11954 *J. Phys. Chem. A* **2009,** *113,* 11954–11962

An Approach to Include the Effects of Diffuse Functions in Potential Energy Surface Calculations[†]

Xinchuan Huang, David W. Schwenke, and Timothy J. Lee*

NASA Ames Research Center, Moffett Field, California 94035-1000

Received: April 20, 2009; Revised Manuscript Received: July 20, 2009

A new approach is proposed and investigated for approximately including the effects of diffuse functions in one-particle basis sets when high accuracy is desired. The method is cost-effective for use in computing quartic force fields (QFFs), global potential energy surfaces (PESs), or other situations when a large part of the PES is needed. It is conservatively estimated that the use of this approximation leads to a computational savings of *a factor of five*, and it is argued that this could be significantly larger if input/output wait times are considered. It can be used when extrapolation to the one-particle basis set limit is performed, or it can be used simply to approximate the effect of diffuse functions for a larger basis set. The new approach is based on scaling the diffuse function effect for a smaller basis set to approximate the effect for a larger basis or an extrapolated energy in which larger basis set(s) are used. The scale factor is written as a function of the geometrical coordinates of the molecule and thus it includes a geometry dependence. We report results where the scale factor is a constant, includes through gradient terms, includes through second derivative terms, and includes through diagonal second and third derivative terms. The method has been tested in the calculation of accurate QFFs, equilibrium structures, and harmonic and fundamental vibrational frequencies for NH_2^-, OH^-, H_2O, and CH_3OH. It is found that including up through diagonal second derivative terms leads to reliable fundamental vibrational frequencies and is cost-effective. It is also concluded that the use of a 5Z-quality basis set is essential if high accuracy is desired for these properties, even with extrapolation to the one-particle basis set limit.

1. Introduction

The calculation of accurate quartic force fields (QFFs) and fundamental vibrational frequencies using coupled-cluster electronic structure methods blossomed in the early 1990s and continues to expand today. (For example, see refs 1−7 and references therein.) Recently, we proposed an approach for computing highly accurate QFFs in which so-called "small" effects were included in the final QFF.[8] In this approach, an initial reference structure is determined at a high level of theory, and this reference structure is used to define a grid of points from which the QFF is fit. Total energies are determined for the grid of points using basis set extrapolation to the one-particle basis set limit including diffuse functions, with other small effects either being included in the calculations directly or added as a correction to the basis set limit energies. Some of these other small effects include core-correlation, scalar relativity, higher-order electron correlation (beyond what was included in the base calculations used to extrapolate to the one-particle basis set limit), and diagonal Born−Oppenheimer corrections.

All of the small effects delineated above can be included as corrections; for example, for core-correlation, one computes electronic energies for a given basis with and without including core-correlation and adds this difference to the one-particle basis extrapolated energies (noting, of course, that the basis set used in the core-correlation steps is designed to treat core-correlation). However, including the effects of diffuse functions is not something that can be computed as a correction because their importance typically becomes smaller as the one-particle basis set becomes larger. In other words, the importance of adding diffuse functions is typically greater for a triple-ζ quality basis set than for a quintuple-ζ basis set. If it is deemed important to include the effects of diffuse functions, such as for a QFF calculation of a molecular anion or just for high accuracy of a neutral species, then one needs to use a series of basis sets that includes diffuse functions to extrapolate to the one-particle basis set limit. Furthermore, there is evidence to suggest that extrapolation to the one-particle basis set limit often behaves differently for a series of basis sets that include diffuse functions compared with a series of basis sets that do not include diffuse functions (e.g., see refs 9 and 10 and references therein). In principle, this is not a problem because, for example, Dunning has designed series of basis sets with[11] (augmented correlation consistent) and without[12] (correlation consistent) diffuse functions. However, basis sets with diffuse functions include a diffuse function for each and every shell, which can significantly increase the size of the one-particle basis for larger molecules. The present study was aimed at finding an alternative to using explicitly the augmented series for all calculations yet including the effects of diffuse functions in an accurate but economical fashion. As we discuss below, if one considers the uncertainty of the extrapolation, then the two types of basis sets give consistent results.

To investigate our new approach, we have included in the present study OH^- and NH_2^-, small molecular anions for which the need for diffuse functions is obvious, and H_2O and CH_3OH, where the need for diffuse functions is less important except for high accuracy. The Theoretical Approach is described in the next section, followed by the Results and Discussion. Our Conclusions are presented in the final section.

[†] Part of the "Walter Thiel Festschrift".

[*] Corresponding author. E-mail: Timothy.J.Lee@nasa.gov.

10.1021/jp9036364 CCC: $40.75

2. Theoretical Approach

Before we describe the approach we have investigated for the inclusion of the effects of diffuse basis functions, we briefly summarize the general approach we proposed in ref 8 for computing QFFs. A reference structure, R_0, designed to be close to the ultimate equilibrium geometry, is determined, usually by optimizing the structure with the largest basis set to be used and explicitly including effects that are important for the structure, such as core correlation. A uniform grid is set up about the reference structure, and electronic energies are evaluated at each point on the grid for all basis sets. Extrapolation to the one-particle basis set limit is achieved using a three-point[13,14] or two-point[15] extrapolation formula. Other effects, such as core-correlation or scalar relativitistic effects, can be included explicitly in each of the electronic energy calculations, or they can be added as a correction to the extrapolated energy. Some effects, such as higher-order correlation, will always be included as a correction. As previously indicated, the effect of including diffuse functions in the one-particle basis set cannot be included as a correction, and thus diffuse functions must always be explicitly included or their effect ignored. In some situations, it is necessary to include diffuse functions explicitly because of their importance. Some examples include computing a QFF for a small molecular anion or for a neutral molecule that possesses an internal motion in which diffuse functions are necessary to describe it properly, such as the hindered rotation in methanol. The present study is aimed at these instances where diffuse functions are potentially important but explicitly including them in all basis sets is prohibitively expensive.

In the present work we use a series of one-particle basis sets designed to enable extrapolation to the one-particle basis set limit. This series is denoted TZ, QZ, and 5Z when diffuse functions are not included in the one-particle basis set and ATZ, AQZ, and A5Z when diffuse are explicitly included. The basis sets without diffuse functions are Dunning's correlation consistent valence basis sets, cc-pVXZ,[12] whereas those including diffuse functions are Dunning's augmented sets, aug-cc-pVXZ.[11] Because the purpose of the present study was to examine the proposed procedure for including the effects of diffuse functions in accurate calculations, in most cases, so-called small corrections were not included, with a few exceptions. For CH_3OH, scalar relativistic effects, via the Douglas−Kroll−Hess method,[16,17] were included as a correction with the TZ basis set. (See ref 8 for details.) For calculations associated with the scalar relativity correction only, the DK contraction coefficients were used; otherwise, the usual contraction coefficients were used. Also, for CH_3OH only, the effects of core-correlation were included as a correction using the Martin−Taylor (MT) basis set.[18] Only spherical harmonic components of the basis functions were included in the calculations. MOLPRO 2002.6[19] was used for all electronic structure calculations, whereas SPECTRO[20] was used in evaluating fundamental vibrational frequencies via standard second-order perturbation theory.[21]

We have investigated several techniques for approximately including the effects of diffuse functions in the final energies, but they all follow the same general strategy of scaling the diffuse function contribution from a relatively small basis set calculation to what is expected for the electronic energy obtained by extrapolation to the one-particle basis set limit. A general formula describing our approach is given in eq 1

$$E_{\mathrm{ABIG}}(\mathrm{R}) = E_{\mathrm{BIG}}(\mathrm{R}) + S_{\mathrm{G}}^{\mathrm{M}}(\mathrm{R})\cdot\Delta_{\mathrm{SMALL}}(\mathrm{R}) \qquad (1)$$

where E_{ABIG} is the electronic energy obtained by a large one-particle basis set with diffuse functions, E_{BIG} is the electronic energy obtained by a large one-particle basis set without diffuse functions, $S_{\mathrm{G}}^{\mathrm{M}}$ is a scale factor described below, and Δ_{SMALL} is the diffuse function contribution obtained using a more economical one-particle basis set. In the present work, we take E_{BIG} to be the result of extrapolating the TZ, QZ, and 5Z basis set results to the one-particle basis set limit[13,14] and Δ_{SMALL} to be the diffuse function contribution computed using the triple-ζ quality basis sets

$$\Delta_{\mathrm{SMALL}}(\mathrm{R}) = E_{\mathrm{ATZ}}(\mathrm{R}) - E_{\mathrm{TZ}}(\mathrm{R}) \equiv \Delta_{\mathrm{TZ}}(\mathrm{R}) \qquad (2)$$

The crux of our method is in choosing $S_{\mathrm{G}}^{\mathrm{M}}$.

