

## Probing the limits of accuracy in electronic structure calculations: Is theory capable of results uniformly better than “chemical accuracy”?

David Feller<sup>a)</sup> and Kirk A. Peterson<sup>b)</sup>*Department of Chemistry, Washington State University, Pullman, Washington 99164-4630*

(Received 6 October 2006; accepted 9 January 2007; published online 16 March 2007)

Current limitations in electronic structure methods are discussed from the perspective of their potential to contribute to inherent uncertainties in predictions of molecular properties, with an emphasis on atomization energies (or heats of formation). The practical difficulties arising from attempts to achieve high accuracy are illustrated via two case studies: the carbon dimer ( $C_2$ ) and the hydroperoxyl radical ( $HO_2$ ). While the  $HO_2$  wave function is dominated by a single configuration, the carbon dimer involves considerable multiconfigurational character. In addition to these two molecules, statistical results will be presented for a much larger sample of molecules drawn from the Computational Results Database. The goal of this analysis will be to determine if a combination of coupled cluster theory with large 1-particle basis sets and careful incorporation of several computationally expensive smaller corrections can yield uniform agreement with experiment to better than “chemical accuracy” ( $\pm 1$  kcal/mol). In the case of  $HO_2$ , the best current theoretical estimate of the zero-point-inclusive, spin-orbit corrected atomization energy ( $\Sigma D_0 = 166.0 \pm 0.3$  kcal/mol) and the most recent Active Thermochemical Table (ATcT) value ( $165.97 \pm 0.06$  kcal/mol) are in excellent agreement. For  $C_2$  the agreement is only slightly poorer, with theory ( $D_0 = 143.7 \pm 0.3$  kcal/mol) almost encompassing the most recent ATcT value ( $144.03 \pm 0.13$  kcal/mol). For a larger collection of 68 molecules, a mean absolute deviation of 0.3 kcal/mol was found. The same high level of theory that produces good agreement for atomization energies also appears capable of predicting bond lengths to an accuracy of  $\pm 0.001$  Å. © 2007 American Institute of Physics. [DOI: 10.1063/1.2464112]

### I. INTRODUCTION

Coupled cluster theory has proven to be one of the most widely applicable and consistently reliable tools for the accurate prediction of spectroscopic and thermochemical properties. Examples of such properties include molecular structures, harmonic and anharmonic vibrational frequencies, excitation energies, and heats of formation (atomization energies), but may also include a wide assortment of other, less commonly selected properties. In most cases, accurate reproduction of high quality experimental data and equally accurate predictions requires the inclusion of electronic excitation levels up to triples. Extensive benchmarking has demonstrated that a single-step, quasiperturbative incorporation of “connected” triple excitations, corresponding to the CCSD(T) method, often yields values in fortuitously close agreement with those obtained from more expensive methods.

Among the various categories of properties that have been the object of theoretical interest over the past several decades, atomization energies have arguably proven to be the most challenging. Relatively low-level, frozen core (FC) coupled cluster calculations with a small augmented valence double zeta quality basis set exhibit a mean absolute deviation,  $\epsilon_{MAD}$ , with respect to experiment of almost

40 kcal/mol (259 comparisons). The analysis was performed with the Computational Results Database (CRDB), which contains approximately 90 000 experimental and theoretical entries covering nearly 300 molecules.<sup>1</sup> Errors of this magnitude render the method essentially useless for practical applications. By way of comparison, bond lengths predicted at the same level of theory exhibit an  $\epsilon_{MAD}$  of  $\sim 0.02$  Å (409 comparisons). For polyatomic molecules, the typical experimental uncertainty in equilibrium bond lengths is on the order of 0.01 Å. Thus, even this modest level of theory appears almost capable of experimental accuracy for bond lengths, in contrast to its very poor showing for atomization energies.

An often-quoted target in computational thermochemistry is the attainment of “chemical accuracy,” a term conventionally associated with errors of  $\pm 1$  kcal/mol (4.184 kJ/mol) or less, when measured with respect to high quality experimental results. This term is commonly found in studies of small molecules where high accuracy is a primary goal and the number of chemical compounds is necessarily limited due to the computationally intensive nature of the methods. The long-term objective, which is frequently unstated, is the desire to find methods capable of chemical accuracy in thermochemical properties for as wide an assortment of chemical systems as possible. This includes molecules that would be judged “large” by present standards, as well as those with special difficulties due to the multiconfigurational nature of their electronic states. The objective of the present work is to further probe the inherent limitations

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

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

of a composite coupled cluster theory approach developed over the past decade in collaboration with Dixon and co-workers.<sup>2-5</sup> This general approach has been applied to more than 400 chemical systems.<sup>2-6</sup> We wish to determine the extent to which our approach is capable of (1) estimating results close to the relativistic one-particle and  $n$ -particle limits and (2) meeting or exceeding the standard of chemical accuracy. Initially, our discussion will focus on an in-depth examination of two prototypical small molecules. They were chosen because of the availability of reliable experimental data and because one is Hartree-Fock dominant and the other is not. The discussion then shifts to a statistical analysis of a much larger collection of molecules. This analysis should prove useful in calibrating the current state of the art in electronic structure methods. As improvements in theoretical and experimental techniques continue to push the limits of accuracy, it is helpful to periodically reassess our understanding of the limitations of each approach.

We have recently examined the leading sources of error in coupled cluster theory calculations of total energies, bond lengths, harmonic frequencies, and atomization energies for first and second row atoms and small molecules.<sup>7</sup> The basic approach incorporates techniques similar to those used by other research groups.<sup>8-11</sup> However, while it shares some common components with other approaches, it differs from them in one key aspect. Rather than rely on a rigidly defined, static recipe or “model chemistry,” involving fixed method/basis set combinations, we prefer to retain the flexibility to adjust the approach to reflect the accuracy requirements, size and composition of the chemical system under study. For example, we might apply the diagonal Born-Oppenheimer correction (DBOC) in high accuracy studies of light systems such as the H–OH bond dissociation energy,<sup>4</sup> but ignore it in a study of BrCl.<sup>12</sup> Similarly, for heavy systems such as BrCl we might include a second order spin-orbit correction that would be completely negligible for lighter molecules. The common goal in all of our studies is to systematically reduce, to the extent possible, all significant sources of error contributing to the uncertainty in predicted properties. Adoption of a model chemistry approach almost inevitably leads to a proliferation of variants in order to meet the sometimes disparate needs of increased accuracy and larger systems. Consequently, the computational chemistry community has witnessed the introduction of G1,<sup>13</sup> G2,<sup>14</sup> and G3 (Ref. 15) model chemistries, as well as a large number of variants.<sup>16</sup> Other families of model chemistries include W1, W2,<sup>17</sup> W3,<sup>10</sup> and W4,<sup>18</sup> as well as HEAT,<sup>8</sup> HEAT345-(Q), HEAT345-Q, HEAT345-Q(P), and HEAT345-QP.<sup>9</sup> In addition, there are the multiple complete basis set (CBS) model chemistries of Petersson and co-workers.<sup>19</sup>

Most of the present discussion is focused on two prototypical small molecules,  $C_2$  ( $^1\Sigma_g^+$ ), the carbon dimer, and  $HO_2$  ( $^2A''$ ), the hydroperoxyl radical. These molecules were chosen as representative of a range of systems, spanning those that are dominated by the Hartree-Fock component of their wave functions, as is the case for  $HO_2$ , or possessing a significant multiconfigurational component, as is the case for  $C_2$ . Their wave functions are approximately given by

$$\Psi(HO_2) \sim 0.94(1a'^2 2a'^2 3a'^2 4a'^2 5a'^2 6a'^2 1a''^2 7a'^2 2a''^1), \quad (1)$$

$$\Psi(C_2) \sim 0.83(1\sigma_g^2 1\sigma_u^2 2\sigma_g^2 1\pi_{ux}^2 1\pi_{uy}^2 2\sigma_u^2) - 0.33(1\sigma_g^2 1\sigma_u^2 2\sigma_g^2 1\pi_{ux}^2 1\pi_{uy}^2 3\sigma_g^2), \quad (2)$$

respectively. In the coupled cluster realm, Lee *et al.* have proposed the  $T_1$  diagnostic as a predictor of the quality of CCSD results.<sup>20</sup> The  $T_1$  values for  $HO_2$  and  $C_2$  are 0.037 and 0.039, respectively. An alternate but related diagnostic, known as  $D_1$ , has been suggested by Janssen and Nielsen.<sup>21</sup> In their study of bond lengths and harmonic frequencies for 29 small molecules,  $C_2$  was reported to have the single largest  $D_1$  diagnostic, although interestingly not the largest errors. Nonetheless, we anticipate that  $C_2$  will present a challenge for CCSD(T), as it is nominally a single reference method. The  $D_1$  value of  $HO_2=0.126$  and of  $C_2=0.087$ .

Previous findings suggest that the fraction of correlation energy recovered by CCSD(T) is sufficiently large to partially compensate for its single reference nature.<sup>3</sup> In a growing number of comparisons with less approximate methods, CCSD(T) has been found to often outperform coupled cluster theory with iterative triples, CCSDT.<sup>22,23</sup> Consequently, it has been adopted as the method of choice for most high accuracy computational studies, especially for molecules in ground or low-lying excited states. Although CCSD(T) is strictly a ground state method, excited states of different symmetries can be treated and the equation of motion CCSD formalism makes other states accessible. Multiple reports in the literature claim theoretical uncertainties as small as 0.2–0.5 kcal/mol for enthalpies of reaction and heats of formation based on the CCSD(T) or CCSDT methods.<sup>11,24,25</sup> By incorporating higher level correlation corrections, uncertainties as small as 0.1 kcal/mol have been reported for molecules such as  $HO_2$ .<sup>26</sup> In the first HEAT paper, an  $\epsilon_{MAD}$  of 0.09 kcal/mol was reported for 11 comparisons with high accuracy experimental measurements.<sup>8</sup> The second HEAT paper, which included several refinements on the original approach, listed an  $\epsilon_{MAD}$  of 0.06 kcal/mol for 18 small molecules.<sup>9</sup> Schuurman *et al.*<sup>11</sup> reported uncertainties of  $\pm 0.2$  kcal/mol for heats of formation of NCO and its isomers by including limited CCSDT corrections. The W4 approach of Karton *et al.* also seeks sub-kJ/mol accuracy by including a higher level correlation correction.<sup>18</sup>

## II. APPROACH

Conceptually, we are attempting to approximate the results of all-electron, full configuration interaction (FCI) wave functions expanded in an effectively CBS, using a relativistic Hamiltonian for the system of interest. Despite the development of massively parallel computers consisting of thousands of processors, FCI remains so computationally intensive that it continues to be a niche method. Even for a simple diatomic, such as  $C_2$ , FCI calculations are limited to relatively small basis sets and the frozen core approximation. Consequently, out of necessity a much more efficient, composite approach has been developed, which replaces a single, intractable calculation with a series of smaller calculations

