

Scaling Factors and Uncertainties for *ab Initio* Anharmonic Vibrational Frequencies

Russell D. Johnson III,^{*,†} Karl K. Irikura,[†] Raghu N. Kacker,[‡] and Rüdiger Kessel[‡]

Chemical and Biochemical Reference Data Division and Mathematical and Computational Sciences Division, National Institute of Standards and Technology, Gaithersburg, Maryland 20899-8320

Received May 10, 2010

Abstract: To predict the vibrational spectra of molecules, *ab initio* calculations are often used to compute harmonic frequencies, which are usually scaled by empirical factors as an approximate correction for errors in the force constants and for anharmonic effects. Anharmonic computations of fundamental frequencies are becoming increasingly popular. We report scaling factors, along with their associated uncertainties, for anharmonic (second-order perturbation theory) predictions from HF, MP2, and B3LYP calculations using the 6-31G(d) and 6-31+G(d,p) basis sets. Different scaling factors are appropriate for low- and high-frequency vibrations. The method of analysis is based upon the *Guide to the Expression of Uncertainty in Measurement*, published by the International Organization for Standardization (ISO). The data used are from the Computational Chemistry Comparison and Benchmark Database (CCCBDB), maintained by the National Institute of Standards and Technology, which includes more than 3939 independent vibrations for 358 molecules.

1. Introduction

One of the most popular uses of computational quantum chemistry models is to predict vibrational spectra. This is generally done in the harmonic approximation, in which the potential energy function (PEF) is taken as a truncated, second-order Taylor series. The neglect of higher-order curvature causes the predictions to deviate from experimental observations of fundamental frequencies. Moreover, the harmonic force constants are distorted by theoretical and numerical approximations inherent in the electronic structure calculations used to compute the PEF. As an approximate correction for these two sources of error, an empirical scaling factor (that is, a multiplicative correction) is usually applied to the theoretical frequencies. The value for the scaling factor is typically determined by least-squares fitting to a set of experimental vibrational frequencies.

It is possible to include higher-order terms in the Taylor series expansion of the PEF, although at significant compu-

tational cost. This is expected to improve the accuracy of the predictions. The anharmonic vibrational problem is usually solved using second-order perturbation theory with a harmonic reference wave function (VPT2), often with some treatment of Fermi resonances.¹ The other popular alternative is vibrational mean-field theory (VSCF) and its more sophisticated derivatives.^{2–4} Compared with VPT2, VSCF does not require a well-behaved Taylor series expansion for the PEF, is resistant to problems arising from near degeneracies, and is computationally expensive. Besides these popular approaches, specialized techniques are used for high-precision predictions for small molecules.^{5,6}

The present study, part of a series directed toward “virtual measurements”,⁷ is restricted to anharmonic vibrational frequencies predicted using VPT2. The goals are (1) to determine empirical scaling factors, along with their associated uncertainties, for some popular computational models and (2) to determine whether such scaling is helpful.

Our analysis is based upon the *Guide to the Expression of Uncertainty in Measurement* (GUM),⁸ which is a de facto international standard for quantifying the uncertainty in all

* Corresponding author e-mail: russell.johnson@nist.gov.

[†] Chemical and Biochemical Reference Data Division.

[‡] Mathematical and Computational Sciences Division.

Table 1. Repeatability of Replicate Vibrational Calculations for the 27 Vibrations of *n*-Propane, Using the 6-31G(d) Basis Sets

theory	HF	HF	B3LYP	B3LYP	B3LYP	B3LYP
convergence ^a	default	tight	tight	tight	tight	tight
grid ^b	n.a.	n.a.	(75, 302) ^c	(99, 590) ^d	(96, 32, 64)	(131, 974)
max. $\sigma(\omega_i)$ ^e	0.65 cm ⁻¹	0.02	0.02	0.02	0.03	0.02
max. $\sigma(\nu_i)$ ^f	11.7 cm ⁻¹	0.03	1.2	0.12	0.14	0.14
repetitions	998	998	998	465	128	176
$(\omega_i - \omega_{i,\text{ref}})$ ^g	(-0.1, 0.1)	0	(-4.4, 1.2)	0	(-0.1, 0.3)	(-0.1, 0.3)
$(\nu_i - \nu_{i,\text{ref}})$ ^h	(-3.6, 3.0)	0	(-6.5, 35.3)	0	(-4.5, 0.0)	(-1.7, 0.5)