We note that the approach proposed here and defined in eq 1 can also be performed simply for a larger basis set and does not require the use of energies extrapolated to the one-particle basis set limit. For example, eq 1 could be written as

$$E_{\mathrm{A5Z}}(\mathrm{R}) = E_{\mathrm{5Z}}(\mathrm{R}) + S_{\mathrm{G}}^{\mathrm{M}}(\mathrm{R})\cdot\Delta_{\mathrm{TZ}}(\mathrm{R}) \qquad (3)$$

where only E_{TZ}, E_{5Z}, and E_{ATZ} energies are used. In the present study, however, we employ only extrapolated energies.

Furthermore, E_{BIG} and Δ_{SMALL} could be evaluated using energies that differed by more than just the one-electron basis set used. For example, E_{BIG} could include the effects of core-correlation, scalar relativity, or other small corrections, whereas Δ_{SMALL} neglected these contributions.

There are several options for choosing the scale factor $S_{\mathrm{G}}^{\mathrm{M}}$ because we have the choice of both what quality of data to use to model the scale factor (the M superscript) and how to represent the geometry dependence (the G subscript). We will model our scale factor at geometry R by using

$$S^{\mathrm{M}}(\mathrm{R}) = [E_{\mathrm{AM}}(\mathrm{R}) - E_{\mathrm{M}}(\mathrm{R})]/\Delta_{\mathrm{TZ}}(\mathrm{R}) \qquad (4)$$

with three choices for M, namely, 2-pt extrapolation of the TZ and QZ basis set results, 2-pt extrapolation of the QZ and 5Z basis set results, and 3-pt extrapolation of the TZ, QZ, and 5Z basis set results. We will denote these choices by TQ, Q5, and TQ5, respectively. Several choices for the geometry dependence have been investigated, with the simplest being a constant. However, because we are interested here in using this technique for computing QFFs or even global PESs, a constant is probably not sufficient, so we investigated the geometric dependence using a series expansion for the scale factor. For the present work, where we are obtaining our results from QFF, a Taylor series expansion about R_0 is sufficient. For global PESs, more sophisticated functional forms will be required, but there is no compelling reason at this point to suspect that functional forms different from that used to fit the energies themselves will be required. The expansion coefficients were determined either by finite difference differentiation of S^{M} or by least-squares fitting to

$$S^{\mathrm{M}} = S_0^{\mathrm{M}} + \sum_{\mathrm{i}} S_i^{\mathrm{M}}\Delta q_i + \sum_{i\geq j} S_{ij}^{\mathrm{M}}\Delta q_i\Delta q_j + \sum_{i\geq j\geq k} S_{ijk}^{\mathrm{M}}\Delta q_i\Delta q_j\Delta q_k + \sum_{i\geq j\geq k\geq l} S_{ijkl}^{\mathrm{M}}\Delta q_i\Delta q_j\Delta q_k\Delta q_l \qquad (5)$$

where Δq_i is the displacement of coordinate i from R_0. The different choices for geometry dependence, G, refer to truncating eq 5 to a specific number of terms; that is, $G = 0$ means use only the constant term, S_0^{M}, $G = 1$ means use only up to S_i^{M}, and so on. Special other cases are $G = \bar{2}$, which means remove the $i \neq j$ terms from $G = 2$, and $G = \bar{3}$, which means that the diagonal cubic terms are added to $G = \bar{2}$.

Examination of eqs 1−5 reveals that only the constant term needs to be included to exactly reproduce the E_{ABIG} energy at the reference structure, R_0, whereas inclusion of the gradient terms will allow the E_{ABIG} energy for the single displaced geometries (e.g., $+\Delta q_i$) to be exactly reproduced, and so on.

Contained in Table 1 are the values of the scaling factor expansion coefficients used in the present work. In all cases, we used symmetry internal coordinates to obtain the geometrical dependence; see footnotes to Table 1 for details of the symmetry internal coordinates. (See ref 22 for the definition of symmetry internal coordinates for CH_3OH.) Therefore, certain constants are exactly zero because of symmetry and thus are not given in Table 1.

The coefficients in Table 1 are presented for completeness, but examination of the values is also instructive. For example, OH^- and NH_2^- behave somewhat differently than H_2O and CH_3OH. The coefficient S_0^{M} changes sign upon going from M = TQ to Q5 or TQ5 for OH^- and for NH_2^- but not for H_2O or CH_3OH. This seems to suggest that the 5Z basis set is important for the small molecular anions. Many of the higher-order constants for OH^- and NH_2^- either change sign or change significantly in magnitude for this same comparison, whereas for H_2O and CH_3OH, this effect is less pronounced. This suggests that at least for the molecules studied here the use of the 5Z basis set in approximating the effect of diffuse functions is important. Overall, the coefficients in Table 1 show that there is a definite geometric dependence to the diffuse function scale factor and that the details are basis-set-dependent, but they also show that the basis set dependence is greater for small molecular anions than for small neutral molecules. In the following Results and Discussion section, we present results and discuss the impact of scaling the contribution from diffuse functions on equilibrium structures and harmonic and fundamental vibrational frequencies.

TABLE 1: Scaling Factors for Diffuse Function Effect for OH^-, H_2O, NH_2^-, and CH_3OH[a]

	M = TQ	M = Q5	M = TQ5
		$OH^{b,c}$	
S_0	0.2043995	−0.0978264	−0.1292385
S_r	0.0542461	0.0869050	0.0914083
S_{rr}	−0.1102239	−0.0427222	−0.0256436
S_{rrr}	0.3124994	−0.1867126	−0.2954506
S_{rrrr}	−0.6338346	−0.0668624	0.2263073
		$H_2O^{b,d}$	
S_0	−0.0808930	−0.1388400	−0.1124170
S_R	0.0810018	−0.0120770	−0.0314152
S_θ	0.0149648	−0.0186153	0.0256765
S_{RR}	−0.0956952	0.0543400	0.0870679
S_{rr}	−0.0910802	0.0319149	0.0591618
$S_{\theta\theta}$	−0.0087758	0.0150253	0.0199855
$S_{R\theta}$	0.0215109	0.0083935	0.0050747
S_{RRR}	−0.0087565	−0.0041202	−0.0464424
$S_{\theta\theta\theta}$	0.1792962	0.0301800	0.0507315
		$NH_2^{-\,e,f}$	
S_0	0.2625284	−0.0613139	−0.0989932
S_R	0.0209561	0.0353970	0.0374950
S_θ	0.0194958	−0.0063018	−0.0120361
S_{RR}	−0.0729322	0.0445371	0.0434761
S_{rr}	−0.0524791	−0.0274035	−0.0206383
$S_{\theta\theta}$	−0.0077832	0.0179672	0.0125481
$S_{R\theta}$	0.0327913	0.0080843	0.0020971
S_{RRR}	0.0115749	−0.0329121	−0.0420823
$S_{\theta\theta\theta}$	0.0274360	0.0016543	−0.0043208
		$CH_3OH^{e,g}$	
S_0	−0.0951171	−0.1132149	−0.0784867
S_1	0.0577553	−0.0394693	0.0605381
S_2	0.0446022	−0.0162777	0.0297649
S_3	0.0210919	0.0061547	−0.0024908
S_4	−0.0125655	−0.0000878	−0.0027963
S_5	−0.0023270	−0.0151012	0.1752281
S_6	−0.0152485	0.0236224	−0.0317718
S_7	−0.0106142	0.0060073	−0.0096351
S_8	0.0107507	−0.0086968	0.0128831
S_{11}	−0.1914436	0.1183639	0.1857589
S_{22}	0.3503178	0.1623448	0.2750400
S_{33}	0.1189931	−0.0237694	−0.0559013
S_{44}	0.0518328	−0.0103309	−0.0243229
S_{55}	0.0298662	−0.0341393	−0.0477244
S_{66}	0.0657316	−0.0091969	−0.0261757
S_{77}	0.0362022	0.0039378	−0.0036494
S_{88}	0.0041270	−0.0168356	−0.0210873
S_{99}	−0.0113872	0.0204745	0.0271155
S_{10-10}	0.0260667	0.0006743	−0.0052121
S_{11-11}	0.0269714	−0.0215684	−0.0320221
S_{12-12}	−0.0056628	−0.0009293	0.0001961

