Frequency and Zero-Point Vibrational Energy Scale Factors for Double-Hybrid Density Functionals (and Other Selected Methods): Can Anharmonic Force Fields Be Avoided?

Manoj K. Kesharwani, Brina Brauer, and Jan M. L. Martin*

Department of Organic Chemistry, Weizmann Institute of Science, 76100 Rehovot, Israel

Supporting Information

**ABSTRACT:** We have obtained uniform frequency scaling factors $\lambda_{\text{harm}}$ (for harmonic frequencies), $\lambda_{\text{fund}}$ (for fundamentals), and $\lambda_{\text{ZPVE}}$ (for zero-point vibrational energies (ZPVEs)) for the Weigend–Ahlrichs and other selected basis sets for MP2, SCS-MP2, and a variety of DFT functionals including double hybrids. For selected levels of theory, we have also obtained scaling factors for true anharmonic fundamentals and ZPVEs obtained from quartic force fields. For harmonic frequencies, the double hybrids B2PLYP, B2GP-PLYP, and DSD-PBE86 clearly yield the best performance at RMSD = 10–12 cm$^{-1}$ for def2-TZVP and larger basis sets, compared to 5 cm$^{-1}$ at the CCSD(T) basis set limit. For ZPVEs, again, the double hybrids are the best performers, reaching root-mean-square deviations (RMSDs) as low as 0.05 kcal/mol, but even mainstream functionals like B3LYP can get down to 0.10 kcal/mol. Explicitly anharmonic ZPVEs only are marginally more accurate. For fundamentals, however, simple uniform scaling is clearly inadequate.

**INTRODUCTION AND THEORETICAL BACKGROUND**

Recently, remarkable progress has been made in the area of high-accuracy computational thermochemistry on small molecules, as witnessed by W4 theory,\textsuperscript{1−3} the HEAT approach,\textsuperscript{4,5} the focal point approximation,\textsuperscript{6} and the Feller–Petterson–Dixon (FPD) approach.\textsuperscript{7} Likewise, a number of options now exist for chemical accuracy (generally defined as ±1 kcal/mol) on medium-sized molecules, such as G0 theory (reviewed in ref \textsuperscript{15}), CBS-QB3,\textsuperscript{16,17} and the ccCA approach,\textsuperscript{18} as well as W1–F12 and W2–F12 theory.\textsuperscript{18} As such, approaches are extended to larger organic and biomolecules, such as benzene\textsuperscript{18} and the amino acids;\textsuperscript{19} it is becoming increasingly clear\textsuperscript{20} that the limiting factor for thermochemical accuracy in such systems may well be the nuclear motion rather than the electronic structure. By far, the largest nuclear motion term is the zero-point vibrational energy (ZPVE); for instance, it reaches 62.14 kcal/mol for benzene\textsuperscript{18} and goes up to about 138.2 kcal/mol for arginine.\textsuperscript{19} This situation will only get worse as applications move to larger systems.

For small molecules, the calculation of accurate anharmonic force fields is a practical option (see, e.g., refs 1 and 21–23). For systems the size of arginine, however, other options clearly need to be sought.

Traditionally, one approximation has been to multiply one-half of the sum of harmonic frequencies by a scaling factor appropriate for the level of theory and basis set.\textsuperscript{24–28} Some theoretical background may be appropriate here.

Limiting ourselves for the moment to asymmetric tops and neglecting higher-order anharmonicity and vibrational resonances, the vibrational energy levels of a polyatomic molecule are given by the expression

$$E(n) = G_0 + \sum \omega_i n_i + \frac{1}{2} \sum_{i<j} X_{ij} (n_i + \frac{1}{2})(n_j + \frac{1}{2}) + O(n^3)$$

(1)

where $G_0$ is the polyatomic equivalent of the Dunham $Y_{00}$ constant,\textsuperscript{30} while the $n_i$ and $X_{ij}$ have their usual meanings (e.g., ref 31) of vibrational quantum numbers, harmonic frequencies, and first-order anharmonicity constants, respectively. For simplicity, we will henceforth drop the remainder term cubic in the vibrational quantum numbers, O($n^3$).