^a Geometry convergence criteria (rms force: default = 3×10^{-4} au, tight = 1×10^{-5} au). ^b Integration grid for density functional calculations written either (radial, angular) or (radial, polar, azimuthal). ^c "Fine" pruned grid; software default. ^d "Ultrafine" pruned grid. ^e Largest standard deviation for a harmonic frequency, in cm⁻¹ units. ^f Largest standard deviation for an anharmonic fundamental frequency, in cm⁻¹ units. ^g Harmonic frequency shifts relative to reference calculation, written as a range (min, max) in cm⁻¹ units. Reference calculations are HF/tight for HF and B3LYP/tight/(99, 590) for B3LYP. ^h Anharmonic frequency shifts relative to reference calculation, as for harmonic frequencies.

kinds of measurements, including those determined from computational models.

2. Methodology

The procedure used here for determining scaling factors and their uncertainties closely follows that used for the scaling of harmonic vibrational frequencies,⁹ as slightly revised in a subsequent study of vibrational zero-point energies.⁷ Thus, it is presented here relatively briefly. Greater detail may be found in the earlier reports.

The scaling factor for a vibrational frequency of interest is obtained from a class of vibrational frequencies for which benchmark values are available. A class of vibrational frequencies is appropriate only if it meets three conditions.¹⁰ (1) The bias for the target frequency is believed to be of similar value to those in the class. (2) The (estimated) biases in the class have an approximately normal and acceptably narrow distribution. (3) The number of vibrational frequencies in the class is reasonably large. When classifying reference frequencies, technical knowledge and understanding are helpful. In the present study, benchmark, gas-phase, experimental values are taken from the Computational Chemistry Comparison and Benchmark Database (CCCB-DB).¹¹ The molecules included in this study are listed in the Supporting Information.

Suppose the value of the measurand (the target fundamental vibrational frequency of interest) is Y_0 and the corresponding value from a theoretical model is X_0 . Suppose the actual computed estimate of X_0 is x_0 with standard uncertainty $u(x_0)$. The *fractional bias* in x_0 is the ratio X_0/Y_0 . When the fractional bias in x_0 is believed to deviate significantly from unity, the GUM recommends that the estimate x_0 be scaled to counter its bias. The GUM requires that the measurement equation, which expresses the mathematical relationship between input quantities and the measurand, include an input variable for each component of the uncertainty. The measurement equation that corresponds to the fractional bias X_0/Y_0 is

$$Y_0 = C_0 X_0 \quad (1)$$

where C_0 is a scaling variable with a probability density function (pdf) representing the state of knowledge about the reciprocal (Y_0/X_0) of the fractional bias. Suppose c_0 and $u(c_0)$

are the expected value and the standard deviation of the state-of-knowledge pdf for C_0 . Then

$$y_0 = c_0 x_0 \quad (2)$$

The expected value $E(C_0) = c_0$ is a scaling factor applied to x_0 . By the common procedure of linear propagation of uncertainty,

$$u_r^2(y_0) \approx u_r^2(x_0) + u_r^2(c_0) \quad (3)$$

where we have taken the correlation coefficient $R(X_0, C_0) = r(x_0, c_0) = 0$ because the probability distributions for X_0 and C_0 are specified independently. The quantity $u_r(q) = u(q)/q$ is the relative standard uncertainty associated with the quantity q .

The uncertainty $u(x_0)$, which represents the dispersion in results that are obtained from nominally equivalent quantum chemistry calculations, arises from a variety of small contributions. For example, nonzero convergence thresholds create dependence upon the choice of the initial geometry and orbitals. We estimated the magnitude of this uncertainty using repeated calculations for propane (C_3H_8), which has 27 vibrational frequencies, starting from randomized initial coordinates (displacements up to $\pm 25\%$ of each z-matrix variable, from a uniform distribution). A calculation was discarded if its energy deviated by more than 10 standard deviations from the mean of all equivalent calculations. The results of this exercise are summarized in Table 1. The Hartree–Fock results indicate that tight geometry convergence criteria are necessary for these anharmonic calculations. The B3LYP results indicate that at least an "ultrafine" integration grid is also needed; values of $u(x_i)$ are only somewhat smaller, but the mean values are substantially different from those of the default grid. These conclusions affirm the original recommendations by Barone.¹²

The standard deviation of the distribution for the frequency x_i is an estimate of the standard uncertainty $u(x_i)$. The calculated anharmonic frequencies range approximately from 220 to 3100 cm⁻¹, so the calculations listed in Table 1 have values of $u_r(x_i)$ less than 0.0006, which can be neglected. This was found earlier for harmonic vibrational frequencies (using default criteria for geometry convergence).⁹ Taking $u_r(x_0) \approx 0$, from eq 3, we have $u(y_0) \approx y_0 u_r(c_0) = (y_0/c_0) u(c_0)$. Using eq 2 we have

$$u(y_0) \approx x_0 u(c_0) \quad (4)$$

Evaluated reference data are used to determine the scaling factor c_0 and the uncertainty $u(c_0)$, as described below. Then, the estimated value of a particular target vibrational frequency, y_0 , and its associated uncertainty, $u(y_0)$, can be determined from eqs 2 and 4, respectively.