[a] See eqs 1−5 and text for details. [b] Least squares fit using quartic expansion. [c] $r_{\mathrm{OH}}(\mathrm{Min}) = 0.96382172$ Å, $\Delta r = r_{\mathrm{OH}} - r_{\mathrm{OH}}(\mathrm{Min})$ in angstroms. [d] Symmetry internal coordinates were used: $\Delta R = r_{\mathrm{OH1}} + r_{\mathrm{OH2}} - 2r_{\mathrm{OH}}(\mathrm{min})$ in angstroms; $\Delta r = |r_{\mathrm{OH1}} - r_{\mathrm{OH2}}|$ in angstroms; $\Delta\theta = \theta - \theta(\mathrm{min})$ in radians. Minimum structures used for scaling factor expansion are: $r_{\mathrm{OH}}(\mathrm{min}) = 0.9575233$ Å, $\theta(\mathrm{min}) = 1.8223030$ rad. [e] All values were computed by finite difference formula. [f] Same as H_2O except $r_{\mathrm{NH}}(\mathrm{min}) = 1.0267018$ Å, $\theta(\mathrm{min}) = 1.7832078$ rad. [g] Symmetry coordinates used for CH_3OH are taken from ref 22. Coordinates 1−8 are A' sym and 9−12 are A'' sym. We changed the torsional coordinate from single dihedral angle $\Delta\tau_1$ to averaged dihedral angle $\Delta\tau = (\Delta\tau_1 + \Delta\tau_2 + \Delta\tau_3)/3$, so we can have stationary points at exact $n(\pi/3)$, $n = -\infty, ..., -1, 0, 1, 2, ...$ Minimum structures used for scaling factor expansion are: $r_{\mathrm{OH}} = 0.957017$ Å, $r_{\mathrm{O\text{-}C}} = 1.417785$ Å, $r_{\mathrm{C\text{-}H1}} = 1.085605$ Å, $r_{\mathrm{C\text{-}H2/H3}} = 1.091355$ Å, $\theta_{\mathrm{C\text{-}O\text{-}H}} = 1.887093$ rad, $\theta_{\mathrm{O\text{-}C\text{-}H1}} = 1.863525$ rad, $\theta_{\mathrm{O\text{-}C\text{-}H2/H3}} = 1.953684$ rad, $\tau_2 = -\tau_3 = 1.0711778$ rad.

3. Results and Discussion

A. OH^- and NH_2^-. Table 2 contains results for OH^-, whereas those for NH_2^- are presented in Table 3. The top three sections of the tables refer to results with different choices of M and then G. The bottom section of results contained in Tables 2 and 3 is defined and discussed below. For convenience, recall that the different choices of modeling the scale factors (eq 4) are denoted M = TQ, Q5, and TQ5. In each section, results are included for which no diffuse functions are included or modeled, the scale factor is approximated as the single constant, $G = 0$, the constant plus gradient terms, $G = 1$, the constant plus gradient plus full second derivative terms, $G = 2$, the constant plus gradient plus diagonal second derivative terms, $G = \bar{2}$, and finally the constant plus gradient plus diagonal second and third derivative terms, $G = \bar{3}$. Examination of the TQ5 results in Table 2 shows that using only the constant term has minimal effect on the minimum geometry and improves the harmonic and fundamental vibrational frequency only marginally, whereas including the gradient term marks a significant improvement in all properties. Now the minimum is given essentially exactly, and the error in the harmonic frequency is reduced by an order of magnitude. The improvement of the fundamental frequency is much less spectacular but still very significant. The inclusion of the second-derivative term (for OH^-, there are no off-diagonal second derivative terms) polishes off the harmonic frequency, giving essentially the exact results that one would expect, whereas the error for the fundamental is almost halved compared with the inclusion of only the gradient term. To obtain

TABLE 2: Equilibrium Structures and Harmonic and Fundamental Vibrational Frequencies for OH^- Obtained Using Various Approximations for Including the Effects of Diffuse Functions[a]

	R(O−H)/Å	ω/cm^{-1}	ν/cm^{-1}
	M = TQ5		
no scaling	-2.5×10^{-3}	42.09	33.52
constant ($G = 0$)	-2.6×10^{-3}	32.90	22.70
gradient ($G = 1$)	-1.1×10^{-6}	−2.85	−7.57
2nd-order ($G = 2$)	2.0×10^{-8}	−0.07	−4.27
3rd-order ($G = 3$)	0.0	0.00	1.10
exact/quartic fit	0.964547	3750.35	3640.97
	M = TQ		
no scaling	-1.6×10^{-3}	−5.49	−10.82
constant ($G = 0$)	-1.6×10^{-3}	9.03	6.22
gradient ($G = 1$)	-5.0×10^{-7}	−11.75	−11.45
2nd-order ($G = 2$)	3.0×10^{-6}	−0.04	2.45
3rd-order ($G = 3$)	0.0	0.0	−1.96
exact/quartic fit	0.963606	3797.93	3685.31
M = TQ − M = TQZ	*-9.4×10^{-4}*	*47.58*	*44.34*
	M = Q5		
no scaling	-2.4×10^{-3}	36.22	28.38
constant ($G = 0$)	-2.5×10^{-3}	29.27	20.20
gradient ($G = 1$)	2.0×10^{-6}	−4.65	8.53
2nd-order ($G = 2$)	2.0×10^{-6}	−0.01	−3.03
3rd-order ($G = 3$)	0.0	0.0	0.32
exact/quartic fit	0.964427	3756.22	3646.11
M = Q5 − M = TQ5	*-1.2×10^{-4}*	*5.87*	*5.14*
	R(O−H)/Å	ω/cm^{-1}	ν/cm^{-1}
VTQZ	-2.9×10^{-3}	6.28	−4.49
VQ5Z	-2.8×10^{-3}	39.29	30.15
VTQ5Z	-2.5×10^{-3}	42.09	33.52
AVTQZ	-1.2×10^{-3}	11.55	6.34
AVQ5Z	-3.4×10^{-4}	3.01	1.70
AVTQ5Z	0.964547	3750.35	3640.97

[a] See the text for details.

wavenumber accuracy for the fundamental frequency, the cubic correction is required.

Before discussing the variation with G of the TQ results in Table 2, we first note that the differences presented for the equilibrium bond distance and the harmonic and fundamental vibrational frequencies are with respect to what is obtained with eq 5 using the full quartic expansion for the scaling factor with M = TQ consistently. The last row in this section is labeled "M = TQ − M = TQ5", and this is the difference between using M = TQ and M = TQ5 using a converged geometrical representation. The first observation, then, is to note that the TQ and TQ5 difference is rather large, indicating that the use of the 5Z basis in determining the diffuse function scale factor is important for OH^-. This is consistent with our earlier observations of the S_0 and derivative terms contained in Table 1. It was hoped that use of the augmented 5Z basis could be avoided, which is why we investigated the M = TQ approach, but it is clear from the large TQ and TQ5 differences and the odd convergence behavior upon going from $G = 0$ to 1 to 2, that this approach does not work well for OH^-.