Eliminating the cumbersome restricted summation from eq 1, we obtain

$$E(n) = G_0 + \sum \omega_i n_i + \frac{1}{2} \omega_i + \sum_{i<j} X_{ij} (n_i + \frac{1}{2}) + \frac{1}{2} \sum_{i<j} X_{ij} (n_i + \frac{1}{2})$$

(2)
The anharmonic ZPVE trivially becomes

\[
ZPVE_{\text{anhar}} \equiv E(0) = G_0 + \frac{1}{2} \sum_i x_i^0 + \frac{1}{4} \sum_{ij} X_{ij}
\]

\[
= G_0 + \sum_i \left( \frac{\omega_i}{2} + \frac{X_i}{4} \right) + \frac{1}{8} \sum_{ij} X_{ij}
\]

\[
E(n) = ZPVE_{\text{anhar}}
\]

\[
= \sum_i \left( \omega_i n_i + X_i n_i^2 + \frac{1}{2} \sum_{j \neq i} X_{ij} n_i n_j \right)
\]

\[
= \sum_i \left( \omega_i n_i + X_i n_i^2 \right) + \sum_{ij} X_{ij} n_i n_j
\]

(4)

where the linear polyatomic Dunham coefficients are given by

\[
\omega_i^0 = \omega_i + X_i + \frac{1}{2} \sum_{j \neq i} X_{ij} + O(n^3)
\]

and obviously the vibrational fundamentals \(\nu_i\) (in the absence of resonances) are \(\nu_i = \omega_i^0 + X_i\). If we now define

\[
ZPVE_{\text{anhar}} = \frac{1}{2} \sum_i \omega_i \quad \text{ZPVE}_{\text{fund}} = \frac{1}{2} \sum_i \nu_i
\]

\[
ZPVE_{\text{linear}} = \frac{1}{2} \sum_i \omega_i^0
\]

(5)

Then it is easily shown that \(27\)

\[
\frac{ZPVE_{\text{harm}} + ZPVE_{\text{linear}}}{2} = \frac{1}{4} \sum_i \left( \omega_i + X_i + \frac{1}{2} \sum_{j \neq i} X_{ij} \right)
\]

\[
+ \frac{1}{4} \sum_i \omega_i = \frac{1}{2} \sum_i \omega_i + \frac{1}{4} \sum_i X_i + \frac{1}{8} \sum_{ij} X_{ij}
\]

\[
= \frac{1}{2} \sum_i \omega_i + \frac{1}{4} \sum_i X_i + \frac{1}{4} \sum X_{ij} = ZPVE_{\text{anhar}} - G_0
\]

\[
\frac{ZPVE_{\text{harm}} + ZPVE_{\text{fund}}}{2} = \frac{ZPVE_{\text{harm}} + ZPVE_{\text{linear}}}{2} + \frac{1}{4} \sum_i X_i
\]

\[
= ZPVE_{\text{anhar}} + \frac{1}{4} \sum_i X_i - G_0
\]

(7)

The anharmonicity constants can be compactly expressed in terms of the quartic force field in reduced normal coordinates (cm\(^{-1}\) units) as

\[
X_{ij} = \frac{\phi_{ij}}{4} - \sum_k \frac{\phi_{ik} \phi_{jk}}{4 \omega_k} - \sum_k \frac{\phi_{ik}^2}{8} [D(k, i, j)
\]

\[
+ D(k, i, -j) + D(k, i, -j) + D(k, -i, j)]
\]

\[
+ \sum \beta_k (\zeta_k^{(a)}) \left( \frac{\alpha_i}{\omega_i} + \frac{\alpha_j}{\omega_j} \right)
\]

(9)

where the \(\omega_i, \phi_{ij},\) and \(\phi_{ijk}\) have their customary meanings (e.g., ref 31) of harmonic frequencies, cubic, and quartic force constants, respectively, the \(B_k\) are the rotational constants about the three principal axes of inertia, and the \(\zeta^{(a)}\) are the Coriolis coupling constants. The inverse denominators \(D\) are defined by \(32,33\)

\[
D(\pm i, \pm j, \pm k) = \frac{1}{\pm\omega_i \pm \omega_j \pm \omega_k}
\]

(10)

The \(D\) array can be precomputed; Fermi resonances will correspond to very large \(D\) elements [e.g., \(2\omega_i \approx \omega_j\) corresponds to a very large \(D(k, -i, -i)\)], and the anharmonicity constants can be easily deperturbed by simply zeroing the affected \(D\) element.

