of arbitrary angular momentum (*s*, *p*, *d*, *f*, etc.). Thus, for example, a so-called valence triple zeta (TZ) basis for Ne was constructed of four *s*-type contractions, three *p*-type contractions, two *d*-type contractions, and a single *f*-type contraction. Earlier work by Jankowski *et al.*<sup>20</sup> reached similar conclusions on the relative importance of the first through third *d* functions versus the first and second *f* functions and the first *g* function.

While the ANO basis set approach represented a breakthrough in thinking about a systematically convergent sequence of basis sets, it was not widely adopted because of the relatively high computational cost associated with the large numbers of underlying Gaussian primitives and the “general” contraction scheme<sup>21</sup> used in creating the ANOs. At the time of Almlöf and Taylor’s publication, computer programs that could efficiently handle general contractions were not widely available. They developed a new integral code specifically for their ANO work. Compared with the

popular 6-31G\* basis set of Hehre *et al.*,<sup>22</sup> which was built from a (10*s*, 4*p*, 1*d*) primitive set for first row elements, the corresponding polarized valence DZ ANO basis set used a much larger (13*s*, 8*p*, 6*d*) primitive space.

Two years later, in 1989, Dunning introduced the correlation consistent basis sets and fundamentally altered the way in which chemists addressed the one-particle basis set problem.<sup>23</sup> He retained the general contraction scheme of Almlöf and Taylor, but abandoned the natural orbital occupation selection scheme in favor of one based on the contribution of each basis function to the atomic configuration interaction singles and doubles (CISD) correlation energy. For first row elements, the resulting mixture of *s*, *p*, *d*, etc., functions was identical to the sets of Almlöf and Taylor, but the underlying primitive spaces began with a modest double zeta set and grew as the quality of the basis set grew. This produced basis sets that were both more economical at the low end and more accurate at the high end. Because the structure of correlation consistent basis sets represented something of a compromise between a strictly segmented contraction scheme and a completely general contraction, they could be used with applications that did not support the latter without paying a large penalty in computer time. In his first paper, Dunning reported on basis sets of double, triple, and quadruple zeta qualities for H and B–Ne, conventionally denoted cc-pVDZ, cc-pVTZ, and cc-pVQZ. This provided a sufficiently attractive starting point for a critical mass of researchers with an interest in higher accuracy calculations to begin adopting the new sets.

Extensions of the original correlation consistent family have grown to include basis sets with one or more shells of diffuse functions for treating anions and weakly bound systems,<sup>24</sup> basis sets for core/valence (CV) correlation recovery,<sup>25</sup> basis sets for relativistic calculations,<sup>16,26</sup> and basis sets for main group elements from Al through Kr (Ref. 27) and first row transition metals.<sup>16</sup> Additionally, basis sets designed for use with relativistic pseudopotentials have been reported for the group 11–18 elements.<sup>28</sup> At the high end, basis sets of 5Z, 6Z, 7Z, and 8Z qualities have been reported.<sup>8,13,29</sup> In total, the correlation consistent family comprises the largest and most comprehensive collection of Gaussian basis sets ever developed. Besides the original family of basis sets, the present work will make use of the diffuse function augmented family, denoted aug-cc-pVnZ (*n* = D, T, Q, 5, etc.). By convention, only the spherical component subsets (e.g., five-term *d* functions, seven-term *f* functions, etc.) of the Cartesian polarization functions are used.

## B. Estimating total energies

During the early development phase of the cc-pV5Z basis sets, it was empirically observed that the convergence pattern of total energies and binding energies of small molecules could be described by an exponential in *n*, the basis set index, *n* = 2(D), 3 (T), 4 (Q), etc. This led to the conjecture that additional basis sets in the correlation consistent sequence might exhibit a similar pattern and that the CBS limit for the total energy could be estimated by means of a simple, three-parameter functional form, such as<sup>30</sup>

$$E(n) = E_{\text{CBS}} + b * \exp(-cn), \quad (1)$$

where  $E_{\text{CBS}}$  is the limiting value. Typically,  $E_{\text{CBS}}$ ,  $b$ , and  $c$  are exactly fitted to the lowest three energies available, but a least-squares fit to four or more energies is also possible. At the same time, there was little direct evidence in support of Eq. (1). However, early comparisons with numerical Hartree-Fock (HF) energies and the exact energy of H<sub>2</sub> were encouraging. Correlated polyatomic calculations with the cc-pV5Z basis set were initially limited to second-order Møller-Plesset perturbation theory (MP2).<sup>31</sup> The ability to routinely estimate molecular properties close to the CBS limit effectively allowed one to unravel the one-particle and *n*-particle errors, leading Woon and Dunning to the notion of the “intrinsic” error for any given theoretical method, i.e., the error with respect to reliable experimental values for any given method at the CBS limit.<sup>32</sup>

The availability of a uniformly convergent basis set sequence that proved amenable to extrapolation techniques led to the introduction of many alternative approaches for estimating energies and properties at the CBS limit. At the same time, there was a growing recognition of the shortcomings of Eq. (1). A comprehensive review of the extensive CBS extrapolation literature is beyond the scope of the present work. In particular, we will not discuss the parametrized methods of Truhlar and co-workers,<sup>33</sup> or the earlier atomic pair natural orbital work of Petersson and co-workers.<sup>34</sup> The goal of achieving high accuracy while simultaneously avoiding the high cost associated with large basis set calculations created a lot of interest in the development of CBS extrapolations.

Even low-level correlated methods, such as MP2, scale as the fifth power of the number of basis functions. Higher-level methods, such as CCSD(T),<sup>35</sup> scale even more severely, as  $N^7$  or worse. For example, a CCSD(T)/aug-cc-pVQZ energy evaluation for benzene, with 756 functions, might require a little over a day on a single-processor high-speed workstation. However, in the event that a larger basis set was required, the next step-up (aug-cc-pV5Z, 1242 functions) would require over 32 days, assuming that sufficient disk and memory resources were available. A recent CCSD(T)/aug-cc-pVQZ calculation on C<sub>8</sub>H<sub>18</sub> using 1468 functions required over 66 h using between 600 and 1400 processors.<sup>15</sup> Incrementing the basis set size to the aug-cc-pV5Z level (2456 functions) would render the calculation well beyond the capabilities of even the most powerful computers available at present and for the foreseeable future. The potential benefit of expressions such as Eq. (1) was so great that work on improved extrapolation techniques remains an active area of research more than a decade later, as will be discussed.

In order to probe the convergence of the one-particle basis set in greater depth, we have developed Ne cc-pV9Z and cc-pV10Z sets as an extension to the previously reported cc-pV8Z set<sup>13</sup> using the approach first outlined by Dunning.<sup>23</sup> The composition of the new basis sets is given in Table I, along with selected other large members of the correlation consistent family. In Fig. 1 the  $\ell$  shell incremental valence correlation energy contributions,  $E_{\text{incr}}(\ell)$ , to the CCSD(T) energy,  $\Delta E_{\text{incr}}$ , are plotted as a function of  $\ell$ , the

TABLE I. Composition of selected correlation consistent basis sets. The correlation consistent basis sets are intended for use with the pure spherical components only, i.e., five-term  $d$  functions, seven-term  $f$  functions, etc.

| Basis set           | Primitives                                     | Contractions                                   |
|---------------------|--|--|
| cc-pV7Z (Ne)        | (18s, 12p, 6d, 5f, 4g, 3h, 2i, 1k)             | [8s, 7p, 6d, 5f, 4g, 3h, 2i, 1k]               |
| cc-pV8Z (Ne)        | (20s, 14p, 7d, 6f, 5g, 4h, 3i, 2k, 1l)         | [9s, 8p, 7d, 6f, 5g, 4h, 3i, 2k, 1l]           |
| cc-pV9Z (Ne)        | (22s, 16p, 8d, 7f, 6g, 5h, 4i, 3k, 2l, 1m)     | [10s, 9p, 8d, 7f, 6g, 5h, 4i, 3k, 2l, 1m]      |
| cc-pV10Z (Ne)       | (24s, 18p, 9d, 8f, 7g, 6h, 5i, 4k, 3l, 2m, 1n) | [11s, 10p, 9d, 8f, 7g, 6h, 5i, 4k, 3l, 2m, 1n] |
| aug-cc-pV6Z (C,N)   | (17s, 11p, 6d, 5f, 4g, 3h, 2i)                 | [8s, 7p, 6d, 5f, 4g, 3h, 2i]                   |
| aug-cc-pV7Z (C,N)   | (19s, 13p, 7d, 6f, 5g, 4h, 3i, 2k)             | [9s, 8p, 7d, 6f, 5g, 4h, 3i, 2k]               |
| aug-cc-pV(6+d)Z (S) | (22s, 15p, 7d, 5f, 4g, 3h, 2i)                 | [9s, 8p, 7d, 5f, 4g, 3h, 2i]                   |
| aug-cc-pV(7+d)Z (S) | (28s, 19p, 8d, 6f, 5g, 4h, 3i, 2k)             | [10s, 9p, 8d, 6f, 5g, 4h, 3i, 2k]              |
| aug-cc-pwCV5Z (C)   | (19s, 13p, 8d, 6f, 4g, 2h)                     | [11s, 10p, 8d, 6f, 4g, 2h]                     |
| aug-cc-pwCV5Z (S)   | (21s, 13p, 9d, 7f, 5g, 3h)                     | [12s, 11p, 9d, 7f, 5g, 3h]                     |
| cc-pCV6Z (N)        | (21s, 15p, 9d, 7f, 5g, 3h, 1i)                 | [12s, 11p, 9d, 7f, 5g, 3h, 2i]                 |

angular momentum index, for the cc-pV6Z through cc-pV10Z basis sets. All coupled cluster calculations in the present work were performed with MOLPRO2002,<sup>36</sup> a development version of PSI3,<sup>37</sup> DALTON2.0,<sup>38</sup> GAUSSIAN03,<sup>39</sup> and NWCHEM4.7.<sup>40</sup> Unless otherwise specified, calculations were performed using the frozen-core (FC) approximation, wherein the  $1s^2$  inner-shell electrons for first row elements ( $1s^2 2s^2 2p^6$  for second row elements) are excluded from the correlation treatment.

Nearly exponential decay, corresponding to a straight line in this logarithmic plot, is observed for  $2 \leq \ell \leq 5$  ( $d-h$

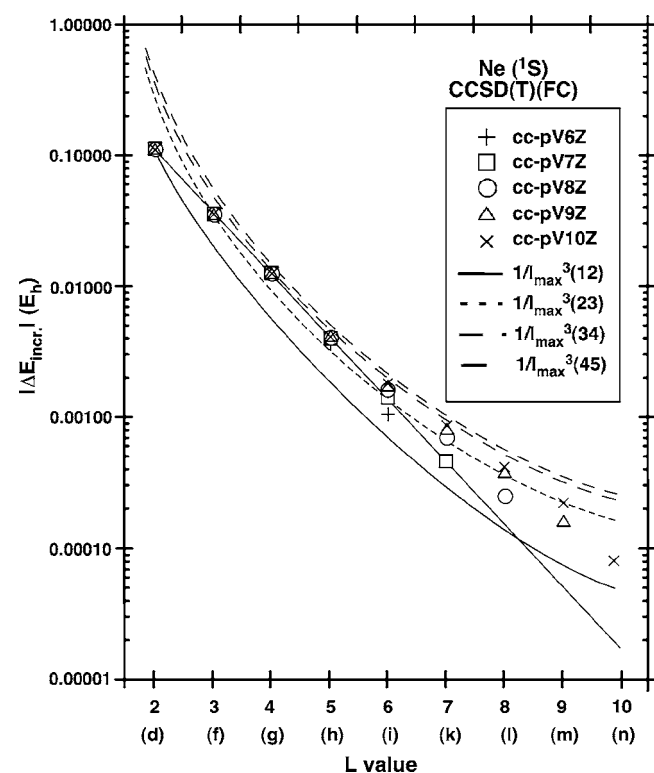


FIG. 1.  $\text{Ne}(1S)$  CCSD(T)(FC)  $\ell$  shell incremental energy lowering,  $E_{\text{incr}}(\ell)$ , from the cc-pV $n$ Z,  $n=6-10$  basis sets. The incremental contribution from a given  $\ell$  shell is defined as the energy lowering resulting from the addition of those functions to the next smaller basis set. For example, the  $\ell=2(d)$  incremental energy lowering is defined as  $E_{\text{CCSD(T)}(spd)} - E_{\text{CCSD(T)}(sp)}$ . Representative CBS extrapolated energies are based on the two-point  $1/\ell_{\text{max}}^3(mm)$  formula using cc-pV9Z energies.

functions) for all basis sets. Perhaps fortuitously, the entire cc-pV7Z basis set convergence sequence through  $k$  functions is exponential in behavior. As the higher  $\ell$  spaces ( $iklm$  functions) become increasingly saturated in the cc-pV8Z through cc-pV10Z basis sets, deviation from an exponential decay pattern becomes more pronounced. A straight line based on  $\ell$  values less than 6 underestimates the correlation contributions from higher  $\ell$  spaces. This behavior was qualitatively predicted in the work by Schwartz, who performed a  $1/Z$  perturbation-theory expansion of the correlation energy for two-electron, He-like systems.<sup>41</sup> Similar conclusions were reached in the multielectron MP2 work of Kutzelnigg and Morgan.<sup>42</sup> These analyses assumed a one-particle basis set that was saturated at every level of  $\ell$ , a condition that is far from being satisfied in the earlier members of the correlation consistent sequence.

The work of Schwartz and Kutzelnigg and Morgan led to basis set convergence formulas expressed in terms of inverse powers of  $\ell_{\text{max}}$ , where  $\ell_{\text{max}}$  is the highest angular momentum present in the basis set,<sup>43-45</sup>

$$E(\ell_{\text{max}}) = E_{\text{CBS}} + b/(\ell_{\text{max}} + \frac{1}{2})^4 \quad (2)$$

and

$$E(\ell_{\text{max}}) = E_{\text{CBS}} + b/(\ell_{\text{max}})^3. \quad (3)$$

In our previous work, we have examined the performance of Eqs. (2) and (3), as well as a mixed Gaussian/exponential function,<sup>46</sup>

$$E(n) = E_{\text{CBS}} + b * \exp[-(n-1)] + c * \exp[-(n-1)^2]. \quad (4)$$

Wilson and Dunning have also examined the expressions of the general form

$$E^{(2)}(\ell_{\text{max}}) = E_{\text{CBS}}^{(2)} + \frac{B}{(\ell_{\text{max}} + d)^m} + \frac{C}{(\ell_{\text{max}} + d)^{m+1}} + \frac{D}{(\ell_{\text{max}} + d)^{m+2}}, \quad (5)$$

for estimating the MP2 CBS limit.<sup>47</sup> Following the MP2 work of Kutzelnigg and Morgan,<sup>42</sup> Klopper advocated using Eq. (3) to fit the CCSD singlet pair energy and  $1/\ell_{\text{max}}^5$  to fit

the triplet pair energy.<sup>48</sup> Equations (1)–(5) share a common goal, namely, to be able to accurately extrapolate finite basis set results to the infinite CBS limit.