Examination of the M = Q5 results in Table 2 shows behavior similar to that of the M = TQ5 results. The use of only the constant term shows only small improvement, whereas the inclusion of the gradient contribution exhibits marked improvement, and the inclusion of the second derivative term shows further improvement. However, whereas the convergence of the Q5 approach with respect to geometrical flexibility may be similar to that of the TQ5 method, the differences between M = Q5 and M = TQ5 are not insignificant when high accuracy is desired. That is, there is a 10^{-4} Å difference for the equilibrium bond distance and more than 5 cm^{-1} change for both the harmonic and fundamental vibrational frequencies.

The last section of Table 2, starting with the row labeled "VTQZ" and ending with the row labeled "AVQ5Z", simply contains the results obtained without the use of the diffuse function scale factors. So, for example, AVTQ5Z means that the QFF was obtained using ATZ, AQZ, and A5Z energies extrapolated to the one-particle basis set limit with the 3-pt formula, whereas VTQZ means that the QFF was obtained using TZ and QZ energies extrapolated to the one-particle basis set limit with the 2-pt formula. Examination of these results clearly shows that diffuse functions are important for an accurate description of the OH^- PES.

Considering the results for NH_2^- (Table 3), we see behavior similar to that found for OH^-, with one exception. In this case, the full second derivative terms include off diagonal elements, and thus we have the option of including the full second derivative scale factors or just the diagonal second derivative terms. The M = TQ5 results contained in Table 3 show that neglecting the off-diagonal second derivative terms has only a small effect on the results, suggesting that including up through the diagonal second derivative terms only is a reasonable compromise between accuracy and computational costs. In all other respects, the M = TQ5 results in Table 3 exhibit a convergence pattern similar to that observed for OH^- in Table 2. That is, the use of only S_0 shows minimal improvement, whereas the inclusion of gradient terms shows marked improvement, and even better agreement is found when the diagonal second derivative terms are included in the diffuse function scale factor. Using up to second derivative terms for the diffuse

TABLE 3: Equilibrium Structures And Harmonic and Fundamental Vibrational Frequencies for NH_2^- Obtained Using Various Approximations for Including the Effects of Diffuse Functions[a]

	R(N−H)/Å	∠HNH/rad	ω_1/cm^{-1}	ω_2/cm^{-1}	ω_3/cm^{-1}	ν_1/cm^{-1}	ν_2/cm^{-1}	ν_3/cm^{-1}
			M = TQ5					
no scaling	-1.9×10^{-3}	1.1×10^{-2}	32.33	−4.00	25.25	27.95	−10.90	22.79
constant ($G = 0$)	-1.4×10^{-3}	4.6×10^{-3}	22.61	4.03	12.32	17.96	−2.95	9.00
gradient ($G = 1$)	-8.5×10^{-6}	2.2×10^{-5}	8.94	5.71	−4.33	5.65	−1.21	−6.51
2nd-order ($G = 2$)	5.6×10^{-8}	-6.0×10^{-8}	−0.10	0.04	−0.13	−5.00	−7.29	−3.93
2nd-order diag ($G = \bar{2}$)	2.9×10^{-7}	-2.0×10^{-6}	−0.09	−0.04	−0.13	−5.01	−7.37	−3.98
3rd-order diag ($G = \bar{3}$)	2.6×10^{-7}	-2.0×10^{-6}	−0.02	−0.06	−0.13	−2.80	−7.15	−3.51
exact/quartic fit	1.0281332	1.7809393	3297.49	1492.00	3383.46	3123.85	1448.55	3191.36
			M = TQ					
no scaling	1.0×10^{-3}	-2.4×10^{-2}	25.85	−24.91	41.47	−36.29	33.79	−45.33
constant ($G = 0$)	-4.6×10^{-4}	-5.4×10^{-3}	−8.91	2.64	−7.18	−9.33	11.58	−8.65
gradient ($G = 1$)	-5.7×10^{-5}	2.1×10^{-4}	−14.28	−3.91	−10.27	−14.41	4.65	−11.42
2nd-order ($G = 2$)	7.8×10^{-8}	-2.7×10^{-5}	−0.13	1.42	−0.12	3.63	10.36	3.30
2nd-order diag ($G = \bar{2}$)	-3.5×10^{-5}	3.4×10^{-5}	0.38	0.11	0.29	3.93	9.28	3.12
3rd-order diag ($G = \bar{3}$)	-3.7×10^{-5}	6.2×10^{-5}	0.41	−1.18	0.31	3.53	6.76	3.46
exact/quartic fit	1.0251672	1.8163071	3364.59	1463.07	3450.18	3188.02	1403.86	3259.50
M = TQ − M = TQ5	*-3.0×10^{-3}*	*3.5×10^{-2}*	*67.10*	*−28.92*	*66.72*	*64.17*	*−44.69*	*68.14*
			M = Q5					
no scaling	-1.6×10^{-3}	7.0×10^{-3}	23.17	−1.52	16.86	19.43	−5.91	14.47
constant ($G = 0$)	-1.2×10^{-3}	2.8×10^{-3}	17.15	3.48	8.85	13.22	−0.95	5.92
gradient ($G = 1$)	-2.9×10^{-6}	-1.4×10^{-5}	4.64	3.80	−5.71	1.94	−0.65	−7.63
2nd-order ($G = 2$)	-2.0×10^{-11}	1.8×10^{-7}	−0.01	0.02	−0.03	−3.28	−4.70	−2.39
2nd-order diag ($G = \bar{2}$)	-3.0×10^{-7}	-4.9×10^{-6}	0.03	−0.31	−0.03	−3.29	−4.99	−2.54
3rd-order diag ($G = \bar{3}$)	-3.3×10^{-7}	-4.9×10^{-6}	0.08	−0.32	−0.03	−1.56	−4.97	−2.10
exact/quartic fit	1.0277672	1.7852822	3306.65	1489.51	3391.84	3132.29	1443.56	3199.69
M = Q5 − M = TQ5	*-3.6×10^{-4}*	*4.3×10^{-3}*	*9.16*	*−2.48*	*8.39*	*8.44*	*−4.99*	*8.33*
	R(N−H)/Å	∠HNH/rad	ω_1/cm^{-1}	ω_2/cm^{-1}	ω_3/cm^{-1}	ν_1/cm^{-1}	ν_2/cm^{-1}	ν_3/cm^{-1}
VTQZ	4.3×10^{-4}	-2.2×10^{-2}	−30.42	19.37	−37.01	−32.09	27.78	−42.14
VQ5Z	-1.8×10^{-3}	7.3×10^{-3}	24.62	−1.45	18.59	20.45	−5.09	15.56
VTQ5Z	-1.9×10^{-3}	1.1×10^{-2}	32.33	−4.01	25.25	27.95	−10.90	22.79
AVTQZ	-7.1×10^{-4}	1.0×10^{-3}	4.35	−1.00	5.16	2.88	1.67	2.89
AVQ5Z	-2.2×10^{-4}	3.1×10^{-4}	1.47	−0.03	1.71	1.06	0.35	1.17
AVTQ5Z	1.0281332	1.7809393	3297.49	1492.00	3383.46	3123.85	1448.55	3191.36

[a] See the text for details.

function scale factor again leads to very small errors for the equilibrium structure and harmonic frequencies, but the fundamental frequencies exhibit somewhat larger errors. The error for the stretching fundamentals is similar to OH^- (~4 cm^{-1}), but the error for the bend is somewhat larger, being slightly over 7 cm^{-1}. In contrast with the OH^- case, however, the addition of diagonal cubic terms has only a small effect on the fundamental frequencies. This is almost certainly due to the richer coupling possibilities in a triatomic molecule that we neglect by only including diagonal geometry dependence. Clearly, if high accuracy is desired for fundamental vibrational frequencies for small molecular anions, it may be necessary to simply include diffuse functions in all calculations. However, if errors in the few inverse centimeters range are acceptable, then the use of the diffuse function scale factor proposed and investigated herein seems to be a reasonable compromise.