Note that typically, the negative cubic terms outweigh the positive quartic term, making the \(X_{ij}\) negative overall. If we take this into account, then the following inequality will generally be obeyed

\[
ZPVE_{\text{harm}} > ZPVE_{\text{anhar}} \approx \frac{ZPVE_{\text{harm}} + ZPVE_{\text{linear}}}{2}
\]

\[
> ZPVE_{\text{linear}} > ZPVE_{\text{fund}}
\]

(11)

While the \(X_{ij}\) are not invariant under deperturbation for Fermi interactions (and neither is \(G_0\)), it can be shown \(22\) that the invariances cancel each other in eq 3, and thus, \(ZPVE_{\text{anhar}}\) is invariant.

Allen and co-workers derived the following expression \(22\)

\[
ZPVE_{\text{anhar}} = \frac{1}{2} \sum_i \omega_i - \sum_{ijk} \frac{\phi_{ijk}^2}{32 \omega_k}
\]

\[
- \frac{1}{48} \sum_{ijk} \omega_{ij} + \omega_j + \omega_k + \sum \frac{\phi_{ijk}}{32} + Z_{\text{kinetic}}
\]

(12)

where the typically small kinetic energy term is

\[
Z_{\text{kinetic}} = -\frac{1}{4} \sum_i \beta_i^{(a)}
\]

\[
\times \left\{ 1 + \sum_{ijk} \left[ \zeta_i^{(a)} \right] \frac{\beta_{ijk}^{(a)} (\omega_i + \omega_j) - \omega_i \omega_j}{\omega_i \omega_j} \right\}
\]

(13)

Equivalent expressions were derived by Vázquez and Stanton (cf. ref 30 in ref 34) as well as by Barone, \(35\) the latter of which is implemented \(36\) in Gaussian 09, revision D.

(Explicit evaluation of eq 12 requires the full cubic force field and a semidiagonal quartic force field. For methods where analytical second derivatives are available, numerical central differ-ences of the Hessian in normal coordinates permit evaluation of all required force constants in at most \(2N + 1\) Hessian evaluations. \(37\))

The relatively high computational cost of determining harmonic force fields early on led to approximations in which force constants determined at a low level of theory were rescaled. While the SQM (scaled quantum mechanical) approach of Pulay and co-workers \(38,39\) employs different scaling factors for different types of force constants and different functional groups, by far, the most popular approach has been the use of a single global scaling factor \(32,37\)

\[
\lambda = 28.29.40.41
\]

When using a simple scaling factor to fit harmonic frequencies \(\omega_{\text{true}} \approx \omega_{\text{calc}}(\text{level}) \lambda_{\text{harm}}(\text{level})\), one seeks to compensate for the
intrinsic biases of the theoretical level. For instance, it has been well-known since the 1970s (see, for example, ref 41) that at the HF level with small basis sets, bonds are too short and stretching frequencies thus (as a special case of Badger’s rule\textsuperscript{32,43}) intrinsically too high. (In fact, it was found empirically that such low-level harmonic frequencies could be brought much closer to experiment by evaluating them at the nonstationary experimental geometry.\textsuperscript{64–66})

The common use of scaled harmonic frequencies to predict approximate fundamental frequencies $\nu_{\text{true}} \approx \theta_{\text{calc}}(\text{level})$, $\lambda_{\text{fund}}$(level) goes a step further in that it also attempts to subsume the generally negative anharmonicity into the scaling factor. In practice, it is indeed seen in the near-IR range that vibrational anharmonicities are roughly proportional to harmonic frequencies, which is why such a crude correction as a scaling factor “has any prayer of working at all”.

