Bond Additivity Corrections for Quantum Chemistry Methods

Carl F. Melius* and Mark D. Allendorf

Sandia National Laboratories, Livermore, California 94551-0969

Received: April 30, 1999; In Final Form: September 7, 1999

New bond additivity correction (BAC) methods have been developed for the G2 method, BAC-G2, as well as for a hybrid density functional theory (DFT) Møller−Plesset (MP)2 method, BAC−hybrid. These BAC methods use a new form of BAC corrections, involving atomic, molecular, and bond-wise additive terms. These terms enable one to treat positive and negative ions as well as neutrals. The BAC-G2 method reduces errors in the G2 method due to nearest-neighbor bonds. The parameters within the BAC-G2 method only depend on atom types. Thus the BAC-G2 method can be used to determine the parameters needed by BAC methods involving lower levels of theory, such as BAC−hybrid and BAC-MP4. The BAC−hybrid method is expected to scale well for large molecules. The BAC−hybrid method uses the differences between the DFT and MP2 predictions as an indication of the method's accuracy, whereas the BAC-G2 method uses its internal methods (G1 and G2MP2) to accomplish this. A statistical analysis of the error in each of the methods is presented on the basis of calculations performed for large sets (more than 120) of molecules.

1. Introduction

The thermochemical properties for many molecular species are not known. This is particularly true for transient intermediates occurring in chemical reactions. These species can be highly reactive and short-lived. Also, the environment in which they are produced can involve high temperatures and pressures that make experimental diagnostic techniques difficult to apply. It has therefore been necessary to develop theoretical quantum chemical techniques that can provide accurate thermochemical data.[1] Efforts have been made since the 1950s to make corrections to ab initio methods.[2] It is important that these computational procedures have predictive capability. That is, it is not sufficient for the calculated results to be accurate on the average, but one must also know when the results may be wrong (i.e., what is the uncertainty in a particular calculated value).

In the 1980s, we developed a bond additivity correction (BAC) procedure for quantum chemical calculations called BAC-Møller−Plesset (MP)4.[3,4] The BAC-MP4 method has proven reliable in calculating the thermochemical properties of molecular species, including radicals as well as stable closed-shell species. It can also treat activation barriers along reaction pathways, although the error uncertainties are somewhat larger. However, this procedure is currently limited to neutral species and cannot adequately treat positive and negative ions with the same level of accuracy. In addition, this procedure is limited to small molecules. When developed over 14 years ago, three- and four-heavy-atom systems could be handled routinely. Today, even with faster computers, we are still limited to routinely doing six- and seven- and occasionally 10-heavy-atom systems, due to the seventh-power dependence of the computational time on the number of basis functions (which typically scales with the number of atoms). Although a subprocedure of BAC-MP4, BAC-MP2, has been used for larger systems,[5] it does not provide the needed accuracy. We also developed two other variations of the BAC-MP4 method. In one case, we used a somewhat larger basis set,[6] that is, BAC-MP4/6-311++G**, which reduces the magnitude of the uncertainty in the calculation. In a second variation, geometry is calculated at the MP2 level[7] rather than at the Hartree−Fock (HF) level. These methods, however, all depend on the same functional form for the BAC corrections and thus contain many of the same limitations inherent in the original BAC-MP4 method.

It is clear that no one procedure can yield the desired accuracy and yet scale to larger molecules. We have therefore developed a new, generic approach to correct the results of ab initio calculations that is applicable to a variety of quantum chemistry methods. This approach is sufficiently flexible that it can be applied to both high and low levels of theory. In particular, we describe the extension of the new correction method to G2 calculations (BAC-G2), which can be used to obtain highly accurate energies for small molecules. These results can be used to establish reference heats of formation. The new approach can also be applied to other methods incorporating high levels of electron correlation, such as quadratic CI (QCI)[8] or coupled-cluster [CCSD(T)][9] theory. Finally, to obtain accurate results for larger molecules, we developed a procedure that combines MP2 and density functional theory (DFT) results, denoted BAC−hybrid.

In Section 2 we describe the new BAC method and define the forms for each of the corrections, including the atomic, molecular, and bond-wise additive terms. In Section 3, we use this method to obtain results of a standard G2 calculation and to the MP2 and DFT levels of theory, which we then combine to yield the BAC−hybrid result. In the remaining sections, we present results of calculations for a broad range of molecules and discuss the accuracy of the new method and its implications.

2. BAC Method for Thermochemical Properties

In this section, we describe the new BAC method for determining thermochemical properties. To determine the thermochemical properties for a given molecular species (e.g., its heat of formation, entropy, heat capacity, and free energy), we need to determine its geometry, vibrational frequencies, and

---

* Author for correspondence. E-mail: melius@ca.sandia.gov; fax: (925) 294-2234.

10.1021/jp9914370 CCC: $19.00

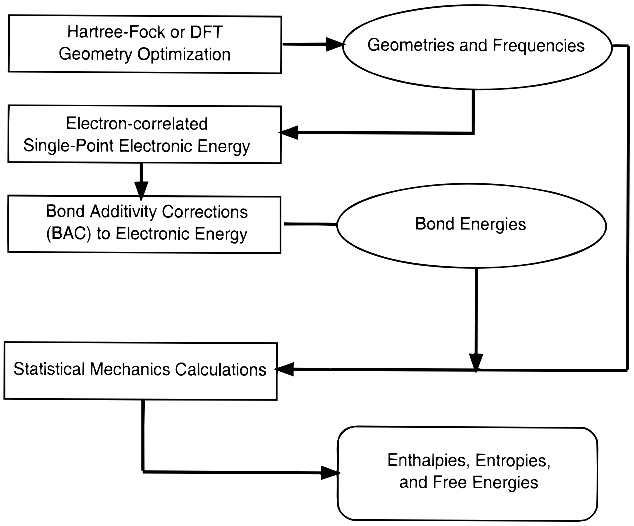

Figure 1. Hierarchical diagram of the bond additivity correction (BAC) method for determining thermochemical properties of molecules. First, a Hartree−Fock or DFT calculation provides the molecular geometry and frequencies. Next, a single-point calculation is performed at a higher level of theory with a larger basis set to determine the ab initio electronic energy, to which BACs are added to obtain a total electronic energy. Finally, statistical mechanics calculations are used to determine the enthalpy, entropy, and free energy of the molecule.

electronic energy. The overall BAC procedure is shown schematically in Figure 1. For purposes of this discussion, the term "BAC method" is used to refer to the entire process shown in Figure 1. In the following subsections we present details of this procedure, with particular emphasis on the form of the various corrections.

2.1. Electronic Structure Calculations for Geometry, Vibrational Frequencies, and Electronic Energy. The BAC method first involves ab initio electronic structure calculations to determine the geometry, vibrational frequencies, and electronic energy. Typically, the equilibrium geometry of the molecule is determined using a lower-level electronic structure method with a small basis set. Having determined the optimized molecular geometry, harmonic vibrational frequencies are calculated[10] at the same level of theory, using the same basis set. Actually, the first and second derivatives of the potential energy surface at the stationary point are calculated. Thus, the harmonic vibrational frequencies along with their corresponding normal modes can be determined for any isotopic form of the molecule. The resulting vibrational frequencies are used subsequently to determine the zero-point energy of the molecule and other properties derived from statistical mechanics. This provides the ability to determine the thermochemistry of isotopically labeled molecules as well as isotopic effects of rate constants using transition-state theory. Next, a single-point electronic energy calculation is performed using a higher level of electronic structure theory that incorporates electron correlation.