For first and second row elements the basis set index  $n$  in Eqs. (1) and (4) is equivalent to  $\ell_{\max}$ . Equation (2) is a truncated version of a three-parameter expression by Martin and Lee.<sup>44</sup> The task of determining which of the many CBS formulas performs best for a particular set of molecules, bonding arrangements, and range of basis sets is a complex one. Instead, we have often taken an *ad hoc* approach and simply adopted an average value from Eqs. (1)–(4) with the spread among the various estimates taken as a crude measure of the intrinsic uncertainty in the extrapolations. Although strictly speaking Eqs. (2) and (3) were intended for the correlation energy component of the energy, we have often used these expressions to extrapolate the total MP2, CCSD(T), or multireference configuration interaction (MRCI) energies, arguing that for basis sets of QZ quality or better, the change in the correlation energy overwhelms the change in the self-consistent-field (SCF) energy. Furthermore, with the exponential formula, Eq. (1), extrapolating the total energy has the added advantage of including partially offsetting errors. For basis sets beyond QZ, the exponential formula tends to overestimate the SCF component of the energy, while underestimating the correlation component.

Empirical evidence provided by “explicitly correlated” methods, which partially avoid the slow one-particle convergence of traditional correlated techniques, favors variations of the  $1/\ell_{\max}^3$  and  $1/\ell_{\max}^5$  expansions.<sup>48,49</sup> Examples include the linear “R12” methods of Kutzelnigg and Klopper.<sup>50</sup> By clever use of closure relationships, the R12 methods avoid the need to explicitly compute expensive multielectron integrals. While a CCSD(T)-R12 implementation has been developed, it requires large uncontracted basis sets, on the order of (19s, 14p, 8d, 6f, 4g, 3h) for first row elements, to achieve high accuracy in the total energy. For comparison purposes, the number of basis functions per first row atom in such an uncontracted basis set (212) is larger than the cc-pV7Z correlation consistent set. Furthermore, the R12 approximation only indirectly affects the perturbative triples correction via modified T1 and T2 amplitudes. As will be discussed, the (T) correction, although much smaller than the CCSD portion of the correlation energy, may require higher angular momentum functions (*ikl*, etc.) if assured accuracy to  $\pm 0.2 mE_h$  in the total energy is desired. Klopper has reported a CCSD(T)-R12/(19s, 14p, 8d, 6f, 4g, 3h) energy for the Ne atom, with conservative error bars of  $\pm 0.5 mE_h$ . In a more recent work, the use of an even larger uncontracted basis set containing multiple sets of *i* functions lowered previously reported CCSD(T)-R12 energies by 0.1–0.2  $mE_h$  (CH<sub>2</sub>, H<sub>2</sub>O, HF, N<sub>2</sub>, CO, and F<sub>2</sub>).<sup>51</sup> The size of the error will increase with the size of the system. Even if we assume an error of only 0.1  $mE_h$  per first row atom, for a molecule the size of C<sub>8</sub>H<sub>18</sub>, the error in the total energy would approach 0.8  $mE_h$ . Thus, emphasizing again that absolute accuracy is very difficult to achieve, using either standard or explicitly correlated methods.

While the  $1/\ell_{\max}$  formulas qualitatively mimic the slower than exponential convergence pattern observed for

very high  $\ell$  basis functions, they appear to be incapable of simultaneously describing the entire range of  $\ell$  values. As seen in Fig. 1, where various  $1/\ell_{\max}^3$  fits of the cc-pV9Z  $\ell$  shell incremental energies are shown, the predicted energy contributions from  $\ell$  shells with  $\ell$  values smaller than those used in the fit are significantly overestimated. This characteristic of the  $1/\ell_{\max}$  formulas has been reported previously.<sup>8</sup> The scarcity of very high accuracy CCSD(T) energies for molecules larger than a triatomic makes drawing general conclusions about the relative performance of competing CBS extrapolation formulas difficult. In this regard, the ability to accurately describe a wide range of known  $E_{\text{incr}}(\ell)$  data points would seem to increase confidence in the ability of any simple functional form to extrapolate to the unknown CBS limit. Although the cc-pV10Z basis set is very large by present day standards, it is still effectively incomplete. As the  $k$ ,  $l$ ,  $m$ , and  $n$  shells grow increasingly saturated in subsequent members of the cc-pVnZ basis set sequence, each shell's energy contribution will monotonically increase, thereby shifting the convergence pattern in the direction of the  $1/\ell_{\max}^3$  curves. Quantitatively measuring the magnitude of the shift must await calculations with still larger basis sets.

Previous hardware and software limitations precluded us from carrying out explicit CCSD(T)  $k$ -function ( $\ell=7$ ) calculations in molecules. In such cases we exploited the uniformity in the  $\Delta E_{\text{incr}}$  convergence to estimate the  $k$ -function contribution to CCSD(T)/aug-cc-pV7Z energies and energy differences.<sup>8,52</sup> Tests performed for the present work on a series of first row atoms and diatomics confirm that a simple exponential function or one of the  $1/\ell_{\max}$  expressions are capable of reproducing the true  $k$ -function contribution to total energies to an accuracy of  $\pm 10^{-5} E_h$ . Energy differences are reproduced to better than 0.05 kcal/mol. As mentioned earlier, the 7Z basis set behavior appears perhaps fortuitously well described by an exponential. When the three-parameter Eq. (1) is used, the  $k$ -function extrapolation is based on  $\ell=4-6$ (*ghi*) energies. When two-parameter fits are used, only the  $\ell=5$  and 6 energies are used. Since CCSD(T)/aug-cc-pV7Z calculations remain computationally very expensive, especially for polyatomic molecules, we continue to rely on  $k$ -function extrapolated values for the CS<sub>2</sub> results reported in the present work. Valeev *et al.* have recently reported on the impact of  $k$  functions on MP2 energies.<sup>49</sup>

In Fig. 2 the CCSD(T) incremental valence correlation energy of the Ne atom is plotted as a function of the cc-pVnZ basis set index. The incremental correlation energy is defined as the difference between successive correlation energies, e.g.,  $\Delta E_{\text{incr}}(\text{TZ}) = E_{\text{corr}}(\text{TZ}) - E_{\text{corr}}(\text{DZ})$ . Exponential decay based on the first two data points is represented as a straight line in this logarithmic plot. As in Fig. 1, the observed convergence pattern shows initial exponential convergence through 5Z, followed by slower convergence for the larger correlation consistent basis sets. The latter effect appears more pronounced in Fig. 2 than in Fig. 1, presumably due to the lack of saturation in the *spdfg* shells of the cc-pVDZ, cc-pVTZ, and cc-pVQZ sets. The long, slowly convergent tail illustrates the widely recognized difficulty that

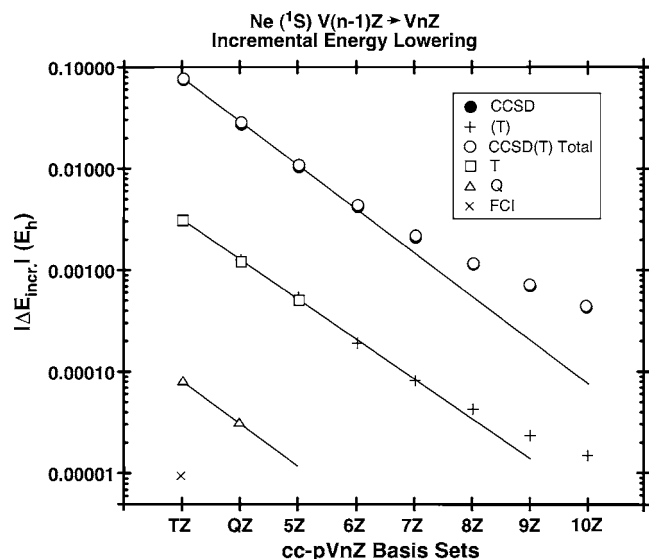


FIG. 2. Ne( $1S$ ) frozen-core CCSD, (T), CCSD(T), T, Q, and FCI energy contributions produced by incrementing the cc-pVnZ basis set index  $n$ . The T correction was based on  $E_{\text{CCSDT}} - E_{\text{CCSD}}$ . The Q correction was based on  $E_{\text{CCSDTQ}} - E_{\text{CCSDT}}$ , etc.

orbital product expansion methods have in reproducing the electron-electron Coulomb cusps of the exact wave function, namely,<sup>53</sup>

$$\lim(r_{12} \rightarrow 0)\Psi(r_{12}) \rightarrow 1 + 0.5r_{12}, \quad (6)$$

where  $r_{12}$  is the interelectronic distance. The cc-pV9Z and cc-pV10Z CCSD(T)(FC) energies differ by  $0.5 mE_h$  ( $0.3 \text{ kcal/mol}$ ) and the latter energy is still  $\sim 1 mE_h$  above various estimates of the CBS limit.

Schwenke has recently proposed<sup>86</sup> the use of a single parameter extrapolation based on the expression

$$E_{\text{CBS}}^C = [E^C(\ell_{\text{max}+1}) - E^C(\ell_{\text{max}})]F^C(\ell_{\text{max}+1}) + E^C(\ell_{\text{max}}), \quad (7)$$

where  $C$ =Hartree-Fock (HF), singlet pair (SP), triplet pair (TP), CCSD, or (T) and  $F^C(\ell_{\text{max}+1})$  is a fitting constant that depends upon the  $\ell_{\text{max}}$  value of the larger of the two basis sets used in the extrapolation. The constants, which are determined via a least-squares fit to accurate energies for seven systems (Ne,  $N_2$ ,  $CH_2$   $^1A_1$ ,  $H_2O$ , CO, HF, and  $F_2$ ), are assumed to be independent of the molecular species and the geometry. The accurate singlet and triplet pair energies were taken from the CCSD-R12 work of Klopper.<sup>48</sup> Because the Klopper HF and (T) energies were judged to be insufficiently accurate, Schwenke computed new values using large uncontracted  $f$ -limit basis sets that were designed to achieve convergence for each  $\ell$  shell. For the Ne atom this led to a ( $23s, 15p, 14d, 11f, 9g, 9h, 7i$ ) basis set, which is likely to be very close to the basis set limit up through  $i$  functions. By way of comparison, the  $f$ -limit basis set produces a CCSD(T) energy that is  $0.28 mE_h$  lower than the cc-pV10Z basis set through  $i$  functions. However, the contribution of functions beyond  $\ell=6$  lowers the Ne CCSD(T)/cc-pV10Z energy  $1.3 mE_h$  below the  $f$ -limit energy. Within the correlation consistent sequence of basis sets, the levels of polarization func-

tions equivalent to those in the  $f$ -limit basis set are not reached until cc-pV15Z.

The performance of several CBS extrapolation formulas applied to the Ne atom singlet/triplet pair energies is shown in Table II, along with the raw data. The singlet pair energy is approximately twice the size of the triplet pair energy and is significantly slower to converge. All of the formulas appear capable of predicting results that are a minimum of one full step-up in basis set quality. In many cases the extrapolated energies based on cc-pVnZ energies are comparable to raw values obtained from a cc-pV( $n+2$ )Z or cc-pV( $n+3$ )Z calculation. The Schwenke formula performs particularly well, as would be expected for one of the systems in its small training set. The raw cc-pV10Z triplet pair energy is within  $0.02 mE_h$  of the CCSD-R12 value, essentially guaranteeing that the (V8Z, V9Z, and V10Z) estimates from the exponential, mixed,  $1/(\ell_{\text{max}}+0.5)^4$ , and  $1/\ell_{\text{max}}^5$  formulas are closely grouped and in good agreement with the Schwenke value. Only minor fluctuations in the extrapolated values are observed beyond the V7Z level.

Unfortunately, in the case of the singlet pair energy the difference between the raw cc-pV10Z value and the CCSD-R12 value is 52 times larger, exceeding  $1 mE_h$ . The mixed(8910) prediction differs from the other, more tightly clustered, formulas by  $\sim 1 mE_h$ . Unlike Eqs. (1)–(4) which are used with basis sets of VQZ quality or better, Schwenke has also provided fitting coefficients for use with VDZ and VTZ basis sets. Even when applied with such low-level basis sets, the Schwenke(DT) estimate is within  $3.6 mE_h$  of the CCSD-R12 value, a level of agreement not reached until the V7Z level in the raw energies. While most formulas approach their limiting value from below, the  $1/\ell_{\text{max}}^3$  expression approaches from above. The exact agreement between the  $1/\ell_{\text{max}}^3$ (910) extrapolated energy and the CCSD-R12 value may be fortuitous, but this functional form clearly does the best job at reproducing the R12/(19s, 14p, 8d, 6f, 4g, 3h) value. The convergence of the exponential, mixed, and  $1/(\ell_{\text{max}}+0.5)^4$  formulas is more erratic, with (789)  $\rightarrow$  (8910) changes of 0.29, 0.28, and  $0.07 mE_h$ , respectively.