Grev et al.\textsuperscript{27} first showed that ZPVE\textsubscript{warm} > ZPVE\textsubscript{inhom} > ZPVE\textsubscript{fund} as well as deriving an expression equivalent to eq 8, and it has been argued\textsuperscript{57} on that basis that the average of $\lambda_{\text{warm}}$(level) and $\lambda_{\text{fund}}$(level) would be a good approximation to $\delta\text{ZPVE}$(level).

A compendia of scale factors were published by Radom and co-workers\textsuperscript{55,56,60} as well as by Truhlar and co-workers,\textsuperscript{28} additional sets of scale factors have been published for the correlation-consistent basis sets,\textsuperscript{57} the polarization-consistent basis sets,\textsuperscript{48} and the Sadlej\textsuperscript{49–52} electrical properties basis set;\textsuperscript{53} the rationale behind this latter study was the performance of these basis sets for Raman activities.

No discussion would be complete without mentioning the more elaborate SQM or SQMFF (scaled quantum mechanical force field) approach\textsuperscript{58} pioneered by Boggs, Pulay, and co-workers.

To the best of our knowledge, no consistent set of scale factors for harmonics, fundamentals, or ZPVEs has ever been proposed for the Weigend–Ahlrichs\textsuperscript{54} def2 basis sets. The def2 family of basis sets not only covers most of the periodic table (presently H–Rn) but holds a middle ground between ab initio optimized basis sets like the correlation-consistent family\textsuperscript{55–58} and DFT-specific basis sets like Jensen’s polarization-consistent\textsuperscript{60–63} family. (For two recent reviews on Gaussian basis sets, see Hill\textsuperscript{64} and Jensen.\textsuperscript{65}) As such, they are especially well-suited for application in conjunction with double-hybrid functionals,\textsuperscript{66–67} where the correlation energy is a hybrid of DFT correlation and MP2 correlation in the basis of Kohn–Sham orbitals.\textsuperscript{68} One purpose of the present paper is to report appropriate frequency ($\lambda_{\text{num}}\lambda_{\text{fund}}$, $\lambda_{\text{ZPVE}}$) scaling factors for the def2 family and a variety of ab initio and DFT methods.

Second, while it has been reported\textsuperscript{38,39} that double-hybrid\textsuperscript{67,69,70} and particularly DSD\textsuperscript{71–73} density functionals perform quite well (between CCSD and CCSD(T) in quality) for calculated vibrational frequencies, no systematic set of scale factors has been proposed (aside from some ad hoc values \textsuperscript{71,73} for $\lambda_{\text{num}}$ with polarization-consistent basis sets). A consistent set of $\lambda_{\text{num}}\lambda_{\text{fund}}\lambda_{\text{ZPVE}}$ for a variety of basis sets will be derived and reported in the present paper.

Third, we will address the question as to whether quartic force fields (perhaps enhanced by appropriate scaling factors) offer an actual advantage over scaled harmonic frequencies. We will show that the answer is generally affirmative for fundamentals, especially if the quartic force field is augmented with higher-level harmonic frequencies, but that for ZPVEs of at least semirigid molecules, ZPVE\textsubscript{num}\textsuperscript{2}\textsuperscript{ZPVE} works almost as well. The reason for the different behaviors is obvious from comparing the coefficients of the anharmonicities in eqs 3 and 4.

A number of basis set correlation methods or basis set–DFT functional combinations overlap with previous studies in the literature; our values for those are not intended to supersede the earlier work but to serve as a “sanity check” on the values obtained for the previously uncovered combinations. The principal resources for comparison are the large compilations of Radom and co-workers\textsuperscript{55} and of Alecu et al.,\textsuperscript{28} additionally, scale factors obtained using a different fitting procedure are available for HF, B3LYP, and MP2 with the correlation-consistent basis sets\textsuperscript{57} and for various DFT functionals (including the B2GP-PLYP double hybrid\textsuperscript{49}) with Jensen’s polarization-consistent basis sets.\textsuperscript{48}

\section{Computational Details}

Most calculations were performed using the Gaussian 09, revision D01 program system\textsuperscript{92} running on the Faculty of Chemistry computing cluster at the Weizmann Institute of Science. Some CCSD(T) frequency calculations were carried out using MOLPRO 2012.1\textsuperscript{74} running on the same hardware platform.