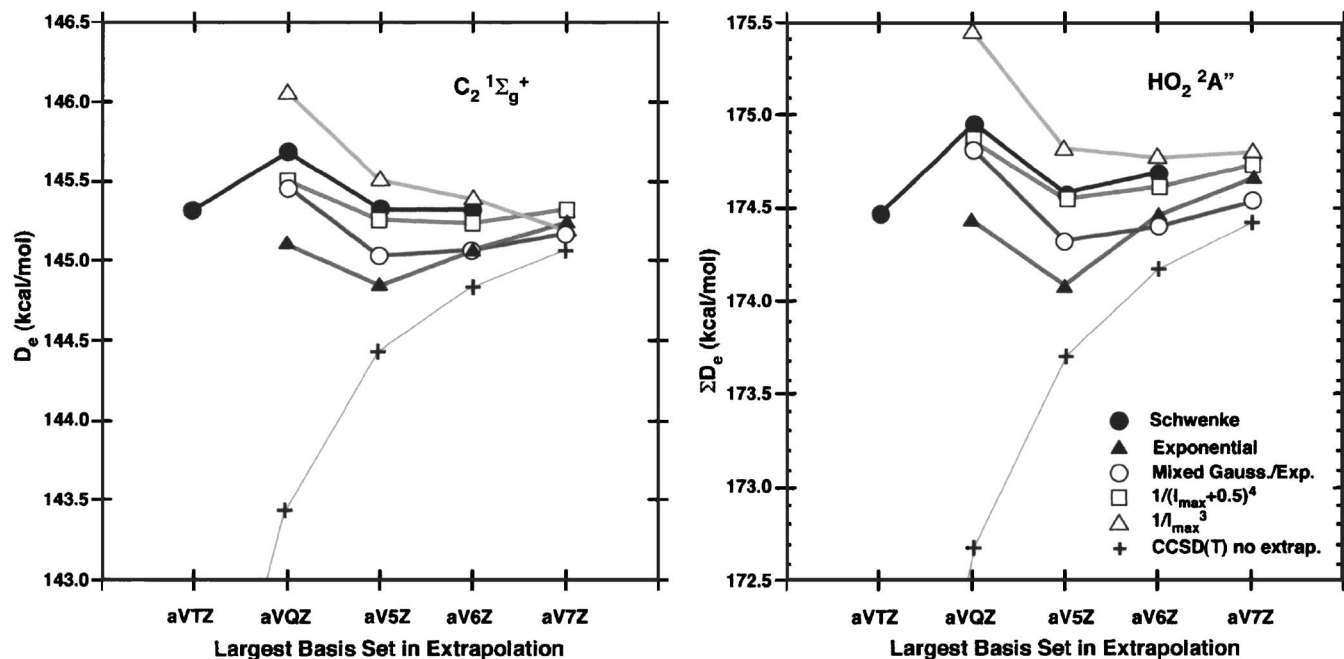


FIG. 1. Convergence of the  $C_2$  and  $HO_2$  extrapolated complete basis set and directly computed CCSD(T)(FC) binding energies (kcal/mol). The open shell  $HO_2$  system and the atomic asymptotes were treated with the R/UCCSD(T) method.

and extrapolations that are designed to minimize individual contributions to the overall error. The use of smaller component calculations permits a degree of fine-tuning in our approach to match the needs of the problem at hand that would be impossible with a single calculation.

The largest source of error in most *ab initio* calculations is the truncation of the one-particle basis set expansion, followed closely by the truncation of the  $n$ -particle expansion. In the present work, one-particle basis sets for  $C_2$  and  $HO_2$  were taken from the correlation consistent family of basis sets. CCSD(T)(FC) energies, geometries and harmonic frequencies were evaluated with the diffuse function augmented sets, aug-cc-pVnZ,  $n=D, T, Q, 5, 6,$  and  $7$ .<sup>27,28</sup> Spectroscopic properties for  $C_2$  were obtained from a sixth degree Dunham fit of the potential energy curve.<sup>29</sup> The largest basis set, denoted aug-cc-pV7Z, consists of a  $(19s, 13p, 7d, 6f, 5g, 4h, 3i, 2k) \rightarrow [9s, 8p, 7d, 6f, 5g, 4h, 3i, 2k]$  contraction. Current hardware limitations did not permit us to explicitly include multiple sets of 15-component  $k$  functions ( $\ell=7$ ) in the  $HO_2$  calculations. The contribution to the total energy of the missing  $k$  functions was estimated by exploiting the uniformity in the correlation energy convergence pattern as a function of  $\ell$ , as described elsewhere.<sup>7,23,28</sup> Tests of this approximation against  $k$  function-inclusive calculations on a series of first row atoms and diatomic molecules established the accuracy of this estimate as  $\pm 10^{-5} E_h$ . Errors in energy differences were  $\leq 0.05$  kcal/mol.

Despite the use of very extended basis sets, the residual one-particle basis set error remains significant in terms of the present discussion's focus on high accuracy. Furthermore, most of the molecules included in the statistical analysis discussion were too large to permit the use of a 7Z quality basis set. For some, aug-cc-pVQZ was the largest affordable basis set. Contributions to properties arising from the remaining basis set incompleteness was estimated by five CBS extrap-

lation formulas, including an exponential,<sup>30</sup> a mixed Gaussian/exponential,<sup>31</sup> two inverse powers of  $\ell_{max}$ , where  $\ell_{max}$  is the highest angular momentum present in the basis set,<sup>24,32</sup> and a recent expression due to Schwenke.<sup>33</sup> Each has been discussed elsewhere.<sup>7</sup> With the first four extrapolation formulas, we have chosen to extrapolate the total CCSD(T) energy, rather than perform separate extrapolations on the self-consistent field (SCF) and correlation components. Tests performed on a variety of molecules treated with basis sets as large as aug-cc-pV7Z showed minimal differences for these two approaches because of the dominance of the correlation component. The sole exception was at the aug-cc-pVQZ basis set level, where differences as large as 0.7 kcal/mol were occasionally observed. While differences of this magnitude are significant, neither approach clearly provided superior estimates of the apparent CBS limits.

As seen in Fig. 1, where the convergence of the CBS estimates for  $D_e(C_2)$  and  $\Sigma D_e(HO_2)$  is plotted alongside the directly computed CCSD(T) values, use of the extrapolations dramatically reduces the one-particle basis set error. An analysis of similar plots for other small molecules leads to several general observations. With basis sets of aug-cc-pVQZ or aug-cc-pV5Z quality, the exponential and mixed formulas tend to underestimate the apparent CBS limit, while the  $\ell_{max}$  formulas tend to overestimate it. Binding energies based on the Schwenke formulas often fall somewhere in between. For molecules consisting of first row elements, similar to the six small molecules in the Schwenke training set, the formula works very well, even when extrapolating from aug-cc-pVDZ and aug-cc-pVTZ basis sets. However, our experience with the Schwenke formulas and second row compounds indicated that further refinement of the approach might be needed when treating elements outside the training set.<sup>7</sup> No single formula clearly works "best" for every chemical system and combination of basis sets. Unless otherwise



noted, we adopt the average of the five formulas as our best approximation to the CBS limit and the spread among the estimates as a crude gauge of the uncertainty in the extrapolations.

Core/valence (CV) correlation effects for  $C_2$  and  $HO_2$  were described with the cc-pCV $nZ$ ,  $n=D-6$ , sets.<sup>34</sup> The CV correction to a particular property is defined as the difference between the frozen core and CV value of the property evaluated with the same CV basis set at the corresponding FC and CV optimized geometries. Because the CV correction usually converges more rapidly than the FC component, our best estimates will often combine a frozen core value from a basis set that is one zeta level higher than the CV calculation used. We will sometimes refer to CV results as “all electron,” as distinguished from “valence electron,” since in the present context there is no possibility of confusion with effective core potential or pseudopotential calculations. In most of our previous work we have assumed that the effects of correlating the inner shell electrons could be introduced as an additive correction to frozen core properties, as a practical expedient aimed at preventing the size of the one-particle expansion from growing excessively large. A discussion of the accuracy of this assumption follows in a subsequent section.

Scalar relativistic (SR) corrections were obtained from Douglas-Kroll-Hess (DKH) calculations<sup>35</sup> using the cc-pV $nZ$ \_DK frozen core basis sets.<sup>36</sup> The statistical analysis discussion involves molecules composed of elements up to  $Z=53$  (I). Basis sets for second row, main group elements (Al–Cl) were taken from the aug-cc-pV( $n+d$ ) $Z$ ,  $n=D-6$ , family, which contain an additional tight d function for the purpose of addressing a known SCF-level deficiency in the aug-cc-pV $nZ$  sets.<sup>37</sup> First row transition metals were described with the recently reported, all-electron aug-cc-pV $nZ$  and aug-cc-pV $nZ$ -DK basis sets.<sup>38</sup> Third and fourth row, main group elements were represented by the aug-cc-pV $nZ$ -PP basis sets.<sup>39</sup>

All CCSD(T) calculations in the present work were performed with MOLPRO 2002.6,<sup>40</sup> MOLPRO 2006.1,<sup>41</sup> and a development version of PSI3.<sup>42</sup> CCSDT (Ref. 43) and CCSDTQ (Ref. 44) calculations were performed with UTCHEM 2004 $\beta$  (Ref. 45) and the MRCC program<sup>46</sup> interfaced to MOLPRO. The statistical analysis was carried out with utilities associated with the CRDB. In the present work, open shell CCSD(T) energies were obtained from three different open shell methods. Most of our results are based on the R/UCCSD(T) method,<sup>47</sup> which begins with restricted open shell Hartree-Fock (ROHF) orbitals, but allows a small amount of spin contamination in the solution of the CCSD equations. 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>48</sup> The third choice, UCCSD(T), employs unrestricted Hartree-Fock (UHF) orbitals. As pointed out by Karton *et al.*,<sup>18</sup> only the open shell UHF-UCCSD(T)

method is uniquely defined. Differences in RCCSD(T) implementations among programs will result in small differences in open shell energies. Note that these differences are in addition to the R/UCCSD(T) versus RCCSD(T) differences within a single program, such as MOLPRO.

As already mentioned, in the present work the  $n$ -particle expansion will be addressed by coupled cluster methods. The majority of the correlation energy will be recovered at the CCSD(T)(FC) level of theory, but higher level corrections to CCSD(T) will require CCSD, CCSDT and CCSDTQ energies. A limited number of FCI calculations will be used for calibration purposes. Due to the time consuming nature of CCSDT and CCSDTQ calculations, the size of the basis sets that can be used with these methods is limited. The various coupled cluster methods formally scale as  $n^6$  (CCSD),  $n^7$  [CCSD(T)],  $n^8$  (CCSDT), and  $n^{10}$  (CCSDTQ), where  $n$  is the number of basis functions. More specifically, CCSD(T) scales as  $n^2N^4N_{it}$  (with a single  $n^3N^4$  step), where  $n$ ,  $N$ , and  $N_{it}$  are the number of occupied and unoccupied molecular orbitals and the number of CCSD iterations. Thus, for a molecule such as  $C_2$ , where  $n$  is fixed, we find that the method scales in practice as  $n^{4.5}$ .

Although CCSDTQ is currently the highest level of theory that can routinely be applied to polyatomic molecules, the residual  $n$ -particle error associated with this method continues to be an issue for high accuracy studies, in much the same way that the remaining one-particle error remained significant with the 6Z and 7Z sets. In earlier work, we adopted a continued fraction (cf) approximant to FCI as an  $n$ -particle analog to the one-particle CBS extrapolation formulas. First introduced by Goodson,<sup>49</sup> the cf approximant was originally formulated using Hartree-Fock, CCSD, and CCSD(T) energies. However, tests involving almost 40 molecules showed this sequence of energies yielded inconsistent results.<sup>5</sup> In particular, the effectiveness of the extrapolation was found to vary widely depending on the “class” of the molecules to which it was applied. For molecules categorized as “class B,” more than 30% of the time the differences between the CCSD(T)-cf and FCI energies were larger than the corresponding differences for CCSD(T) without extrapolation.