Suppose the values for the fundamental vibrational frequencies in the specified class are Y_i and the corresponding values from a theoretical model are X_i , where $1 \leq i \leq m$, with m being the number of frequencies in the class. Estimates x_i (for X_i) are taken from actual calculations. Benchmark, experimental values z_i (estimates for Y_i) and their associated standard uncertainties $u(z_i)$ are taken from the CCCBDB.¹¹ The ratio $b_i = x_i/z_i$ is an estimate for the fractional bias X_i/Y_i , and the corresponding estimated scaling factor is

$$c_i = 1/b_i = z_i/x_i \quad (5)$$

Let C_i be a scaling variable for the fractional bias X_i/Y_i . The expected value and standard deviation of C_i are approximately c_i and $u(c_i)$, respectively. Taking $u(x_i) \approx 0$, as discussed above, we have

$$u(c_i) \approx u(z_i)/x_i \quad (6)$$

Define a nominal scaling variable $C_N = \sum_i k_i C_i$, where the k_i are normalized weights described below. Then the expected value of a state-of-knowledge pdf for C_N is

$$c_N = \sum_i k_i c_i \quad (7)$$

and the standard deviation is

$$u(c_N) = [\sum_i k_i^2 u^2(c_i) + \sum_i \sum_{j \neq i} k_i k_j u(c_i) u(c_j) r(c_i, c_j)]^{1/2} \quad (8)$$

When the state-of-knowledge pdf's of C_1, \dots, C_m are uncorrelated, the standard deviation (8) reduces to

$$u(c_N) = [\sum_i k_i^2 u^2(c_i)]^{1/2} \quad (9)$$

Since the bias in the unknown, target frequency is thought to be similar to the biases of the selected class of frequencies, it is reasonable to estimate the target scaling factor c_0 as $c_0 = c_N$. However, c_0 is likely to differ from c_N as suggested by the scatter of the estimated correction factors c_1, \dots, c_m about c_N . Therefore, the uncertainty $u(c_0)$ associated with $c_0 = c_N$ should be larger than the uncertainty $u(c_N)$. To determine $u(c_0)$, we need the following measurement equation for the correction variable C_0

$$C_0 = C_N + \delta C_0 \quad (10)$$

where δC_0 is a variable with a state-of-knowledge pdf for $C_0 - C_N$. A pdf for δC_0 is specified independently of the pdf for C_N ; therefore, they are uncorrelated. Suppose the expected value and standard deviation of δC_0 are δc_0 and $u(\delta c_0)$, respectively. Then

$$c_0 = c_N + \delta c_0 \quad (11)$$

and

$$u(c_0) = [u^2(c_N) + u^2(\delta c_0)]^{1/2} \quad (12)$$

The available information for specifying a state-of-knowledge pdf for δC_0 is the set of deviations of the estimated scaling factors c_i from the nominal scaling factor c_N . A useful class of pdf's that could be assigned to δC_0 are discrete probability distributions that assign a non-negative, normalized weight h_i to each deviation $(c_i - c_N)$.¹³ Then, the expected value is

$$\delta c_0 = \sum_i h_i (c_i - c_N) = \sum_i h_i c_i - c_N = \sum_i h_i c_i - \sum_i k_i c_i \quad (13)$$

and the variance is

$$u^2(\delta c_0) = \sum_i h_i [(c_i - c_N) - \delta c_0]^2 \quad (14)$$

We wish to use the nominal scaling factor c_N as the correction factor c_0 ; so from eq 11, we require that $\delta c_0 = 0$. From eq 13, this means that $\sum_i (h_i - k_i) c_i = 0$, which is satisfied if $h_i = k_i$ for all $i = 1, \dots, m$. Making these substitutions into eq 14 yields