Aside from conventional CCSD(T),\textsuperscript{75} we considered the following levels of theory:

- regular MP2 and spin-component-scaled SCSCS-SCS-MP2\textsuperscript{76,77}
- on rung two of the “Jacob’s Ladder of DFT”,\textsuperscript{78} the GGA functionals BP86\textsuperscript{79} and PBE\textsuperscript{80}
- on rung three, the meta-GGA functionals TPSS\textsuperscript{81} and M06L\textsuperscript{82}
- on rung four, the hybrid GGA B3LYP\textsuperscript{83,84} PBE0\textsuperscript{85} and B97-1\textsuperscript{86,87} and hybrid meta-GGAs TPSSO\textsuperscript{88,89} M06\textsuperscript{82} and M06-2X\textsuperscript{82}
- on rung five, the double hybrids (see Grimme\textsuperscript{67} for a very recent review) B2PLYP\textsuperscript{66} B2G-PLYP\textsuperscript{69} and the spin-component-scaled double-hybrid DSD-PBEPP6-D\textsubscript{2}\textsuperscript{90}
- the range-separated hybrid oB97X-D\textsuperscript{91}

As the “training set”, we considered a slight modification of the HFREQ2014 data set,\textsuperscript{96} which was compiled by two of us in a paper on explicitly correlated harmonic frequencies; for details of the experimental and “semi-experimental” reference data, the reader is referred to p 2086 of that paper. The modifications consist of deleting F\textsubscript{2} (which is a statistical outlier for all DFT and lower-level ab initio methods owing to severe static correlation), HNO (which has an anomalously high anharmonicity in the H–N stretch), and CF\textsubscript{2} and compensating by adding the open-shell diatomics S\textsubscript{2} and SO (reference data taken from Huber and Herzberg\textsuperscript{93}).

In ref 92, we found that the valence CCSD(T) limit still has a root-mean-square deviation (RMSD) of 4.7 cm\textsuperscript{−1} from the reference data, which actually increases to 7.0 cm\textsuperscript{−1} if inner-shell correlation is accounted for; the remaining error is primarily due to neglect of higher-order correlation effects (particularly connected quadruples). This means in practice that a RMSD of 7 cm\textsuperscript{−1} represents the lower limit of what we can reasonably expect.

The following basis sets have been considered:

- (a) From the correlation-consistent family,\textsuperscript{55–58} cc-pVDZ, cc-pV(T+d)Z, cc-pV(Q+d)Z, and the latter two’s augmented counterparts aug-cc-pV(T+d)Z and aug-cc-pV(Q+d)Z, where the “+d” refers to the addition of a high-exponent d function on second-row elements; this has repeatedly been found\textsuperscript{88,94,95} to be essential for the description of the 3d orbital, which in high oxidation states lies low enough to become a back-donation acceptor from oxygen and fluoride.
- (b) From the Weigend–Ahlrichs family, the def2-SVP, def2-TZV, def2-TZVP, def2-QZV, and def2-QZV basis sets,\textsuperscript{54} as well as their diffuse-augmented variants\textsuperscript{46} def2-SVVP, def2-TZVPD, 1703 dx.doi.org/10.1021/jp508422u.j.phy.chem.a.2015.119.1701–1714
def2-QZVPD, and def2-QZVPP, respectively. (We note that def2-QZVPP is equivalent to def2-QZVP for the elements presently under consideration.)

(c) From the Jensen polarization-consistent family, the aug-\textsuperscript{pc}2+2d and aug-\textsuperscript{pc}3+d basis sets, which consist of pc-2 and pc-3, respectively, on hydrogen combined with aug-\textsuperscript{pc}2 and aug-\textsuperscript{pc}3, respectively, on nonhydrogen atoms and where second-row basis sets have been augmented with two and one set of high-exponent d functions, respectively, the exponent being obtained by successively multiplying the highest \textit{d} exponent already present by a factor of 2.5. These basis set combinations should be close to the basis set limit for DFT calculations and were employed in previous DFT benchmarking papers by our group (see, for example, refs 69, 71, 73, 90, 97, and 98).