2.2. BAC Corrections. Ab initio electronic structure methods involving electron correlation, such as the Møller−Plesset perturbation methods (MP2, MP4) and coupled-cluster approaches and using typical basis sets (such as double-ζ and triple-ζ), do not provide the level of accuracy required for useful heats of formation (± 2−3 kcal mol^{-1}, depending on the application). However, we find that when using these methods and basis sets, the resulting errors in the electronic energy are quite systematic. To correct for these errors, we developed a set of empirically derived expressions that we collectively refer to as the bond additivity correction (BAC), $E_{\mathrm{BAC-correction}}$. This correction is composed of three parts: (a) a term due to the individual atoms making up the molecule; (b) a term due to the molecule as a whole; and (c) a term due to identity of the individual bonds composing the molecule:

TABLE 1: BAC-G2 Parameters (Energies in kcal mol^{-1})

	$K_{\text{elec pair}} = 0.860$		
atom	A_{atom}	B_{atom}	A_{ii}
H	0.485	−0.146	1.462
C	1.081	0.051	0.0
N	1.498	−0.010	2.281
O	−0.501	−0.010	114.3
F	-1.942	0.215	373.1
Si	0.097	0.008	297.4
P	0.277	0.042	1121.8
S	0.155	0.003	921.4
Cl	−0.776	0.087	1433.7
Br	−1.703	−0.106	−0.106

$$E_{\mathrm{BAC-correction}}(\mathrm{total}) = \sum_k E_{\mathrm{BAC-atom}}(A_k) + E_{\mathrm{BAC-molecule}} + \sum_{ij} E_{\mathrm{BAC-bond}}(A_iA_j) \quad (1)$$

where ij is summed over chemical bonds in the molecule and k is summed over the atoms in the molecule. This correction (and each of its constituent terms) has units of energy. $E_{\mathrm{BAC-correction}}$ is added to the calculated electronic energy to yield a corrected value that is used to determine the atomization energy and, ultimately, the molecular heat of formation. In the following subsections, we describe the empirical expressions representing each of the terms in eq 1.

2.2.1. Atomic Corrections. We define a correction that depends on the identity of the atoms that make up the molecule:

$$E_{\mathrm{BAC-atom}} = \sum_k E_{\mathrm{BAC-atom}}(A_k) \quad (2)$$

where the sum runs over all the atoms in the molecule and the value of $E_{\mathrm{BAC-atom}}(A_k)$ depends on the atom type and A_k is an adjustable parameter. Values of A_k for various elements are given in Table 1. The parameters were determined by minimizing the difference between BAC-G2 predictions and established experimental values for a set of 143 molecules (see further discussion below in Section 4).[24]

We note that in this new BAC method the atomic parameters are adjustable, that is, the calculated heat of formation for an atom is not forced to agree with accepted experimental values. The need for an atomic correction arises from errors due to intra-atomic electron correlation, which are the result of differences between atomic electronic configurations, charge, and spin−orbit coupling, as well as core−valance interactions and other relativistic effects. For many elements, these corrections are small (Table 1), but can become significant with increasing numbers of atoms in a molecule. The availability of this correction in the new method allows us to address errors in calculations for individual atoms as well as molecules. In contrast, the original BAC-MP4 method had no error associated with the atoms, because we fixed atomic heats of formation at experimental values. As a result, the new method improves the accuracy of molecular predictions by shifting some of the systematic error into the calculation of the constituent atoms, rather than artificially fixing their error at zero.

2.2.2. Molecular Corrections. The second type of BAC correction arises from errors in the overall electronic structure of the molecule. The BAC correction for this molecular term is given by

$$E_{\text{BAC}-\text{molecule}} = E_{\text{BAC}-\text{elec pair}} + E_{\text{BAC}-\text{S2}} \tag{3}$$

where $E_{\text{BAC}-\text{elec pair}}$ depends on the difference between the spin of the molecule and the sum of the spins of the constituent atoms:

$$E_{\text{BAC}-\text{elec pair}} = K_{\text{elec pair}}(\text{spin}_{\text{molecule}} - \sum_{atom} \text{spin}_{\text{atom}}) \tag{4}$$

where $K_{\text{elec pair}}$ is an empirically adjusted parameter for a given BAC method and spin refers to the S quantum number. $E_{\text{BAC}-\text{S2}}$ depends on the amount of spin contamination (if it exists) in the electronic wave function and depends on the choice of wave function used in the calculation.

$E_{\text{BAC}-\text{elec pair}}$ arises from the basic nature of the interactions between electrons and occurs in atoms as well as molecules. This error is systematic, resulting from the way electronic structure codes treat electron correlation. In general, it is difficult to determine the functional form for the systematic errors due to the interaction between electrons. In the original BAC method, we used bond-wise additive terms to correct for electron pairing caused by bond formation. In that case, the correction term exponentially increased in value as two atoms approached each other to form a bond. However, there are other situations in which the electron interaction error can arise. For example, simply bringing an electron up to a molecule (forming a negative ion), without forming a bond. Similarly, removing a proton from a molecule removes a bond but does not remove the electron, which still interacts with the remaining molecule. Finally, the size of the correlation error depends on the spin state of the molecule: the error is much larger when electrons are paired than when they are unpaired. Thus, the molecular correction (eq 4) describes the more general case of electron pairing. Note that this term is very similar to the empirical correction term found within the G2 method.[8]

Note that this term destroys the size consistency of ab initio methods. That is, the sum of the calculated energies for two individual hydrogen atoms (two doublets) is not the same as that obtained from a single calculation of the energy for a hydrogen molecule in which the atoms are fixed at infinity (a singlet state; note there is also a triplet state).

The second term in eq 3, $E_{\text{BAC}-\text{S2}}$, corrects for an error that arises when open-shell methods based on the unrestricted Hartree−Fock approach (UHF) are used. This error is a result of spin contamination in the wave function (e.g., a doublet state may have quartet-state spin contamination). The resulting correlated energy [obtained from an unrestricted (U)MP4 calculation, for instance] still contains contributions from the contaminated spin states. We therefore allow for the inclusion of a spin-contamination correction.[4] Unfortunately, not all ab initio methods make this energy available. Thus, its use depends on the particular method and it is currently applied only to MP methods such as MP2 and MP4 (see Section 3.2 below).

A separate "spin-dependent" correction was used in the BAC-MP4 method for closed shell systems.[3,4] This correction used the spin contamination predicted by an unrestricted calculation of the singlet state to indicate the need for a multireference wave function. For larger basis sets, this systematic error is smaller. DFT methods tend not to have an unrestricted wave function instability. For these reasons, we eliminated the unrestricted wave function spin correction in the new BAC method. However, the contribution of this wave function instability to the overall error estimate, indicating greater multireference contributions to the energy, will be lost.

Another error due to electron correlation is related to the spatial extent of the electrons in a molecule. In general, the amount of error follows the trend cations < neutrals < anions. Although we investigated the possibility of using additional parameters to correct for systematic "net charge" dependence, we find that the electron pair change correction (eq 4 above) addresses most of the error. No additional systematic correction was found that substantially reduces the discrepancies between experimental energies for positive and negative ions and predicted values.

2.2.3. Bond-Wise Corrections. The third type of BAC correction depends on the formation of chemical bonds. In this instance, one must distinguish between bonds and pairwise interactions. A bond is taken to mean the formation of an electron pair between the atoms. This correction addresses systematic errors arising from electron pairing not covered by eq 4. The correction for each bond A−B in the molecule having neighbors C and D (e.g., C−A−B−D) is given by

$$E_{\text{BAC}-\text{bond}}(AB) = A_{\text{AB}}\, \text{e}^{-\alpha R_{\text{AB}}} + \sum_{C} B_{\text{CA}} + \sum_{\text{D}} B_{\text{DB}} \tag{5}$$

where the first term is the correction for the bond alone, and the corrections for its nearest neighbors are treated as a sum of corrections for each neighbor of the form

$$B_{\text{CA}} = B_{\text{C}} + B_{\text{A}} \tag{6}$$

The B_{A}s are constants that depend only on the type of atom. The bond-distance dependence exists only in the first term for the bond itself. Furthermore, α no longer depends on the type of bond, as it did in the original BAC method.[4] Note that the bond-wise corrections do not go to zero at infinity, because of the terms $\sum B_{\text{CA}} + \sum B_{\text{DB}}$ defined by eq 5.