$$u(\delta c_0) = \left[\sum_i k_i (c_i - c_N)^2 \right]^{1/2} \quad (15)$$

The usual practice when determining scaling factors is to fit the linear model

$$z_i = c_N x_i + e_i \quad (16)$$

by minimizing the least-squares objective function

$$\Delta^2 = \sum_i e_i^2 = \sum_i (z_i - c_N x_i)^2 \quad (17)$$

The well-known solution is

$$c_N = \frac{\sum_i x_i z_i}{\sum_i x_i^2} = \frac{\sum_i x_i^2 c_i}{\sum_i x_i^2} \quad (18)$$

which corresponds to eq 7 with weights

$$k_i = \frac{x_i^2}{\sum_i x_i^2} \quad (19)$$

To summarize, we have

$$c_0 = \frac{\sum_i x_i^2 c_i}{\sum_i x_i^2} \quad (20)$$

and, by combining eqs 9, 12, 15, and 19,

$$u(c_0) = \left[\frac{\sum_i x_i^4 u^2(c_i)}{(\sum_i x_i^2)^2} + \frac{\sum_i x_i^2 (c_i - c_0)^2}{\sum_i x_i^2} \right]^{1/2} \quad (21)$$

where c_i and $u(c_i)$ are given by eqs 5 and 6, respectively.

Electronic Structure Calculations. Anharmonic vibrational frequencies were computed using a partial quartic force field computed by using finite differences of analytical second derivatives.¹ All open-shell calculations were spin-unrestricted. Core orbitals were frozen, i.e., uncorrelated, in all MP2 calculations, including those involving group 1 (alkali) and group 2 (alkaline-earth) metals. Upon the basis of the results in Table 1, all calculations were done using tight geometry convergence criteria, and B3LYP calculations were done using the “ultrafine” grid (99 radial and 590 angular points). Because of software limitations, molecules were excluded if they possessed a 3-fold or higher symmetry axis. All computations were performed using the Gaussian 03 software package.^{14,15}

Where convenient, we use the symbol ω_{calc} to denote a computed, unscaled harmonic vibrational frequency and the symbol ν_{calc} to denote a computed, unscaled anharmonic frequency. The symbol ν_{expt} denotes an observed, experimental fundamental frequency.

3. Results and Discussion

Classification of Vibrational Frequencies. As stated in section 2, the scaling factor for a particular target frequency should be determined using reference data for an appropriate class of frequencies. When determining scaling factors for (harmonic) vibrational frequencies, it is traditional to consider all frequencies as a single class. Figure 1 shows a histogram of the estimated biases b_i , from eq 5, for anharmonic frequencies from the inexpensive HF/6-31G(d) model. Unlike the corresponding distribution for harmonic scaling factors,⁹ the histogram in Figure 1 is flat-topped. Thus, this class fails to satisfy the second criterion presented in section 2; that is, the distribution does not appear normal and reasonably narrow. We found it helpful to use two scaling factors, one for X–H stretches and one for all other vibrations. However, a simpler and nearly equivalent classification is between low and high (harmonic) frequencies. Earlier studies of harmonic frequencies have noted this distinction,^{16–22} often attributing it to anharmonicity. A plot of anharmonic bias against harmonic frequency, Figure 2, suggests that the boundary between “low” and “high” harmonic frequencies be set near 2700 cm^{−1}. Thus, we have chosen to use the value $\omega_{\text{calc}} = 2700$ cm^{−1} as the boundary for HF/6-31G(d). Histograms for both frequency ranges, displayed as line plots, are included in Figure 1. Analogous plots (see the Supporting Information) for the frequencies obtained from the other five models [B3LYP/6-31G(d), MP2/6-31G(d), HF/6-31+G(d,p), MP2/6-31+G(d,p), and B3LYP/6-31+G(d,p)] suggest setting the boundary between low and high frequencies at 2500 cm^{−1} for MP2 and 2600 cm^{−1} for B3LYP calculations. The MP2 and B3LYP models display smaller discrepancies between low and high frequencies than do the HF models. This suggests that the discrepancy in HF/6-31G(d) scaling

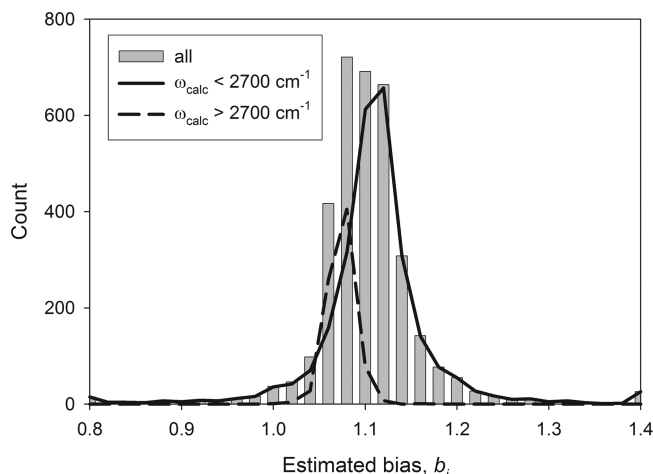


Figure 1. Histograms of estimated biases b_i for anharmonic HF/6-31G(d) calculations of 3446 frequencies of 215 molecules. The solid line is for low frequencies, the dashed line for high frequencies, and the bars for all frequencies combined.