Figure 3 shows the convergence of the CCSD(T)(FC)/CBS extrapolated energies for Ne ( $1S$ ), CS ( $1\Sigma^+$ ),  $CS_2$  ( $1\Sigma_g^+$ ), and  $C_6H_6$  ( $1A_{1g}$ ). The energies being extrapolated were obtained from the cc-pVnZ (Ne), aug-cc-pVnZ ( $C_6H_6$ ), and aug-cc-pVnZ(C)/aug-cc-pV( $n+d$ )Z(S) (CS and  $CS_2$ ) basis sets. As the size of the chemical system increases, the spread among the extrapolated energies at the low end (DZ, TZ, and QZ) expands. For example, while the QZ estimates for Ne span a range of about  $10 mE_h$ , for  $CS_2$  the range is  $12 mE_h$  and for  $C_6H_6$  it is  $25 mE_h$ . When larger basis set calculations are affordable, the spread among the CBS formulas is less dependent on the size of the system. For Ne the (789) CBS estimates are within  $\sim 1.5 mE_h$  of each other. As was the case for the singlet/triplet pair energies, the mixed(789) estimate lies slightly outside the cluster of energies produced by the other four formulas. In order to make the comparisons in Fig. 3 as straightforward and unbiased as possible, the total CCSD(T) energies reported for Eqs. (1)–(4) were obtained by summing separate extrapolations of the SCF, CCSD cor-

TABLE II. Computed and CBS extrapolated Ne ( $^1S$ ) CCSD singlet/triplet pair energies ( $E_h$ ).

| Basis sets <sup>a</sup> | Computed <sup>b</sup> | Exp       | Singlet   |                         |                        | Schwenke <sup>c</sup> |
|-------------------------|-----------------------|-----------|-----------|-------------------------|------------------------|-----------------------|
|                         |                       |           | Mixed     | $1/(\ell_{\max}+0.5)^4$ | $1/\ell_{\max}^3$      |                       |
| VDZ,VTZ                 | -0.170 53             |           |           |                         |                        | -0.207 03             |
| (VDZ),VTZ,VQZ           | -0.192 73             | -0.209 51 | -0.205 95 | -0.205 54               | -0.208 93              | -0.209 76             |
| (VTZ),VQZ,V5Z           | -0.201 63             | -0.207 57 | -0.206 81 | -0.208 85               | -0.210 96              | -0.210 41             |
| (VQZ),V5Z,V6Z           | -0.205 47             | -0.208 40 | -0.207 71 | -0.209 52               | -0.210 75 <sup>d</sup> | -0.210 53             |
| (V5Z),V6Z,V7Z           | -0.207 48             | -0.209 69 | -0.208 65 | -0.210 09               | -0.210 90              |                       |
| (V6Z),V7Z,V8Z           | -0.208 58             | -0.209 92 | -0.209 23 | -0.210 28               | -0.210 82              |                       |
| (V7Z),V8Z,V9Z           | -0.209 19             | -0.209 94 | -0.209 39 | -0.210 28               | -0.210 63              |                       |
| (V8Z),V9Z,V10Z          | -0.209 58             | -0.210 23 | -0.209 67 | -0.210 35               | -0.210 61              |                       |
| CCSD-R12/B <sup>c</sup> | -0.210 61             |           |           |                         |                        |                       |
| Basis sets <sup>a</sup> | Computed <sup>b</sup> | Exp       | Triplet   |                         |                        | Schwenke <sup>c</sup> |
|                         |                       |           | Mixed     | $1/(\ell_{\max}+0.5)^4$ | $1/\ell_{\max}^5$      |                       |
| VDZ,VTZ                 | -0.095 81             |           |           |                         |                        | -0.105 01             |
| (VDZ),VTZ,VQZ           | -0.101 95             | -0.103 87 | -0.105 21 | -0.105 50               | -0.103 86              | -0.104 80             |
| (VTZ),VQZ,V5Z           | -0.103 86             | -0.104 73 | -0.104 97 | -0.105 42               | -0.104 79              | -0.104 85             |
| (VQZ),V5Z,V6Z           | -0.104 43             | -0.104 68 | -0.104 77 | -0.105 04               | -0.104 82 <sup>d</sup> | -0.104 86             |
| (V5Z),V6Z,V7Z           | -0.104 67             | -0.104 84 | -0.104 80 | -0.104 97               | -0.104 87              |                       |
| (V6Z),V7Z,V8Z           | -0.104 78             | -0.104 89 | -0.104 84 | -0.104 96               | -0.104 90              |                       |
| (V7Z),V8Z,V9Z           | -0.104 83             | -0.104 85 | -0.104 85 | -0.104 90               | -0.104 88              |                       |
| (V8Z),V9Z,V10Z          | -0.104 85             | -0.104 86 | -0.104 85 | -0.104 88               | -0.104 87              |                       |
| CCSD-R12/B <sup>c</sup> | -0.104 87             |           |           |                         |                        |                       |

<sup>a</sup>Range of basis sets used in the CBS extrapolation. Three basis sets are required for the exponential and mixed extrapolations.

<sup>b</sup>Raw value obtained with the largest basis set in this group.

<sup>c</sup>Produced with the singlet/triplet pair energy coefficients from Table VII and Eq. (8) of Schwenke (Ref. 86). The Ne atom was one of the seven systems in the training set for the Schwenke CBS procedure.

<sup>d</sup>The  $1/\ell_{\max}^3$  and  $1/\ell_{\max}^5$  CBS values obtained from the cc-pV5Z and cc-pV6Z basis sets were previously reported by Klopper (Ref. 48).

<sup>e</sup>Obtained with the (19s,14p,8d,6f,4g,3h) uncontracted basis set.

relation energy [ $E_{\text{CCSD}(\text{corr})}$ ], and (T) correction, in a manner analogous to the procedure followed by the Schwenke formulas.

For Ne, the SCF energies obtained with the cc-pV8Z, cc-pV9Z, and cc-pV10Z basis sets are  $-128.547\,096\,9$ ,  $-128.547\,097\,6$ , and  $-128.547\,097\,8\,E_h$ , respectively. This compares to a numerical, finite element value of  $-128.547\,098\,0\,E_h$ .<sup>54</sup> The present best estimate of the Ne CCSD(T)(FC) energy is  $-128.8687 \pm 0.0005\,E_h$ , which is based on the average of the five CBS extrapolations, Eqs. (1)–(4) and (7), applied to the total CCSD(T) energies. The uncertainty is based on the spread in the CBS energies. An alternative estimate, obtained from separate extrapolations of the (9Z/10Z) singlet/triplet CCSD pair energies with  $1/\ell_{\max}^3$  and  $1/\ell_{\max}^5$  expressions and the (T) correction with  $1/\ell_{\max}^3$  yields a value of  $-128.8691\,E_h$ , which falls within the upper range of the previously mentioned best estimate. Both of these values are in essentially exact agreement with the “estimated limit” frozen-core value of  $-128.869 \pm 0.001\,E_h$  reported by Halkier *et al.*<sup>45</sup> and based in part on a CCSD(T)-R12B/(18s,13p,11d,9f,7g,5h) energy of  $-128.8690\,E_h$ . A later, slightly higher CCSD(T)-R12B/(19s,14p,8d,6f,4g,3h) energy of  $-128.86891\,E_h$  was given by Klopper.<sup>48</sup>

After inclusion of the extrapolated core/valence energies using the same five formulas, we arrive at a best estimate for the all-electron CCSD(T) energy of  $-128.9372 \pm 0.0008\,E_h$ ,

whose error bars just include the  $-128.938 \pm 0.001\,E_h$  value of Halkier *et al.* Reducing the error bars associated with these estimates via direct computation presents a formidable challenge. A single 506 basis function CCSD(T)/cc-pV10Z Ne atom calculation with Dalton required 41 wall clock hours on a 2 GHz Opteron processor. The cc-pV11Z basis set, which is the next set in the correlation consistent sequence, includes 650 basis functions. At this point in the standard basis set approach, we are in a regime where enormous efforts must be expended to make modest progress.

For  $N_2$ , the Hartree-Fock limit at the optimal restricted Hartree-Fock (RHF) bond length ( $1.0654\,\text{\AA}$ ) is  $-108.996\,558\,9\,E_h$ . Our best estimate of the CCSD(T)(FC)/CBS energy at the optimal bond length ( $1.0989\,\text{\AA}$ ) is  $-109.4213 \pm 0.0007\,E_h$ . Following the procedure established for the Ne atom, this estimate is based on the average of the energies obtained from the five CBS formulas and the uncertainty is taken from the spread in values. The exponential and mixed extrapolated values lie in the upper end of this range, i.e., their energies are higher than the inverse  $\ell_{\max}$  values. CBS energies obtained from the exponential formula decrease by  $-0.6\,mE_h$  between extrapolations performed with the aug-cc-pV6Z and aug-cc-pV7Z basis sets, whereas the  $1/\ell_{\max}^3$  energies increase by  $1.8\,mE_h$ . A second estimate was obtained by (1) extrapolating the SCF energy with Eq. (1) and (2) extrapolating the (aV6Z/aV7Z) singlet/triplet CCSD pair energies, and (3) extrapolating the (T) correction,

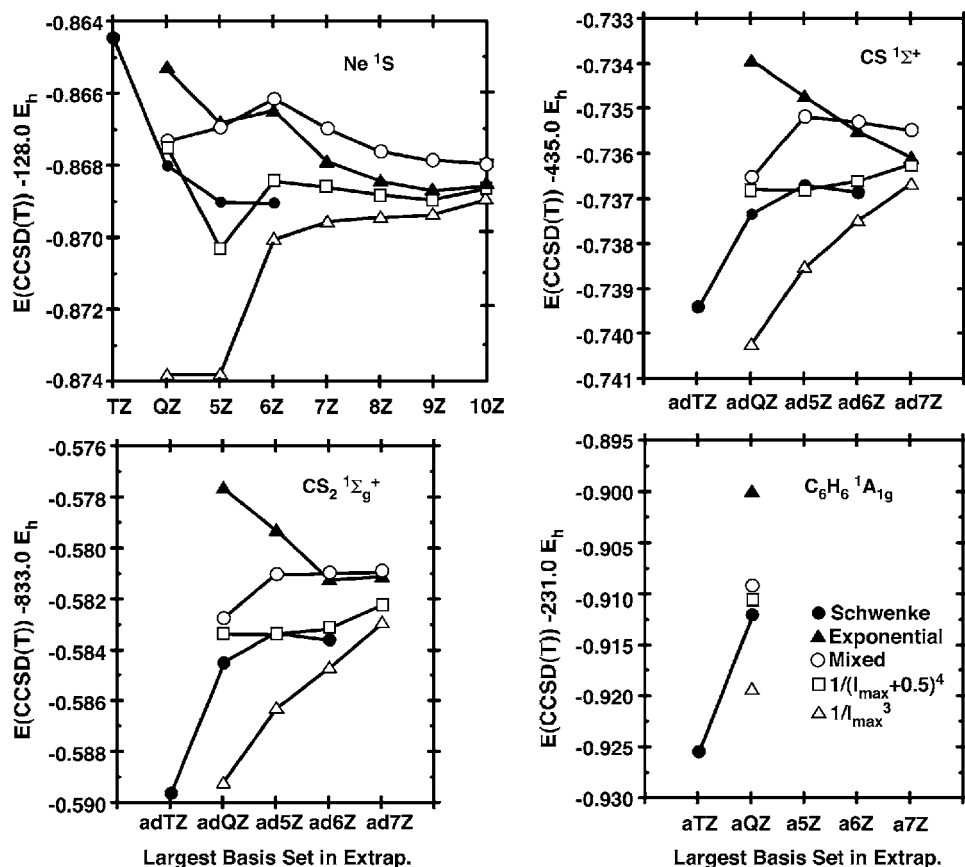


FIG. 3. Convergence of the  $\text{Ne}(1S)$ ,  $\text{CS}(1\Sigma^+)$ ,  $\text{CS}_2(1\Sigma_g^+)$ , and  $\text{C}_6\text{H}_6(1A_{1g})$  CCSD(T)(FC)/CBS energies for the cc-pVnZ(Ne), aug-cc-pVnZ(C), and aug-cc-pV(n+d)Z(S) basis sets. The exponential, mixed,  $1/(\ell_{\max} + 0.5)^4$ , and  $1/\ell_{\max}^3$  extrapolations were applied to the total energy. The Schwenke estimate was based on separate extrapolations of the SCF, CCSD correlation energy, and the (T) correction.

which yields a slightly lower value of  $-109.4218 E_h$ . The current best estimate compares to a CCSD(T)/R12/B value of  $-109.4204 E_h$  and a CBS value of  $-109.4224$  reported by Halkier *et al.*<sup>45</sup> at  $r=1.0977 \text{ \AA}$ . Their estimate was based on a value slightly lower than the R12/B results. The same authors reported a core/valence CCSD(T)/CBS value of  $-109.541 \pm 0.002 E_h$  and noted that their value was  $4 mE_h$  lower than the  $-109.5369 E_h$  estimate of Peterson *et al.*<sup>55</sup> This difference was attributed to the use of the exponential formula [Eq. (1)] by Peterson *et al.* However, while the exponential formula does tend to underestimate the contribution from higher angular momentum basis functions,  $\sim 2 mE_h$  of the difference is due to the use of a least-squares fitting of the CVDZ through CV6Z energies by Peterson *et al.* Since there is no variational principle involved, the prediction of a lower CBS energy does not guarantee a better estimate. Because they were unable to carry out a CCSD(T)/R12 all-electron calculation on  $\text{N}_2$ , Halkier *et al.*<sup>45</sup> constructed their estimate by adding the [CCSD/cc-pCV(56)Z estimated limit—CCSD/cc-pCV6Z] difference to the CCSD(T)/cc-pCV6Z value.<sup>56</sup> Our current best estimate of the CCSD(T)(CV)/CBS energy is  $-109.5401 \pm 0.0012 E_h$ , where the uncertainty is taken as the sum of the uncertainties associated with extrapolating the separate CV and FC energies.

For sulfur we used the aug-cc-pV(n+d)Z family of basis sets, which contain an additional tight  $d$  function for the purpose of partially ameliorating a known SCF-level deficiency in the aug-cc-pVnZ sets for second row elements.<sup>57</sup> We have previously discussed the benefits of the tight  $d$  func-

tions on the atomization energy of  $\text{CS}_2$ .<sup>11</sup> Since Schwenke coefficients were not reported for the aug-cc-pV(n+d)Z basis set, we elected to use the aug-cc-pVnZ coefficients. Despite the lack of any sulfur compounds among the seven systems that comprised the Schwenke training set, Eq. (7) performed well for CS and  $\text{CS}_2$ . In the latter case the Schwenke(56) CBS estimate appears to be  $1\text{--}2 mE_h$  lower than the other formulas. Unlike some of the other formulas, which in the four cases shown in Fig. 3 approach a limiting value from either above (exponential) or below ( $1/\ell_{\max}^3$ ), the manner in which the Schwenke formula approaches its limiting value is system dependent.

### C. Energy differences

While high absolute accuracy in the total energy is an attractive, long-term goal, it is unnecessary for achieving correspondingly high accuracy in thermochemical or spectroscopic properties. Nonetheless, many CBS studies have focused on absolute accuracy in total energies. Instead, we have chosen to stress energy differences, arguing that this perspective is more in-line with the way electronic structure methods are used in practice. In most cases, ensuring high absolute accuracy in the total energy is prohibitively expensive for all but very small systems, even with explicitly correlated methods. Thus, practically all computational approaches targeted at small molecules, as defined in this work, have abandoned absolute accuracy as a goal in favor of methods possessing balanced errors that can be calibrated for

representative classes of molecules. In the following section we shift our focus away from absolute accuracy in the total energy to accuracy in energy differences.