This approach is different from the original BAC-MP4 method,[3,4] in which the correction for each bond A−B in the molecule having neighbors C and D (e.g., C−A−B−D) was of the form

$$E_{\text{BAC}-\text{bond}}(AB) = (A_{\text{AB}}\, \text{e}^{-\alpha_{\text{AB}} R_{\text{AB}}})\Pi f_{\text{CAB}} \Pi f_{\text{DBA}} \tag{7}$$

where the first term in parentheses is the correction for the bond by itself, whereas the correction for nearest neighbors is treated as the product of corrections, f_{CAB}, for each neighboring bond.

This new form for the correction results from two factors. First, most of the error in the bond energy has already been corrected by the spin-pairing term (eq 4). In the old BAC-MP4 functional form (eq 7), the bond-wise correction increases substantially with decreasing bond distance in large part because multiple bonds, resulting in additional spin pairings, are being formed. For example, the C−C bond distance in triplet ethylene is close to that of ethane, because it has only one set of paired electrons in a σ bond. Pairing the unpaired electrons to form a π bond of ethylene causes the bond distance to become substantially shorter. In the old form of the BAC correction, this leads to an exponential increase in the magnitude of the correction. Errors caused by this spin pairing are now no longer included in the bond-wise correction. Second, the new method assumes larger basis sets are being used, including diffuse functions that have been added to treat ionic character. Thus, the corrections themselves are smaller. Indeed, for the more ionic bonds, the correction might even be negative. Thus, the multiplicative form of eq 7 (which reduces the size of the correction) is no longer meaningful. In general, the new bond-wise corrections have the opposite effect of the spin-pairing correction. The correction for electron correlation error due to the pairing of electrons is reduced by bond formation, because of the ability to correlate the electrons of the pair on different atoms.

We note at this point that one must first define what constitutes a bond before a correction can be applied. This definition is, in fact, rather subjective. In our method, hydrogen bonds do not have an associated BAC correction because they have no effect on electron pairings. On the other hand, dative bonds, such as in N_2O, do have a correction, because their formation involves a rearrangement of electron pairings. Hydrogen−boron bonds are particularly problematic. In diborane (B_2H_6), for example, the two bridging hydrogens are treated as having two bonds each, whereas the boron−boron interaction is not considered to be a bond. For transition-state structures, both the bonds being formed and the bonds being broken receive corrections. In the older BAC-MP4 method, in contrast, the stretching of the bond distance for the breaking and forming of bonds in the transition state decreases the size of the correction for each bond, which compensates for having only one bond defined for the reactants or products.

2.3. Heats of Formation, Entropies, and Free Energies. Having obtained the geometry, vibrational energies, ab initio electronic energy, and the BAC energy corrections, we can then calculate the enthalpies, entropies, and free energies of a molecule using statistical mechanics (see Figure 1). The equations used in these calculations are an extension of the subroutines in the Gaussian codes.[11] These subroutines use standard expressions for an ideal gas in the canonical ensemble to compute the entropy, heat capacity, and internal energy.

The corrected heat of formation at 0 K (ΔH°_{f0}) can now be obtained from the calculated electronic energy. First, the electronic energy is added to the zero-point energy. Next, the resulting energy is subtracted from the electronic energies of the atoms to give an electronic heat of atomization:

$$E_{\text{atomization}} = \sum_i^n E_i(\text{atoms}) - [E_{\text{ab initio}}(\text{molecule}) + E_{\text{ZPE}}] \quad (8)$$

Referencing this energy against the experimental ΔH°_{f0} of the atoms in the gas phase yields the uncorrected ΔH°_{f0}:

$$\Delta H^\circ_{f0,\text{uncorrected}} = \sum_{atoms} \Delta H^\circ_{f0,\text{atoms}} - E_{\text{atomization}} \quad (9)$$

Subtracting the BAC corrections from this energy finally yields ΔH°_{f0} at 0 K:

$$\Delta H^\circ_{f0,\text{BAC}} = \Delta H^\circ_{f0,\text{uncorrected}} - E_{\text{BAC-correction}} \quad (10)$$

Statistical mechanical calculations then provide the thermochemical properties of the molecule at a given temperature. The values of these properties as a function of temperature are then fit to a polynomial compatible with the Chemkin thermodynamic database.[12]

3. Applications of the New BAC Method

To illustrate the new BAC method, we now define two new procedures that we will denote BAC-G2 and BAC−hybrid. In the former, we apply the BAC method to the results of a standard G2 calculation, whereas in the latter we take advantage of the combined power of MP2 and DFT to achieve an estimate of the calculational error. The following subsections also present detailed, worked-out examples of the application of each method to the calculation of the heat of formation of CF_4.

3.1. BAC-G2 Method. The BAC-G2 method applies the BAC corrections to the standard G2 method.[8] The electronic structure calculations to determine the geometry, vibrational frequencies, and electronic energies (Section 2.1) are the same as those in the G2 method. Specifically, the geometry and vibrational frequencies in the BAC-G2 method are obtained from a HF calculation (restricted Hartree−Fock,[13] RHF, for closed-shell molecules and UHF[14] for open-shell molecules) using the 6-31G(d) split-valence basis set with polarization functions on the heavy atoms.[15,16] At this level of theory, vibrational frequencies are systematically too large compared to experimental values. We therefore scale the HF harmonic frequencies downward by 12%.[17] The electronic energies at the QCI, MP4, and MP2 levels of theory, as well as the collective G1, G2MP2, and G2 electronic energies, are taken directly from the output of the G2 method. The basis sets are the same as those defined in the standard G2 method. The geometry used in the single-point calculations is obtained by reoptimizing the HF geometry at the MP2 level, again as defined in the G2 method.

The BAC corrections for the BAC-G2 method are those defined in Section 2.2. The parameters for each of the corrections are given in Table 1. The atomic corrections (eq 2) are straightforward. For the molecular corrections (eq 3), the spin-contamination correction, $E_{\text{BAC-S2}}$ is not used; in the QCI method, the error is negligible, whereas the energies obtained from MP calculations are used in the form of differences between pairs of calculations. Thus, the systematic spin-contamination errors from the MP terms tend to cancel. For the bond-wise corrections (eq 5), the α exponent is (somewhat arbitrarily) taken to be 3.0 Å^{-1}, whereas the pre-exponential coefficient A_{AB} is taken to be the geometric mean of the individual atom types, that is,

$$A_{\text{AB}} = -(A_{\text{AA}}A_{\text{BB}})^{1/2} \quad (11)$$

Equation 5 also includes contributions from the nearest-neighbor B_{ij} terms (defined by eq 6). The accuracy of the parameters comprising these terms (see Table 1) is difficult to assess because of their small size. This is due to the fact that to date we have only applied the BAC-G2 method to small molecules (less than seven heavy, i.e., nonhydrogen, atoms), for which accurate experimental thermodynamic data exist. However, these terms become quite significant for larger molecules. Unfortunately, given the limited accuracy of experimental data for larger nonhydrocarbon, unsaturated gas-phase species it will remain difficult to establish the accuracy of the B_{atom} terms.

Using an ad hoc expression similar to that formulated for the earlier BAC-MP4 method,[4] we can obtain an estimate of the error (or confidence level) in the BAC-G2 method. In this case, we use the similarities between the G1 and G2-MP2 methods and the G2 method itself as an indication of the error:

$$\text{Error(BAC-G2)} = \text{Sqrt}[1.0\ \text{kcal mol}^{-1} + (\Delta H_{\text{BAC-G2}} - \Delta H_{\text{BAC-G2MP2}})^2 + (\Delta H_{\text{BAC-G2}} - \Delta H_{\text{BAC-G1}})^2] \quad (12)$$

Heats of formation, entropies, and free energies (Section 2.3) are obtained using the same procedure as in the original BAC-MP4 method. Thus, for finite temperatures, the raw G2 energies (without BAC corrections) obtained from the BAC-G2 method

TABLE 2: Electronic Energies (in hartrees) and BAC Parameters for BAC-MP2 and BAC-DFT Methods for CF_4 (in kcal mol^{-1})

	G2	MP2	DFT	experiment (0 K)
energies:				
C	−37.78430	−37.75737	−37.85745	169.98[a]
F	−99.63281	−99.60010	−99.76165	18.47[b]
CF_4	−437.06631	−436.93813	−437.64957	
ZPE		0.01713	0.01713	
ΔH°_{f0} CF_4				−221.6
parameters:				
$K_{\text{elec pair}}$	0.860	−5.70	−4.24	
$K_{\text{spin cont}}$	0.0	1.317	0.0	
A_{atom}				
C	1.081	0.266	2.797	
F	−1.942	−0.584	−0.071	
B_{atom}				
C	0.051	0.047	0.771	
F	0.215	−0.415	1.482	
A_{ij}				
C−F	0.0	−336.4	−8.896	

[a] Ref 18. [b] Baulch, D. L.; Cox, R. A.; Hampson, R. F.; Kerr, J. A.; Troe, J.; Watson, R. T. *J. Phys. Chem. Ref. Data* **1984**, *13*, 1259.

do not correspond to those from the output of a Gaussian G2 calculation, because hindered rotors are included in the BAC procedure.