(d) Several Pople-family basis sets that are routinely used for zero-point and thermal corrections in various composite ab initio thermochemistry schemes, such as 6-31G(d) in G2 and G3 theory, GTB3s (effectively 6-31G(2df)) in G4 and G4MP2 theory, and CBS\textsuperscript{7} (effectively 6-311G(d) on first row and 6-311G(2d) on second row) in CBS-QB3. In addition, we also considered the N07D basis set, which has been advocated by Barone and co-workers\textsuperscript{97} as a good compromise between accuracy and computational cost for DFT anharmonic force fields.

Finally, the scaling factors were obtained by regression through the origin according to eqs (6.1.5) through (6.1.8) in ref 100, as implemented in the LINESS function of Microsoft Excel.

\section*{RESULTS AND DISCUSSION}

Approximate Harmonic Frequencies to Exact Harmonic Frequencies. RMSDs from the HFREQ2014 reference values\textsuperscript{92} for unscaled harmonic frequencies can be found in the Supporting Information; optimal scaling factors \(\omega_{\text{true}} \approx \omega_{\text{ex}}(\text{level})\) for \(\omega_{\text{true}} \approx \omega_{\text{ex}}(\text{level})\) are given in Table 1, while the corresponding RMSD values are given in Table 2.

First of all, the largest basis sets of each of the three families, that is, aug-cc-pV(Q+d)\textit{Z}, aug-\textsuperscript{pc}3+d, and def2-QZVPP, all yield similar error statistics for both raw and scaled harmonic frequencies. As a probe for performance near the basis set limit of a given method, we shall single out def2-QZVPP on the grounds of not being specifically biased toward either ab initio or DFT methods.

Of the DFT functionals considered, M06, M06-2X, and (barely) \(\omega_{\text{B97X-D}}\) have RMSDs larger than the \(30 \text{ cm}^{-1}\) of MP2;
SCS-MP2 outperforms not just MP2 itself but all non-double-hybrid DFT functionals. Somewhat surprising, perhaps, are the low RMSDs for the simple GGAs BP86 and PBE. DSD-PBE86, at 9.9 cm\(^{-1}\), is the best performer in the table short of CCSD(T) itself.

Not surprisingly, RMSDs for (meta)GGA and hybrid DFT functionals display much less basis set sensitivity than the correlated position but are apparently closer to the ab initio than to the DFT end of the spectrum.

Among the small basis set DFT results, NO7D appears to be clearly superior to de2-SVP, 6-31G(d), and even CBSB7.

With the double hybrids, on the other hand, 6-31G(d) and NO7D are clearly inferior to de2-SVP and CBSB7. Here, however, the extra computational effort for a larger basis set clearly pays off; for instance, for B2PLYP, replacing de2-SVP by de2-TZVP nearly cuts RMSD in half, and de2-TZVP actually does better still. Still larger basis sets appear to be past the point of diminishing returns. Comparing de2-TZVP to similar-sized basis sets from the other families, it would seem that the pc-n family is less suited to double-hybrid frequency calculations than the other options.

After scaling, B2PLYP, B2GP-PLYP, and DSD-PBE86 deliver similar performances, with a slight edge for DSD-PBE86. For unscaled harmonic frequencies, B2GP-PLYP is clearly inferior to the two other double hybrids.

What about the scaling factors themselves? Unsurprisingly, those for CCSD(T) are effectively within statistical uncertainty from unity; near the basis set limit, the pure (meta)GGAs BP86, PBE, and TPSS all have scaling factors significantly greater than unity, while B3LYP, B97-1, B2PLYP, DSD-PBE86-D2, and M06 have scale factors close to unity, and a number of other functionals (such as TPSS0) appear to overestimate frequencies for unscaled frequencies, B2PLYP is clearly inferior to the two other double hybrids.

Optimal scaling factors for the DFT methods are rather weakly dependent on the basis set, to the point that scaling factors can be considered transferable to similar-sized basis sets not listed in the table. Obviously, there is more variation for MP2, with smaller scaling factors, not significantly greater than unity; near the basis set limit, the pure (meta)GGAs BP86, PBE, and TPSS all have scaling factors significantly greater than unity, while B3LYP, B97-1, B2PLYP, DSD-PBE86-D2, and M06 have scale factors close to unity, and a number of other functionals (such as TPSS0) appear to overestimate frequencies for unscaled frequencies, B2PLYP is clearly inferior to the two other double hybrids.