The M = TQ results for NH_2^- again show a large deviation from the M = TQ5 values and erratic convergence behavior because higher-order derivative terms are included in the diffuse function scale factor. It would appear that for small molecular anions, the use of the augmented 5Z basis set, at least for the scale factor, is a necessity.

Convergence of the equilibrium structure and harmonic and fundamental frequencies with respect to the inclusion of higher-order derivative terms for the M = Q5 results is similar to that found with the M = TQ5 approach, but the absolute differences are also similar to what was observed for OH^-, namely, the M = Q5 and M = TQ5 difference is significant if high accuracy is desired, being more than 8 cm^{-1} for both the symmetric and antisymmetric N−H stretch fundamentals.

The last section of Table 3 shows the importance of including diffuse functions in the one-particle basis for small molecular anions such as NH_2^- and is consistent with OH^-. In Figure 1, we show that extrapolation to the one-particle basis set limit behaves differently when using basis sets that include diffuse functions versus basis sets that do not include diffuse functions. Figure 1 refers specifically to NH_2^-, although similar plots for OH^- and H_2O exhibit the same pattern, albeit to a lesser extent (especially for H_2O). As noted in the Introduction, this is consistent with previous studies.[9,10]

It is instructive to try to understand why the augmented and standard basis sets extrapolate to different energies by trying to quantify the uncertainty introduced by the extrapolation. It is hard to determine the uncertainty quantitatively, so we will make the simple assumption that the uncertainty is the fraction δ of the difference between the best directly calculated value and the extrapolated value. We can then ask what value of δ is required for the two kinds of basis sets to agree. In other words, we solve

$$E_{TQ5} + \delta(E_{TQ5} - E_5) = E_{ATQ5} - \delta(E_{ATQ5} - E_{A5}) \tag{6}$$

for δ. For the example of NH_2^- shown in Figure 1, this results in $\delta = 0.18$. This value is small compared with unity, and thus we conclude that it is very likely that the two kinds of basis sets are

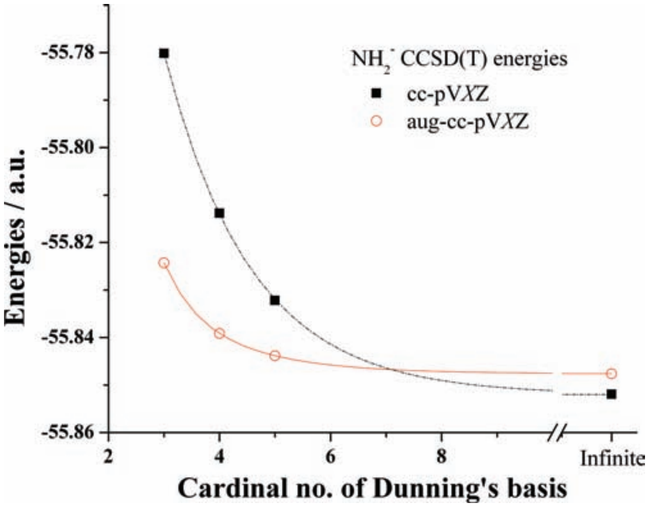

Figure 1. Plot of NH_2^- electronic energies using increasingly larger one-particle basis sets, including extrapolation to the one-particle basis set limit. The series with black squares does not contain diffuse functions, whereas the series with empty circles contains diffuse functions. See the text for further details.

converging to the same energy, but we do not accurately know the functional form to use in the extrapolation. In some sense, the difference between E_{TQ5} and E_{ATQ5} is a measure of the error in the extrapolation formula used. We note, however, that even though we do not know the exact functional form of the extrapolation formula, any error associated with the extrapolation is likely to cancel largely for a QFF or PES. In other words, the error will be similar for all grid points used to fit the QFF or PES. Finally, note that S_0 (Table 1) is negative, which results from E_{TQ5} being lower in energy than E_{ATQ5}, except for the TQ 2-pt extrapolations for OH^- and NH_2^-. This indicates that at least for the reference structure, E_{ATQ} is lower in energy than E_{TQ}, and this may also explain the erratic behavior found for the results from the TQ extrapolation for both OH^- and NH_2^-.

B. H_2O. Small molecular anions are a stringent test for the need of diffuse functions, but it is also of interest to investigate the effect for small neutral molecules that are composed of atoms with significant differences in their electronegativities, such as H_2O. In this case, the use of diffuse functions may be needed when high accuracy is desired. Presented in Table 4, then, are results for H_2O. (We note that the AVTQ5Z results contained in Table 4 do not agree with our recent work[8] because core-correlation has not been included in the present work but was previously.) Examining the M = TQ5 results first, it is interesting to note that the harmonic and fundamental vibrational frequencies actually appear better with only the constant term compared with the inclusion of the gradient terms. This is certainly a fortuitous cancellation of errors. The inclusion of up to the gradient terms seems to overcorrect for diffuse functions. However, the inclusion of the diagonal second derivative term reduces the error essentially to zero for all of the harmonic frequencies, whereas it reduces the error to essentially 1 cm^{-1} for the fundamental stretching frequencies and to half that for the bend fundamental. The inclusion of the diagonal cubic terms ($G = \bar{3}$) has essentially no effect on the equilibrium geometry and harmonic frequencies but does improve the fundamental frequencies somewhat. The main difference between the results for NH_2^- and H_2O is that the inclusion of diffuse functions is much less important for H_2O. Indeed, the effect on the fundamental vibrational frequencies, obtained by comparison of the AVTQ5Z and VTQ5Z results (i.e., diffuse functions not included), is always less than 2 cm^{-1}.

For NH_2^- and OH^-, the M = TQ results were found to be severely erratic, but for H_2O, the pattern for the M = TQ results is similar to the pattern for the M = TQ5 results, although the errors are somewhat larger. The comparison of the M = Q5 and M = TQ5 results shows little difference, although, unlike what was found for NH_2^- and OH^-, in this case, the former are slightly more accurate for the fundamental vibrational frequencies.

Examination of the last section of Table 4 shows that once the 5Z basis set has been included, either with or without diffuse functions, the fundamental vibrational frequencies determined from a QFF evaluated from an extrapolated set of energies are consistent to within ~2 cm^{-1}. This shows that the use of the 5Z basis set in the extrapolation procedure is more important for H_2O than the inclusion of diffuse functions. Nonetheless, if high accuracy is desired, then the comparison of the M = TQ5 and AVTQ5Z results for the equilibrium structure, harmonic vibrational frequencies, and fundamental vibrational frequencies of H_2O shows that approximating the effect of diffuse functions by inclusion up through the second-order diagonal terms is a cost-effective approach.

C. CH_3OH. As a final example, we include results for the methanol molecule in Table 5. For CH_3OH, only a limited subset of approaches has been investigated on the basis of our assessment of the results for H_2O, NH_2^-, and OH^- previously discussed. Specifically, we have investigated the use of a constant term, the inclusion of gradient terms, and finally the inclusion of up through diagonal second-order terms; a full second derivative correction has not been investigated. These corrections have been computed according to eq 1, except that E_{BIG} is without 5Z energies. That is, we do not yet have a full QFF computed with the 5Z basis set, so we have performed a 3-pt extrapolation using the DZ, TZ, and QZ basis sets, plus we have added corrections for scalar relativity and core-correlation on which the effect of diffuse functions is added. Whereas we have included corrections for scalar relativity and for core-correlation, the lack of a full QFF using the 5Z basis set limits the desired accuracy; results comparable to experiment will be reported in due course.[23] For purposes of deducing the quality of our proposed approach for approximating the effects of diffuse functions, the results in Table 5 are more than adequate.