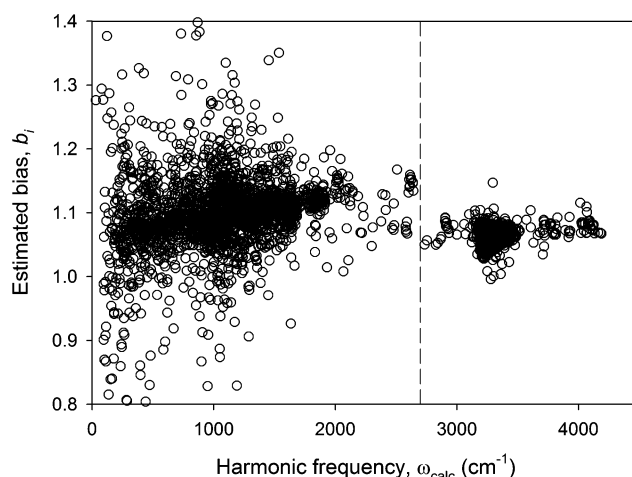


Figure 2. Distribution of estimated biases b_i for anharmonic HF/6-31G(d) calculations of 3446 frequencies of 215 molecules, plotted against calculated harmonic frequency, ω_{calc} . The dashed vertical line is drawn at 2700 cm^{−1}.

factors for high and low frequencies is primarily due to missing electron correlation. That is, X–H stretching vibrations are less affected by electron correlation than are other molecular vibrations, presumably because of their lower electron density.²³ As noted by a reviewer, the smaller values of $u(c_0)$ for high frequencies in all models (Table 2) can be explained by the greater similarity among the X–H stretching motions.

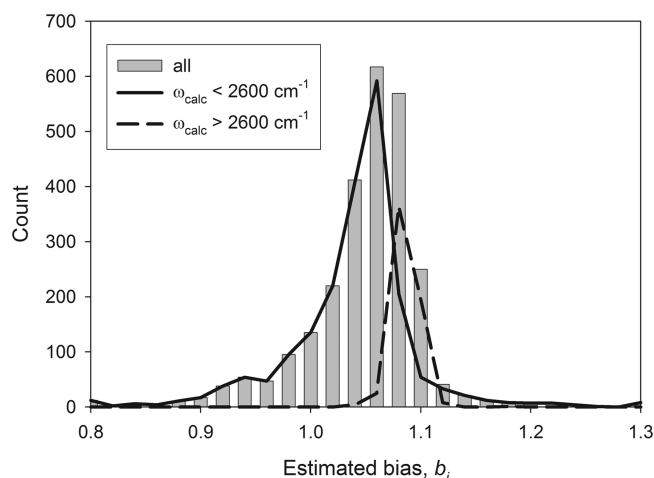
For MP2/6-31+G(d,p) calculations, the histogram of estimated bias for low-frequency vibrations is asymmetric (Figure 3). This indicates that more detailed classification is appropriate. However, we have not pursued this further in the present study.

Experimental Uncertainties Are Negligible. The first term of eq 21 is the contribution from the experimental uncertainties $u(z_i)$, via eq 6. As shown below, and as found previously for scaling harmonic frequencies,⁹ this term is negligible. We consider the popular B3LYP/6-31G(d) quantum chemistry model. Our data set for this model contains

Table 2. Anharmonic^a Scaling Factors, c_0 , and Their Associated Standard Uncertainties, $u(c_0)$

	$\omega_{\text{calc}} < \omega_{\text{high}}$	$\omega_{\text{calc}} > \omega_{\text{high}}$
HF/6-31G(d) ^b	0.9047 ± 0.0366 (2659)	0.9398 ± 0.0126 (783)
MP2/6-31G(d) ^c	0.9675 ± 0.0493 (2291)	0.9797 ± 0.0142 (673)
B3LYP/6-31G(d) ^d	0.9863 ± 0.0339 (2299)	1.0015 ± 0.0131 (675)
HF/6-31+G(d,p) ^e	0.9117 ± 0.0351 (2311)	0.9438 ± 0.0159 (675)
MP2/6-31+G(d,p) ^f	0.9822 ± 0.0352 (1998)	0.9698 ± 0.0117 (598)
B3LYP/6-31+G(d,p) ^g	0.9988 ± 0.0332 (2196)	1.0005 ± 0.0110 (647)