Atomic energies are needed for computing  $D_e$  for diatomics and total atomization energies,  $\Sigma D_e$ , for polyatomics. In the present work,  $D_e$  and  $\Sigma D_e$  were obtained from a combination of three different open-shell methods. Most of our results are based on the R/UCCSD(T) method, which utilizes restricted open-shell Hartree-Fock (ROHF) orbitals, but allows a small amount of spin contamination in the solution of the CCSD equations. It is requested in MOLPRO with the keyword "UCCSD(T)." Full atomic symmetry was imposed on the orbitals. There are two other commonly used open-shell techniques. RCCSD(T), which is requested in MOLPRO with the keyword "RCCSD(T)," also begins with ROHF orbitals, but subsequently imposes a restriction on the coupled cluster amplitudes such that the linear part of the wave function becomes a spin eigenfunction.<sup>58</sup> The third choice, UCCSD(T), employs unrestricted Hartree-Fock (UHF) orbitals.

As we have previously noted, the choice of open-shell treatment for the isolated atoms has consequences for  $D_e$  and  $\Sigma D_e$ .<sup>11</sup> At the frozen-core level of theory,  $D_e$  for a diatomic molecule such as  $N_2$  can vary by 0.4–0.5 kcal/mol. When core/valence correlation effects are included, energy differences obtained from R/UCCSD(T) and UCCSD(T) atoms largely disappear, in agreement with the comments of Ruden *et al.*<sup>59</sup> However, differences between R/UCCSD(T) and RCCSD(T) energies remain on the order of 0.1–0.2 kcal/mol per first row atom. Similarly, the imposition of symmetry equivalencing in the Hartree-Fock calculation can increase the atomic energies by another 0.1 kcal/mol per first row atom. Thus, for a system such as  $C_8H_{18}$  different open-shell methods can affect the total atomization energies by 1 kcal/mol or more, an amount equal to our target accuracy. For single configuration dominant systems, evidence is beginning to accumulate that CCSD(T) generally underestimates binding energies predicted by even higher-level methods, such as CCSDTQ or FCI. Therefore, it might be argued that the use of RCCSD(T) with symmetry-equivalenced atoms would result in a better agreement with the FCI result. The present choice of R/UCCSD(T) as our primary method for describing isolated atoms is based on our experience with potential-energy surfaces where this method was found to be better behaved than the completely restricted alternative.<sup>60</sup>

$N_2$  spectroscopic constants ( $r_e$ ,  $\omega_e$ , and  $\omega_e x_e$ ) and dissociation energies obtained at the CCSD(T)(FC)/aug-cc-pVnZ,  $n=D-7$ , are listed in Table III. The first three properties were obtained from a sixth degree Dunham fit of the potential-energy curve.<sup>61</sup> If convergence is measured as a percentage of the DZ to 7Z change in property value, the four measurables are 94% ( $r_e$ ), 84% ( $\omega_e$ ), 100% ( $\omega_e x_e$ ), and 88% ( $D_e$ ) converged at the aug-cc-pVQZ basis set level. The aug-cc-pVQZ value of the binding energy is more than 3 kcal/mol smaller than the estimated CBS limit. Nevertheless, it is converging much faster than the total energy. For example, at the 7Z level the CCSD(T)(FC) energy of Ne, an

eight valence electron system, remains  $>3 mE_h$  above the CBS limit, while  $D_e(N_2)$ , involving ten valence electrons, is converged to  $\sim 0.8 mE_h$  (0.5 kcal/mol).

Core/valence corrections to the properties exhibit little variation with respect to basis set improvement beyond the cc-pCVQZ level. The CV corrections to  $D_e$ , assuming R/UCCSD(T) atoms, are 0.72 (CVQZ), 0.77 (CV5Z), and 0.79 (CV6Z) kcal/mol. These findings are consistent with other investigations of the basis set dependence of CV corrections in polyatomic systems.<sup>7,8</sup> Most of our previous thermochemical studies have relied on CVQZ calculations. Relativistic corrections to the properties in Table III were obtained from spin-free, one-electron Douglas-Kroll-Hess (DKH) CCSD(T) calculations.<sup>62</sup> For  $N_2$ , these corrections are considerably smaller than the CV corrections and are essentially converged at the cc-pVTZ\_DK basis set level to the number of digits shown.

Complete basis set extrapolation dramatically improves the raw  $D_e$  values, although the convergence of the estimates as a function of the underlying basis sets is not always uniform, as seen in Fig. 4. We have previously shown how the CCSD(T)/CBS dissociation energies for  $N_2$ ,  $O_2$ , CO, and  $S_2$  vary as a function of the underlying basis sets up through aug-cc-pV7Z( $k$  estimate) when using Eqs. (1)–(4).<sup>52</sup> In that work the separated atoms were described with the UCCSD(T) and RCCSD(T) methods. All four extrapolation formulas proved capable of predicting dissociation energies only attainable from explicit calculations with basis sets that were far larger than those used in the extrapolation. For example, the exponential (DTQ) and  $1/(\ell_{\max} + 0.5)^4$ (TQ) CBS estimates for  $D_e$  are 226.9 and 227.1 kcal/mol, respectively. In order to achieve similar values via direct calculation, i.e., without resorting to extrapolation techniques, basis sets of better than aug-cc-pV7Z quality would be required. With the largest basis sets available, the span in the CBS values of  $D_e$  is 0.4 kcal/mol.

$N_2$  is another of the seven systems in the Schwenke training set. Consequently, the Schwenke extrapolation formula might be expected to perform well for  $N_2$  and we found this to be the case. As the quality of the underlying basis sets improves, the Schwenke CBS estimates show little variation, provided sets of QZ quality or better are used. When smaller basis sets are used, such as in the Schwenke(DT) extrapolation, the predicted binding energy (225.6 kcal/mol) differs by several kcal/mol from the estimates obtained with larger basis sets. Despite this difference, the Schwenke(DT) estimate still represents a substantial improvement over the raw aug-cc-pVTZ value of 218.0 kcal/mol. The average of the five CBS formulas, using the best available energies, is 227.1 kcal/mol with a spread of  $\pm 0.2$  kcal/mol. Due to the relatively close agreement among the various CBS values for  $D_e$ , it is difficult to single out one as being clearly superior to the others. As stated previously, one of the assumptions made in the Schwenke approach is that the fitting coefficients are independent of geometry. We tested that assumption by extrapolating the energies at seven bond lengths along the  $N_2$  potential-energy curve and subsequently performing a Dunham analysis. The resulting CBS estimates for  $r_e$ ,  $\omega_e$ , and  $\omega_e x_e$  were in good agreement with the estimates obtained

TABLE III.  $N_2 \ ^1\Sigma_g^+$  theoretical results. The units are  $r_e$  in Å,  $\omega_e$  and  $\omega_e x_e$  in  $\text{cm}^{-1}$ , electronic dissociation energy ( $D_e$ ) in kcal/mol, and total energies ( $E$ ) at  $r_e$  in hartrees. The CCSD(T)(FC)/aug-cc-pVnZ dissociation energies and CCSD(T)/cc-pCVnZ and CCSD(T)-DK(FC)/cc-pVnZ dissociation energy corrections are with respect to the R/UCCSD(T) atomic energies obtained with symmetry equivalencing and with the removal of high angular momentum contaminants. The CCSDTQ and FCI dissociation energy corrections are with respect to RHF/ROHF-UCCSDTQ and ROHF-FCI atoms. Total energies are reported at the optimal bond length for each level of theory, with the exception of the FCI/cc-pVTZ energy reported by Ruden *et al.* (Ref. 59) which was obtained at the fixed bond length  $r=1.1040$  Å. The “composite” values are the result of combining all of the corrections associated with each level of basis set, or in the case where explicit corrections could not be determined, e.g., the FCI(FC)  $\Delta$  correction with the cc-pVQZ basis set, the corrections from the best available basis set.

| Basis              | Method                     | $r_e$   | $\omega_e$ | $\omega_e x_e$ | $D_e$                             | $E$          |
|--------------------|----------------------------|---------|------------|----------------|-----------------------------------|--------------|
| aug-cc-pVDZ        | CCSD(T)(FC)                | 1.1209  | 2319.0     | 14.4           | 201.8                             | -109.295 320 |
| cc-pCVDZ           | CCSD(T) $\Delta^a$         | -0.0005 | 3.2        | 0.0            | 0.4                               | -109.356 338 |
| cc-pVDZ            | CCSD(T) DKH $\Delta^b$     | -0.0001 | -1.0       | 0.0            | -0.1                              | -109.334 659 |
| cc-pVDZ            | CCSDTQ(FC) $\Delta^c$      | 0.0009  | -10.9      | 0.2            | 0.81                              | -109.278 113 |
| cc-pVDZ            | FCI(FC) $\Delta^{d,e}$     | 0.0003  | -4.8       | 0.0            | 0.14                              | -109.278 340 |
| DZ                 | Composite                  | 1.1215  | 2305.5     | 14.6           | 203.0                             |              |
| aug-cc-pVTZ        | CCSD(T)(FC)                | 1.1040  | 2339.6     | 13.8           | 218.0                             | -109.380 845 |
| cc-pCVTZ           | CCSD(T) $\Delta^a$         | -0.0016 | 7.9        | 0.0            | 0.7                               | -109.482 426 |
| cc-pVTZ            | CCSD(T) DKH $\Delta^b$     | -0.0002 | -0.9       | 0.0            | -0.1                              | -109.432 244 |
| cc-pVTZ            | CCSDTQ(FC) $\Delta^c$      | 0.0007  | -9.5       | 0.5            | 0.41                              | -109.375 242 |
| cc-pVTZ            | FCI(FC) $\Delta^f$         | 0.0002  | -3.9       |                | 0.05                              | -109.375 4   |
| TZ                 | Composite                  | 1.1029  | 2333.2     | 14.3           | 219.1                             |              |
| aug-cc-pVQZ        | CCSD(T)(FC)                | 1.1005  | 2354.5     | 13.8           | 223.8                             | -109.407 243 |
| cc-pCVQZ           | CCSD(T) $\Delta^a$         | -0.0020 | 9.2        | 0.0            | 0.7                               | -109.519 838 |
| cc-pVQZ            | CCSD(T) DKH $\Delta^b$     | -0.0002 | -0.9       | 0.0            | -0.1                              | -109.462 840 |
| QZ                 | Composite <sup>g</sup>     | 1.0992  | 2349.4     | 14.3           | 224.9                             |              |
| aug-cc-pV5Z        | CCSD(T)(FC)                | 1.0995  | 2359.3     | 13.8           | 225.6                             | -109.415 507 |
| cc-pCV5Z           | CCSD(T) $\Delta$           | -0.0020 | 9.7        | 0.0            | 0.8                               | -109.531 509 |
| cc-pV5Z            | CCSD(T) DKH $\Delta^b$     | -0.0002 | -0.9       | 0.0            | -0.1                              | -109.472 686 |
| 5Z                 | Composite <sup>g</sup>     | 1.0982  | 2354.7     | 14.3           | 226.8                             |              |
| aug-cc-pV6Z        | CCSD(T)(FC)                | 1.0993  | 2360.6     | 13.8           | 226.3                             | -109.418 384 |
| cc-pCV6Z           | CCSD(T) $\Delta^a$         | -0.0021 | 10.0       | 0.0            | 0.8                               | -109.535 855 |
| 6Z                 | Composite <sup>g,h</sup>   | 1.0979  | 2356.3     | 14.3           | 227.5                             |              |
| aug-cc-pV7Z        | CCSD(T)(FC)                | 1.0991  | 2361.3     | 13.8           | 226.7                             | -109.419 714 |
| 7Z                 | Composite <sup>g,h,i</sup> | 1.0977  | 2357.0     | 14.3           | 227.9                             |              |
| CBS                | CCSD(T)(FC) <sup>j</sup>   | 1.0989  | 2362.0     | 13.8           | 227.2±0.2                         | -109.421 3   |
| CBS                | Composite                  | 1.0975  | 2358.1     | 14.3           | 228.4±0.2<br>(228.7) <sup>k</sup> |              |
| Expt. <sup>l</sup> |                            | 1.0977  | 2358.6     | 14.3           | 228.4                             |              |

<sup>a</sup>Difference between the frozen-core and non-frozen-core values, i.e.,  $X-X_{\text{FC}}$ . Frozen-core calculations excluded the  $N(1s)$  electrons from the correlation treatment. In  $D_{\infty h}$  symmetry, this corresponds to  $1\sigma_g^2 1\sigma_u^2$ . Non-frozen-core calculations correlated all 14 electrons.

<sup>b</sup>Difference between the CCSD(T)(FC) Douglas-Kroll-Hess and nonrelativistic values, i.e.,  $X-X_{\text{DKH}}$ .

<sup>c</sup>RHF( $N_2$ )/ROHF(N)-UCCSDTQ/cc-pVnZ correction to the R( $N_2$ )/R/U(N)CCSD(T)/cc-pVnZ values. The cc-pVDZ results were obtained from a seven-point Dunham fit. The cc-pVTZ results were taken from Ruden *et al.* (Ref. 59).

<sup>d</sup>Full CI/cc-pVDZ correction to the CCSDTQ/cc-pVDZ values. The FCI results were taken from Leininger *et al.* (Ref. 87).

<sup>e</sup>CCSDTQ5/cc-pVDZ correction to the CCSDTQ/cc-pVDZ values are  $r_e$  0.0003 Å,  $\omega_e$ -3.9  $\text{cm}^{-1}$ , and  $D_e$  0.1 kcal/mol taken from Ruden *et al.* (Refs. 59 and 79).

<sup>f</sup>Based on the approximate FCI/cc-pVTZ energy of Ruden *et al.* (Ref. 59) which has a stated uncertainty of  $\pm 0.0001 E_h$ . This implies an uncertainty of  $\pm 0.06$  kcal/mol in the FCI correction to  $D_e$ . The  $r_e$  and  $\omega_e$  corrections are based on the continued fraction approximation.

<sup>g</sup>Includes the CCSDTQ/cc-pVTZ and FCI/cc-pVTZ ( $D_e$ ) corrections.

<sup>h</sup>Includes the CCSD(T)/cc-pV5Z DKH correction.

<sup>i</sup>Includes the CCSD(T)/cc-pCV6Z correction.

<sup>j</sup>The CBS values of  $r_e$ ,  $\omega_e$ , and  $\omega_e x_e$  are based on an exponential extrapolation of the aV5Z, aV6Z, and aV7Z values.  $D_e$  was based on an average of the five CBS estimates discussed in the text. The uncertainty is taken from the spread in the CBS estimates.