Before discussing the results in Table 5, we note that the column labeled "no scaling" indicates the use of VDTQZ energies (i.e., 3-pt extrapolation using the DZ, TZ, and QZ basis sets) plus corrections for scalar relativity and core-correlation according to the procedure outlined in ref 8, or in other words, the effect of diffuse functions has been neglected. The column labeled M = TQ5, $G = \bar{2}$ indicates that the effect of diffuse functions has been approximated according to eq 1 with the one change noted in the previous paragraph, and this column represents the most accurate results presented herein for CH_3OH. The remaining columns in Table 5 should be self-explanatory.

Table 5 contains data for the equilibrium structure and fundamental vibrational frequencies for methanol. Examination of the results in the column labeled "M = TQ5, G = $\bar{2}$ − no scaling" gives an indication of the importance of diffuse functions, and we see not surprisingly that the importance of diffuse functions for methanol is similar to that found for H_2O. The largest effect on a fundamental vibrational frequency is ~2 cm^{-1}, although there are several modes for which the effect is less than 1 cm^{-1}. Examination of the results in the column labeled "M = TQ5, G = $\bar{2}$ − M = TQ5, G = 0" shows that the use of the constant term alone in eq 5 is not as reliable,

TABLE 4: Equilibrium Structures and Harmonic and Fundamental Vibrational Frequencies for H_2O Obtained Using Various Approximations for Including the Effects of Diffuse Functions[a]

	R(O−H)/Å	∠HOH/rad	ω_1/cm^{-1}	ω_2/cm^{-1}	ω_3/cm^{-1}	ν_1/cm^{-1}	ν_2/cm^{-1}	ν_3/cm^{-1}
			M = TQ5					
no scaling	3.7×10^{-4}	2.7×10^{-3}	−2.63	−2.28	−1.95	−1.57	−1.62	−1.44
constant ($G = 0$)	1.3×10^{-4}	1.5×10^{-3}	0.80	0.32	0.98	1.19	0.90	0.87
gradient ($G = 1$)	-1.2×10^{-6}	-2.6×10^{-6}	3.53	2.15	2.40	3.93	2.85	2.28
2nd-order ($G = 2$)	1.0×10^{-10}	-5.7×10^{-8}	0.00	0.04	−0.01	−0.96	0.45	−1.26
2nd-order diag ($G = \bar{2}$)	-2.0×10^{-8}	-4.7×10^{-7}	0.00	0.00	−0.01	−0.96	0.44	−1.28
3rd-order diag ($G = \bar{3}$)	-2.2×10^{-8}	-4.6×10^{-7}	0.01	−0.02	−0.01	−0.45	−0.06	−1.04
exact/quartic fit	0.9582033	1.8235386	3835.62	1649.32	3947.06	3658.27	1595.86	3756.53
			M = TQ					
no scaling	-2.3×10^{-4}	2.3×10^{-4}	0.89	−1.42	−0.38	1.93	−2.74	0.96
constant ($G = 0$)	-4.0×10^{-4}	-6.5×10^{-4}	3.36	0.45	1.73	3.92	−0.92	2.62
gradient ($G = 1$)	2.3×10^{-6}	-1.0×10^{-6}	−3.86	−1.22	−3.65	−3.20	−2.74	−2.69
2nd-order ($G = 2$)	-1.0×10^{-8}	-3.9×10^{-7}	0.04	−0.10	0.07	2.26	−1.44	2.47
2nd-order diag ($G = \bar{2}$)	-3.4×10^{-7}	-2.8×10^{-6}	0.06	−0.24	0.08	2.27	−1.55	2.41
3rd-order diag ($G = \bar{3}$)	-3.1×10^{-7}	-3.0×10^{-6}	0.00	−0.17	0.08	0.38	−0.94	1.86
exact/quartic fit	0.9588042	1.8260404	3832.10	1648.46	3945.49	3654.77	1596.98	3754.13
M = TQ − M = TQ5	6.0×10^{-4}	2.5×10^{-3}	*−3.52*	*−0.86*	*−1.57*	*−3.50*	*1.12*	*−2.40*
			M = Q5					
no scaling	3.3×10^{-4}	2.6×10^{-3}	−2.92	−2.84	−2.47	−1.66	−2.51	−1.64
constant ($G = 0$)	2.8×10^{-5}	$1.1 \times 10^{-3.}$	1.31	0.37	1.15	1.75	0.61	1.22
gradient ($G = 1$)	-8.0×10^{-7}	-2.0×10^{-6}	2.21	1.58	1.30	2.68	1.89	1.38
2nd-order ($G = 2$)	1.2×10^{-9}	4.0×10^{-8}	0.00	0.03	0.00	−0.37	0.14	−0.58
2nd-order diag ($G = \bar{2}$)	-3.7×10^{-8}	-7.7×10^{-7}	0.01	−0.03	0.00	−0.37	0.09	−0.61
3rd-order diag ($G = \bar{3}$)	-3.8×10^{-8}	-7.6×10^{-7}	0.01	−0.04	0.00	−0.31	−0.20	−0.51
exact/quartic fit	0.9582454	1.8236283	3835.91	1649.88	3947.58	3658.36	1596.75	3756.72
M = Q5 − M = TQ5	4.2×10^{-5}	*9.E-5*	*0.29*	*0.56*	*0.52*	*0.09*	*0.89*	*0.19*
	R(O−H)/Å	∠HOH/rad	ω_1/cm^{-1}	ω_2/cm^{-1}	ω_3/cm^{-1}	ν_1/cm^{-1}	ν_2/cm^{-1}	ν_3/cm^{-1}
VTQZ	-1.4×10^{-3}	5.3×10^{-4}	11.40	2.51	8.58	8.49	1.45	6.28
VQ5Z	-2.9×10^{-6}	2.9×10^{-3}	−0.18	−2.06	0.06	0.05	−1.77	−0.07
VTQ5Z	3.7×10^{-4}	2.8×10^{-3}	−2.63	−2.28	−1.95	−1.57	−1.62	−1.44
AVTQZ	-1.2×10^{-3}	3.0×10^{-4}	10.54	3.85	9.04	6.58	4.14	5.39
AVQ5Z	-3.3×10^{-4}	2.1×10^{-4}	2.76	0.84	2.53	1.72	0.79	1.58
AVTQ5Z	0.95820334	1.8235386	3835.62	1649.32	3947.06	3658.27	1595.86	3756.53

[a] See the text for details.

with the error for the O−H stretch fundamental vibrational frequencies greater than 3 cm^{-1}. Similar to what was found for H_2O, the use of up through only gradient terms seems to overcorrect at least some modes, and the error for the O−H stretch is now greater than 4 cm^{-1}. Again similar to H_2O, the examination of the results in the column labeled "M = TQ5,G = $\bar{2}$ − M = Q5,G = $\bar{2}$" shows little difference between the use of M = TQ5 and Q5 for approximating the effects of diffuse functions, whereas results in the final column, labeled "M = TQ5, G = $\bar{2}$ − M = TQ, G = $\bar{2}$", show that it is important to perform explicit calculations with the 5Z basis set when high accuracy is desired, even when extrapolating to the one-particle basis set limit. In summary, the CH_3OH results presented in Table 5 are entirely consistent with those reported in Table 4 for the H_2O molecule, indicating that approximating the effects of including diffuse functions in one-particle basis sets is achieved to high accuracy by using up through diagonal second derivative terms in eq 5 for neutral but polar molecules.