As an example of the BAC-G2 method, we evaluate the heat of formation for CF_4 at 0 K using the equations defined above and the parameters given in Tables 1 and 2. The reference states for the atoms are the G2 energies. Combining the calculated G2 energy of CF_4 with the calculated G2 energies of the atoms yields the heat of atomization for CF_4 (eq 8):

$$E_{\text{atomization}} = \sum_i^n E_i(\text{atoms}) - [E_{\text{ab initio}}(\text{molecule}) + E_{\text{2PE}}] = (-37.78430 + 4 \times -99.63281) - (-437.066309) = 0.75075 \text{ hartree} = 471.12 \text{ kcal mol}^{-1}$$

Using the experimental heats of formation of C and F,[18] the uncorrected heat of formation for CF_4 is (eq 9):

$$\Delta H^\circ_{f0,\text{uncorrected}} = \sum_{atoms} \Delta H^\circ_{f0,\text{atoms}} - E_{\text{atomization}} = 169.98 + 4 \times 18.47 - 471.12 = -227.26 \text{ kcal mol}^{-1}$$

The BAC corrections consist of three terms: atomic, molecular, and bond-wise. The atomic correction (eq 2) is

$$E_{\text{BAC-atom}} = \sum_k E_{\text{BAC-atom}}(A_k) = 1.081 + 4 \times (-1.942) = -6.687 \text{ kcal mol}^{-1}$$

The molecular correction (eqs 3 and 4) is

$$E_{\text{molecular}} = E_{\text{BAC-elec pair}} = K_{\text{elec pair}}(\text{spin}_{\text{molecule}} - \sum_{atom} \text{spin}_{\text{atom}}) = 0.860 \times [0.0 - (1.0 + 4 \times 0.5)] = -2.580 \text{ kcal mol}^{-1},$$

where the spin of C is 1 and the spin of F is 1/2. For the bond corrections, each of the four C−F bonds has three nearest neighbors. Thus eq 5 becomes

$$E_{\text{BAC-bond}}(\text{CF}) = -(0.0 \times 373.1)^{1/2}\, e^{-3.0 \times 1.3292} + 3 \times (0.051 + 0.215) = 0.798 \text{ kcal mol}^{-1}$$

The total BAC correction (eq 1) is therefore

$$E_{\text{BAC-correction}} = \sum_{ij} E_{\text{BAC-bond}}(A_iA_j) + \sum_k E_{\text{BAC-atom}}(A_k) + E_{\text{BAC-molecule}} + \sum_{ij} E_{\text{BAC-bond}}(A_iA_j) = (-6.687) + (-2.580) + 4 \times (0.798) = -6.075 \text{ kcal mol}^{-1}$$

Thus, the resulting BAC-G2 heat of formation at 0 K is (eq 10)

$$\Delta H^\circ_{f0,\text{BAG-G2}} = \Delta H^\circ_{f0,\text{uncorrected}} - E_{\text{BAC-correction}} = -227.26 - (-6.07) = -221.19 \text{ kcal mol}^{-1}$$

For comparison, the experimental heat of formation reported in the JANAF Tables is −221.6 kcal mol^{-1}.[18]

3.2. The BAC−Hybrid Method. Although the BAC-G2 method provides a highly accurate procedure for determining heats of formation, it is impractical for larger molecular systems, that is, those with more than eight heavy atoms. The extension to larger molecules involves more than just the scaling of the electronic energy expression. It must also address the number of conformers that may exist as well as the time required to search for stationary points (i.e., the reaction surface for a chemical species becomes combinatorially more complex as the number of atoms increases). In this section we describe a hybrid BAC method (BAC−hybrid) that uses both the BAC-MP2 and BAC-DFT methods (described in detail below). These two methods are computationally practical for large molecules. However, they do not involve a series of calculations at successively higher levels of theory from which a convergence trend can be observed, making it difficult to determine their accuracy. Thus, the objective of the BAC−hybrid method is to provide an estimate of the error in the BAC-MP2 and BAC-DFT methods, thereby increasing confidence in their accuracy when applied to larger, uncharacterized molecules.

In the BAC−hybrid method, we first obtain the heat of formation from a BAC-MP2 calculation. We then estimate its error through the use of an ad hoc expression that combines the predictions of BAC-MP2 and BAC-DFT:

$$\text{Error (BAC-hybrid)} = \text{Sqrt}[1.0 \text{ kcal mol}^{-1} + (\Delta H_{\text{BAC-MP2}} - \Delta H_{\text{BAC-DFT}})^2] \quad (13)$$

These two methods, which use fundamentally different approaches to calculate the electronic energy, tend not to bias their error in the same direction. Thus, a small value gives one confidence that the calculated value is correct. We therefore gain the ability to estimate the accuracy of the BAC-MP2 method without significantly increasing the cost of the calculation. Note that, in the MP2 method itself, there are no other perturbation terms to use as error estimates, unlike the original BAC-MP4 and BAC-G2 methods.

3.2.1 BAC-MP2 Method. In this subsection, we describe a new BAC-MP2 method that uses the new correction procedure. The molecular geometry used in the single-point calculation of the electronic energy and the vibrational frequencies are determined using B3LYP (a hybrid density functional method mixed with HF theory).[19−21] The restricted RB3LYP21 method is used for closed-shell systems, whereas the unrestricted UB3LYP21 method is used for open-shell systems; the 6-31G-

(d) basis set is used in all cases. No scaling of the frequencies is applied. In fact, limits on the quality of the geometry optimization due to the use of HF are not present in the BAC-MP2 method as a result of the progress made in density functional methods, which provide geometries and frequencies that are comparable in accuracy with those obtained from the MP2 method.

The single-point calculation of the electronic energy (see Figure 1) is performed at the MP2 level of theory. An MP2 calculation with a large basis set, 6-311++G(2df,2pd), is used. This corresponds to a triple-ζ basis set with balanced polarization functions (i.e., double d functions, single f functions) on the heavy atoms and analogous polarization functions on the hydrogen atoms. In addition, diffuse functions are added to treat ionic effects. This electronic energy method is denoted MP2/6-311++G(2df,2pd). The overall procedure, BAC-MP2/6-311++G(2df,2pd)//B3LYP/6-31G(d), is denoted BAC-MP2.[25]

The corrections for the BAC-MP2 method are again those defined in Section 2.2. Within the molecular corrections (eq 3), the spin-contamination correction,

$$E_{\mathrm{BAC-S2}} = K_{\mathrm{spin\ cont}}(\mathrm{PUMP2} - \mathrm{UMP2}) \qquad (14)$$

is used for open-shell molecules. $K_{\mathrm{spin\ cont}}$ is an adjustable parameter that corrects for the spin contamination. If $K_{\mathrm{spin\ cont}}$ equals one, this term is equivalent to using the PUMP2 energy instead of the MP2 energy for open-shell systems. The value of this parameter and all others used in the method (Table 4) were obtained by minimizing the error between BAC-MP2 predictions and heats of formation for a 122-molecule test set. (The larger list of 143 compounds used for the BAC-G2 method includes species known to require multireference configurations, for which DFT performs poorly.) For the bond-wise corrections (eq 5), the α exponent is taken to be 3.0 Å^{-1}, as in the BAC-G2 method. Heats of formation, entropies, and free energies (Section 2.3) are obtained using the same procedure as in the original BAC-MP4 method, except that the geometries and frequencies are obtained using the B3LYP method. No scaling of the frequencies is needed with the B3LYP method.