We have recently reinvestigated a modified version of the continued fraction approximant, in which the original sequence of energies was replaced with CCSD, CCSDT, and CCSDTQ energies.<sup>7</sup> This substitution was found to significantly improve the reliability of the extrapolation, albeit at the cost of substantially more expensive energies. Our initial evaluation involved only a limited number of comparisons with explicit FCI results. In the present work, further calibration of the method was undertaken by expanding the number of comparisons to include cc-pVDZ and cc-pVTZ basis set results for a series of first row diatomic molecules. In some cases, the CCSDTQ properties were already very close to the equivalent FCI values and the extrapolation had no effect. In other cases, where the CCSDTQ and FCI values differed by an appreciable amount, the cf method was consistently found to be effective in improving the agreement with FCI. Dissociation energies presented the greatest challenge. Across all comparisons, the cf approximant recovered an average of 80% of the FCI correction to CCSDTQ dissociation energies.

The worst performance was observed for  $C_2$  with the cc-pVDZ basis set, where the extrapolation recovered only 50% of the actual FCI correction of 0.4 kcal/mol. With a larger cc-pVTZ+diffuse(*s,p*) basis set the continued fraction recovered increased to 71%. FCI bond lengths and harmonic frequencies proved to be more easily reproduced by the cf estimated FCI method.

In general, our approach employs molecular energies evaluated at the corresponding optimized geometries, e.g., CCSD(T)/aug-cc-pVDZ energies at CCSD(T)/aug-cc-pVDZ optimized geometries. In situations where that becomes prohibitively expensive, the use of an alternative high quality structure, obtained from either theory or experiment, has been found to introduce errors on the order of 0.1–0.2 kcal/mol for molecules with four or fewer first row atoms.

### III. RESULTS AND DISCUSSION

#### A. $C_2$

Theoretical spectroscopic properties for  $C_2$  are organized in Table I according to the basis set level, from DZ to 7Z. The “composite” entries represent our best estimate of the scalar relativistic, all-electron FCI properties at each basis set level. While the convergence of the frozen core CCSD(T) bond length is seen to be comparatively rapid, with the aug-cc-pVTZ value falling within 0.006 Å of the CBS limit, the convergence of the harmonic frequency,  $\omega_e$ , is noticeably slower. The aug-cc-pVTZ value is almost 19  $cm^{-1}$  smaller than the CBS limit. At the CCSD(T)(FC)/CBS limit,  $r_e$  is  $\sim 0.001$  Å longer than the experimental value reported by Huber and Herzberg<sup>50</sup> and  $\omega_e$  is too large by  $\sim 5$   $cm^{-1}$ .

Our best estimate of the frozen core CCSD(T) binding energy is  $145.1 \pm 0.1$  kcal/mol when dissociating to R/UCCSD(T) atomic asymptotes and including atomic spin-orbit effects. This estimate is based on the average of four CBS extrapolation formulas (exponential, mixed, and two inverse powers of  $\ell_{max}$ ) applied to total energies up to aug-cc-pV7Z. The Schwenke formula was not used due to the lack of extrapolation coefficients for the aug-cc-pV7Z basis set. Estimates of the CBS limit in  $D_e$  obtained by averaging over multiple extrapolation formulas proved to be remarkably stable with respect to changes in the underlying basis sets. If the largest basis set is reduced to aug-cc-pV6Z, the average remains unchanged, but the error bars increase slightly,  $D_e(CBS) = 145.1 \pm 0.2$  kcal/mol (five formulas). Still smaller basis sets yielded averages of  $145.1 \pm 0.4$  (aug-cc-pV5Z) and  $145.6 \pm 0.5$  (aug-cc-pVQZ). At least for  $C_2$ , the adoption of the spread among the various formulas as a crude measure of the uncertainty in the CBS extrapolation procedure produces error bars that behaved as expected; i.e., they decrease as the size of the basis sets increase. Similar trends have been observed for other chemical systems, such as  $N_2$ , CS,  $CS_2$ , CO, and  $O_2$ .

The use of UCCSD(T) for treating the carbon atom asymptotes decreases the binding energy by  $\sim 0.2$  kcal/mol, whereas the RCCSD(T) method increases it by a comparable

amount. As noted previously,<sup>7</sup> the difference between the UCCSD(T) and R/UCCSD(T) open shell methods for the atoms largely disappears when core/valence correlation corrections are incorporated for computing  $D_e$ . However, the difference between UCCSD(T) or R/UCCSD(T) and RCCSD(T) remains on the order of 0.1–0.2 kcal/mol per first row atom. Thus, for large closed shell systems the choice of open shell method for treating the atoms can have a significant effect on the atomization energy.

The present CCSD(T)(FC)/CBS dissociation energy overlaps the low end of the available experimental measurements,<sup>50–53</sup> but an interim Active Thermochemical Table (ATcT) value from Ruscic,<sup>54</sup>  $146.67 \pm 0.13$  kcal/mol, is more than 1.6 kcal/mol larger in magnitude. The ATcT combines experimental and theoretical determinations in a thermochemical network approach to arrive at a best currently available consensus.<sup>55</sup> Application of CV and DKH corrections, although relatively minor for a small first row diatomic such as  $C_2$ , is expected to improve the agreement with experiment. The size of both corrections scales with the size of the system. For example, in *n*-octane,  $C_8H_{18}$ , the CV and DKH corrections approach 9 and  $-2.3$  kcal/mol, respectively.<sup>56</sup> For  $C_2$ , Table I shows that the CV correction approaches a CBS value of  $\sim 0.99$  kcal/mol, but basis sets of at least CVQZ quality are required to converge the correction to within 0.1 kcal/mol of the limit. A similar conclusion was reached for  $N_2$  in an earlier study.<sup>7</sup> The scalar relativistic correction converges to a limiting value of  $-0.17$  kcal/mol, and is essentially converged at the VTZ\_DK basis set level.

As previously mentioned, our approach normally relies on a series of frozen core calculations followed by the incorporation of several smaller corrections that are assumed to be additive. The practical advantage of this approach is that it facilitates the use of very large, diffuse function augmented basis sets at the FC level. If we were required to combine these sets with the additional “tight” functions needed for CV correlation, the resulting basis sets would become prohibitively expensive. For  $C_2$ , a test of the error introduced by assuming additivity in the CV correction was performed at the aug-cc-pCV6Z basis set level. The resulting errors (0.000 06 Å in  $r_e$ , 0.01 kcal/mol in  $D_e$ , and 0.3  $cm^{-1}$  in  $\omega_e$ ) are negligible compared to other sources of error in our calculations. We often take the decomposition of the problem a step further by performing the CV calculation with a basis set that is one basis set level smaller than the one used for the FC calculation. For example, if a CCSD(T)(FC) calculation was performed with the aug-cc-pV6Z basis set, it might be combined with a CCSD(T)(CV)/cc-pCV5Z correction. By exploiting the different rates of convergence in the FC, CV, and DKH portions of the problem, even greater savings in computer resources are possible with minimal increase in uncertainty. The errors resulting from approximating the full (CV)/aug-cc-pCV6Z calculation by (FC)/aug-cc-pV6Z + (CV)/cc-pCV5Z are 0.00015 Å ( $r_e$ ),  $-0.03$  kcal/mol ( $D_e$ ), and  $-0.5$   $cm^{-1}$  ( $\omega_e$ ).

As indicated by the CCSD(T) and FCI entries in Table I, corrections to the three spectroscopic properties arising from higher order excitations are seen to be significant if uncertainties of  $\leq 0.001$  Å ( $r_e$ ),  $\leq 1$   $cm^{-1}$  ( $\omega_e$ ), and  $\leq 1$  kcal/mol

TABLE I.  $C_2$  ( $^1\Sigma_g^+$ ) theoretical results. Units are  $r_e$  in Å,  $\omega_e$  in  $\text{cm}^{-1}$ ,  $D_e$  in kcal/mol, and total energies ( $E$ ) at the optimized bond length 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 R/UCCSD(T) atomic energies obtained with symmetry equivalencing and with the removal of high angular momentum basis set contaminants. All theoretical  $D_e$  values have been decreased by 0.17 kcal/mol to account for atomic spin-orbit effects.

Basis	Method	$r_e$	$\omega_e$	$D_e$	$E$
aug-cc-pVDZ	CCSD(T)(FC)	1.2720	1814.3	128.57	-75.734 752
cc-pCVDZ	CCSD(T) $\Delta^a$	-0.0009	3.7	0.35	-75.803 987
cc-pVDZ_DK	CCSD(T) DKH $\Delta^b$	-0.0001	-0.7	-0.19	-75.757 396
cc-pVDZ	CCSDTQ(FC) $\Delta^c$	0.0017	-12.2	0.43	-75.729 196
cc-pCVDZ	CCSDTQ $\Delta^d$	0.0000	-0.5	0.12	-75.805 525
cc-pVDZ	FCI(FC) $\Delta^e$	0.0005	-3.0	0.41	-75.729 853
DZ	Composite <sup>f</sup>	1.2732	1801.6	129.69	
aug-cc-pVTZ	CCSD(T)(FC)	1.2508	1840.7	139.72	-75.786 381
cc-pCVTZ	CCSD(T) $\Delta^c$	-0.0022	7.5	0.84	-75.883 828
cc-pVTZ_DK	CCSD(T) DKH $\Delta^c$	-0.0001	-0.5	-0.18	-75.812 806
cc-pVTZ	CCSDTQ(FC) $\Delta^d$	0.0016	-12.8	0.23	-75.784 632
cc-pCVTZ	CCSDTQ $\Delta^h$	-0.0001	-0.5	-0.02	
cc-pVTZ	FCI(FC) $\Delta^h$	0.0004	-3.9	0.37	-75.785 419
TZ	Composite <sup>i</sup>	1.2504	1830.5	140.96	
aug-cc-pVQZ	CCSD(T)(FC)	1.2460	1854.5	143.28	-75.802 144
cc-pCVQZ	CCSD(T) $\Delta^c$	-0.0030	9.5	0.89	-75.908 749
cc-pVQZ_DK	CCSD(T) DKH $\Delta^d$	-0.0001	-0.5	-0.17	-75.830 539
cc-pVQZ	CCSDTQ(FC) $\Delta^i$	0.0015	-12.1	0.09	
VQZ	Composite <sup>f</sup>	1.2447	1847.0	144.46	
aug-cc-pV5Z	CCSD(T)(FC)	1.2448	1858.4	144.28	-75.806 730
cc-pCV5Z	CCSD(T) $\Delta^c$	-0.0032	10.6	0.96	-75.916 084
cc-pV5Z_DK	CCSD(T) DKH $\Delta^d$	-0.0001	-0.5	-0.17	-75.835 858
V5Z	Composite <sup>f</sup>	1.2433	1852.0	145.53	
aug-cc-pV6Z	CCSD(T)(FC)	1.2444	1859.5	144.67	-75.808 324
cc-pCV6Z	CCSD(T) $\Delta^c$	-0.0033	10.9	0.98	-75.918 788
V6Z	Composite <sup>f</sup>	1.2428	1853.4	145.94	
aug-cc-pV7Z	CCSD(T)(FC)	1.2443	1860.1	144.91	-75.809 041
V7Z	Composite <sup>f</sup>	1.2427	1854.0	146.18	
CBS	CCSD(T)(FC)	1.2443	1860.1	145.1±0.1	-75.809 9
CBS	Composite <sup>f</sup>	1.2427	1854.0	146.4±0.3	
Expt.		1.2425 <sup>j</sup>	1854.7 <sup>j</sup>	146.67±0.13 <sup>k</sup> 147.3±0.3 <sup>l</sup> 147.5±0.4 <sup>m</sup> 147.9±0.5 <sup>o</sup> 144.4±0.9 <sup>p</sup> 145.8±4.6	

<sup>a</sup>Difference between the frozen core and all-electron values, i.e.,  $X-X_{FC}$ . Frozen core calculations excluded the C(1s) electrons from the correlation treatment.