^a Using “tight” geometry optimization and “ultrafine” DFT integration grid. The number of vibrations in each data set is printed between parentheses. ^b $\omega_{\text{high}} = 2700 \text{ cm}^{-1}$. If all vibrations are considered a single class, $c_0 = 0.9285 \pm 0.0284$ (3442 frequencies). ^c $\omega_{\text{high}} = 2600 \text{ cm}^{-1}$. If all vibrations are considered a single class, $c_0 = 0.9759 \pm 0.0304$ (2964 frequencies). ^d $\omega_{\text{high}} = 2500 \text{ cm}^{-1}$. If all vibrations are considered a single class, $c_0 = 0.9967 \pm 0.0230$ (2974 frequencies). ^e $\omega_{\text{high}} = 2700 \text{ cm}^{-1}$. If all vibrations are considered a single class, $c_0 = 0.9335 \pm 0.0281$ (2986 frequencies). ^f $\omega_{\text{high}} = 2600 \text{ cm}^{-1}$. If all vibrations are considered a single class, $c_0 = 0.9735 \pm 0.0223$ (2596 frequencies). ^g $\omega_{\text{high}} = 2500 \text{ cm}^{-1}$. If all vibrations are considered a single class, $c_0 = 1.0000 \pm 0.0205$ (2843 frequencies).

**Figure 3.** Histograms of estimated biases b_i for anharmonic frozen-core MP2/6-31+G(d,p) calculations of 2599 frequencies of 176 molecules. The solid line is for low frequencies, the dashed line for high frequencies, and the bars for all frequencies combined.

$m = 495$ frequencies for which experimental uncertainties are known, primarily from the compilation by Shimanouchi.^{21,22} From eqs 20 and 21, we obtain $c_0 = 1.0001$ and $u(c_0) = 0.0203$ when all 495 vibrations are considered together. If the first term of eq 21 is ignored, we obtain the simpler, more approximate expression

$$u(c_0) \approx \left[\frac{\sum_i x_i^2 (c_i - c_N)^2}{\sum_i x_i^2} \right]^{1/2} \quad (22)$$

This yields a value of $u(c_0)$ that is smaller by only 0.000002. Similar results are obtained when only high or only low frequencies are considered. Thus, we conclude that eq 22 is adequate for estimating $u(c_0)$. Since uncertainties are not available for all of the experimental frequencies employed, eq 22 is used in all of the calculations described below.

Recommended Scaling Factors. Table 2 lists the estimated corrections for bias c_0 and estimated, associated

Table 3. Number of Anharmonic Frequency Predictions Improved (Worsened) Significantly^a by Scaling^b

	$\omega_{\text{calc}} < \omega_{\text{high}}$	$\omega_{\text{calc}} > \omega_{\text{high}}$	all ω_{calc} together
HF/6-31G(d)	2283 (218)	767 (9)	3129 (182)
MP2/6-31G(d)	1359 (492)	562 (81)	2002 (473)
B3LYP/6-31G(d)	889 (447)	0 (0)	55 (71)
HF/6-31+G(d,p)	1987 (198)	648 (16)	2687 (172)
MP2/6-31+G(d,p)	1060 (385)	556 (34)	1506 (598)
B3LYP/6-31+G(d,p)	0 (0)	0 (0)	0 (0)

^a By at least 10 cm^{-1} . ^b Computational details, ω_{high} values, and scaling factors from Table 2 and its footnotes.

standard uncertainties $u(c_0)$ for six theoretical models, i.e., combinations of theory and a one-electron basis set. In preparing the table, vibrations were excluded when the computed anharmonic correction exceeded 50%, that is, when $|\nu_{\text{calc}} - \omega_{\text{calc}}|/\omega_{\text{calc}} > 0.5$. Such large corrections are unusual and may indicate situations where vibrational perturbation theory is unreliable. If a single scaling factor is desired for the entire frequency range, it may be found among the footnotes to Table 2.