D. Computational Savings. The computational savings incurred using the new method proposed and investigated in the present study are dependent on the size of the molecule being studied, and dramatically increase as the molecular size increases. Contained in Table 6 is a compilation of the number of displacements needed for a full QFF as well as for the various approximations studied here, that is, $G = 1$, $G = 2$, $G = \bar{2}$, and $G = \bar{3}$, for the molecules investigated in this study. Examination of Table 6 shows that as the molecule increases in size, the percentage of displacements for which the larger augmented basis set (i.e., including diffuse functions) must be used explicitly in calculations decreases significantly. For example, using the $G = 2$ approximation means that the CCSD(T)/A5Z energy must be computed for 50% of the displacements for OH^-, 18% of the displacements for H_2O/NH_2^-, and only 3.4% of the displacements for CH_3OH. Similarly, these percentages are 50% (50%), 12% (18%), and 1.1% (1.4%) using $G = \bar{2}$ ($G = \bar{3}$) for OH^-, H_2O/NH_2^-, and CH_3OH, respectively. This demonstrates that for molecules with three or more atoms, the number of full CCSD(T)/A5Z calculations needed will be significantly reduced when using the scaling method proposed herein.

The full computational savings, however, must take into account the fact that a full set of displacement calculations must be performed at the CCSD(T)/5Z level of theory. To determine the full computational cost for the four molecules studied here, presented in Table 7 is the actual CPU time (in seconds) needed for a typical CCSD(T) calculation for all basis sets needed in our procedure. Because most of the computational savings will result from calculations with the A5Z basis set, we initially focus our attention on the last row of Table 7. Explicitly using the augmented basis sets will result in 74 displacement calculations for NH_2^- (and H_2O) using the A5Z basis set or 74(192.9) = 14 274.6 s. Using the $G = 2$ approximation leads to 74 calculations using the 5Z basis set plus 13 calculations using the A5Z basis set or 74(39.9) + 13(192.9) = 5460.3 s. Therefore, for the largest basis set alone only 38% (i.e., 5460.3/14 274.6) of the computational time is needed. A similar analysis

TABLE 5: CH_3OH Equilibrium Geometries (angstroms/radians) and Vibrational Fundamentals (inverse centimeters) Computed with a QFF Obtained Various Approximations for Including the Effects of Diffuse Functions[a]

Å/rad	no scaling	M = TQ5, $G = \bar{2}$	M = TQ5, $G = \bar{2}$ − no scaling	M = TQ5, $G = \bar{2}$ − M = TQ5, $G = 0$	M = TQ5, $G = \bar{2}$ − M = TQ5, $G = 1$	M = TQ5, $G = \bar{2}$ − M = Q5, $G = \bar{2}$	M = TQ5, $G = \bar{2}$ − M = TQ, $G = \bar{2}$
$r_{\text{O-H}}$	0.955767	0.955599	-1.7×10^{-4}	-4.1×10^{-5}	-2.2×10^{-6}	-8.0×10^{-5}	-3.6×10^{-4}
$r_{\text{C-O}}$	1.410446	1.414624	4.2×10^{-3}	-4.4×10^{-4}	-4.2×10^{-6}	-2.7×10^{-4}	-1.3×10^{-3}
$r_{\text{C-H1}}$	1.086268	1.086385	1.2×10^{-4}	1.6×10^{-4}	-2.2×10^{-7}	2.8×10^{-5}	1.4×10^{-4}
$r_{\text{C-H2/H3}}$	1.091676	1.091747	7.1×10^{-5}	7.4×10^{-5}	-7.0×10^{-8}	-1.3×10^{-5}	-4.2×10^{-5}
$\angle_{\text{C-O-H}}$	1.892576	1.891797	-7.8×10^{-4}	4.7×10^{-5}	-1.2×10^{-5}	-1.6×10^{-5}	-2.0×10^{-4}
$\angle_{\text{C-O-H1}}$	1.865528	1.864328	-1.2×10^{-3}	5.1×10^{-4}	9.6×10^{-7}	1.7×10^{-4}	8.5×10^{-4}
$\angle_{\text{C-O-H2/H3}}$	1.955825	1.955254	-5.7×10^{-4}	8.3×10^{-4}	-2.0×10^{-6}	4.0×10^{-4}	1.9×10^{-3}
$\angle_{\text{H2-C-H3}}$	1.901559	1.902923	1.3×10^{-3}	-2.0×10^{-4}	9.9×10^{-7}	-1.2×10^{-4}	-6.2×10^{-4}
$\angle_{\text{H1-C-H2/H3}}$	1.891385	1.891882	5.0×10^{-4}	-1.0×10^{-3}	1.1×10^{-6}	-4.5×10^{-4}	-2.2×10^{-3}
$\tau_{\text{H-O-C-H2/H3}}$	1.072004	1.072472	4.7×10^{-4}	4.9×10^{-4}	-8.1×10^{-7}	2.2×10^{-4}	1.0×10^{-3}

vibrational fundamental/cm^{-1}	no scaling	M = TQ5, $G = \bar{2}$	M = TQ5, $G = \bar{2}$ − no scaling	M = TQ5, $G = \bar{2}$ − M = TQ5, $G = 0$	M = TQ5, $G = \bar{2}$ − M = TQ5, $G = 1$	M = TQ5, $G = \bar{2}$ − M = Q5, $G = \bar{2}$	M = TQ5, $G = \bar{2}$ − M = TQ, $G = \bar{2}$
A' O−H str	3691.02	3689.05	−1.97	−3.23	−4.29	−0.61	−2.77
A' CH_3-asym str	3011.41	3011.30	−0.11	−0.24	0.83	0.34	1.14
A' CH_3-sym str	2846.90	2847.19	0.29	2.02	1.74	1.37	6.80
A' CH_3-asym bend	1475.05	1476.56	1.51	0.46	0.44	0.43	2.90
A' CH_3-sym bend	1460.97	1460.35	−0.61	2.13	1.50	1.35	6.12
A' COH bend	1334.30	1336.41	2.11	1.64	1.63	0.85	2.88
A' CH_3-rock	1070.22	1071.99	1.77	0.64	0.53	0.44	1.91
A' C−O str	1042.15	1044.03	1.88	0.91	−0.43	0.69	2.82
A'' CH_3-asym str	2967.52	2968.99	1.47	0.39	0.38	0.77	3.14
A'' CH_3-asym bend	1471.85	1473.80	1.95	1.08	0.53	0.71	2.52
A'' CH_3-rock	1158.34	1159.02	0.68	0.64	1.09	0.37	1.72
A'' O−H torsion	272.55	271.49	−1.06	1.38	1.09	1.28	−3.61

[a] See the text for details.

TABLE 6: Number of Unique Displacement Geometries Required for Different Approximations for Scaling the Effect of Diffuse Functions[a]

	OH^-	H_2O/NH_2^-	CH_3OH
gradient ($G = 1$)	2	4	16
2nd-order ($G = 2$)	4	13	204
2nd-order diag ($G = \bar{2}$)	4	9	68
3rd-order diag ($G = \bar{3}$)	4	13	84
exact/quartic fit	8	74	6064

[a] Reference geometry (center of the grid) is excluded.

for $G = \bar{2}$ ($G = \bar{3}$) shows that only 33% (38%) of the computational effort is required. For CH_3OH, the computational savings is significantly greater. Performing the analogous analysis for CH_3OH shows that for $G = 2$, $G = \bar{2}$, and $G = \bar{3}$, only 23, 21, and 21% of the computational time is required. In other words, for CH_3OH, using virtually any of the approximations proposed and investigated in the present study reduces the computational effort by about a factor of five. Examination of the A5Z/5Z ratios contained in Table 7 shows that these are fairly consistently around five for the molecules studied here. This suggests that in the limit of a very large molecule, using any of the approximations proposed and investigated in this study will lead to a required CPU time of about one-fifth that required without the use of scaling the diffuse function effect and that this limit is approached rapidly relative to the molecular size, as evidenced by our analysis of the computational savings for CH_3OH. However, this analysis has been strictly limited to CPU time, and it is well known that for large basis set CCSD(T) calculations, the input/output (I/O) wait time increases much more rapidly than the CPU time; see ref 24 for a more detailed discussion. Therefore, the actual savings in computational resources will be larger and could be much larger depending on the relative sizes of the molecular A5Z and 5Z basis sets. In addition, this analysis has not taken into account the fact that there will be an analogous savings associated with the QZ/AQZ basis sets. (See Table 7.) In summary, for molecules as large as CH_3OH and larger, we conservatively estimate, as an upper bound, that only 20% of the computational resources would be needed when scaling the effect of diffuse functions compared with not scaling, whereas the actual savings could be much greater if I/O wait times are considered.