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

<sup>c</sup>CCSDTQ(FC) correction to the CCSD(T)(FC) properties. The CCSDTQ dissociation energy correction is with respect to ROHF-CCSDTQ atoms without the imposition of symmetry equivalencing of the atomic orbitals.

<sup>d</sup>CCSDTQ(CV) correction to the CCSDTQ(FC) properties. The CCSDTQ dissociation energy corrections are with respect to ROHF-CCSDTQ atoms without the imposition of symmetry equivalencing of the atomic orbitals.

<sup>e</sup>FCI(FC) correction to the CCSDTQ(FC) values.

<sup>f</sup>Composite values represent an approximation to the scalar relativistic, core/valence FCI result at each basis set level. Composite values were obtained by combining the frozen core, core/valence, scalar relativistic, and higher order corrections.

<sup>g</sup>Estimated by combining CCSDT(CV)/cc-pCVTZ corrections with CCSDTQ(CV)/cc-pCVTZ(no  $f$ ) corrections.

<sup>h</sup>Continued fraction estimated FCI/cc-pVTZ corrections for  $r_e$  and  $\omega_e$ . The correction for  $D_e$  is based on an FCI/cc-pVTZ(+diff.s,p) calculation at  $r_e(\text{Expt.})=1.2425$  Å by Gan and Harrison. (Ref. 57)

<sup>i</sup>CCSDTQ/cc-pVQZ properties were estimated as follows:  $r_e$  and  $\omega_e$  were based on CCSDT/cc-pVQZ+CCSDTQ/cc-pVQZ(no  $g$ ) energies. The dissociation energy included an additional correction for the missing  $g$  functions at that CCSDTQ level, as described in the text.

<sup>j</sup>Experimental  $r_e$  and  $\omega_e$  values are taken from Huber and Herzberg (Ref. 50).

<sup>k</sup>Ruscic, based on Active Thermochemical Tables, ATcT Version 1.35 and the Core (Argonne) Thermochemical Network Version 1.062, 2006 (Ref. 59).

<sup>l</sup>Ruscic, based on Active Thermochemical Tables, ATcT Versin 1.25 and the Core (Argonne) Thermochemical Network Version 1.056, 2006 (Ref. 54).

<sup>m</sup>Burcat and Ruscic (Ref. 51).

<sup>n</sup>Urdahl *et al.* (Ref. 52).

<sup>o</sup>NIST/JANAF (Ref. 53).

<sup>p</sup>Huber and Herzberg (Ref. 50).

TABLE II.  $C_2$  Convergence of the CCSDT(FC) and CCSDTQ(FC) corrections to  $D_e$  (kcal/mol). Values were obtained at the optimal CCSDT and CCSDTQ bond lengths, unless otherwise specified.

Basis	CCSDT <sup>a</sup>	$\Delta_{\text{CCSDT}}^b$	CCSDTQ <sup>c</sup>	$\Delta_{\text{CCSDTQ}}$	Total <sup>d</sup>
cc-pVDZ	-1.234		1.663		0.43
cc-pVTZ	-1.918	0.69	2.148	0.46	0.23
cc-pVQZ	-2.170	0.25	2.258 <sup>e</sup>	0.11	0.09
cc-pV5Z	-2.195	0.03			
CBS	-2.198		2.290		0.09

<sup>a</sup> $D_e(\text{CCSDT}) - D_e[\text{CCSD(T)}]$ . The CCSD(T) values were determined with respect to R/UCCSD(T) symmetry equivalenced atoms.

<sup>b</sup>Difference between successive values, e.g., CCSDT(VTZ) - CCSDT(VDZ).

<sup>c</sup> $D_e(\text{CCSDTQ}) - D_e(\text{CCSDT})$ . Symmetry equivalencing of the atoms was not enforced.

<sup>d</sup>CCSDT+CCSDTQ.

<sup>e</sup>CCSDTQ/cc-pVQZ(no g)+correction for missing g functions, as described in text.

( $D_e$ ) are desired. With the cc-pVDZ basis set, higher order correlation lengthens  $r_e$  by 0.0022 Å, decreases  $\omega_e$  by 15.7  $\text{cm}^{-1}$ , and increases  $D_e$  by 1.0 kcal/mol, relative to the CCSD(T) values. In the latter case, over 40% of the effect comes from the FCI correction to CCSDTQ. The size of the higher order corrections to  $r_e$  and  $\omega_e$  are relatively insensitive to further increase in the size of the basis set. However, the correction for  $D_e$  generally decreases with increasing basis set size, in accord with our earlier findings for  $N_2$ .<sup>7</sup> Because we were unable to carry out explicit FCI/cc-pVTZ calculations or to find comparable results in the literature, the cc-pVTZ corrections for  $r_e$  and  $\omega_e$  in Table I are based on the cf approximant. For  $D_e$  we used the FCI/cc-pVTZ(+diff.s,p) result from Gan and Harrison,<sup>57</sup> which was obtained at the experimental bond length ( $r=1.2425$  Å).

Local software limitations prevented us from performing CCSDTQ/cc-pVQZ calculations. Estimates of the CCSDTQ/cc-pVQZ corrections were obtained by combining CCSDT/cc-pVQZ+CCSDTQ/cc-pVQZ(no g) components with a small correction for the missing g functions. The last of these involved calculations in which the g functions were present, but the two sets of f functions were removed. Combining CCSD(T)→CCSDT and CCSDT→CCSDTQ corrections, where the T correction is obtained with a larger basis set than the Q correction, rests upon the assumption that the former converges more rapidly, with respect to the one-particle basis set, than the latter. Each successively higher level of excitation converges roughly an order of magnitude faster, than the next lower order.<sup>7</sup> However, care must be exercised, since the CCSDT and CCSDTQ corrections are often of opposite signs and may individually be three to four times larger than the total correction, as seen in Table II, where data for the higher order corrections to the dissociation energy of  $C_2$  are shown. In the HEAT procedure, the CCSDT correction is extrapolated to the CBS limit, but the CCSDTQ correction is obtained from a cc-pVDZ calculation.<sup>8</sup> In the case of  $C_2$ , this would seem to lead to an error on the order of 0.5 kcal/mol in the higher order correction.

We have also investigated the effects of connected quadruple excitations on the CV correction. In our previous studies, it was assumed that CV contributions beyond CCSD(T) would be insignificant. That assumption appears to have been correct, as seen in the results in Table I. With the

cc-pCVDZ basis set, no correction was found for  $r_e$  and the adjustment in  $\omega_e$  was only 0.5  $\text{cm}^{-1}$ .  $D_e$  increases by just 0.12 kcal/mol. Software limitations prevented us from performing calculations at the CCSDTQ(CV)/cc-pCVTZ level of theory. An estimate of the effect was obtained by combining CCSDT(CV)/cc-pCVTZ+CCSDTQ(CV)/cc-pCVTZ (no f) corrections. At the CVTZ level, the size of the CCSDTQ(CV) corrections to  $r_e$  and  $\omega_e$  are essentially the same as the CVDZ values. The correction to  $D_e$  decreased to a negligible -0.02 kcal/mol. In the W4 paper, Karton *et al.*<sup>18</sup> reported an UCCSDT(Q)/cc-pCVTZ correction to the binding energy of 0.08 kcal/mol.

The DBOC was evaluated at the CISD/aug-cc-pVTZ level of theory. As was the case in our earlier study of  $N_2$  and CS,<sup>7</sup> the magnitude of the correction was negligibly small, at 0.03 kcal/mol. Electron correlation has little impact on this value. The RHF/aug-cc-pVTZ value was essentially the same, as has been reported by Karton *et al.*<sup>18</sup>

Our best final estimates, labeled “CBS Composite” at the bottom of Table I, are in excellent agreement with experiment. The theoretical  $D_e$  values include a -0.17 kcal/mol correction for atomic spin-orbit effects.<sup>58</sup> Combining the CBS composite atomization energy with a -2.64 kcal/mol anharmonic zero point vibrational correction<sup>50</sup> yields  $D_0=143.7\pm 0.3$  kcal/mol when dissociating to R/UCCSD(T) atoms. Theory and experiment are in essentially exact agreement on the zero point energy. This value compares well with the most recent ATcT value of  $144.03\pm 0.13$  kcal/mol due to Ruscic.<sup>59</sup> Over time, the experimental value has tended to drift lower. As recently as early 2006, the best available experimental value ( $144.7\pm 0.3$  kcal/mol) (Ref. 54) lay slightly beyond the theoretical error bars. The 1998 NIST/JANAF value ( $141.8\pm 0.9$  kcal/mol) appears to be several kcal/mol too small. We believe our error bars, which are based on the uncertainty in the frozen core CCSD(T)/CBS estimate, plus a contribution due to the uncertainty in the higher order correction, to be conservative. If the RCCSD(T) method is used for the atomic asymptotes, the theoretical estimate increases slightly, to  $144.2\pm 0.3$  kcal/mol and is in somewhat better agreement with experiment.



TABLE III. HO<sub>2</sub> <sup>2</sup>A'' previous theoretical and experimental atomization energies (kcal/mol).

		Total									
$\Sigma D_e(\text{FC})$	$\Delta E_{\text{CV}}$	$\Sigma D_e(\text{FC}+\text{CV})$	$\Delta E_{\text{SR}}$	$\Delta E_{\text{SO}}$	$\Delta E_{\text{HO}}$	$\Sigma D_e$	ZPE	$\Sigma D_0$	$\Delta H_f(298.15 \text{ K})$	Reference	
174.46 <sup>a</sup>	0.22 <sup>b</sup>	174.68	-0.37 <sup>c</sup>	-0.43 <sup>d</sup>		173.88	-8.61 <sup>e</sup>	165.27	3.6	Dixon <i>et al.</i> (Ref. 60)	
		175.07 <sup>f</sup>	-0.29 <sup>g</sup>	-0.39	0.51 <sup>h</sup>	174.90	-8.85 <sup>i</sup>	166.05±0.1	2.96±0.1	Flowers <i>et al.</i> (Ref. 26) and Tajti <i>et al.</i> (Ref. 8)	
								165.4+1.0/-0.5	3.5+1.0/-0.5	Shum <i>et al.</i> (Ref. 61)	
								165.1±1.2	3.8±1.2	Fischer and Armentrout (Ref. 62)	
								165.4±3	3.5±3	Holmes <i>et al.</i> (Ref. 63)	
								166.1±0.5	2.8±0.5	DeMore <i>et al.</i> (Ref. 64)	
								168.4±2.1	0.5±2.1	NIST-JANAF (Ref. 53)	
								165.6±0.8	3.3±0.5	Litorja and Ruscic (Ref. 65)	
								165.6±0.8	3.3±0.8	Sander <i>et al.</i> JPL (Ref. 66)	
								165.7±0.5	3.2±0.5	Raymond <i>et al.</i> (Ref. 67)	
								165.85±0.21	3.06±0.21	Ruscic (Ref. 68)	
								165.97±0.06	2.94±0.06	Ruscic <i>et al.</i> (Ref. 69)	

<sup>a</sup>Frozen core CBS estimate (mixed formula) based on R/UCCSD(T) energies obtained with the aug-cc-pVnZ, n=D, T and Q, basis sets.

<sup>b</sup>Core/valence correction based on CCSD(T)/cc-pCVQZ calculations.

<sup>c</sup>Scalar relativistic correction based on CISD/cc-pVTZ mass-velocity + one electron Darwin terms.

<sup>d</sup>Based on atomic splittings given in the tables of Moore.