<sup>k</sup>Dissociation energy based on the RCCSD(T) atomic asymptotes with equivalenced orbitals.

<sup>l</sup>Experimental values are taken from Huber and Herzberg (Ref. 4). The uncertainty in  $D_e$  is  $\pm 0.01$  kcal/mol. The NIST/JANAF value for  $D_e$  is  $228.4 \pm 0.4$  kcal/mol.

from an exponential extrapolation. For example, the Schwenne(TQ) formula predicted values of  $r_e=1.0987$  Å,  $\omega_e=2363.6$   $\text{cm}^{-1}$ , and  $\omega_e x_e=13.7$   $\text{cm}^{-1}$ , compared to the exponential values of  $r_e=1.0989$  Å,  $\omega_e=2362.4$   $\text{cm}^{-1}$ , and  $\omega_e x_e=13.8$   $\text{cm}^{-1}$ . In the case of the anharmonicity, the Schwenne(TQ) CBS estimate is slightly too small, since the raw, directly computed values appear stable at  $13.8$   $\text{cm}^{-1}$  for basis sets of aug-cc-pVTZ quality or better.

Theoretical results for CS and  $\text{CS}_2$  are shown in Tables IV and V, along with the available experimental data. The spectroscopic constants for CS were taken from Huber and Herzberg.<sup>4</sup> We have elected to apply an adjustment to the experimental dissociation energy in order to account for atomic spin-orbit effects. This adjustment, which increases the observed experimental value, arises from the failure of most electronic structure programs to properly treat the

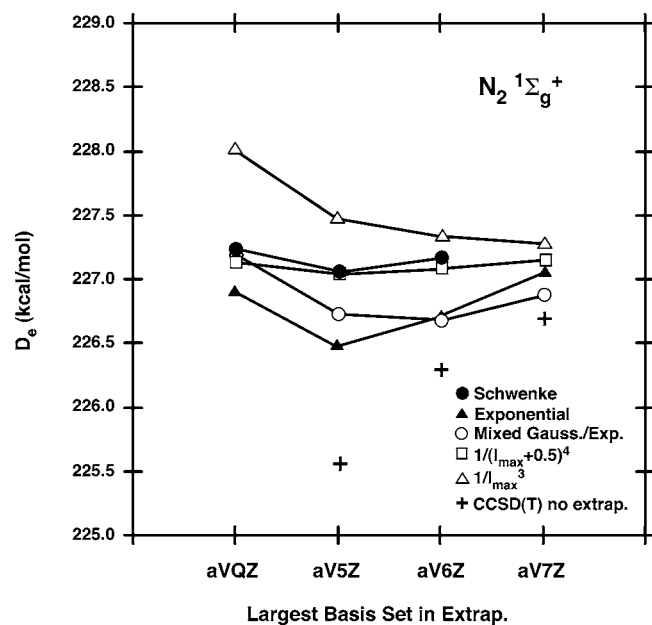


FIG. 4. Convergence of the  $N_2$  extrapolated complete basis set and raw CCSD(T)(FC)  $D_e$  values (kcal/mol) as a function of the quality of the underlying basis sets. The separated atoms are described with the R/UCCSD(T) method.

lowest-energy multiplet of the dissociated atoms. The atomic spin-orbit corrections were based on the tables of Moore,<sup>63</sup> and are as follows (in kcal/mol): 0.083 (C) and 0.545 (S). The experimental bond length for  $CS_2$  was taken from the work by Maki and Sams.<sup>64</sup> The vibrational fundamentals,  $\nu_i$ , are from Shimanouchi.<sup>65</sup> The vibrationless atomization energy was derived from the  $\Sigma D_0$  value reported in the *NIST/JANAF Tables*<sup>5</sup> and a theoretical anharmonic ZPE of 4.30 kcal/mol. The latter result was obtained from a near-equilibrium potential function based on CCSD(T)/aug-cc-pV(T+d)Z energies. A total of 29 symmetry unique points were fitted to polynomials in  $D_{2h}$  symmetry coordinates. The coefficients of this fit were then used to evaluate the harmonic frequencies and anharmonicity constants via the usual second-order perturbation-theory formulas. However, for the final calculation of the ZPE the aug-cc-pV(5+d)Z harmonic frequencies were substituted for the aug-cc-pV(T+d)Z values. Both the fitting and spectroscopic constant analyses were carried out with the program SURFIT.<sup>66</sup>

In general, the one-particle basis set convergence patterns for CS and  $CS_2$  appear qualitatively similar to what was observed for  $N_2$ . Differences between DZ values and the CBS limits for bond lengths ( $\sim 0.02$  Å), harmonic frequencies ( $25$ – $45$   $cm^{-1}$ ), and dissociation energies ( $\sim 26$  kcal/mol) span approximately the same range, with the exception of  $D_e(CS)$ , which at 14 kcal/mol is only about half the size of the  $N_2$  and  $CS_2$  values. The presence of additional tight  $d$  functions in the aug-cc-pV( $n+d$ )Z sulfur basis sets appreciably improves the computed properties, an effect that is most noticeable at the VDZ level. For atomization energies, tight  $d$  functions increase  $\Sigma D_e$  by up to 5 kcal/mol. Because the aug-cc-pVnZ and aug-cc-pV( $n+d$ )Z basis set sequences both lead to a complete basis set, the properties obtained from the two sequences must ultimately converge. At the 6Z

level, the dissociation energies obtained with and without tight  $d$  functions differ by less than 0.5 kcal/mol.

For comparison purposes, we have also carried out density-functional theory (DFT) calculations of  $D_e(CS)$  and  $\Sigma D_e(CS_2)$  using the hybrid B3LYP method.<sup>67</sup> DFT is known to converge more rapidly with respect to the one-particle basis set than conventional *ab initio* methods.<sup>68,69</sup> For example, the difference between the TZ and QZ values of  $\Sigma D_e(CS_2)$  is 0.5 kcal/mol (B3LYP) versus 6.5 kcal/mol [CCSD(T)]. We find, in agreement with the comments of Bauschlicher and Partridge,<sup>68</sup> that tight  $d$  functions are important for obtaining atomization energies that are close to the Kohn-Sham limit. The difference between the aug-cc-pVDZ and aug-cc-pV( $D+d$ )Z atomization energies for  $CS_2$  is 5.5 kcal/mol, a value which exceeds the 4.8 kcal/mol found at the CCSD(T) level of theory. The agreement with experiment is somewhat variable. For CS the B3LYP/aug-cc-pV(5+d)Z value of  $D_e=166.6$  kcal/mol can be compared to an adjusted experimental value of 171.9 kcal/mol. In the case of  $CS_2$ , B3LYP did a better job, 277.1 (B3LYP) versus 279.1 (expt.).

Figure 5 shows the variation in the CBS estimates of  $D_e$  and  $\Sigma D_e$  for CS and  $CS_2$ . As was the case for  $N_2$ , all of the formulas considered in this paper produced binding energies that were markedly closer to the apparent CBS limit than the raw, directly computed values, indicated by the + symbols in the figure. The Schwenke(DT) CBS extrapolated results are 2–3 kcal/mol smaller than the relatively tight grouping of values obtained with basis sets of QZ quality or better. In situations where TZ calculations represent the best available data, application of the Schwenke(DT) extrapolation appears to be effective at improving the raw binding energy. For  $\Sigma D_e(CS_2)$  the raw TZ binding energy (268.1 kcal/mol) is improved by over 7 kcal/mol. The Schwenke results vary so little with basis set size, at least with QZ or better basis sets, that it is questionable whether calculations at the 5Z or 6Z level are worthwhile. In the case of the other five formulas, expanding the basis set beyond the QZ level often results in an improvement in the CBS estimate. This is most noticeable in Fig. 5 for the  $1/l_{\max}^3$  results.

#### IV. $N$ -PARTICLE BASIS SET CONSIDERATIONS

Having addressed the error arising from the truncation of the one-particle expansion, we now turn our attention to the  $n$ -particle expansion. Our discussion so far has focused on the CCSD(T) method in which triple excitations are handled in an approximate, noniterative step. In the CCSDT method, triple excitations are treated in the same iterative fashion as the singles and doubles. Compared to CCSD(T), whose computational expense grows as  $\sim n^2 N^4 N_{it}$  (with a single  $n^3 N^4$  step), CCSDT exhibits a  $\sim n^3 N^5 N_{it}$  scaling, where  $n$ ,  $N$ , and  $N_{it}$  are the numbers of occupied and unoccupied molecular orbitals and the number of CCSD or CCSDT iterations, respectively.<sup>70</sup> In multiple studies we have examined the performance of CCSDT relative to CCSD(T), FCI, and experiment.<sup>11,52,71,72</sup> Overall, CCSDT corrections to CCSD(T) were found to be significant in magnitude, relative to a target accuracy of  $\pm 1$  kcal/mol, but the correction fre-

TABLE IV. CS  $^1\Sigma^+$  theoretical results. The units are  $r_e$  in Å,  $\omega_e$  in  $\text{cm}^{-1}$ , the electronic dissociation energy ( $D_e$ ) in kcal/mol, and total energies ( $E$ ) at  $r_e$  in hartrees. The CCSD(T) dissociation energies are with respect to R/UCCSD(T) atomic energies obtained with symmetry equivalencing and with the removal of high angular momentum contaminants. The notation “aug-cc-pV( $n+d$ )Z” indicates the use of the mixed aug-cc-pVnZ(C)/aug-cc-pV( $n+d$ )Z(S) basis set.

| Basis                            | Method                    | $r_e$   | $\omega_e$ | $\omega_e x_e$ | $D_e$                | $E$          |
|----------------------------------|---------------------------|---------|------------|----------------|----------------------|--------------|
| aug-cc-pV( $D+d$ )Z              | CCSD(T)(FC)               | 1.5620  | 1259.7     | 6.6            | 156.8                | -435.629 154 |
| aug-cc-pwCVdZ                    | CCSD(T) $\Delta^a$        | -0.0010 | 2.8        | 0.0            | 0.3                  | -435.809 255 |
| cc-pVDZ_DK                       | CCSD(T) DKH $\Delta^b$    | -0.0005 | -1.2       | 0.0            | -0.1                 | -436.701 008 |
| cc-pVDZ                          | CCSDTQ(FC) $\Delta^c$     | 0.0015  | -7.7       | 0.2            | 0.64                 | -435.607 085 |
| cc-pVDZ                          | cf FCI(FC) $\Delta^d$     | 0.0001  | -0.3       | 0.0            | 0.04                 | -435.607 152 |
| DZ                               | Composite                 | 1.5621  | 1253.3     | 6.8            | 157.6                |              |
| aug-cc-pV( $T+d$ )Z              | CCSD(T)(FC)               | 1.5459  | 1270.3     | 6.4            | 164.9                | -435.701 616 |
| aug-cc-pwCVTZ                    | CCSD(T) $\Delta^a$        | -0.0037 | 6.2        | 0.0            | 0.7                  | -436.021 923 |
| cc-pVTZ_DK                       | CCSD(T) DKH $\Delta^b$    | -0.0005 | -1.2       | 0.0            | -0.3                 | -436.789 452 |
| cc-pVTZ                          | CCSDTQ(FC) $\Delta^c$     | 0.0015  | -6.2       | 0.0            | 0.48                 | -435.694 927 |
| cc-pVTZ                          | cf FCI(FC) $\Delta^d$     | 0.0002  | -0.8       |                | 0.08                 | -435.695 052 |
| TZ                               | Composite                 | 1.5434  | 1268.3     | 6.4            | 165.9                |              |
| aug-cc-pV( $Q+d$ )Z              | CCSD(T)(FC)               | 1.5414  | 1280.5     | 6.3            | 168.6                | -435.723 920 |
| aug-cc-pwCVQZ                    | CCSD(T) $\Delta^a$        | -0.0044 | 7.0        | 0.0            | 0.8                  | -436.097 303 |
| cc-pVQZ_DK                       | CCSD(T) DKH $\Delta^b$    | -0.0005 | -1.2       |                | -0.2                 | -436.816 919 |
| QZ                               | Composite <sup>e</sup>    | 1.5382  | 1279.3     | 6.3            | 169.7                |              |
| aug-cc-pV( $5+d$ )Z              | CCSD(T)(FC)               | 1.5397  | 1284.0     | 6.3            | 169.9                | -435.731 024 |
| aug-cc-pwCV5Z                    | CCSD(T) $\Delta^a$        | -0.0047 | 7.4        | 0.0            | 0.8                  | -436.124 548 |
| cc-pV5Z_DK                       | CCSD(T) DKH $\Delta^b$    | -0.0005 | -1.2       |                | -0.2                 | -436.826 498 |
| 5Z                               | Composite <sup>e</sup>    | 1.5362  | 1283.2     | 6.3            | 171.0                |              |
| aug-cc-pV( $6+d$ )Z              | CCSD(T)(FC)               | 1.5391  | 1284.8     | 6.3            | 170.4                | -435.733 738 |
| 6Z                               | Composite <sup>e, f</sup> | 1.5356  | 1284.0     | 6.3            | 171.5                |              |
| aug-cc-pV( $7+d$ )Z <sup>g</sup> | CCSD(T)(FC)               | 1.5389  | 1284.8     | 6.3            | 170.6                | -435.734 734 |
| 7Z                               | Composite <sup>e, f</sup> | 1.5354  | 1284.0     | 6.3            | 171.7                |              |
| CBS                              | CCSD(T)(FC) <sup>h</sup>  | 1.5388  | 1284.8     | 6.3            | 170.8±0.2            | -435.736 2   |
| CBS                              | Composite <sup>e, f</sup> | 1.5353  | 1284.0     | 6.3            | 171.9±0.2            |              |
|                                  |                           |         |            |                | (172.1) <sup>i</sup> |              |
| Expt. <sup>j</sup>               |                           | 1.5349  | 1285.1     | 6.5            | 172.0                |              |
| Expt. <sup>k</sup>               |                           |         |            |                | 171.8±6              |              |

<sup>a</sup>Difference between the frozen-core and non-frozen-core values, i.e.,  $X - X_{\text{FC}}$ . Frozen-core calculations excluded the 8 C(1s) and S(1s, 2s, 2p) electrons from the correlation treatment. In  $C_{\infty v}$  symmetry, this corresponds to  $1\sigma^2 2\sigma^2 3\sigma^2 4\sigma^2 1\pi^4$ . Non-frozen-core calculations excluded two electrons, corresponding to the S(1s) or  $1\sigma^2$  electrons.