As an example of the new BAC-MP2 method, we evaluate the heat of formation at 0 K for $\mathrm{CF_4}$ using the equations defined above, along with the parameters given in Table 2. The reference states for the atoms are the PUMP2 energies. From the calculated MP2 energy of $\mathrm{CF_4}$ as well as the calculated MP2 energies of the atoms, the heat of atomization for $\mathrm{CF_4}$ is (eq 8):

$$E_{\mathrm{atomization}} = \sum_i^n E_i(\mathrm{atoms}) - [E_{\mathrm{ab\ initio}}(\mathrm{molecule}) + E_{\mathrm{ZPE}}] = (-37.75737 + 4 \times -99.60010) - (436.93813 + 0.01713) = 0.76323\ \mathrm{hartree} = 478.95\ \mathrm{kcal\ mol^{-1}}$$

Using the experimental C and F heats of formation,[18] the uncorrected $\mathrm{CF_4}$ heat of formation is (eq 9):

$$\Delta H^{\circ}_{f0,\mathrm{uncorrected}} = \sum_{atoms} \Delta H^{\circ}_{f0,\mathrm{atoms}} - E_{\mathrm{atomization}} = (169.98 + 4 \times 18.47) - 478.95 = -235.09\ \mathrm{kcal\ mol^{-1}}$$

The BAC corrections consist of three terms, atomic, molecular, and bond-wise. The atomic correction (eq 2) is

$$E_{\mathrm{BAC-atom}} = \sum_k E_{\mathrm{BAC-atom}}(A_k) = 0.266 + 4 \times (-0.584) = -2.07\ \mathrm{kcal\ mol^{-1}}$$

The molecular correction (eqs 3 and 4) is

$$E_{\mathrm{molecular}} = E_{\mathrm{BAC-elec\ pair}} + E_{\mathrm{BAC-S2}} = K_{\mathrm{elec\ pair}}(\mathrm{spin_{molecule}} - \sum_{atom} \mathrm{spin_{atom}}) + K_{\mathrm{spin\ cont}}(\mathrm{PUMP2} - \mathrm{UMP2}) = -5.70 \times [0.0 - (1.0 + 4 \times 0.5)] + 1.317 \times (-436.93813 - -436.93813) \times 627.53 = 17.10\ \mathrm{kcal\ mol^{-1}}$$

For the bond corrections, each of the four C−F bonds has three nearest neighbors. Thus eq 5 is

$$E_{\mathrm{BAC-bond}}(\mathrm{CF}) = -336.4\ \mathrm{e}^{-3.0 \times 1.3294} + 3 \times (0.047 - 0.415) = -7.34\ \mathrm{kcal\ mol^{-1}}$$

The total BAC correction (eq 1) is therefore

$$E_{\mathrm{BAC-correction}} = \sum_{ij} E_{\mathrm{BAC-bond}}(A_iA_j) + \sum_k E_{\mathrm{BAC-atom}}(A_k) + E_{\mathrm{BAC-molecule}} + \sum_{ij} E_{\mathrm{BAC-bond}}(A_iA_j) = (-2.07) + (17.10) + 4 \times (-7.34) = -14.33\ \mathrm{kcal\ mol^{-1}}$$

Thus, the resulting BAC-MP2 heat of formation at 0 K is

$$\Delta H^{\circ}_{f0,\mathrm{BAC-MP2}} = \Delta H^{\circ}_{f0,\mathrm{uncorrected}} - E_{\mathrm{BAC-correction}} = -235.09 - (-14.33) = -220.76\ \mathrm{kcal\ mol^{-1}}$$

3.2.2. BAC-DFT Method. In this subsection, we present the BAC-DFT method. The electronic structure calculations for the geometry and vibrational frequencies (Section 2.1) use the same B3LYP method as the BAC-MP2 method. The electronic energy method is denoted B3LYP/6-31G(D). The single-point calculation of the electronic energy (Figure 1) is performed at the same level of theory as the geometry optimization and frequencies, but with a larger basis set, 6-311++G(2DF,2PD). This basis set is the same as that used for the BAC-MP2 method. The resulting single-point electronic energy method is denoted B3LYP/6-311++G(2DF,2PD). The BAC corrections are those defined in Section 2.2. For the molecular corrections (eq 3), no spin-contamination correction, $E_{\mathrm{BAC-S2}}$ is used. For the unrestricted DFT methods, the spin contamination is smaller, so its effect is less serious. The α exponent in eq 5 is taken to be 0.0; that is, the bond-dependent error does not depend on the bond distance. The heats of formation, entropies, and free energies (Section 2.3) are obtained using the same procedure as in the BAC-MP2 method, using the geometries and frequencies obtained from the B3LYP/6-31G(D) method. No frequency scaling is used. We denote this overall procedure, BAC−B3LYP/6-311++G(2DF,2PD)//B3LYP/6-31G(d), as the BAC-DFT method. The value of all parameters used in the method (Table 2) were obtained by minimizing the error between BAC-MP2 predictions and heats of formation for a 122-molecule test set. The larger list of 143 compounds used for the BAC-G2 method includes species known to require multireference configurations, for which DFT performs poorly.

As an example of the BAC-DFT method, we evaluate the heat of formation at 0 K for $\mathrm{CF_4}$ using the equations defined above, along with the parameters given in Table 2. From the calculated DFT energy of $\mathrm{CF_4}$ as well as the calculated DFT energies of the atoms, the heat of atomization for $\mathrm{CF_4}$ is (eq 8)

$$E_{\text{atomization}} = \sum_{i}^{n} E_i(\text{atoms}) - [E_{\text{ab initio}}(\text{molecule}) + E_{\text{ZPE}}] = (-37.85745 + 4 \times -99.76165) - -437.64957 - 0.01713 = 0.72838\ \text{hartree} = 457.08\ \text{kcal mol}^{-1}$$

Using the experimental C and F heats of formation, the uncorrected CF_4 heat of formation is (eq 9)

$$\Delta H^{\circ}_{f0,\text{uncorrected}} = \sum_{atoms} \Delta H^{\circ}_{f0,\text{atoms}} - E_{\text{atomization}} = 169.98 + 4 \times 18.47 - 457.08 = -213.22\ \text{kcal mol}^{-1}$$

The BAC corrections again consist of atomic, molecular, and bond-wise terms. The atomic correction (eq 2) is

$$E_{\text{BAC-atom}} = \sum_{k} E_{\text{BAC-atom}}(A_k) = 2.797 + 4 \times (-0.071) = 2.513\ \text{kcal mol}^{-1}$$

The molecular correction (eqs 3 and 4) is

$$E_{\text{molecular}} = E_{\text{BAC-elec pair}} + E_{\text{BAC-S2}} = K_{\text{elec pair}}(\text{spin}_{\text{molecule}} - \sum_{atom} \text{spin}_{\text{atom}}) = -4.24 \times [0.0 - (1.0 + 4 \times 0.5)] = 12.720\ \text{kcal mol}^{-1}$$

For the bond corrections, each of the four C$-$F bonds has three nearest neighbors. Thus eq 5 is

$$E_{\text{BAC-bond}}(\text{CF}) = -8.896\ e^{-0.0 \times 1.3294} + 3 \times (0.771 + 1.482) = -2.14\ \text{kcal mol}^{-1}$$

The total BAC correction (eq 1) is

$$E_{\text{BAC-correction}} = \sum_{ij} E_{\text{BAC-bond}}(A_iA_j) + \sum_{k} E_{\text{BAC-atom}}(A_k) + E_{\text{BAC-molecule}} + \sum_{ij} E_{\text{BAC-bond}}(A_iA_j) = (2.513) + (12.720) + 4 \times (-2.137) = 6.685\ \text{kcal mol}^{-1}$$