<sup>e</sup>Zero point vibrational energy based on the average of the CCSD(T)/aug-cc-pVDZ harmonic frequencies and the experimental fundamentals.

<sup>f</sup>Based on separate extrapolations of the UHF and UCCSD(T) correlation energies. An exponential formula was used for the UHF component and a  $1/\ell^3_{\text{max}}$  formula for the UCCSD(T) component. The basis sets were taken from the aug-cc-pVnZ, n=D, T, Q and 5, sequence.

<sup>g</sup>Scalar relativistic correction based on CCSD(T)(full)/aug-cc-pCVTZ mass-velocity + one electron Darwin terms.

<sup>h</sup>Higher order correction obtained from a  $1/\ell^3_{\text{max}}$  CBS extrapolation of the CCSDT-CCSD(T) cc-pVTZ and cc-pVQZ corrections -0.15 kcal/mol plus a CCSDTQ/cc-pVDZ (0.60 kcal/mol) and CCSDTQ5/cc-pVDZ (0.06 kcal/mol) correction.

<sup>i</sup>Anharmonic zero point vibrational energy obtained from CCSD(T)(full)/cc-pVQZ calculations.

## B. HO<sub>2</sub>

A summary of previously reported theoretical and experimental atomization energies for HO<sub>2</sub> is shown in Table III. Dixon *et al.*<sup>60</sup> reported  $\Sigma D_0=165.27$  kcal/mol and  $\Delta H_f(298.15 \text{ K})=3.6$  kcal/mol, based on CCSD(T)/aug-cc-pVDZ through CCSD(T)/aug-cc-pVQZ calculations, followed by a single CBS extrapolation and CV and SR corrections. Several years later, Flowers *et al.*<sup>26</sup> arrived at slightly different values [ $\Sigma D_0=166.05\pm 0.1$ ,  $\Delta H_f(298)=2.96\pm 0.1$  kcal/mol] using the larger aug-cc-pV5Z basis set and incorporating corrections for higher order correlation. HO<sub>2</sub> was also included in the initial HEAT paper by some of the same authors.<sup>8</sup> The atomization energy remained unchanged. Experimental  $\Delta H_f$  values ranged from 3.5+1.0/-0.5 to 2.940.06 kcal/mol.<sup>53,61-69</sup>

Hardware limitations prevented us from carrying out CCSD(T)/aug-cc-pV7Z calculations with the two sets of *k* functions present on both oxygen atoms. The small contribution from the missing *k* functions was estimated by performing an exponential extrapolation of the incremental correlation energy lowering associated with each  $\ell$  value.

The spectroscopic constants for HO<sub>2</sub> obtained in the current work are listed in Table IV. Convergence of the properties, with respect to the one-particle expansion, follows a similar pattern to what was observed for C<sub>2</sub>, despite the difference in the degree of multiconfigurational character in the two systems. At the CCSD(T)(FC)/aug-cc-pVDZ basis set level, both systems are approximately 17 kcal/mol below the CBS atomization limit. At the aug-cc-pVQZ level the error has fallen to 2 kcal/mol. The convergence in bond lengths is somewhat faster in the case of HO<sub>2</sub>.

The CCSDTQ/cc-pVDZ correction to the atomization energy is actually larger for HO<sub>2</sub> (0.95 versus 0.43 kcal/mol) than for C<sub>2</sub>. Unlike the situation with C<sub>2</sub>, the CCSD(T) → CCSDT and CCSDT → CCSDTQ corrections are of the same sign, 0.32 and 0.63 kcal/mol, respectively. Due to the decrease in molecular symmetry, relative to C<sub>2</sub>, an explicit CCSDTQ/cc-pVTZ calculation proved to be prohibitively expensive. The CCSDTQ/cc-pVTZ entry in Table IV was estimated by combining CCSDT/cc-pVTZ+CCSDTQ/cc-pVTZ(no *f* O, no *d* H) corrections.

Combining the CBS composite atomization energy from Table IV with a -0.43 kcal/mol correction for atomic spin-orbit effects<sup>58</sup> and a -8.85 kcal/mol anharmonic zero point vibrational correction,<sup>26</sup> leads to a  $\Sigma D_0=166.0\pm 0.3$  kcal/mol. This value is in essentially exact agreement with the values reported by Flowers *et al.*<sup>26</sup> and Tajti *et al.*,<sup>8</sup> although the present approach differs in several details. It is also in excellent agreement with the 165.97±0.06 kcal/mol ATcT value reported by Ruscic *et al.*<sup>69</sup> The RCCSD(T) open shell method produces a slightly larger estimate,  $\Sigma D_0=166.2\pm 0.3$  kcal/mol.

## C. Other systems

In addition to C<sub>2</sub> and HO<sub>2</sub>, statistics have been compiled for a much larger collection of molecules. Due to the high cost of coupled cluster or configuration interaction techniques when combined with extended basis sets, many previous thermochemical surveys have been limited to a relatively small number of molecules. For example, the seminal work of Klopper *et al.*,<sup>25</sup> which reported an  $\epsilon_{\text{MAD}}$  of 1 kJ/mol (0.24 kcal/mol) for a set of reaction enthalpies,



TABLE IV. HO<sub>2</sub> (<sup>2</sup>A'') theoretical results. Units are  $r_{\text{OO}}$  and  $r_{\text{OH}}$  in Å,  $\angle\text{HOO}$  in degrees,  $\omega_i$  in cm<sup>-1</sup>, the electronic atomization energy ( $\Sigma D_e$ ) in kcal/mol, and total energies ( $E$ ) at the optimized geometry 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 R/UCCSD(T) atomic energies obtained with symmetry equivalencing and with the removal of high angular momentum contaminants. The CCSDTQ dissociation energy corrections are with respect to UHF-CCSDTQ atoms.

Basis	Method	$r_{\text{OO}}$	$r_{\text{OH}}$	$\angle\text{HOO}$	$\omega_1$	$\omega_2$	$\omega_3$	$\Sigma D_e$	$E$
aug-cc-pVDZ	R/UCCSD(T)(FC) <sup>a</sup>	1.3484	0.9789	103.9	1076.3	1424.6	3618.7	157.17	-150.600 937
cc-pVDZ	R/UCCSD(T) $\Delta^b$	-0.0008	-0.0005	0.0	0.5	1.3	3.1	0.25	-150.640 416
cc-pVDZ_DK	R/UCCSD(T) DKH $\Delta^c$	0.0006	0.0001	0.0	-1.9	-0.9	-2.3	-0.35	-150.662 499
cc-pVDZ	RCCSDTQ(FC) $\Delta$	0.0042	0.0005	-0.1				0.95	-150.560 592
cc-pVDZ	cf FCI(FC) $\Delta^d$							0.09	-150.560 748
DZ	Composite <sup>e</sup>	1.3524	0.9790	103.8	1074.9	1425.0	3619.5	158.11	
aug-cc-pVTZ	R/UCCSD(T)(FC) <sup>f</sup>	1.3376	0.9741	104.2	1121.3	1428.0	3641.5	168.93	-150.726 674
cc-pVTZ	R/UCCSD(T) $\Delta^b$	-0.0019	-0.0007	0.0	3.0	1.8	4.2	0.27	-150.827 580
cc-pVTZ_DK	R/UCCSD(T) DKH $\Delta^c$	0.0006	0.0001	0.1	-1.9	-0.6	-2.4	-0.26	-150.816 794
cc-pVTZ	RCCSDTQ(FC) $\Delta^g$							0.48	
TZ	Composite <sup>e</sup>	1.3405	0.9740	104.2	1122.4	1429.2	3643.3	169.51	
aug-cc-pVQZ	R/UCCSD(T)(FC)	1.3304	0.9715	104.4	1139.4	1438.4	3659.2	172.68	-150.765 403
cc-pVQZ	R/UCCSD(T) $\Delta^b$	-0.0021	-0.0008	0.1				0.24	-150.881 684
cc-pVQZ_DK	R/UCCSD(T) DKH $\Delta^c$	0.0005	0.0000	0.0				-0.26	-150.864 231
VQZ	Composite <sup>e</sup>	1.3330	0.9712	104.4	1140.5	1439.6	3661.0	173.23	
aug-cc-pV5Z	R/UCCSD(T)(FC)	1.3290	0.9711	104.4	1144.2	1444.2	3662.9	173.71	-150.778 045
cc-pV5Z	R/UCCSD(T) $\Delta^b$	-0.0022	-0.0008	0.1				0.24	-150.899 215
cc-pV5Z_DK	R/UCCSD(T) DKH $\Delta^c$	0.0005	0.0000	0.0				-0.27	-150.880 327
V5Z	Composite <sup>e</sup>	1.3315	0.9708	104.4	1145.3	1445.4	3664.7	174.25	
aug-cc-pV6Z	R/UCCSD(T)(FC)	1.3287	0.9710	104.4				174.16	-150.782 390
cc-pV6Z	R/UCCSD(T) $\Delta^b$	-0.0022	-0.0008	0.1				0.25	-150.905 312
V6Z	Composite <sup>e</sup>	1.3312	0.9707	104.4				174.71	
aug-cc-pV7Z	R/UCCSD(T)(FC)	1.3286	0.9710	104.4				174.41	-150.784 50 <sup>h</sup>
V7Z	Composite <sup>e</sup>	1.3311	0.9707	104.4				174.96	
CBS	R/UCCSD(T)(FC)	1.3286	0.9710	104.4	1145.9	1451.5	3663.9	174.7±0.1	-150.7869
CBS	Composite <sup>e</sup>	1.3311	0.9707	104.4	1147.0	1452.7	3665.7	175.3±0.3	

<sup>a</sup>For comparison purposes, the UCCSD(T)/aug-cc-pVDZ values are  $r_{\text{OO}}=1.3482$  Å;  $r_{\text{OH}}=0.9789$  Å;  $\angle\text{HOO}=103.9^\circ$ ;  $\omega_i=1080.3, 1426.1, 3620.3$  cm<sup>-1</sup>;  $\Sigma D_e=157.1$  kcal/mol (without symmetry equivalencing of the atoms). The RCCSD(T) values are  $r_{\text{OO}}=1.3471$  Å;  $r_{\text{OH}}=0.9787$  Å;  $\angle\text{HOO}=103.9^\circ$ ;  $\omega_i=1079.5, 1426.8, 3620.4$  cm<sup>-1</sup>;  $\Sigma D_e=156.7$  kcal/mol.

<sup>b</sup>Difference between the frozen core and nonfrozen core values, i.e.,  $X-X_{\text{FC}}$ . Frozen core calculations excluded the O(1s) electrons from the correlation treatment.

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

<sup>d</sup>Estimated full CI/cc-pVDZ correction to the CCSDTQ/cc-pVDZ values, based on a continued fraction extrapolation of the CCSD, CCSDT, and CCSDTQ energies.

<sup>e</sup>Composite values represent an approximation to the scalar relativistic, core/valence FCI result at each basis set level. Composite values were obtained by combining the frozen core, core/valence, scalar relativistic, and higher order corrections. Whenever a correction is unavailable for a given basis set, e.g., FCI/cc-pVQZ, the next available lower basis set correction is substituted.

<sup>f</sup>For comparison purposes, the UCCSD(T)/aug-cc-pVTZ values are  $r_{\text{OO}}=1.3375$  Å;  $r_{\text{OH}}=0.9741$  Å;  $\angle\text{HOO}=104.2^\circ$ ;  $\Sigma D_e=168.8$  kcal/mol (without symmetry equivalencing of the atoms). The RCCSD(T) values are  $r_{\text{OO}}=1.3361$  Å;  $r_{\text{OH}}=0.9739$  Å;  $\angle\text{HOO}=104.2^\circ$ ;  $\omega_i=1123.2, 1430.2, 3642.0$  cm<sup>-1</sup>;  $\Sigma D_e=168.6$  kcal/mol.