<sup>b</sup>Difference between the CCSD(T)(FC) Douglas-Kroll-Hess and nonrelativistic values, i.e.,  $X - X_{\text{DKH}}$ .

<sup>c</sup>RHF(CS)/ROHF(C,S)-UCCSDTQ/cc-pVnZ correction to the R(CS)/R/U(C,S)CCSD(T)/cc-pVnZ values.

<sup>d</sup>Continued fraction estimates of the FCI corrections.

<sup>e</sup>Includes the CCSDTQ/cc-pVTZ corrections.

<sup>f</sup>Includes CCSD(T)/aug-cc-pwCV5Z corrections.

<sup>g</sup>Values of  $r_e$ ,  $\omega_e$ , and  $\omega_e x_e$  were obtained without  $k$  functions.

<sup>h</sup>The CBS values of  $r_e$ ,  $\omega_e$ , and  $\omega_e x_e$  are based on an exponential extrapolation of the aV( $5+d$ )Z, aV( $6+d$ )Z, and aV( $7+d$ )Z values.  $D_e$  was based on an average of the five CBS estimates discussed in the text. The uncertainty is taken from the spread in the CBS estimates.

<sup>i</sup>Dissociation energy based on RCCSD(T) atomic asymptotes with equivalenced orbitals.

<sup>j</sup>Experimental values are taken from Huber and Herzberg (Ref. 4). The experimental  $D_e$  has been adjusted by 0.63 kcal/mol in order to account for atomic spin-orbit effects in order to allow a more direct comparison with the current theoretical values which were obtained from atomic calculations that correspond to the average of spin multiplets.

<sup>k</sup>Experimental values are taken from the NIST/JANAF Tables (Ref. 5).

quently produces worse agreement with experiment and FCI. Similar conclusions were reached in more recent studies.<sup>59,73</sup> Thus, although CCSD(T) involves additional approximations beyond CCSDT, in a large number of cases results from this level of theory lie fortuitously closer to FCI and experiment than the latter. Due to the scale used in Fig. 2, the iterative triples contribution to the correlation energy of the Ne atom appears to be nearly indistinguishable from the (T) contribution. This figure shows the incremental energy lowering as-

sociated with the cc-pVnZ sequence. Both T and (T) are approximately an order of magnitude smaller than the double excitation contribution.

The next method within the coupled cluster hierarchy is CCSDTQ, which incorporates connected quadruple excitations. Computationally, it scales as  $n^4 N^6 N_{\text{it}}$ , where  $n$ ,  $N$ , and  $N_{\text{it}}$  are defined as before, and therefore its scope is much more restrictive than CCSD(T). The quadruple excitation contribution to the correlation energy is approximately an

TABLE V.  $\text{CS}_2$   $^1\Sigma_g^+$  theoretical results. Units are  $r_e$  in Å,  $\omega$  in  $\text{cm}^{-1}$ , the total atomization energy ( $\Sigma D_e$ ) in kcal/mol, and total energy ( $E$ ) in hartrees. The CCSD(T) dissociation energies are with respect to R/UCCSD(T) atomic energies obtained without the imposition of symmetry equivalencing for the orbitals.

| Basis              | Method                    | $R_{\text{CS}}$       | $\omega(\pi_u)$      | $\omega(\sigma_g)$   | $\omega(\sigma_u)$    | $\Sigma D_e$             | $E$          |
|--------------------|---------------------------|-----------------------|----------------------|----------------------|-----------------------|--------------------------|--------------|
| aug-cc-pV(D+d)Z    | CCSD(T)(FC)               | 1.5767                | 370.4                | 660.5                | 1536.0                | 252.3                    | -833.395 658 |
| aug-cc-pwCVDZ      | CCSD(T) $\Delta^a$        | -0.0010               | 0.6                  | 1.4                  | 3.4                   | 1.0                      | -833.723 905 |
| cc-pVDZ_DK         | CCSD(T) DKH $\Delta^b$    | -0.0001               | -0.3                 | -1.1                 | -3.0                  | -0.5                     | -835.530 271 |
| cc-pVDZ            | CCSDTQ(FC) $\Delta^c$     | 0.0014                |                      |                      |                       | 0.98                     | -835.356 980 |
| cc-pVDZ            | cf FCI(FC) $\Delta^d$     | 0.0002                |                      |                      |                       | 0.1                      | -833.357 170 |
| DZ                 | Composite                 | 1.5772                | 370.7                | 660.8                | 1536.4                | 253.8                    |              |
| aug-cc-pV(T+d)Z    | CCSD(T)(FC)               | 1.5617                | 394.8                | 665.7                | 1549.0                | 268.1                    | -833.523 308 |
| aug-cc-pwCVTZ      | CCSD(T) $\Delta^a$        | -0.0033               | 2.4                  | 2.0                  | 6.3                   | 1.4                      | -834.115 327 |
| cc-pVTZ_DK         | CCSD(T) DKH $\Delta^b$    | -0.0002               | -0.4                 | -1.1                 | -3.1                  | -0.6                     | -835.685 548 |
| cc-pVTZ            | CCSDTQ(FC) $\Delta^c$     | (0.0014) <sup>e</sup> |                      |                      |                       | (0.7) <sup>e</sup>       |              |
| TZ                 | Composite                 | 1.5596                | 396.8                | 666.6                | 1552.2                | 270.0                    |              |
| aug-cc-pV(Q+d)Z    | CCSD(T)(FC)               | 1.5580                | 397.5                | 670.2                | 1555.3                | 274.6                    | -833.561 391 |
| aug-cc-pwCVQZ      | CCSD(T) $\Delta^a$        | -0.0040               | 2.2                  | 4.1                  | 6.9                   | 1.5                      | -834.255 298 |
| cc-pVQZ_DK         | CCSD(T) DKH $\Delta^b$    | -0.0002               | -0.4                 | -0.9                 | -2.7                  | -0.6                     | -835.732 841 |
| QZ                 | Composite <sup>f</sup>    | 1.5552                | 399.3                | 673.4                | 1559.5                | 276.3                    |              |
| aug-cc-pV(5+d)Z    | CCSD(T)(FC)               | 1.5566                | 397.3                | 671.4                | 1558.5                | 276.9                    | -833.573 514 |
| 5Z                 | Composite <sup>f, g</sup> | 1.5538                | 399.1                | 674.6                | 1562.7                | 278.9                    |              |
| aug-cc-pV(6+d)Z    | CCSD(T)(FC)               | 1.5561                |                      |                      |                       | 277.7                    | -833.578 207 |
| 6Z                 | Composite <sup>f, g</sup> | 1.5533                |                      |                      |                       | 279.7                    |              |
| aug-cc-pV(7+d)Z    | CCSD(T)(FC)               | (1.5559) <sup>h</sup> |                      |                      |                       | 278.1                    | -833.579 940 |
| 7Z                 | Composite <sup>f, g</sup> | 1.5531                |                      |                      |                       | 280.1                    |              |
| CBS                | Composite                 | 1.5531                | 399.1                | 674.6                | 1562.7                | 280.5±0.3                |              |
|                    |                           |                       | (397.8) <sup>i</sup> | (659.3) <sup>i</sup> | (1539.3) <sup>i</sup> | (280.9±0.3) <sup>j</sup> |              |
| Expt. <sup>k</sup> |                           | 1.553                 | 397                  | 658                  | 1535                  | 279.1±0.2                |              |

<sup>a</sup>Difference between the frozen-core and non-frozen-core values, i.e.,  $X-X_{\text{FC}}$ . Frozen-core calculations excluded the 22 C(1s) and S(1s, 2s, 2p) electrons from the correlation treatment. In  $C_{\infty v}$  symmetry, this corresponds to  $(1-4)\sigma_g^2 1\pi_u^4 (1-3)\sigma_u^2 1\pi_g^4$ . Non-frozen-core calculations excluded four electrons, corresponding to the S(1s) electrons, i.e.,  $1\sigma_g^2 1\sigma_u^2$ .

<sup>b</sup>Difference between the CCSD(T)(FC) Douglas-Kroll-Hess and nonrelativistic values, i.e.,  $X-X_{\text{DKH}}$ .

<sup>c</sup>CCSDTQ correction to the CCSD(T) values.

<sup>d</sup>Continued fraction estimates of the FCI correction.

<sup>e</sup>The cc-pVTZ correction for  $r_{\text{CS}}$  is taken by analogy from the correction in CS. The correction to  $\Sigma D_e$  is estimated by scaling the CCSDTQ/cc-pVDZ correction for CS by 0.75. See text.

<sup>f</sup>Includes the CCSDTQ/cc-pVTZ estimated correction for  $\Sigma D_e$ .

<sup>g</sup>Includes CCSD(T)/aug-cc-pwCVQZ and CCSD(T) DKH/cc-pVQZ\_DK corrections.

<sup>h</sup>Estimated from an exponential fit of the aV(Q+d)Z, aV(5+d)Z, and aV(6+d)Z bond lengths.

<sup>i</sup>Estimated theoretical fundamentals obtained by adding the aug-cc-pV(T+d)Z anharmonic corrections.

<sup>j</sup>Dissociation energy based on RCCSD(T) atomic asymptotes with equivalenced orbitals.

<sup>k</sup>Experimental values are taken from  $R_{\text{CS}}$  Maki and Sams (Ref. 64), vibrational fundamentals,  $\nu_i$  (Ref. 65), and  $\Sigma D_e$  (NIST/JANAF) (Ref. 5) with a 1.17 kcal/mol correction for atomic spin-orbit effects.

order of magnitude smaller than the triple excitation contribution, as seen in Fig. 2. Interestingly, the linear approximations to the data sets associated with CCSD, CCSDT [or CCSD(T)], and CCSDTQ all possess similar slopes. CCSDTQ5/cc-pVDZ results have been reported for  $\text{N}_2$ ,  $\text{F}_2$ , and  $\text{HF}$ .<sup>59,74</sup> However, while the  $n^5 N^7 N_{\text{it}}$  scaling of CCSDTQ5 is less severe than FCI, it is nonetheless currently impractical for all but the smallest chemical systems and basis sets. For  $\text{N}_2$ , the CCSDTQ5/cc-pVDZ contribution to  $D_e$  (0.1 kcal/mol, 0.5 kJ/mol) was roughly an order of magnitude smaller than the CCSDTQ contribution (0.9 kcal/mol, 3.7 kJ/mol). For particularly difficult, multireference cases, such as  $\text{C}_2$   $^1\Sigma_g^+$ , Boese *et al.* have reported a CCSDTQ5 correction as large as 0.32 kcal/mol.<sup>75</sup> Among the CCSDTQ5 corrections for small molecules reported by Tajti *et al.*, the largest have been on the order of 0.1 kcal/mol ( $\text{C}_2\text{H}_2$  and CCH), although most are less than half that size.<sup>76</sup> The Ne atom FCI correction shown in Fig. 2 is also an order of magnitude less sensitive to the basis set size than the Q cor-

rection, suggesting that CCSDTQ5 is quite close to the FCI limit. Due to the decreased sensitivity of the CCSDTQ5 and FCI corrections to the size of the basis set, it appears unlikely that basis sets beyond cc-pVTZ would be needed to achieve near quantitative agreement with the CBS limit.

While the CCSD and CCSDT differential correlation energies for Ne [ $E_{\text{corr}}^{\text{diff}}(\text{CCSD})=E_{\text{CCSD}}-E_{\text{SCF}}$  and  $E_{\text{corr}}^{\text{diff}}(\text{CCSDT})=E_{\text{CCSDT}}-E_{\text{CCSD}}$ ] display strongly monotonic convergence to the CBS limit, the same is not true of the CCSDTQ differential correlation energy [ $E_{\text{corr}}^{\text{diff}}(\text{CCSDTQ})=E_{\text{CCSDTQ}}-E_{\text{CCSDT}}$ ]. We find that the cc-pVDZ basis set yields a larger value of  $E_{\text{corr}}^{\text{diff}}(\text{CCSDTQ})$  than the cc-pVTZ basis set, 150  $\mu E_h$  (VDZ) versus 66  $\mu E_h$  (VTZ). The cc-pVQZ value lies between the cc-pVDZ and cc-pVTZ values, with  $E_{\text{corr}}^{\text{diff}}(\text{CCSDTQ})=96 \mu E_h$ . Although we were unable to perform a FCI/cc-pVQZ calculation on the Ne atom, the cc-pVDZ and cc-pVTZ trends follow the CCSDTQ pattern, 189  $\mu E_h$  (VDZ) versus 80  $\mu E_h$  (VTZ).