Thus, the resulting BAC-DFT heat of formation at 0 K is (eq 10)

$$\Delta H^{\circ}_{f0,\text{BAC-DFT}} = \Delta H^{\circ}_{f0,\text{uncorrected}} - E_{\text{BAC-correction}} = -213.22 - (6.685) = -219.90\ \text{kcal mol}^{-1}$$

4. Results and Discussion

The BAC-G2 heats of formation have been compared with experimental data for 143 neutral compounds. A statistical analysis of these predictions (given in Table 3) indicates that the BAC-G2MP2 and BAC-G1 methods give consistently excellent agreement with experiment. Even the worst cases, that is, those yielding the maximum error given in Table 3, are in relatively good agreement with experiment. Because the G2 method itself consists of a series of methods (i.e., G2, G2MP2, G1, QCI, MP4, etc.), individual BAC corrections can be applied to each of them. Table 3 displays the same statistical analysis for each of these component methods; for purposes of calculating the error estimate, we use only the first three methods (see eq 12). Clearly, these provide better results than the MP4 and QCI methods. For comparison, in Table 4 we present some results obtained from the old BAC-MP4 method, which shows that the new correction approach yields more accurate results with fewer parameters.

TABLE 3: Accuracy of Various BAC Procedures within G2 Method (Energies in kcal mol^{-1})

method	avg. error	RMS error	max. error
BAC-G2	0.69	0.90	2.68
BAC-G2MP2	0.74	0.95	3.02
BAC-G1	0.75	0.98	2.88
BAC-QCI/6-311G(d)	0.82	1.11	3.92
BAC-MP4/6-311+G(d,p)	0.95	1.32	4.3

143 reference compounds were used in the comparison.

TABLE 4: Comparison of $\Delta H^{\circ}_{f\ 298}$ for BAC-MP4 versus Experiment (Energies in kcal mol^{-1})

comound class	number of compounds	average energy difference
C_xH_y	27	1.3
$C_xO_yH_z$	23	1.4
$C_xN_yH_z$	10	0.9
$C_xN_yO_zH_w$	8	1.3
$N_xO_yH_z$	12	1.2
fluorinated species	13	1.2
total	93	1.25

TABLE 5: Accuracy of Various BAC Procedures for MP2 and DFT (Energies in kcal mol^{-1})

method	avg. error	RMS error	max. error
BAC-MP2/6-311++G(2df,2pd)[a]	0.78	1.13	3.69
BAC$-$B3LYP/6-311++G(2df,2pd)[b]	0.93	1.28	4.04
BAC$-$B3LYP/6-31G(d)	1.10	1.54	5.43
BAC$-$HF6/DFT4/6-311++G(2df,2pd)	1.28	1.90	7.26

122 reference compounds were used in the comparison. [a] BAC-MP2 method (Sect. 3.2.1), used for energy in BAC$-$Hybrid [b] BAC-DFT method (Sect. 3.2.2), used for error estimate in BAC$-$Hybrid

The BAC-MP2 method yields somewhat better results than the BAC-DFT method. Table 5 presents a statistical analysis (similar to that presented in Table 3) of BAC-DFT and BAC-MP2 results for 122 neutral compounds. Biradical species, which were included in the BAC-G2 analysis, were eliminated from this analysis, because of known deficiencies in the ability of the MP2 and DFT methods to treat such species. The average and root mean-square errors shown in Table 5 are only slightly larger than those shown in Table 3 for the BAC-G2 method; however, the maximum deviations are significantly larger. These results thus suggest that, for most species, BAC-MP2 or BAC-DFT, which are computationally more efficient than BAC-G2, provide sufficient accuracy for most thermochemical needs. However, caution should be exercised in interpreting the results because the potential for larger errors exists.

Before establishing the BAC-DFT method just described, we investigated a number of other DFT methods and basis sets to determine whether a BAC approach was feasible. Results of these investigations are also shown in Table 5. The double-ζ basis set [i.e., 6-31G(d)], although sufficiently accurate to provide molecular geometries, yields results that are inferior to those obtained from the triple-ζ basis set. The accuracy of the hybrid HF6/DFT4 method, with 60% HF and 40% DFT, is significantly worse than that obtained from the B3LYP, particularly for positive and negative ions. Neither method is as accurate as BAC$-$B3LYP/6-311++G(2df,2pd), and thus they were not pursued further.

The results of Tables 3 and 5 show that, as expected, the BAC-G2 method gives the best results. In fact, for the majority of species, the G2 method itself is sufficiently accurate without application of the BAC method. Furthermore, the BAC-G2 method requires far fewer parameters than the other methods.

TABLE 6: Differences between Theory and Experiment (Energies in kcal mol^{-1}) for Molecules with Highest Deviation for BAC-G2, BAC-MP2, and BAC-DFT Methods

molecule	BAC G2[a]	BAC MP2	BAC DFT	BAC MP4
carbon containing:				
C	1.5 (0.4)	−6.0	13.0	−8.6
CCl_4	0.1 (2.9)	−0.8	1.8	2.7
CF_4	0.4 (−5.6)	0.8	1.7	−0.3
^2CH	−1.3 (−0.6)	1.6	8.2	4.8
^4CH	2.3 (3.9)	−0.3	8.2	0.0
HNCO	−3.9 (−4.7)	−1.9	−3.8	−3.7
CH_2S	4.0 (3.4)	5.3	5.5	1.6
CH_3O	1.8 (0.5)	4.3	−2.4	2.4
$(NH_2)_2CO$	2.8 (1.8)	5.2	5.0	1.1
CN	2.4 (3.3)	6.0	5.2	4.2
C_2F_4	−1.3 (−7.8)	−0.7	−0.5	1.7
$C_2H_5NO_2$	−2.4 (−4.1)	−1.6	−0.2	−0.4
cyclopropene	1.3 (2.7)	4.5	0.1	1.8
cyclopropane	−0.5 (0.6)	1.4	−5.3	−1.1
methylenecyclopropane	−2.3 (−0.6)	1.6	−5.3	−1.1
cyclobutene	3.6 (5.3)	4.9	3.5	3.2
benzene	0.6 (3.6)	−0.4	1.7	−2.8
non-carbon-containing:				
ClO	1.0 (2.1)	6.5	−2.6	4.8
ClOO	5.2 (6.4)	−19.8	−0.5	12.3
OClO	−0.5 (4.8)	0.1	0.8	13.6
FNO	−5.2 (−7.7)	−5.2	−2.5	−1.2
HNO	1.5 (0.7)	3.0	4.0	−0.5
^3HNO	−0.5 (−0.4)	−5.9	−3.4	0.0
O_2NNO_2	1.0 (−2.6)	−8.6	−1.7	−5.5
HNNH	−3.1 (−3.5)	−1.0	−1.1	−3.9
NO_3	3.9 (1.5)	−8.2	−1.5	5.7
O_2	1.6 (2.4)	−8.4	2.0	0.0
O_3	−1.6 (−1.1)	−16.5	7.9	0.4
SO_2	0.7 (5.0)	−1.6	0.7	−1.6
SO_3	−0.6 (6.8)	1.1	0.3	2.8
$HOSO_2$	1.6 (8.1)	5.0	−1.9	−24.3 (−3.2)[b]

[a] G2 raw energies in parentheses. [b] BAC-MP4/MP2 method in parentheses (geometry optimized using the MP2 method).

Although this method is the most accurate of the ones presented in this paper, it is also computationally the most expensive. In contrast, the BAC-MP2 and BAC-DFT are computationally relatively inexpensive and are expected to scale better for larger molecules.