<sup>g</sup>Estimated CCSDTQ/cc-pVTZ correction to the CCSD(T) atomization energy based on CCSDT/cc-pVTZ+CCSDTQ/cc-pVTZ(no *f* O, no *d* H).

<sup>h</sup>Estimated by extrapolating the contribution of the *k* functions on oxygen, as described in the text.

was based on just 15 examples. Two molecules from their original set were eliminated because they either involved excessive multireference character or the accuracy of the experimental data was called into question. Other studies, such as those discussing the HEATx methods,<sup>8,9</sup> have limited themselves to molecules with available experimental heats of formation whose uncertainties are <0.1 kcal/mol.

Our goal is somewhat different. We wish to determine if currently affordable levels of coupled cluster theory are capable of reproducing experimental heats of formation (or atomization energies) to an accuracy of  $\pm 1$  kcal/mol or better,

without imposing an overly restrictive criterion for inclusion of molecules in the test set. Besides thermochemical comparisons, we wish to evaluate the performance of theoretical models for molecular structures and normal mode frequencies. However, in this work, we will only examine the first two. Statistical results will be presented for two basic levels of theory. The first involves the CCSD(T)(FC) method with basis sets ranging from aug-cc-pVDZ up to aug-cc-pV6Z. At the frozen core level of theory, comparisons will range over the entire list of molecules in the CRDB, regardless of the accompanying uncertainty in the experimental data. Conse-

quently, those statistics will involve molecules such as  $B_2$  ( $^3\Sigma_g^-$ ) with potentially less reliable experimental information, e.g.,  $D_0(\text{Expt.})=75.6\pm 8$  kcal/mol. The number of atomization energy comparisons will range from 252 for the aug-cc-pVDZ basis set to 155 for the aug-cc-pV5Z set and 71 for the aug-cc-pV6Z set.

The second, or “Best,” level for which statistics will be gathered involves a CBS extrapolation plus corrections for core/valence and scalar relativistic effects, at minimum. For most entries in the Best category, an additional correction for higher order correlation via CCSDTQ, cf estimated FCI or FCI, is also included. Where required by the nature of the elements composing the molecule, there are also corrections for the DBOC and second order spin-orbit effects. A relatively loose criterion of  $\pm 5$  kcal/mol (experimental uncertainty) was used for including molecules in the test set for this category. In light of this, conclusions drawn from the statistics should be tempered by an understanding of the inherent uncertainty in the reference data. At present, only 68 comparisons are available in the CRDB for this category. Nonetheless, these systems include elements through the fourth row of the periodic table. More than half of the CCSD(T)(FC)/CBS estimates were based on calculations through aug-cc-pV6Z [aug-cc-pV(6+d)Z for second row elements] energies. An additional nine estimates incorporated aug-cc-pV7Z energies. Core/valence corrections were generally obtained with the cc-pCV5Z or cc-pwCV5Z (Ref. 70) basis sets and most scalar relativistic corrections were obtained from DKH CCSD(T)/cc-pVTZ\_DK calculations.

As discussed in our earlier studies of sources of error in electronic structure calculations,<sup>3,7</sup> we prefer to use accurate theoretical, anharmonic zero point energies (ZPEs) whenever possible, but such values are seldom available for systems with more than five or six atoms. Instead, we have often estimated the anharmonic ZPE by performing a 1:1 averaging of theoretical harmonic frequencies and experimental fundamentals, as suggested by Grev *et al.*<sup>71</sup> More recently, Barone<sup>72</sup> demonstrated that approximate, theoretically determined fundamentals are adequate in many cases for this purpose. In a comparison of ZPEs based on the average of high quality, CCSD(T) harmonic frequencies and experimental fundamentals against accurate anharmonic ZPEs, we found mean absolute deviations ranging from 0.16 kcal/mol (aug-cc-pVDZ, 47 comparisons) to 0.08 kcal/mol (aug-cc-pVQZ, 27 comparisons).

Mean absolute deviations, standard deviations and maximum positive, and negative errors are shown in Fig. 2. The CCSD(T)(FC)  $\epsilon_{\text{MAD}}$  values range from 49.5 kcal/mol with the small aug-cc-pVDZ basis set to 2.8 kcal/mol with the much larger aug-cc-pV6Z basis set. At the Best level, incorporation of the CV, SR, and higher order corrections proved effective at driving down the MAD to  $\sim 0.3$  kcal/mol, but at sharply increased computational expense due principally to the notoriously poor scaling of higher order correlation methods. Experimentalists tend to prefer quoting an uncertainty corresponding to a 95% confidence interval (approximately two standard deviations). The standard deviation for the Best theory is 0.9 kcal/mol. Karton *et al.*<sup>18</sup> report an  $\epsilon_{\text{MAD}}$  value of 0.066 kcal/mol for the W4 protocol compared

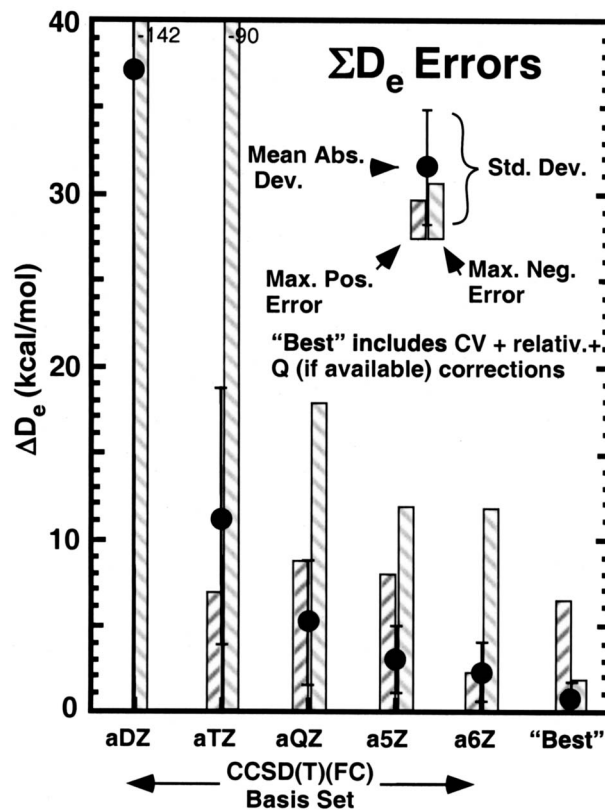


FIG. 2. Errors (kcal/mol) with respect to experimental atomization energies. The data labeled “Best” includes the effects of a CBS extrapolation, CV and SR corrections, and a CCSDTQ or FCI correction for higher order correlation recovery (when available).

to ATcT for 19 first row compounds. Using the same ATcT  $\Sigma D_0$  values, we find  $\epsilon_{\text{MAD}}=0.1$  kcal/mol following the procedure outlined in this work.

In only three cases did the Best error exceed chemical accuracy ( $\pm 1$  kcal/mol). The worst positive error, +6.6 kcal/mol, was observed for MnH ( $^7\Sigma^+$ ) which has an experimental  $D_0=29.3\pm 4.4$  kcal/mol.<sup>73</sup> The current theoretical value was taken from Balabanov and Peterson.<sup>38</sup> The largest negative error,  $-1.1$  kcal/mol, occurred in TiH ( $^4\Phi$ ) where the experimental  $D_0=48.0\pm 2.1$  kcal/mol.<sup>74</sup> Corrections for higher order correlation effects were not included. While most higher order correlation corrections were less than 1 kcal/mol for the systems examined in this study, the correction for  $O_3$  was 3.2 kcal/mol, reflecting the need for including some estimate of the remaining error at the CCSD(T) level of theory if a consistent agreement to  $\pm 1$  kcal/mol is desired. Similar conclusions were reached in the W4 study.<sup>18</sup>

Although the small size of the MAD for atomization energies is reassuring, the possibility exists that theoretical predictions based on the approach outlined in this work may eventually prove to be even more accurate than the current statistics indicate as better experimental and/or ATcT estimates are reported. For example, the Best theoretical zero point inclusive atomization energy for HOF is  $149.1\pm 0.3$  kcal/mol, after including a  $-0.60$  kcal/mol correction for atomic spin-orbit effects. Recent experimental values include the NIST/JANAF (Ref. 53) value of

TABLE V. Selected structural comparisons with experiment for polyatomic molecules. Bond lengths are in angstroms and bond angles are in degrees. The estimated experimental uncertainties are  $\pm 0.001 \text{ \AA}$  and  $0.1^\circ$ . The theoretical values are based on complete basis set extrapolations of the frozen core CCSD(T) parameters using up to aug-cc-pV5Z (ethylene  $\text{C}_2\text{H}_4$ ) and aug-cc-pVQZ (butadiene  $\text{C}_4\text{H}_6$  and cyclopropane  $\text{C}_3\text{H}_6$ ) basis sets, with core/valence and Douglas-Kroll relativistic corrections applied.

Ethylene	$r_{\text{CC}}$	$r_{\text{CH}}$	$\angle\text{CCH}$							Reference
Theory	1.3305	1.0804	121.4							This work
Expt.	1.3305	1.0805	121.5							Craig <i>et al.</i>
1,3-Butadiene	$r_{\text{CC}}$	$r_{\text{CC}}$	$\angle\text{CCC}$	$r_{\text{CH}}$	$\angle\text{CCH}$	$r_{\text{CH}}$	$\angle\text{CCH}$	$r_{\text{CH}}$	$\angle\text{CCH}$	Reference
Theory	1.3377	1.4548	123.5	1.0823	121.0	1.0799	121.5	1.0846	119.8	This work
Expt.	1.3376	1.4539	122.6	1.0819	121.0	1.0793	121.5	1.0847	119.9	Craig <i>et al.</i>
Cyclopropane	$r_{\text{CC}}$	$r_{\text{CH}}$	$\angle\text{HCH}$							Reference
Theory	1.5024	1.0791	114.8							This work
Expt.	1.5030	1.0786	115.0							Gauss <i>et al.</i>

$151.8 \pm 0.6 \text{ kcal/mol}$  and the Burcat/Ruscic<sup>51</sup> value of  $151.5 \text{ kcal/mol}$ , both of which are more than  $2 \text{ kcal/mol}$  larger than theory. However, the very recent ATcT value of  $148.96 \pm 0.15 \text{ kcal/mol}$  due to Ruscic<sup>59</sup> falls within  $0.1 \text{ kcal/mol}$  of theory.

One of the 68 chemical systems treated at the highest or Best levels of theory is  $\text{SO}_2$  ( $^1A_1$ ). In their W4 paper, Karton *et al.*<sup>18</sup> commented on potential basis set problems arising from what they refer to as “inner polarization” effects. They speculate that the aug-cc-pV( $n+d$ )Z sequence of basis sets might not adequately describe this effect and suggest the use of alternative basis sets.<sup>75</sup> With the approach outlined in the present work, we find no evidence of a basis set problem with the aug-cc-pV( $5+d$ )Z or aug-cc-pV( $6+d$ )Z sets. Our final  $\Sigma D_0 = 254.6 \pm 0.5 \text{ kcal/mol}$  value is in very close agreement with the ATcT value of  $254.46 \pm 0.02 \text{ kcal/mol}$  quoted in the W4 paper. Our value incorporates the same  $4.38 \text{ kcal/mol}$  zero point energy as W4, as well as CCSDT and CCSDTQ corrections. CBS extrapolations were averaged over five methods. Reducing the maximum size of the basis set to aug-cc-pV( $5+d$ )Z worsens the agreement with the ATcT value by only  $0.1 \text{ kcal/mol}$ .