In addition to the CCSDT and CCSDTQ methods, our

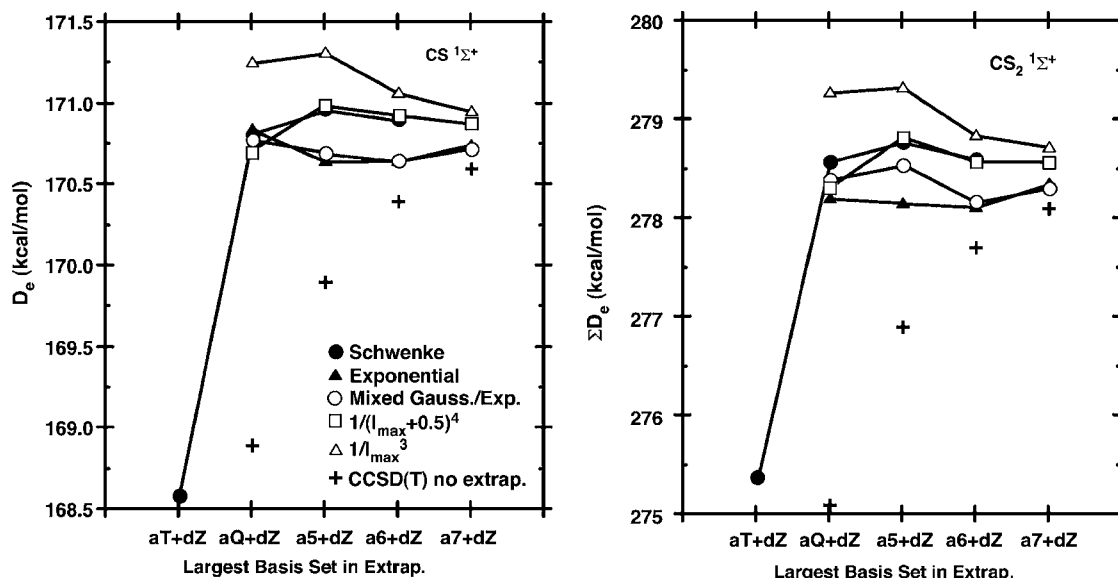


FIG. 5. Convergence of the CS and  $CS_2$  extrapolated complete basis sets and raw CCSD(T)(FC) binding energies (kcal/mol). The aug-cc-pVnZ basis set was used for C and the aug-cc-pV(n+d)Z basis set was used for S.

search for a practical higher-order technique capable of reliably improving upon the CCSD(T) level of theory has led us to examine the CCSD(TQ) method.<sup>77,78</sup> As with our CCSDT studies, a comparison was made with FCI, estimated FCI, and experimental results.<sup>71</sup> Although computationally faster than CCSDTQ, this method was found to often overestimate the effect of higher-order correlation and was, therefore, not recommended for general use. Similar conclusions were reached for the Brueckner doubles with perturbative triples and quadruples [BD(TQ)] method.<sup>77</sup>

Table III shows the CCSDTQ(FC)/cc-pVDZ and CCSDTQ(FC)/cc-pVTZ corrections to selected CCSD(T)  $N_2$  spectroscopic constants. Calculations at the CCSDTQ(FC)/cc-pVQZ level of theory were prohibitively expensive. The cc-pVTZ corrections for  $r_e$  and  $\omega_e$  were first reported by Ruden *et al.*<sup>79</sup> The relative importance of the CCSDTQ corrections is somewhat subjective. For example, relative to the residual one-particle error at the CCSD(T)/VTZ level of theory ( $r_e$ ,  $-0.0049$  Å;  $\omega_e$ , 16.4; and  $D_e$ , 10.4 kcal/mol), the CCSDTQ corrections are small and of opposite sign for the first two properties ( $r_e$ , 0.0007 Å and  $\omega_e$ ,  $-9.1$ ). For  $D_e$  the CCSDTQ correction (0.4 kcal/mol) is of the same sign as the basis set expansion correction. These results are in-line with the general trend that connects increased correlation recovery with an increasing bond length and decreasing harmonic frequency. Although the sizes of the CCSDTQ corrections are small for  $N_2$ , for other diatomics they have been shown to be larger. Ruden *et al.*<sup>79</sup> reported a CCSDTQ(FC)/cc-pVTZ bond length correction of 0.0034 Å in  $F_2$ , nearly five times larger than in  $N_2$ . Among the three small molecules examined in this study, the largest VTZ correction to  $D_e$  is  $\sim 0.3$  kcal/mol. For  $N_2$ , the CCSDTQ/cc-pVDZ corrections are larger than the core/valence and relativistic corrections for  $r_e$ ,  $\omega_e x_e$ , and  $D_e$ . At the VTZ level, only the correction  $\omega_e$  is larger. The present cc-pVTZ correction for  $D_e$  (0.4 kcal/mol) is larger than the 0.13 kcal/mol (0.54 kJ/mol) correction reported by Ruden *et al.*<sup>79</sup> presum-

ably due to differences in bond lengths. We used the optimal CCSD(T) and CCSDTQ bond lengths, whereas they used the experimental bond length.  $N_2$  is a good illustration of the shortcomings of the CCSDT method for estimating the higher-order correlation corrections to CCSD(T). Contrary to the small CCSDTQ increase in  $D_e$ , CCSDT produces a  $-0.8$  kcal/mol decrease.

The FCI corrections in Table III account for the small, residual  $n$ -particle error surviving at the CCSDTQ level of theory. Explicit FCI/cc-pVTZ calculations are extremely time consuming, even in the case of a homonuclear diatomic. To the best of our knowledge, the only available FCI/cc-pVTZ energy for  $N_2$  is the one reported by Ruden *et al.*<sup>59</sup> Goodson has proposed the use of an empirically motivated, coupled cluster continued fraction approximation [CCSD(T)-cf] as a way to approximate the FCI energy.<sup>80</sup> His approach requires only HF, CCSD, and CCSD(T) energies, and makes use of a simple extrapolation formula. The accuracy of this approach depends upon the nature of the chemical system being studied. Systems characterized by a monotonic convergence of perturbation-theory expansions were labeled “class A” and all other systems were labeled “class B.”<sup>80,81</sup> In a test of its effectiveness, we applied the CCSD(T)-cf method to almost 40 molecules. For class A chemical systems, the CCSD(T)-cf method was capable of predicting total energies that were always closer to the FCI energy than CCSD(T), but the improvement was marginal in many cases. For class B systems the performance was noticeably worse. In 6 out of 19 cases the difference between the CCSD(T)-cf and FCI energies was larger than the corresponding difference for CCSD(T).<sup>14</sup>

In the present work we revisited the continued fraction approximation to FCI, replacing the HF, CCSD, and CCSD(T) sequence with CCSD, CCSDT, and CCSDTQ. In the case of  $N_2$ , where exact FCI results are available, the results were in excellent agreement. At the VDZ level, the estimate was within  $8 \mu E_h$  of the exact FCI energy and it

accurately reproduced the FCI spectroscopic constants. At the VTZ level, the cf estimate matched the single available FCI energy exactly,  $-109.375\ 40$  (cf) versus  $-109.3754 \pm 0.0001$  (FCI). Consequently, cc-pVTZ cf-based, estimated FCI corrections were adopted for  $r_e$  and  $\omega_e$ .

The “composite”  $N_2$  entries in Table III, created by adding the CV, relativistic, CCSDTQ, and FCI corrections to the CCSD(T)(FC) results, represent our best estimates for each basis set level. The set of final CBS composite properties is seen to be in excellent agreement with the experimental values reported by Huber and Herzberg.<sup>4</sup> For the dissociation energy, the *NIST/JANAF Tables*<sup>5</sup> adopted a value ( $228.4 \pm 0.4$  kcal/mol) that is essentially the same as the value from Huber and Herzberg. The two final theoretical dissociation energies,  $228.4 \pm 0.2$  and  $228.7 \pm 0.2$  kcal/mol, based on R/UCCSD(T) and RCCSD(T) atomic treatments, sandwich the experimental value and fall within the NIST/JANAF experimental uncertainty. Thus, for  $N_2$  the use of RCCSD(T) atoms increases the CBS dissociation energy by 0.3 kcal/mol, relative to R/UCCSD(T) or UCCSD(T). Approximate error bars are based on the uncertainty in the CBS energies. We believe that the error arising from the uncertainty in the  $n$ -particle expansion is  $<0.1$  kcal/mol. The present best estimate for  $D_e$  represents a slight improvement over previous CBS coupled cluster estimates,<sup>7,8</sup> but is nearly identical to the value reported by Ruden *et al.*<sup>59</sup>

CCSDTQ and continued fraction approximate FCI corrections for CS and  $CS_2$  are presented in Tables IV and V. CCSDTQ/cc-pVTZ calculations for  $CS_2$  proved intractable, as a single-point, serial CCSDTQ/cc-pVDZ energy evaluation for  $CS_2$  required 25 days on a 1.6 GHz Athlon processor running UTCHEM.<sup>82</sup> We estimated the CCSDTQ/cc-pVTZ correction to the  $CS_2$  atomization energy, shown in Table V, by scaling the  $CS_2$  CCSDTQ/cc-pVDZ correction (0.98 kcal/mol) by the ratio of the CS VTZ/VDZ corrections. The size of the CS and  $N_2$  corrections might be expected to be similar in light of the fact that the two systems possess the same number of valence electrons. Table IV reveals this to be the case. As was true for  $N_2$ , the CCSDT correction to the CCSD(T) dissociation energy of CS is opposite in sign to the CCSDTQ correction and much larger in magnitude. Thus, the present results further reinforce the conclusions drawn in our earlier studies regarding the usefulness of CCSDT as a method for estimating higher-order [i.e., beyond CCSD(T)] correlation effects. For many chemical systems, CCSD(T) does a better job of mimicking CCSDTQ and FCI properties than does CCSDT, even though its treatment of triple excitations is more approximate. Continued fraction FCI estimates were used in lieu of explicit FCI results, since we were unable to perform a FCI/cc-pVDZ calculation on CS in 16 GB of memory.

The final composite spectroscopic constants for CS are in very good agreement with the experiment. The NIST/JANAF dissociation energy carries a large uncertainty of  $\pm 6$  kcal/mol. The final composite  $CS_2$  bond distance, anharmonic frequencies, and atomization energy are also in good agreement with the experiment, although our best theoretical value for  $\Sigma D_e$  [ $280.5 \pm 0.3$  kcal/mol with R/UCCSD(T) atoms] lies slightly outside the tight experimental error bars.

## V. SMALLER CONTRIBUTIONS

In previous work we have investigated the contribution of the diagonal Born-Oppenheimer correction (DBOC) to the dissociation energy of OH, the atomization energy of  $H_2O$ , and the H–OH bond dissociation energy,<sup>12</sup> as well as to the  $H_3$  potential-energy surface.<sup>83</sup> This correction, which accounts for the finite mass of the nuclei present in a molecule, is generally much smaller than the uncertainties associated with the factors already discussed.<sup>84</sup> For  $D_e(OH)$  the correction at the complete active space self-consistent field (CASSCF)/aug-cc-pVTZ level of theory was a mere  $-0.01$  kcal/mol. For  $\Sigma D_e(H_2O)$  it was 0.10 kcal/mol. Tajti *et al.*<sup>76</sup> have recently reported a RHF/aug-cc-pVTZ DBOC for  $N_2$  of 0.02 kcal/mol. We calculate that at the CISD/aug-cc-pVTZ level of theory the corrections for  $N_2$  are quite small:  $+0.01$  kcal/mol,  $+1.0 \times 10^{-5}$  Å, and  $+0.1$   $cm^{-1}$  for  $D_e$ ,  $r_e$ , and  $\omega_e$ , respectively. The corresponding CISD corrections for CS are  $-0.03$  kcal/mol,  $+1.0 \times 10^{-5}$  Å, and  $0.0$   $cm^{-1}$ . Although the DBOC can be significant for very high accuracy studies of molecules containing one or more hydrogen atoms, e.g.,  $H_2O$ ,  $NH_3$ , and  $H_2O_2$ , Tajti *et al.* found that the magnitude of the corrections to  $D_e$  and  $\Sigma D_e$  never exceeded 0.15 kcal/mol. Valeev and Sherrill have discussed the impact of using correlated wave functions to evaluate the DBOC.<sup>85</sup>

## VI. CONCLUSION

Coupled cluster calculations with very large correlation consistent basis sets graphically illustrate the slow convergence of the total energy with respect to completeness in the one-particle expansion, as implemented in traditional Gaussian function-based approaches. The largest basis set developed for this study (cc-pV10Z) includes 506 functions, ranging up to  $\ell_{\max}=10$  ( $n$  functions). Despite the sophistication of the basis set, the raw CCSD(T)/cc-pV10Z energy of the Ne ( $^1S$ ) atom remains  $\sim 0.001 E_h$  (0.6 kcal/mol) above the estimated CBS limit and energies obtained from explicitly correlated R12 methods. Fortunately, energy differences and the observables they depend upon converge much more rapidly, so that the dissociation energy of  $N_2$  is converged to  $\sim 0.0008 E_h$  at the aug-cc-pV7Z basis set level. A number of simple functional forms can be used to exploit the systematic convergence properties of the correlation consistent basis set families, effectively improving the computed properties by three or more step-ups in basis set quality. Identifying a single functional form as superior to all others is complicated by the relatively small spread in values and the variability in the extrapolated results as a function of the underlying basis sets and the nature of the chemical system being studied.

Due to fortuitous cancellation of error, CCSD(T) is often found to produce results in better agreement with FCI than the less approximate CCSDT. After careful accounting for the one-particle,  $n$ -particle, core/valence, and relativistic errors, the dissociation energy of  $N_2$  is estimated to be  $228.4 \pm 0.2$  kcal/mol [ $228.7 \pm 0.2$  kcal/mol with RCCSD(T) atoms], in good agreement with the available experimental measurements. Through the efforts of many individuals over the course of the past 50 years, the uncertainty in the theo-

retical value has gradually dropped four orders of magnitude, from over 200 to  $\sim 0.2$  kcal/mol. Results of similar quality were found for CS and CS<sub>2</sub>.

## ACKNOWLEDGMENTS

Dr. Ed Valeev and Dr. C. David Sherrill are thanked for assistance with PSI3. Dr. Karol Kowalski and Dr. Theresa Windus are thanked for help with selected CCSDTQ calculations performed with NWCHEM. Dr. David Dixon is thanked for early access to his results on C<sub>8</sub>H<sub>18</sub>. Dr. Wim Klopper is thanked for an informative discussion about the CCSD(T)-R12 method and early access to his latest large basis set results.