An additional significant advantage of the BAC-G2 method is that its parameters depend only on the type of atomic elements contained in the molecule. The old BAC-MP4 method and the BAC−hybrid methods, on the other hand, use parameters that depend on pairs of element types (i.e., bonds). This can lead to inaccurate values of the BAC parameters because they may be forced to agree with experiment when the amount of data available for compounds containing a given bond type is very limited. Thus, by using the BAC-G2 method, only a few reference compounds containing a given element are needed to determine the parameters for all compounds that contain that element. The BAC-G2 method can then be used to determine the BAC parameters for each bond type of interest for the BAC−hybrid method, for which experimental data are either not available or are not trusted. Of course, care should be taken because the G2 method uses similar methodologies in its procedure, that is, the MP perturbation theory expansions, so errors inherent in G2 could propagate.

There are certain species for which the accuracy of the G2 method is not sufficient (typically ± 2−3 kcal mol^{-1}, although this may vary depending on the application). Table 6 lists the species from the 143-molecule test set for which the BAC-G2 method yields the poorest agreement with experiment; results from the old BAC-MP4 method are included for comparison. This table also includes molecules that were not contained in the original test set of 143 species. As is evident from the table, species involving halogens and oxides are particularly problematic. A major part of this error may be due to molecular size. The errors in the G2 method tend to grow as the number of nonhydrogen atoms in the molecule increases. Unfortunately, G2 is rather limited in the number of heavy atoms it can handle, so it is difficult to determine the exact dependence of the error on molecular size. The BAC method attempts to correct for this size dependence with nearest-neighbor bond corrections (see eq 5), though the atomic (see eq 2) and molecular (see eq 4) corrections also depend on the size. As one can see from Table 6, the BAC-G2 results represent a significant improvement over the uncorrected values for the larger molecules.

Although the BAC-G2 method is highly accurate on average, the ability to accurately predict thermochemistry for unknown species is highly desirable. Thus, it is very important to reduce the maximum error (the last column of Tables 3 and 5) as much as possible. A small number here means not only that the BAC-G2 method yields accurate results, but also that it does so for all compounds, not just on the average. The 143 species used in the accuracy evaluation include not only stable compounds, but also represent different types of chemical bonds and resonant structures. Ions are not included because of uncertainty in the accuracy of these data. The heats of formation of all 143 neutral species are quite well established experimentally. This strongly suggests that for the largest disagreements between the BAC-G2 method and experiment (Table 6), that is, those that differ from experiment by more than 3 kcal mol^{-1}, the BAC-G2 values are more likely to be correct than the corresponding experimental values. This is particularly likely if the other BAC methods give similar results.

Table 6 also lists worst-case results for the BAC-MP2 and BAC-DFT methods (note that Table 6 includes results for the multireference cases that were not included in the set of 122 compounds used to optimize the parameters for this method). The results show that BAC-MP2 and BAC-DFT often perform poorly for the same types of species as the BAC-G2 method, although the size of the deviation can be much larger. It is encouraging, however, that the BAC-MP2 and BAC-DFT methods tend to not give the same incorrect results; for example, the BAC-MP2 deviation for molecular oxygen is −8.4 kcal mol^{-1}, whereas the BAC-DFT deviation is 2.0 kcal mol^{-1}. This may be due to the different approaches used by these methods to treat electron-pair correlation. Thus, by combining BAC-MP2 and BAC-DFT together within the BAC−hybrid method, one can use the difference between the two methods as an indicator of the accuracy (eq 13). It should also be noted that many of the highest-deviation cases represent unique electron correlation issues that will not extend to the larger molecules for which the BAC−hybrid method would be used. When in doubt about a chemical moiety, one should find a smaller molecular prototype and perform the calculation at a higher level of theory. One can then use the differences in the BAC methods for the smaller molecule as an additional BAC correction factor, that is, a chemical moiety correction factor for a specific chemical moiety.

$$E_{\mathrm{BAC}}(\mathrm{total}) = \sum_k E_{\mathrm{BAC-atom}}(A_k) + E_{\mathrm{BAC-molecule}} + \sum_{ij} E_{\mathrm{BAC-bond}}(A_iA_j) + E_{\mathrm{BAC-chemical\ moiety}} \qquad (1a)$$

This correction has been used to correct the results of the old BAC-MP2 method for large aromatic hydrocarbons using the BAC-MP4 results on a small aromatic species.

TABLE 7: Activation Energies ΔE_f Predicted by BAC-G2, BAC-MP2, BAC-DFT, and BAC-MP4 Methods for Various Reactions

reaction	BAC_{G2}[a]	BAC_{hybrid}	BAC_{DFT}	BAC_{MP4}
abstraction:				
$H_2 + H \leftrightarrow H + H_2$	9.2 (8.9)	8.7	7.6	8.8
	59.7	60.9	59.7	60.9
$CH_4 + OH \leftrightarrow CH_3 + H_2O$	5.0 (4.7)	1.6	−0.5	3.2
	−3.8	−6.7	−9.8	−5.1
$H_2 + O \leftrightarrow H + OH$	12.1 (11.9)	9.8	4.9	10.5
	70.6	69.4	66.0	70.3
$NH_3 + O \leftrightarrow NH_2 + OH\,^3A'$	12.2 (12.2)	11.4	2.8	12.2
	61.5	59.8	53.8	61.1
$NH_3 + O \leftrightarrow NH_2 + OH\,^3A''$	10.1 (10.1)	10.5	1.2	9.8
	59.4	58.9	52.2	58.7
addition:				
$C_2H_4 + H \leftrightarrow C_2H_5$	2.3 (1.9)	1.9	1.2	2.9
	66.2	67.7	66.1	67.3
$H + CO \leftrightarrow HCO$	2.4 (2.4)	2.4	−4.6	2.4
	26.9	27.3	23.0	25.7
shift:				
$C_2H_5 \leftrightarrow C_2H_5$ (1,2 H-shift)	40.6 (40.1)	39.9	38.7	42.3
	69.9	69.5	66.9	71.1
$HCN \leftrightarrow HNC$	44.4 (44.1)	47.9	47.1	45.1
	75.1	78.3	78.9	76.9
elimination:				
$C_2H_5NO_2 \leftrightarrow C_2H_4 + HONO$	48.3 (48.1)	46.4	41.3	48.5
	21.5	20.4	16.7	23.7
$C_2H_5Cl \leftrightarrow C_2H_4 + HCl$	59.0 (59.1)	61.8	52.2	70.5
	32.1	34.9	24.0	43.3

Energies are given for the forward direction (energies in kcal mol^{-1}). The heat of formation (298 K) for the transition-state structure is given in the second line for each reaction. [a] G2 raw energies in parentheses.

So far, we have shown how well this new form of correction works for stable species. Because the new form involves additive constants rather than terms that go to zero as bond lengths increase, it is necessary to establish the method's performance for transition states (i.e., reactions). In Table 7, we show the application of the BAC methods to the determination of activation energies for a variety of reaction types. The activation energy represents the difference between two heats of formation, that of the transition state structure and that of the reactants. Because errors can occur (or cancel) because of the accuracy of either the transition-state structure or the reactants, we provide both sets of information, that is, the difference is presented in the first row and the absolute value of the transition-state heat of formation is presented in the second row. These data help define the limitations of the methods and indicate classes of compounds for which the present procedures may be inadequate. The results in Table 7 suggest that the method is working at least as well as the original BAC-MP4 method, which we have found from experience yields reasonably accurate activation energies. Because the BAC-G2 method involves higher levels of theory than the other methods, we expect its predictions to be better than the set in Table 7. It is also encouraging that the differences between BAC-G2 and G2 (numbers in parentheses) are small; this is again due to cancellation of the terms in the BAC corrections.