In 1998 we examined the performance of the CCSD(T)  $\times$ (FC)/aug-cc-pVnZ,  $n=D, T$ , and  $Q$ , levels of theory in reproducing bond lengths and bond angles for the Gaussian-2 (G2) collection of molecules.<sup>76</sup> With the largest basis set, mean absolute deviations for bond lengths of  $0.009 \text{ \AA}$  (AH) and  $0.007 \text{ \AA}$  (AB) were reported. Although core/valence ge-

ometry optimizations were prohibitively expensive for most of the G2 systems at that time, it was noted that earlier work on first row diatomics indicated that changes of  $\pm 0.002 \text{ \AA}$  could be expected.<sup>77</sup> Subsequent studies explored the impact of correlation recovery beyond CCSD(T) for first row diatomics.<sup>22,23</sup> More recently, Ruden *et al.*<sup>78</sup> extended these earlier studies by applying the CCSDTQ(FC) method to four diatomics (HF,  $\text{N}_2$ ,  $\text{F}_2$ , and CO). Other studies include the work of Heckert *et al.*<sup>79,80</sup>

Comparisons between theoretical and experimental  $r_e$  bond lengths in diatomics are relatively straightforward compared to the situation with polyatomic molecules. In the latter case, typical “bottom of the well” values obtained from electronic structure methods are not directly comparable to the assortment of values ( $r_0, r_g, r_\alpha, r_z$ , etc.) obtained from spectroscopic or diffraction-based experimental measurements. Recently, this situation has begun to change with the adoption of a technique that adjusts the experimental rotational constants with vibration-rotation constants obtained from high quality calculations. The resulting “semiexperimental”  $r_e$  structures are often in excellent agreement with coupled cluster theory results, as seen in Table V. The semiexperimental results are taken from the work of Craig *et al.*<sup>81</sup> and Gauss *et al.*<sup>82</sup> A similar agreement was found for 1,1-difluorocyclopropane.<sup>83</sup>

CCSD(T)(FC) theory  $\epsilon_{\text{MAD}}$  values for bond lengths between hydrogen and nonhydrogen atoms (AH), between two nonhydrogens (AB) and bond angles (ABC) are presented in

TABLE VI. Mean absolute deviations in bond lengths ( $\text{\AA}$ ) and bond angles (degrees).

Method	$\epsilon_{\text{MAD}}(\text{AH})$	AH	$\epsilon_{\text{MAD}}(\text{AB})$	AB	$\epsilon_{\text{MAD}}(\text{HAB})$	HAB
CCSD(T)(FC)/aug-cc-pVDZ	0.0153	181	0.0278	225	1.0	50
CCSD(T)(FC)/aug-cc-pVTZ	0.0088	164	0.0099	190	0.6	38
CCSD(T)(FC)/aug-cc-pVQZ	0.0062	149	0.0058	175	0.5	37
CCSD(T)(FC)/aug-cc-pV5Z	0.0050	78	0.0058	100	0.3	13
CCSD(T)(FC)/aug-cc-pV6Z	0.0019	29	0.0033	42	0.5	8
“Best” <sup>a</sup>	0.0014	34	0.0007	52	0.2	8

<sup>a</sup>“Best” results are based on frozen core calculations performed with basis sets of at least aug-cc-pV5Z quality and including corrections for core/valence, scalar relativistic and higher order correlation recovery (when available). “AH” indicates bond distances involving a hydrogen atom. “HAB” indicates a bond angle involving a hydrogen atom and two non-hydrogen atoms. AH and AB indicate the number of comparisons.



Table VI, along with findings comparable to the Best results discussed previously for atomization energies. The majority of results were obtained from aug-cc-pV6Z frozen core calculations, adjusted by cc-pwCV5Z/cc-pCV5Z and cc-pVTZ\_DK corrections. Higher order corrections were also applied, where feasible, with CCSDT/cc-pVTZ and CCSDTQ/cc-pVDZ calculations. A small exponential extrapolation in the frozen core internal coordinates was performed in order to obtain the Best results. This extrapolation typically shortened bond distances by 0.0002–0.0003 Å, in agreement with the CCSD(T)-R12 findings of Heckert *et al.*<sup>80</sup> In Table VI the errors in bond lengths are seen to monotonically decline as the quality of the basis set improves. At the Best level of theory, the errors are very similar to the reported experimental uncertainties, i.e.,  $\epsilon_{\text{MAD}} = \pm 0.0014$  Å (AH, 34 comparisons) and  $\pm 0.0007$  Å (AB, 51 comparisons). Heckert *et al.*<sup>80</sup> reported a mean absolute deviation of 0.0003 Å for a smaller collection of 20 bond lengths, but relativistic corrections were not included in this value.

#### IV. CONCLUSION

The requirements for obtaining high accuracy spectroscopic and thermodynamic properties from coupled cluster theory calculations are illustrated for two prototypical small molecules. In accord with previous findings, various complete basis set extrapolations were found to be effective in accelerating the one-particle convergence. Nevertheless, expensive, higher order techniques beyond CCSD(T) are needed in order to consistently obtain thermodynamic properties to better than chemical accuracy ( $\pm 1$  kcal/mol). Success in minimizing theoretical uncertainty requires meticulous attention to detail in all steps. In particular, the treatment of CCSDT and CCSDTQ corrections must be carefully handled, since the individual corrections are often of opposite signs and much larger than the overall higher order correction.

The approach described in this and previous articles as well as the related approaches developed by other groups consist of many parts. For example, if we include the DBOC correction, there were nine contributions to the final composite  $D_0$  value for  $C_2$ . If carbon was replaced by a much heavier element, additional corrections might have been required, such as one to account for second order spin-orbit effects. Strictly speaking, in order to claim convergence in the theoretical properties to some level, e.g.,  $\pm 0.24$  kcal/mol ( $\pm 1$  kJ/mol), it is necessary to demonstrate convergence in each of the component pieces to an even smaller value. Due to the likelihood of at least partial cancellation of error, close statistical agreement with a small body of experimental data should not be construed as implying the same inherent level of overall uncertainty in the calculations.

In the overwhelming majority of cases studied to date, coupled cluster theory is capable of better than chemical accuracy in predicting atomization energies. For well-behaved systems possessing highly accurate experimental or ATcT values, the  $\epsilon_{\text{MAD}} = 0.1$  kcal/mol, similar to what was recently reported for the W4 protocol. For a larger collection of 68

molecules, the  $\epsilon_{\text{MAD}}$  increases to 0.3 kcal/mol ( $\epsilon_{\text{rms}} = 0.9$  kcal/mol). The same high level of theory is capable of predicting bond lengths to an accuracy of  $\pm 0.001$  Å.

#### ACKNOWLEDGMENT

Dr. B. Ruscic is thanked for early access to his most recently revised ATcT heats of formation values for  $C_2$ ,  $HO_2$ , and HOF and for helpful discussions about the ATcT approach.

- <sup>1</sup>D. Feller, *J. Comput. Chem.* **17**, 1571 (1996).
- <sup>2</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); D. Feller and D. A. Dixon, *ibid.* **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, K. A. Peterson, W. A. de Jong, and D. A. Dixon, *J. Chem. Phys.* **118**, 3510 (2003); D. Feller, D. A. Dixon, and J. S. Francisco, *J. Phys. Chem. A* **107**, 1604 (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>3</sup>D. Feller and D. A. Dixon, *J. Chem. Phys.* **115**, 3484 (2001).
- <sup>4</sup>B. Ruscic, A. F. Wagner, L. B. Harding *et al.*, *J. Phys. Chem. A* **106**, 2727 (2002).
- <sup>5</sup>D. Feller and D. A. Dixon, *J. Phys. Chem. A* **107**, 9641 (2003).
- <sup>6</sup>D. Feller and D. A. Dixon, *J. Phys. Chem. A* **107**, 10419 (2003).
- <sup>7</sup>D. Feller, K. A. Peterson, and T. D. Crawford, *J. Chem. Phys.* **124**, 054107 (2006).
- <sup>8</sup>A. Tajti, P. G. Szalay, A. G. Császár, M. Kállay, J. Gauss, E. F. Valeev, B. A. Flowers, J. Vázquez, and J. F. Stanton, *J. Chem. Phys.* **121**, 11599 (2004).
- <sup>9</sup>Y. J. Bomble, J. Vázquez, M. Kállay, C. Michauk, P. G. Szalay, A. G. Császár, J. Gauss, and J. F. Stanton, *J. Chem. Phys.* **125**, 064108 (2006).
- <sup>10</sup>A. D. Boese, M. Oren, O. Atasoylu, J. M. L. Martin, M. Kállay, and J. Gauss, *J. Chem. Phys.* **120**, 4129 (2004).
- <sup>11</sup>M. S. Schuurman, S. R. Muir, W. D. Allen, and H. F. Schaefer III, *J. Chem. Phys.* **120**, 11586 (2004).
- <sup>12</sup>B. C. Shepler and K. A. Peterson, *J. Phys. Chem. A* **110**, 12321 (2006).
- <sup>13</sup>J. A. Pople, M. Head-Gordon, D. J. Fox, K. Raghavachari, and L. A. Curtiss, *J. Chem. Phys.* **90**, 5622 (1989); L. A. Curtiss, C. Jones, G. W. Trucks, K. Raghavachari, and J. A. Pople, *ibid.* **93**, 2537 (1990).
- <sup>14</sup>L. A. Curtiss, K. Raghavachari, G. W. Trucks, and J. A. Pople, *J. Chem. Phys.* **94**, 7221 (1991).
- <sup>15</sup>L. A. Curtiss, K. Raghavachari, P. C. Redfern, and J. A. Pople, *J. Chem. Phys.* **109**, 7764 (1998).
- <sup>16</sup>C. W. Bauschlicher, Jr. and H. Partridge, *J. Chem. Phys.* **103**, 1788 (1995); L. A. Curtiss, P. C. Redfern, K. Raghavachari, V. Rassolov, and J. A. Pople, *ibid.* **110**, 4703 (1999); L. A. Curtiss, P. C. Redfern, K. Raghavachari, and J. A. Pople, *ibid.* **114**, 108 (2001); D. J. Henry, M. B. Sullivan, and L. Radom, *ibid.* **118**, 4849 (2003).
- <sup>17</sup>J. M. L. Martin and D. G. Oliveira, *J. Chem. Phys.* **111**, 1843 (1999).
- <sup>18</sup>A. Karton, E. Rabinovich, J. M. L. Martin, and B. Ruscic, *J. Chem. Phys.* **125**, 144108 (2006).
- <sup>19</sup>J. W. Ochterski, G. A. Petersson, and J. A. Montgomery, Jr., *J. Chem. Phys.* **104**, 2598 (1996); J. A. Montgomery, Jr., J. W. Ochterski, and G. A. Petersson, *ibid.* **101**, 5900 (1994).
- <sup>20</sup>T. J. Lee, J. E. Rice, G. E. Scuseria, and H. F. Schaefer III, *Theor. Chim. Acta* **75**, 81 (1989); T. J. Lee and P. R. Taylor, *ibid.* **23**, 199 (1989).
- <sup>21</sup>C. L. Janssen and I. M. B. Nielsen, *Chem. Phys. Lett.* **290**, 423 (1998).
- <sup>22</sup>D. Feller and J. A. Sordo, *J. Chem. Phys.* **112**, 5604 (2000).
- <sup>23</sup>D. Feller and J. A. Sordo, *J. Chem. Phys.* **113**, 485 (2000).
- <sup>24</sup>J. M. L. Martin, *Chem. Phys. Lett.* **259**, 669 (1996).
- <sup>25</sup>W. Klopper, K. L. Bak, P. Jørgensen, J. Olsen, and T. Helgaker, *J. Phys. B* **32**, R103 (1999).
- <sup>26</sup>B. A. Flowers, P. G. Szalay, J. F. Stanton, M. Kállay, J. Gauss, and A. G. Császár, *J. Phys. Chem. A* **108**, 3195 (2004).
- <sup>27</sup>T. H. Dunning, Jr., *J. Chem. Phys.* **90**, 1007 (1989); R. A. Kendall, T. H.