4. Conclusions

A new approach is proposed and investigated for approximating the effects of including diffuse functions in one-particle basis sets for use in situations where (1) high accuracy is desired and (2) a significant portion of a PES is needed, such as when computing a QFF or even a global PES. The new approach can be used in situations where extrapolation to the one-particle basis set limit is performed or in cases where extrapolation is not performed but the QFF or PES is needed for a larger one-particle basis set. The new method has been tested on two small molecular anions, NH_2^- and OH^-, and two polar neutral molecules, H_2O and CH_3OH.

The new method is based on scaling the effect of diffuse functions from a smaller one-particle basis set to obtain the effect for a larger one-particle basis set or for obtaining the effect for an extrapolated electronic energy in which larger one-particle basis sets were used. It is well known that the importance of diffuse functions becomes less as the one-particle basis set is increased and hence the need to scale the effect. The new approach also takes into account the geometrical dependence of the scale factor. We have investigated including up through second derivative terms ($G = 2$) or diagonal second derivative plus diagonal third derivative terms ($G = \bar{3}$) and find that including up through just the diagonal second derivative terms

TABLE 7: Actual Total CPU Timings (in seconds) for OH^-, H_2O, NH_2^-, and CH_3OH Hartree−Fock + CCSD(T) Calculations on a Single Geometry around the Global Minimum[a]

	OH^-		H_2O		NH_2^-		CH_3OH	
	total time	ratio	time	ratio	time	ratio	time	ratio
ATZ/TZ	0.98/0.27	3.6	1.94/0.47	4.1	2.07/0.48	4.3	143.5/26.0	5.5
AQZ/QZ	7.96/1.75	4.5	20.92/4.25	4.9	23.6/4.5	5.2	1779/338	5.3
A5Z/5Z	61.4/13.8	4.4	179.1/38.3	4.7	192.9/39.9	4.8	15808/3118	5.1

[a] C_{2v} symmetry is adopted in the $OH^-/H_2O/NH_2^-$ calculations. C_s symmetry is adopted in the CH_3OH calculations. The platform is an Intel Xeon 5160 at 3.0 GHz + OpenSuse 10.1 + Molpro 2002.6 (compiled with Intel C/Fortran Compiler 10.1). (A)XZ refers to the (aug-)cc-pVXZ basis, X = T, Q, and 5.

($G = \bar{2}$) is an excellent compromise between accuracy and computational costs.

For small molecular anions, diffuse functions are imperative to obtain a reliable QFF, equilibrium structure, and harmonic and fundamental vibrational frequencies. Scaling the TZ-level contribution from diffuse functions proposed and investigated herein significantly reduces the error of not including any diffuse functions, provided that the scale factor includes up through diagonal second derivative terms. In fact, for equilibrium structures and harmonic frequencies, the error is essentially eliminated, which is a function of including through diagonal second derivative terms. For polar neutral molecules, the new method (including up through diagonal second derivative terms) provides highly accurate results and is a cost-effective approach when the use of a series of basis sets that includes diffuse functions (for extrapolation to the one-particle basis set limit) would be prohibitively expensive.

The results included in Tables 1−4 also show that the use of a 5Z basis is important for highly accurate QFFs, equilibrium structures, and harmonic and fundamental vibrational frequencies, even with extrapolation to the one-particle basis set limit. This is true for the geometry dependence (eq 5) as well as for the model used (eq 4) in the scaling procedure. For electronic structure methods that explicitly include r_{12} in the wave function, this latter conclusion may be different.

A detailed analysis of the computational savings that results when scaling the effect of diffuse functions using any of the approximations proposed and investigated herein is presented. It is conservatively estimated for molecules as large as CH_3OH or larger that at most only 20% of the computational resources will be required when using any of the approximations (i.e., $G = 2$, $G = \bar{2}$, and $G = \bar{3}$) compared with not scaling the effect of diffuse functions but that the actual savings could be much larger if I/O wait times are considered. In other words, the use of one of the approximations proposed herein leads to a computational savings of at least a factor of five for molecules as large as CH_3OH.

Acknowledgment. We gratefully acknowledge support from the NASA Herschel GO Program, Cycle 0 TR/LA PID 1022, and the NASA Spitzer GO Program, Cycle 4, AR-4 Science Program PID 40645. X.H. acknowledges the support by an appointment to the NASA Postdoctoral Program at the Ames Research Center, administered by Oak Ridge Associated Universities through a contract with NASA.

References and Notes

(1) Lee, T. J.; Scuseria, G. E. *J. Chem. Phys.* **1990**, *93*, 489.
(2) Rendell, A. P.; Lee, T. J.; Taylor, P. R. *J. Chem. Phys.* **1990**, *92*, 7050.
(3) Martin, J. M. L.; Lee, T. J. *Chem. Phys. Lett.* **1992**, *200*, 502.
(4) Császár, A. G. *J. Phys. Chem.* **1992**, *96*, 7898.
(5) Wong, A. T.; Bacskay, G. B. *Mol. Phys.* **1993**, *79*, 819.
(6) Martin, J. M. L.; Taylor, P. R.; Lee, T. J. *Chem. Phys. Lett.* **1993**, *205*, 535.
(7) Lee, T. J.; Scuseria, G. E. Achieving Chemical Accuracy with Coupled-Cluster Theory. In *Quantum Mechanical Electronic Structure Calculations with Chemical Accuracy*, Langhoff, S. R., Ed.; Kluwer Academic Publishers: Dordrecht, The Netherlands, **1995**; pp 47−108.
(8) Huang, X.; Lee, T. J. *J. Chem. Phys.* **2008**, *129*, 044312.
(9) Schwenke, D. W. *J. Chem. Phys.* **2005**, *122*, 014107.
(10) Feller, D.; Peterson, K. A.; Crawford, T. D. *J. Chem. Phys.* **2006**, *124*, 054107.
(11) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.
(12) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007.
(13) Martin, J. M. L.; Lee, T. J. *Chem. Phys. Lett.* **1996**, *258*, 136.
(14) Martin, J. M. L. *Chem. Phys. Lett.* **1996**, *259*, 169.
(15) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. *J. Chem. Phys.* **1997**, *106*, 9639.
(16) Douglas, M.; Kroll, N. M. *Ann. Phys.* **1974**, *82*, 89.
(17) Hess, B. A. *Phys. Rev. A* **1985**, *32*, 756; **1986**, *33*, 3742.
(18) Martin, J. M. L.; Taylor, P. R. *Chem. Phys. Lett.* **1994**, *225*, 473.
(19) Werner, H.-J.; Knowles, P. J.; Lindh, R.; Manby, F. R.; Schütz, M.; et al. *MOLPRO, A Package of Ab Initio Programs*, version 2002.6. http://www.molpro.net.
(20) Gaw, J. F.; Willets, A.; Green, W. H.; Handy, N. C. *SPECTRO*, version 3.0; **1996**.
(21) Mills, I. M. In *Molecular Spectroscopy: Modern Research*; Rao, K. N., Mathews, C. W., Ed.; Academic Press: New York, **1972**.
(22) Miani, A.; Hanninen, V.; Horn, M.; Halonen, L. *Mol. Phys.* **2000**, *98*, 1737.
(23) Huang, X.; Schwenke, D. W.; Lee, T. J., to be published.
(24) Lee, T. J.; Huang, X.; Dateo, C. E. *Mol. Phys.* **2009**, *107*, 1139.

JP9036364