Is Scaling Helpful? Some of the scaling factors in Table 2 are within one standard uncertainty of unity, suggesting that scaling is not helpful. To examine this, we count the number of individual frequency predictions that are improved (or worsened) significantly by scaling, that is, the number of vibrations for which $|\nu_{\text{calc}} - \nu_{\text{exptl}}|$ is less (or greater) than $|\nu_{\text{calc}} - \nu_{\text{exptl}}|$ by at least 10 cm^{-1} . The results are listed in Table 3. Where the scaling factors differ noticeably from unity, many more frequencies are improved significantly by scaling than are degraded. We conclude that scaling is usually helpful. For B3LYP/6-31G(d) at higher frequencies ($\omega_{\text{calc}} > 2500 \text{ cm}^{-1}$) and for B3LYP/6-31+G(d,p) at all frequencies, the scaling factors are essentially unity, $c_0 = 1$. However, the values of $u(c_0)$ should not be approximated as zero. For example, high-frequency predictions using the B3LYP/6-31G(d) model may be scaled by the trivial factor of 1, with the resulting prediction having a relative standard uncertainty of 1.3%.

Effects of Basis Set. Although only two basis sets were studied, their comparison suggests that average scaling factors do not depend strongly upon the basis set, as was found earlier for harmonic scaling factors.⁹ The MP2 method is expected to be most sensitive because it uses virtual orbitals, whose number increases with basis-set size. The data of Table 2 substantiate this expectation. When the basis set is enlarged from 6-31G(d) to 6-31+G(d,p), low-frequency scaling factors increase by 0.007, 0.015, and 0.013 for HF, MP2, and B3LYP theories, respectively. The associated uncertainties change only for MP2. The corresponding high-frequency scaling factors change noticeably only for MP2, for which they decrease by 0.010.

Discrepancies between Software Versions. After this manuscript was completed, we learned that different versions of the software^{14,15} sometimes yield different results. To test whether this affects our conclusions, HF/6-31G(d) calculations were run on 215 molecules (3442 vibrational frequencies) using two software versions, B05 and E01. As before, frequencies were discarded when the anharmonic correction exceeded 50% of the harmonic frequency. Differences in ν_{calc}

between the different software versions ranged from -192 (for ν_2 of CH_2S) to $+247\text{ cm}^{-1}$ (for ν_8 of CH_2F_2), with a mean value of -0.8 cm^{-1} and standard deviation of 17.6 cm^{-1} . Values for 650 frequencies (19% of the set) differed by 10 cm^{-1} or more. However, the resulting scaling factors, 0.9047 ± 0.0366 and 0.9037 ± 0.0381 for low frequencies and 0.9398 ± 0.0126 and 0.9418 ± 0.0129 for high frequencies, do not differ significantly. Thus, our conclusions are unaffected.

Although tangential to the present study, we investigated this discrepancy somewhat further, as suggested by both reviewers. All harmonic frequencies are in excellent agreement between the two software versions. However, version E01 treats both vibrations (mentioned above) as affected by Fermi resonances, while version B05 does not. Moreover, there are some differences in cubic and quartic force constants. We were unable to devise any combination of options that would allow one version to reproduce the results of the other. As an independent test, the calculations on CH_2S and CH_2F_2 were repeated using the ACES2-MAB package.²⁴ This agreed with Gaussian03 for all harmonic frequencies and agreed with version B05 for ν_2 (CH_2S) and ν_8 (CH_2F_2). However, it disagreed seriously for ν_1 of both molecules [with B05, discrepancies of 193 cm^{-1} (CH_2S) and -247 cm^{-1} (CH_2F_2); with E01, discrepancies of 173 cm^{-1} (CH_2S) and -256 cm^{-1} (CH_2F_2)]. It is clear that the results are affected by choices made within different implementations. It is not clear whether any implementation is more “correct” than another.

4. Conclusions

Multiplicative scaling is an effective strategy for improving predictions from anharmonic vibrational calculations. The associated uncertainties reveal that the scaling factors have only two significant figures. However, we have reported four figures to conform to current, common practice for reporting harmonic scaling factors.

Slightly different scaling factors are recommended for low and high frequencies. The difference is greatest for HF calculations and smallest for B3LYP calculations. The uncertainties associated with low-frequency scaling factors are about 3 times as large as those for high-frequency scaling factors. For B3LYP/6-31+G(d,p) calculations, both high- and low-frequency scaling factors are negligibly different from unity.

Scaling factors depend only weakly upon the basis set. The scaling factors for high frequencies (X–H stretches) change less with the basis set than do the scaling factors for low frequencies.