- Dunning, Jr., and R. J. Harrison, *ibid.* **96**, 6796 (1992); D. E. Woon, K. A. Peterson, and T. H. Dunning, Jr., (unpublished); A. K. Wilson, T. van Mourik, and T. H. Dunning, Jr., *J. Mol. Struct.: THEOCHEM* **388**, 339 (1996).
- <sup>28</sup> D. Feller and K. A. Peterson, *J. Chem. Phys.* **110**, 8384 (1999).
- <sup>29</sup> J. L. Dunham, *Phys. Rev.* **41**, 713 (1932).
- <sup>30</sup> D. Feller, *J. Chem. Phys.* **96**, 6104 (1992); D. Feller, *ibid.* **98**, 7059 (1993).
- <sup>31</sup> K. A. Peterson, D. E. Woon, and T. H. Dunning, Jr., *J. Chem. Phys.* **100**, 7410 (1994).
- <sup>32</sup> J. M. L. Martin and T. J. Lee, *Chem. Phys. Lett.* **258**, 136 (1996); T. Helgaker, W. Klopper, H. Koch, and J. Nago, *J. Chem. Phys.* **106**, 9639 (1997); A. Halkier, T. Helgaker, P. Jorgensen, W. Klopper, H. Koch, J. Olsen, and A. K. Wilson, *Chem. Phys. Lett.* **286**, 243 (1998).
- <sup>33</sup> D. W. Schwenke, *J. Chem. Phys.* **122**, 014107 (2005).
- <sup>34</sup> D. E. Woon and T. H. Dunning, Jr., *J. Chem. Phys.* **103**, 4572 (1995); A. K. Wilson and T. H. Dunning, Jr. (unpublished).
- <sup>35</sup> M. Douglas and N. M. Kroll, *Ann. Phys. (N.Y.)* **82**, 89 (1974); G. Jansen and B. A. Hess, *Phys. Rev. A* **39**, 6016 (1989).
- <sup>36</sup> W. A. de Jong, R. J. Harrison, and D. A. Dixon, *J. Chem. Phys.* **114**, 48 (2001).
- <sup>37</sup> T. H. Dunning, Jr., K. A. Peterson, and A. K. Wilson, *J. Chem. Phys.* **114**, 9244 (2001).
- <sup>38</sup> N. B. Balabanov and K. A. Peterson, *J. Chem. Phys.* **123**, 064107 (2005).
- <sup>39</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).
- <sup>40</sup> R. D. Amos, A. Bernhardsson, A. Berning *et al.*, MOLPRO 2002.6, a package of *ab initio* programs designed by H.-J. Werner and P. J. Knowles, Universität Stuttgart, Stuttgart, Germany, Cardiff University, Cardiff, United Kingdom, 2002.
- <sup>41</sup> H.-J. Werner, P. J. Knowles, R. Lindh, <http://www.molpro.net>; MOLPRO 2006.1, a package of *ab initio* programs, Universität Stuttgart, Stuttgart, Germany, Cardiff University, Cardiff, United Kingdom, 2006.
- <sup>42</sup> T. D. Crawford, C. D. Sherrill, E. F. Valeev, J. T. Fermann, R. A. King, M. L. Leininger, S. T. Brown, C. L. Janssen, E. T. Seidl, J. P. Kenny, and W. D. Allen, *PSI VERSION 3.2*, 2005.
- <sup>43</sup> J. Noga and R. J. Bartlett, *J. Chem. Phys.* **86**, 7041 (1987); G. E. Scuseria and H. F. Schaefer III, *Chem. Phys. Lett.* **152**, 382 (1988); J. D. Watts and R. J. Bartlett, *J. Chem. Phys.* **93**, 6104 (1990).
- <sup>44</sup> S. A. Kucharski and R. J. Bartlett, *Theor. Chim. Acta* **80**, 387 (1991); S. A. Kucharski and R. J. Bartlett, *J. Chem. Phys.* **97**, 4282 (1992); N. Oliphant and L. J. Adamowicz, *ibid.* **94**, 1229 (1991).
- <sup>45</sup> T. Yanai, M. Kamiya, Y. Kawashima *et al.*, *UTCHEM 2004  $\beta$* .
- <sup>46</sup> M. Kállay, MRCC, a string-based quantum chemical program suite, Budapest University of Technology and Economics, 2001.
- <sup>47</sup> P. J. Knowles, C. Hampel, and H.-J. Werner, *J. Chem. Phys.* **99**, 5219 (1994); P. J. Knowles, C. Hampel, and H.-J. Werner, *ibid.* **112**, 3106 (2000); G. E. Scuseria, *Chem. Phys. Lett.* **176**, 27 (1991).
- <sup>48</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>49</sup> D. Z. Goodson, *J. Chem. Phys.* **116**, 6948 (2002); D. Z. Goodson, *Int. J. Quantum Chem.* **92**, 35 (2003).
- <sup>50</sup> K. P. Huber and G. Herzberg, *Molecular Spectra and Molecular Structure IV: Constants of Diatomic Molecules* (Van Nostrand Reinhold Co. Inc., New York, 1979).
- <sup>51</sup> A. Burcat and B. Ruscic, (Argonne National Laboratory, Report No. TAE 960, 2005 (unpublished)).
- <sup>52</sup> R. S. Urdahl, Y. Bao, and W. M. Jackson, *Chem. Phys. Lett.* **178**, 425 (1991).
- <sup>53</sup> M. W. Chase, Jr., NIST-JANAF Tables, 4th Ed. *J. Phys. Chem. Ref. Data, Mono.* **9**, Suppl. 1 (1998).
- <sup>54</sup> B. Ruscic, (interim private communication) [unpublished data based on Active Thermochemical Tables, ATcT Version 1.25 and the Core (Argonne) Thermochemical Network Version 1.056 2006].
- <sup>55</sup> B. Ruscic, R. E. Pinzon, M. L. Morton, G. v. Laszewski, S. Bittner, S. G. Nijssure, K. A. Amin, M. Minkoff, and A. F. Wagner, *J. Phys. Chem. A* **108**, 9979 (2004); B. Ruscic, R. E. Pinzon, G. v. Laszewski, D. Kodeboyina, A. Burcat, D. Leahy, D. Montoya, and A. F. Wagner, *J. Phys.: Conf. Ser.* **16**, 561 (2005).
- <sup>56</sup> L. Pollack, T. L. Windus, W. A. de Jong, and D. A. Dixon, *J. Phys. Chem. A* **109**, 6934 (2005).
- <sup>57</sup> Z. Gan and R. H. Harrison (unpublished).
- <sup>58</sup> C. E. Moore, *Atomic Energy Levels*, Natl. Stand. Ref. Data Ser., Natl. Bur. Stand. (U.S.) Circ. No. 35, (U.S. GPO, Washington, D. C., 1971).
- <sup>59</sup> B. Ruscic (private communication) [unpublished data based on Active Thermochemical Tables, ATcT Version 1.35 and the Core (Argonne) Thermochemical Network Version 1.062 2006].
- <sup>60</sup> D. A. Dixon, D. Feller, and K. A. Peterson, *J. Phys. Chem. A* **106**, 3191 (2002).
- <sup>61</sup> L. G. Schum and S. W. Benson, *J. Phys. Chem.* **87**, 3479 (1983); L. G. Schum and S. W. Benson, *Int. J. Chem. Kinet.* **15**, 323 (1983).
- <sup>62</sup> E. R. Fisher and P. B. Armentrout, *J. Phys. Chem.* **94**, 4396 (1990).
- <sup>63</sup> J. L. Holmes, F. P. Lossing, and P. M. Mayer, *J. Am. Chem. Soc.* **113**, 9723 (1991).
- <sup>64</sup> W. B. DeMore, S. P. Sander, D. M. Golden, R. F. Hampson, M. J. Kurylo, C. J. Howard, A. R. Ravishankara, C. E. Kolb, and M. J. Molina, Jet Propulsion Laboratory (JPL) Report No. 97-4, 1977 (unpublished).
- <sup>65</sup> M. Litorja and B. Ruscic, *J. Electron Spectrosc. Relat. Phenom.* **97**, 131 (1998).
- <sup>66</sup> S. P. Sander, R. R. Friedl, D. M. Golden *et al.*, Jet Propulsion Laboratory (JPL), Report No. 02-25, 2003 (unpublished).
- <sup>67</sup> T. M. Raymond, S. J. Blanksby, S. Kato, V. M. Bierbaum, G. E. Davico, R. L. Schwatz, W. C. Lineberger, and G. B. Ellison, *J. Phys. Chem. A* **106**, 9641 (2002).
- <sup>68</sup> B. Ruscic (private communication); [B. A. Flowers, P. G. Szalay, J. F. Stanton, M. Kállay, J. Gauss, and A. G. Császár, *J. Phys. Chem. A* **108**, 3195 (2004)].
- <sup>69</sup> B. Ruscic, R. E. Pinzon, M. L. Morton, N. K. Srinivasan, M.-C. Su, J. W. Sutherland, and J. V. Michael, *J. Phys. Chem. A* **110**, 6592 (2006).
- <sup>70</sup> K. A. Peterson and T. H. Dunning, Jr., *J. Chem. Phys.* **117**, 10548 (2002).
- <sup>71</sup> R. S. Grev, C. L. Janssen, and H. F. Schaefer III, *J. Chem. Phys.* **95**, 5128 (1991).
- <sup>72</sup> V. Barone, *J. Chem. Phys.* **120**, 3059 (2004).
- <sup>73</sup> R. Georgiadis, E. R. Fisher, and P. B. Armentrout, *J. Am. Chem. Soc.* **111**, 4251 (1989).
- <sup>74</sup> Y.-M. Chen, D. E. Clemmer, and P. B. Armentrout, *J. Chem. Phys.* **95**, 1228 (1991).
- <sup>75</sup> J. M. L. Martin, *J. Chem. Phys.* **108**, 2791 (1998).
- <sup>76</sup> D. Feller and K. A. Peterson, *J. Chem. Phys.* **108**, 154 (1998).
- <sup>77</sup> K. A. Peterson and T. H. Dunning, Jr., *J. Mol. Struct.: THEOCHEM* **400**, 93 (1997).
- <sup>78</sup> T. A. Ruden, T. Helgaker, P. Jørgensen, and J. Olsen, *J. Chem. Phys.* **121**, 5874 (2004).
- <sup>79</sup> M. Heckert, M. Kállay, and J. Gauss, *Mol. Phys.* **103**, 2109 (2005).
- <sup>80</sup> M. Heckert, M. Kállay, D. P. Tew, W. Klopper, and J. Gauss, *J. Chem. Phys.* **125**, 044108 (2006).
- <sup>81</sup> N. C. Craig, P. Groner, and D. C. McKean, *J. Phys. Chem. A* **110**, 7461 (2006).
- <sup>82</sup> J. Gauss, D. Cremer, and J. F. Stanton, *J. Phys. Chem. A* **104**, 1319 (2000).
- <sup>83</sup> N. C. Craig, D. Feller, P. Groner, H. Y. Hsin, D. C. McKean, and D. J. Nemchick, *J. Phys. Chem. A* (in press).