Finally, we note that the new BAC methods include bond corrections (see eqs 5 and 6) that are not size-consistent, that is, they do not go to zero as the bond is stretched to infinity. Likewise, the molecular correction itself (see eq 3) is not size-consistent (see discussion in Section 2.2.2). Despite this, Table 7 indicates that the BAC method provides useful results for transition states. The table indicates that, in general, the BAC-G2 and BAC−hybrid methods give comparable results with the BAC-MP4 method. Although one cannot compare the absolute heats of formation of transition states directly with tables of experimental values, one expects the BAC-G2 results to be the most reliable. Because most of the BAC corrections cancel in the determination of the activation energy and because the values of the B_{atom} term in the BAC-G2 method are small (see Table 1), it is not surprising that the raw G2 results are nearly identical to the BAC-G2 results. In contrast, the BAC-DFT method tends to give activation energies that are too low compared with the DFT method itself. This effect results from setting the parameter α equal to zero in the bond correction term (eq 5). A negative α would improve the agreement in Table 7 (as well as the overall agreement in Table 6) because it would result in a BAC correction that would go to negative infinity as the bond distance goes to infinity. Thus, the use of the BAC-MP2 method for determining the BAC−hybrid energy is particularly preferred when investigating reaction energies. Compared with the BAC-MP4 method, the new BAC−hybrid method greatly improves the activation energy for elimination of HCl from C_2H_5Cl because of the presence of diffuse functions that can represent the ionic nature of the transition state structure. Overall, the BAC−hybrid method and BAC-G2 method appear to work well for a wide variety of reaction mechanisms.

5. Summary and Conclusions

A new, generic method for applying empirical corrections to ab initio electronic structure calculations has been developed that extends the capabilities of the earlier BAC-MP4 method. The method uses a parametrization that can treat positive and negative ions as well as neutrals. It can be applied to virtually any ab initio method, including electron correlation methods such as MP perturbation and coupled-cluster theories, hybrid methods such as G2, and DFT. In this paper, we apply the corrections to G2, MP2, and DFT to develop new BAC methods for these levels of theory. The approaches used are designed to balance the computational time between optimization searching, frequency evaluation, and single-point energy calculations. To permit better treatment of ionic character, diffuse basis functions are used. The BAC-G2 method provides high accuracy for small species. The new BAC−hybrid method uses the DFT procedure for improved geometry optimization and frequency evaluation and both the MP2 and DFT procedures for the single-point calculations. The BAC-DFT and BAC-MP2 methods provide complementary treatments of electron correlation, providing an error estimate at these low levels of theory. The DFT and MP2 methods are well suited for massively parallel and distributed computational platforms and should scale well for large molecules.

The new BAC parametrization should be readily extendible to other quantum chemistry procedures being developed. It is also possible to develop other hybrid methods, for example, by mixing density functional geometry optimizations with coupled-cluster methods. It should thus be possible to derive parameters for a BAC−CCSD(T) method, in which the geometry is determined by the DFT method, whereas the single-point calculation is performed using the coupled-cluster with triples method, CCSD(T).[9] Application of the new BAC method to CBS procedures,[22] analogous to the implementation of the BAC-G2 method, or to a combined CCSD(T)/CBS[23] method, should also be feasible. Finally, we anticipate that the BAC method is applicable to new algorithms such as quadratically scaling local correlation methods. Errors introduced in the local correlation approximation are likely to be systematic, so the BAC corrections should be applicable. These new BAC methods can be used to determine the parameters needed for BAC methods using the lower levels of theory. To date, the new methods have been applied primarily to small molecules. Thus, additional work is required to determine the magnitude of nonlinear errors that accumulate as the size of a molecule increases.

Acknowledgment. We are deeply indebted to W. A. Goddard for the breadth of contributions that he has made to the field of computational chemistry. Although the number of computational tools has expanded and computer power has increased enormously over the years, it still takes the knowledge of the chemist to identify which reactions should be investigated, and where one should look on the potential energy surface. Goddard's contributions to understanding the nature of the chemical bond will continue to be of particular significance in our ability to unravel the complexities of chemical reaction pathways. This work was supported by the Laboratory Research and Development Program at Sandia National Laboratories.

References and Notes

(1) Irikura, K. K.; Frurip, D. J. In *Computational Thermochemistry*; ACS Symposium Series 677; American Chemical Society: Washington, DC, 1998.

(2) Arai, T. *Rev. Mod. Phys.* **1960**, *32*, 370.

(3) Ho, P.; Melius, C. F. *J. Phys. Chem.* **1990**, *94*, 5120.

(4) Melius, C. F. In *Chemistry and Physics of Energetic Materials*; Bulusu, S. N., Ed.; Kluwer Academic Publishers: Dorderecht, 1990; p 21.

(5) Melius, C. F.; Colvin, M. E.; Marinov, N. M.; Pitz, W. J.; Senkan, S. M. In *Twenty-Sixth Symposium (International) on Combustion*, The Combustion Institute: Pittsburgh, 1996; p 685.

(6) Zachariah, M. R.; Melius, C. F. *J. Phys. Chem.* **1997**, *101*, 913.

(7) Garrett, B. C.; Koszykowski, M. L.; Melius, C. F.; Page, M. *J. Phys. Chem.* **1990**, *94*, 7096.

(8) Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. *J. Chem. Phys.* **1991**, *94*, 7221.

(9) Watts, J. D.; Gauss, J.; Bartlett, R. J. J. *J. Chem. Phys.* **1993**, *98*, 8718.

(10) Pople, J. A.; Krishnan, R.; Schlegel, H. B.; Binkley, J. S. *Int. J. Quantum Chem.* **1979**, *S13*, 225.

(11) Frisch, M. J.; Trucks, G. W.; Head-Gordon, M.; Gill, P. M. W.; Wong, M. W.; Foresman, J. B.; Johnson, B. G.; Schlegel, H. B.; Robb, M. A.; Replogle, E. S.; Gomperts, R.; Andres, J. L.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C.; Martin, R. L.; Defrees, D. J.; Baker, J.; Stewart, J. J. P.; Pople, J. A. *Gaussian 94*, revision D.3; Gaussian, Inc.: Pittsburgh, 1992.

(12) Kee, R. J.; Rupley, F. M.; Miller, J. A. *The CHEMKIN Thermodynamic Data Base*; Report SAND-87-8215; Sandia National Laboratories: Albuquerque, NM, 1987.

(13) Roothaan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69.

(14) Pople, J. A.; Nesbet, R. K. *J. Chem. Phys.* **1954**, *22*, 571.

(15) Hariharan, P. c.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213.

(16) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. *J. Chem. Phys.* **1982**, *77*, 3654.

(17) Pople, J. A.; Schlegel, H. B.; Krishnan, R.; DeFrees, D. J.; Binkley, J. S.; Frisch, M. J.; Whiteside, R. A. *Int. J. Quantum Chem.* **1981**, *S15*, 269.

(18) Chase, M. W.; Davies, C. A.; Downey, J. R.; Frurip, D. J.; McDonald, R. A.; Szverud, A. N. *J. Phys. Chem. Ref. Data* **1985**, *1985*, 14.

(19) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(20) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

(21) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.

(22) Petersson, G. A.; Irikura, K. K.; Frurip, D. J. In *Computational Thermochemistry*; ACS Symposium Series; Amercian Chemical Society: Washington, DC, 1998; p 237.

(23) Martin, J. M. L. *Chem. Phys. Lett.* **1996**, *259*, 669.

(24) Detailed results for all molecules examined in this study can be found on the following web page: http://z.ca.sandia.gov/~melius.

(25) If there is ambiguity in what the BAC method refers to, additional information should be added to the naming scheme. The full naming scheme is BAC(Version)−Method/Basis Set//Optimization Method/Basis Set. The version refers to the BAC parametrization version for a particular method and basis set. Thus, the full name of the new BAC-MP2 method would be BAC(2.0)-MP2/6-311++G(2df,2pd)//B3LYP/6-31G(d), whereas the old BAC-MP2 method would be BAC(1.0)-MP2/6-31G(d,p)//HF/6-31G(d). However, BAC(2)-MP2 and BAC(1)-MP2 would be sufficient to distinguish these two methods if no other BAC-MP2 methods were being discussed. Thus, the notations in Table 5 are sufficient to distinguish between the two B3LYP methods. The BAC-DFT notation can be used to refer to the BAC-(2.0)-B3LYP/6-311++G(2df,2pd)//B3LYP/6-31G(d) method if it does not cause confusion with some other DFT method. The BAC-hybrid notation can be used instead of BAC(2.0)-MP2 to denote that the BAC(2.0)-B3LYP/6-311++G(2df,2pd)//B3LYP/6-31G(d) has been used for error estimates.