Acknowledgment. Many of the calculations were performed on the “Biowulf” Linux cluster at the National Institutes of Health, Bethesda, MD (<http://biowulf.nih.gov>).

Supporting Information Available: Lists of molecules (one table) and histograms of estimated biases (four figures) and distributions of estimated biases (five figures) for theoretical models considered in this study. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References

- (1) Barone, V. *J. Chem. Phys.* **2005**, *122*, 014108.
- (2) Matsunaga, N.; Chaban, G. M.; Gerber, R. B. *J. Chem. Phys.* **2002**, *117*, 3541–3547.
- (3) Christiansen, O. *Phys. Chem. Chem. Phys.* **2007**, *9*, 2942–2953.
- (4) Daněček, P.; Bouř, P. *J. Comput. Chem.* **2007**, *28*, 1617–1624.
- (5) Bowman, J. M.; Carter, S.; Huang, X. C. *Int. Rev. Phys. Chem.* **2003**, *22*, 533–549.
- (6) McCoy, A. B. *Int. Rev. Phys. Chem.* **2006**, *25*, 77–107.
- (7) Irikura, K. K.; Johnson, R. D., III; Kacker, R. N.; Kessel, R. *J. Chem. Phys.* **2009**, *130*, 114102.
- (8) *Guide to the Expression of Uncertainty in Measurement*, 2nd ed.; International Organization for Standardization (ISO): Geneva, Switzerland, 1995.
- (9) Irikura, K. K.; Johnson, R. D., III; Kacker, R. N. *J. Phys. Chem. A* **2005**, *109*, 8430–8437.
- (10) Irikura, K. K.; Johnson, R. D., III; Kacker, R. N. *Metrologia* **2004**, *41*, 369–375.
- (11) Johnson, R. D., III *NIST Computational Chemistry Comparison and Benchmark Database*, version 14; NIST Standard Reference Database Number 101, September 2006; National Institute of Standards and Technology: Gaithersburg, MD, 2006. <http://srdata.nist.gov/cccbdb/> (accessed March 31, 2009).
- (12) Barone, V. *J. Chem. Phys.* **2004**, *120*, 3059–3065.
- (13) Kacker, R. N.; Datta, R. U.; Parr, A. C. *J. Res. Natl. Inst. Stand. Technol.* **2003**, *108*, 439–446.
- (14) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*; Gaussian, Inc.: Pittsburgh, PA, 2003.
- (15) Certain commercial materials and equipment are identified in this paper in order to specify procedures completely. In no case does such identification imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the material or equipment identified is necessarily the best available for the purpose.
- (16) Scott, A. P.; Radom, L. *J. Phys. Chem.* **1996**, *100*, 16502–16513.
- (17) Bauschlicher, C. W.; Langhoff, S. R. *Spectrochim. Acta A* **1997**, *53*, 1225–1240.

- (18) Halls, M. D.; Velkovski, J.; Schlegel, H. B. *Theor. Chem. Acc.* **2001**, *105*, 413–421.
- (19) Yoshida, H.; Takeda, K.; Okamura, J.; Ehara, A.; Matsuuura, H. *J. Phys. Chem. A* **2002**, *106*, 3580–3586.
- (20) Sinha, P.; Boesch, S. E.; Gu, C. M.; Wheeler, R. A.; Wilson, A. K. *J. Phys. Chem. A* **2004**, *108*, 9213–9217.
- (21) Shimanouchi, T. *Tables of Molecular Vibrational Frequencies, Consolidated Vol. I*; U.S. GPO: Washington, DC, 1972.
- (22) Shimanouchi, T. *J. Phys. Chem. Ref. Data* **1977**, *6*, 993–1102.
- (23) Irikura, K. K. *J. Phys. Chem. A* **1998**, *102*, 9031–9039.
- (24) Stanton, J. F.; Gauss, J.; Watts, J. D.; Szalay, P. G.; Bartlett, R. J.; Auer, A. A.; Bernholdt, D. E.; Christiansen, O.; Harding, M. E.; Heckert, M.; Heun, O.; Huber, C.; Johsson, D.; Jusélius, J.; Lauderdale, W. J.; Metzroth, T.; Michauk, C.; Price, D. R.; Ruud, K.; Schiffmann, F.; Tajti, A.; Varner, M. E.; Vázquez, J.; Almlöf, J.; Taylor, P. R.; Helgaker, T.; Jensen, H. J. A.; Jørgensen, P.; Olsen, J. *ACES II, vers. Mainz-Austin-Budapest*; Johannes Gutenberg-Universität: Mainz, Germany, 2005.

CT100244D