- <sup>1</sup>C. W. Sherr, *J. Chem. Phys.* **23**, 569 (1955).
- <sup>2</sup>A. C. Hurley, *Proc. Phys. Soc., London, Sect. A* **69**, 767 (1956).
- <sup>3</sup>E. Thulstrup and A. Andersen, *J. Phys. B* **8**, 965 (1975).
- <sup>4</sup>K. P. Huber and G. Herzberg, *Constants of Diatomic Molecules*, Molecular Spectra and Molecular Structure Vol. IV (Van Nostrand Reinhold, New York, 1979).
- <sup>5</sup>M. W. Chase, Jr., *J. Phys. Chem. Ref. Data Monogr.* **9**, 1 (1998).
- <sup>6</sup>L. A. Curtiss, K. Raghavachari, G. W. Trucks, and J. A. Pople, *J. Chem. Phys.* **94**, 7221 (1991).
- <sup>7</sup>D. Feller and K. A. Peterson, *J. Chem. Phys.* **108**, 154 (1998).
- <sup>8</sup>D. Feller and K. A. Peterson, *J. Chem. Phys.* **110**, 8384 (1999).
- <sup>9</sup>D. Feller, *J. Comput. Chem.* **17**, 1571 (1996).
- <sup>10</sup>D. Feller, D. A. Dixon, and K. A. Peterson, *J. Phys. Chem.* **102**, 7053 (1998); D. A. Dixon and D. Feller, *J. Phys. Chem. A* **102**, 8209 (1998); D. A. Dixon, D. Feller, and G. Sandrone, *ibid.* **103**, 4744 (1999); D. Feller and D. A. Dixon, *ibid.* **103**, 6413 (1999); **104**, 3048 (2000); D. Feller and J. A. Franz, *ibid.* **104**, 9017 (2000); B. Ruscic, D. Feller, D. A. Dixon, K. A. Peterson, L. B. Harding, R. L. Asher, and A. F. Wagner, *ibid.* **105**, 1 (2001); D. A. Dixon, W. A. de Jong, K. A. Peterson, and J. S. Francisco, *ibid.* **106**, 4725 (2002); D. Feller, D. A. Dixon, and J. S. Francisco, *ibid.* **107**, 1604 (2003); D. Feller and D. A. Dixon, *ibid.* **107**, 10419 (2003); N. B. Balabanov and K. A. Peterson, *ibid.* **107**, 7465 (2003); D. A. Dixon, D. Feller, K. O. Christe, W. W. Wilson, A. Vij, V. Vij, H. D. B. Jenkins, R. Olson, and M. S. Gordon, *J. Am. Chem. Soc.* **126**, 834 (2004).
- <sup>11</sup>D. Feller and D. A. Dixon, *J. Chem. Phys.* **115**, 3484 (2001).
- <sup>12</sup>B. Ruscic, A. F. Wagner, L. B. Harding *et al.*, *J. Phys. Chem. A* **106**, 2727 (2002).
- <sup>13</sup>D. Feller, K. A. Peterson, W. A. de Jong, and D. A. Dixon, *J. Chem. Phys.* **118**, 3510 (2003).
- <sup>14</sup>D. Feller and D. A. Dixon, *J. Phys. Chem. A* **107**, 9641 (2003).
- <sup>15</sup>L. Pollack, T. L. Windus, W. A. de Jong, and D. A. Dixon, *J. Phys. Chem. A* **109**, 6934 (2005).
- <sup>16</sup>N. B. Balabanov and K. A. Peterson, *J. Chem. Phys.* **123**, 064107 (2005).
- <sup>17</sup>R. S. Grev, C. L. Janssen, and H. F. Schaefer III, *J. Chem. Phys.* **95**, 5128 (1991).
- <sup>18</sup>S. Huzinaga, *Comput. Phys. Rep.* **2**, 279 (1985); E. R. Davidson and D. Feller, *Chem. Rev. (Washington, D.C.)* **86**, 681 (1986); D. Feller and E. R. Davidson, *Basis Sets for Ab Initio Molecular Orbital Calculations and Intermolecular Interactions* (VCH, New York, 1990).
- <sup>19</sup>J. Almlöf and P. R. Taylor, *J. Chem. Phys.* **86**, 4070 (1987).
- <sup>20</sup>K. Jankowski, R. Becherer, P. Scharf, H. Schiffer, and R. Alhrichs, *J. Chem. Phys.* **82**, 1413 (1985).
- <sup>21</sup>R. C. Raffanetti, *J. Chem. Phys.* **58**, 4452 (1973).
- <sup>22</sup>W. J. Hehre, R. Ditchfield, and J. A. Pople, *J. Chem. Phys.* **56**, 2257 (1972); P. C. Hariharan and J. A. Pople, *Theor. Chim. Acta* **28**, 213 (1973).
- <sup>23</sup>T. H. Dunning, Jr., *J. Chem. Phys.* **90**, 1007 (1989).
- <sup>24</sup>R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, *J. Chem. Phys.* **96**, 6796 (1992); D. E. Woon and T. H. Dunning, Jr., *ibid.* **100**, 2975 (1994).
- <sup>25</sup>D. E. Woon and T. H. Dunning, Jr., *J. Chem. Phys.* **103**, 4572 (1995); K. A. Peterson and T. H. Dunning, Jr., *ibid.* **117**, 10548 (2002); J. Koput and K. A. Peterson, *J. Phys. Chem. A* **106**, 9595 (2002).
- <sup>26</sup>W. A. de Jong, R. J. Harrison, and D. A. Dixon, *J. Chem. Phys.* **114**, 48 (2001).
- <sup>27</sup>D. Woon and T. H. Dunning, Jr., *J. Chem. Phys.* **98**, 1358 (1993); A. K. Wilson, D. E. Woon, K. A. Peterson, and T. H. Dunning, Jr., *ibid.* **110**, 7667 (1999).
- <sup>28</sup>K. A. Peterson, *J. Chem. Phys.* **119**, 11099 (2003); K. A. Peterson, D. Figgen, E. Goll, H. Stoll, and M. Dolg, *ibid.* **119**, 11113 (2003); K. A. Peterson and C. Puzzarini, *Theor. Chem. Acc.* **114**, 283 (2005).
- <sup>29</sup>A. K. Wilson, T. van Mourik, and T. H. Dunning, Jr., *J. Mol. Struct.: THEOCHEM* **388**, 339 (1996); T. van Mourik and T. H. Dunning, Jr., *Int. J. Quantum Chem.* **76**, 205 (1999).
- <sup>30</sup>D. Feller, *J. Chem. Phys.* **96**, 6104 (1992); S. S. Xantheas and T. H. Dunning, Jr., *J. Phys. Chem.* **97**, 18 (1993); D. Feller, *J. Chem. Phys.* **98**, 7059 (1993).
- <sup>31</sup>C. Moller and M. S. Plesset, *Phys. Rev.* **46**, 618 (1934).
- <sup>32</sup>D. E. Woon and T. H. Dunning, Jr., *J. Chem. Phys.* **99**, 1914 (1993).
- <sup>33</sup>D. G. Truhlar, *Chem. Phys. Lett.* **294**, 45 (1998); P. L. Fast, M. L. Sánchez, and D. G. Truhlar, *J. Chem. Phys.* **111**, 2921 (1999).
- <sup>34</sup>G. A. Petersson, A. Bennett, T. G. Tensfeldt, M. A. Al-Laham, W. A. Shirley, and J. Mantzaris, *J. Chem. Phys.* **89**, 2193 (1988); G. A. Petersson and M. A. Al-Laham, *ibid.* **94**, 6081 (1989); G. A. Petersson, in *Computational Thermochemistry: Predictions and Estimation of Molecular Thermodynamics*, ACS Symposium Series Vol. 667 edited by K. K. Irikura and D. J. Frurip (ACS, Washington, DC, 1998).
- <sup>35</sup>K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, *Chem. Phys. Lett.* **157**, 479 (1989).
- <sup>36</sup>R. D. Amos, A. Bernhardsson, A. Berning *et al.*, MOLPRO2002.6, a package of *ab initio* programs (Universität Stuttgart, Stuttgart, Germany/University of Birmingham, Birmingham, United Kingdom, 2002).
- <sup>37</sup>T. D. Crawford, C. D. Sherrill, E. F. Valeev *et al.*, PSI, version 3.2 (2005).
- <sup>38</sup>C. Angeli, K. L. Bak, V. Bakken *et al.*, DALTON2.0 (2005).
- <sup>39</sup>M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, GAUSSIAN 03, revision C.02, Gaussian, Inc., Wallingford, CT, 2004.
- <sup>40</sup>E. Aprà, T. L. Windus, T. P. Straatsma *et al.*, NWCHEM4.7, a computational chemistry package for parallel computers (Pacific Northwest National Laboratory, Richland, Washington, 2005).
- <sup>41</sup>C. Schwartz, *Phys. Rev.* **126**, 1015 (1962); in *Methods in Computational Physics*, edited by B. J. Alder, S. Fernbach, and M. Rotenberg (Academic, New York, 1963), Vol. 2, p. 262.
- <sup>42</sup>W. Kutzelnigg and J. D. Morgan, *J. Chem. Phys.* **96**, 4484 (1992).
- <sup>43</sup>J. M. L. Martin, *Chem. Phys. Lett.* **259**, 669 (1996); T. Helgaker, W. Klopper, H. Koch, and J. Nago, *J. Chem. Phys.* **106**, 9639 (1997).
- <sup>44</sup>J. M. L. Martin and T. J. Lee, *Chem. Phys. Lett.* **258**, 136 (1996).
- <sup>45</sup>A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen, and A. K. Wilson, *Chem. Phys. Lett.* **286**, 243 (1998).
- <sup>46</sup>K. A. Peterson, D. E. Woon, and T. H. Dunning, Jr., *J. Chem. Phys.* **100**, 7410 (1994).
- <sup>47</sup>A. K. Wilson and T. H. Dunning, Jr., *J. Chem. Phys.* **106**, 8718 (1997).
- <sup>48</sup>W. Klopper, *Mol. Phys.* **99**, 481 (2001).
- <sup>49</sup>E. F. Valeev, W. D. Allen, R. Hernandez, C. D. Scherrill, and H. F. Schaefer III, *J. Chem. Phys.* **118**, 8594 (2003).
- <sup>50</sup>W. Kutzelnigg, *Theor. Chim. Acta* **68**, 445 (1985); W. Klopper and W. Kutzelnigg, *Chem. Phys. Lett.* **134**, 17 (1987); *J. Phys. Chem.* **94**, 5625 (1990); W. Kutzelnigg and W. Klopper, *J. Chem. Phys.* **94**, 1985 (1991).
- <sup>51</sup>W. Klopper and J. Gauss (private communication).
- <sup>52</sup>D. Feller and J. A. Sordo, *J. Chem. Phys.* **113**, 485 (2000).
- <sup>53</sup>T. Kato, *Pure Appl. Math.* **10**, 151 (1957).
- <sup>54</sup>J. R. Flores, *Phys. Rev. A* **46**, 6063 (1992).
- <sup>55</sup>K. A. Peterson, A. K. Wilson, D. E. Woon, and T. H. Dunning, Jr., *Theor. Chem. Acc.* **97**, 251 (1997).
- <sup>56</sup>W. Klopper and T. Helgaker (private communication).
- <sup>57</sup>T. H. Dunning, Jr., K. A. Peterson, and A. K. Wilson, *J. Chem. Phys.* **114**, 9244 (2001).
- <sup>58</sup>C. Hampel, K. A. Peterson, and H. J. Werner, *Chem. Phys. Lett.* **190**, 1 (1990); M. J. O. Deegan and P. J. Knowles, *ibid.* **227**, 321 (1994); P. J. Knowles, C. Hampel, and H. J. Werner, *J. Chem. Phys.* **99**, 5219 (1993).
- <sup>59</sup>T. A. Ruden, T. Helgaker, P. Jørgensen, and J. Olsen, *Chem. Phys. Lett.* **371**, 62 (2003).
- <sup>60</sup>K. A. Peterson and T. H. Dunning, Jr., *J. Phys. Chem. A* **101**, 6280 (1997).
- <sup>61</sup>J. L. Dunham, *Phys. Rev.* **41**, 713 (1932).
- <sup>62</sup>M. Douglas and N. M. Kroll, *Ann. Phys.* **82**, 89 (1974); G. Jansen and B. A. Hess, *Phys. Rev. A* **39**, 6016 (1989).
- <sup>63</sup>C. E. Moore, *Atomic Energy Levels*, Natl. Bur. Stand. (U.S.) Circ. No. 37 (U.S. GPO, Washington, D.C., 1971).
- <sup>64</sup>A. G. Maki and R. L. Sams, *J. Mol. Spectrosc.* **52**, 233 (1974).
- <sup>65</sup>T. Shimanouchi, *Tables of Molecular Vibrational Frequencies Consoli-*

- dated* (National Bureau of Standards, Washington, D.C., 1972).
- <sup>66</sup> J. Senekowitsch, Ph.D. thesis, Universität Frankfurt, 1988.
- <sup>67</sup> A. D. Becke, *J. Chem. Phys.* **98**, 5648 (1993); C. Lee, W. Yang, and R. G. Parr, *Phys. Rev. B* **37**, 785 (1988).
- <sup>68</sup> C. W. Bauschlicher, Jr. and H. Partridge, *Chem. Phys. Lett.* **240**, 533 (1995).
- <sup>69</sup> K. S. Raymond and R. A. Wheeler, *J. Comput. Chem.* **20**, 207 (1999); A. D. Boese and J. M. L. Martin, *J. Chem. Phys.* **119**, 3005 (2003); N. X. Wang and A. K. Wilson, *ibid.* **121**, 7632 (2004); N. X. Wang and A. K. Wilson, *Mol. Phys.* **103**, 345 (2005).
- <sup>70</sup> R. J. Bartlett and J. F. Stanton, *Reviews in Computational Chemistry* (VCH, New York, 1994), Vol. 5.
- <sup>71</sup> D. Feller, *J. Chem. Phys.* **111**, 4373 (1999).
- <sup>72</sup> D. Feller and J. A. Sordo, *J. Chem. Phys.* **112**, 5604 (2000).
- <sup>73</sup> P. A. Denis, *Chem. Phys. Lett.* **382**, 65 (2003).
- <sup>74</sup> T. Helgaker, W. Klopper, A. Halkier, K. L. Bak, P. Jørgensen, and J. Olsen (Kluwer, Dordrecht, 2001), Vol. 22, p. 1.
- <sup>75</sup> A. D. Boese, M. Oren, O. Atasoylu, J. M. L. Martin, M. Kallay, and J. Gauss, *J. Chem. Phys.* **120**, 4129 (2004).
- <sup>76</sup> A. Tajti, P. G. Szalay, A. G. Császár, M. Kallay, J. Gauss, E. F. Valeev, B. A. Flowers, J. Vázquez, and J. F. Stanton, *J. Chem. Phys.* **121**, 11599 (2004).
- <sup>77</sup> K. Raghavachari, J. A. Pople, E. S. Replogle, and M. Head-Gordon, *J. Phys. Chem.* **94**, 5579 (1990).
- <sup>78</sup> R. J. Bartlett, J. D. Watts, S. A. Kucharski, and J. Noga, *Chem. Phys. Lett.* **165**, 513 (1990).
- <sup>79</sup> T. A. Ruden, T. Helgaker, P. Jørgensen, and J. Olsen, *J. Chem. Phys.* **121**, 5874 (2004).
- <sup>80</sup> D. Z. Goodson, *J. Chem. Phys.* **116**, 6948 (2002).
- <sup>81</sup> D. Cremer and Z. He, *J. Phys. Chem.* **100**, 6173 (1996).
- <sup>82</sup> T. Yanai, M. Kamiya, Y. Kawashima *et al.*, *UTCHEM 2004 B* (2004).
- <sup>83</sup> S. L. Mielke, D. W. Schwenke, and K. A. Peterson, *J. Chem. Phys.* **122**, 224313 (2005).
- <sup>84</sup> N. C. Handy, Y. Yamaguchi, and H. F. Schaefer III, *J. Chem. Phys.* **84**, 4481 (1986).
- <sup>85</sup> E. F. Valeev and C. D. Scherrill, *J. Chem. Phys.* **118**, 3921 (2003).
- <sup>86</sup> D. W. Schwenke, *J. Chem. Phys.* **122**, 014107 (2005).
- <sup>87</sup> M. L. Leininger, W. D. Allen, H. F. Schaefer III, and C. D. Sherrill, *J. Chem. Phys.* **112**, 9213 (2000).