What about the scaling factors themselves? Unsurprisingly, those for CCSD(T) are effectively within statistical uncertainty from unity; near the basis set limit, the pure (meta)GGAs BP86, PBE, and TPSS all have scaling factors significantly greater than unity, while B3LYP, B97-1, B2PLYP, DSD-PBE86-D2, and M06 have scale factors close to unity, and a number of other functionals (such as TPSS0) appear to overestimate frequencies for unscaled frequencies, B2PLYP is clearly inferior to the two other double hybrids.

Optimal scaling factors for the DFT methods are rather weakly dependent on the basis set, to the point that scaling factors can be considered transferable to similar-sized basis sets not listed in the table. Obviously, there is more variation for MP2, with smaller scaling factors, not significantly greater than unity; near the basis set limit, the pure (meta)GGAs BP86, PBE, and TPSS all have scaling factors significantly greater than unity, while B3LYP, B97-1, B2PLYP, DSD-PBE86-D2, and M06 have scale factors close to unity, and a number of other functionals (such as TPSS0) appear to overestimate frequencies for unscaled frequencies, B2PLYP is clearly inferior to the two other double hybrids.
The scale factors for the PBE, PBE0, B97-1, BP86, TPSS, B3LYP, and MP2 methods with the 6-31G(2df,p) basis set taken from ref 92. For CCSD(T) with smaller basis sets, RMSD slightly rises to reach 28.8 cm⁻¹ for cc-pV(T+d)Z.

Interestingly, a number of functional/basis set combinations have RMSDs considerably lower than 24.9 cm⁻¹, for example, B3LYP, TPSS0, and B97-1 with def2-TZVP or larger basis sets. Clearly a “better answer for the wrong reasons” is obtained here. SCS-MP2 is now clearly inferior to most DFT functionals, unlike the case for harmonic frequencies.

For the double hybrids, error statistics with sufficiently large basis sets are comparable to those of CCSD(T) and about 2.5 times worse than can be achieved for harmonic frequencies. Clearly, uniformly scaling harmonics to reproduce fundamentals easize any advantage that such functionals have for harmonic frequencies.

Interestingly, the scaling factors for the simple BP86 and PBE functionals are close to unity, and again, RMSDs in the 25 cm⁻¹ range can be obtained with def2-TZVP or better basis sets. (We note that the corresponding scaling factors for harmonic frequencies are in the 1.03—1.04 range.) The low cost of these methods (if density fitting $^{34-130}$ is employed) obviously makes them a viable option for large systems.

The scale factors for the PBE, PBE0, B97-1, BP86, TPSS, B3LYP, and MP2 methods with the 6-31G(2df,p) basis set agree to within overlapping uncertainties with the values obtained presently of 0.9959, 0.9865, and 0.9949, respectively.

### Approximate Harmonic Frequencies to Exact Fundamental Frequencies.

The compiled experimental fundamentals can be found in the Supporting Information. For the diatomics Cl₂, CIF, CO, CS, S₂, HCl, HF, N₂, SiO, and SO, the data were taken from Huber and Herzberg. While detailed source references are as follows: C₂H₂ $^{101,102}$, C₂H₄ $^{103-107}$, CH₂OH $^{108}$, H₂CS $^{109,110}$, CH₂ $^{111}$, CiC $^{112}$, CO₂ $^{114}$, H₂CO $^{115-117}$, H₂O $^{118}$, H₂S $^{119,120}$, HCN $^{114}$, HCl $^{121}$, N₂O $^{123}$, NH₃ $^{124-127}$, OCS $^{128}$, Ph $^{129,130}$, SO₂ $^{131}$, CS₂ $^{122}$, BH₃ $^{126}$, and CCl₂ $^{133}$.

Scalings factors are given in Table 3, while RMSDs are given in Table 4. As a sanity check, we may attempt to fit experimentally extracted harmonics to experimentally observed fundamentals. For the data set $^{1}$ considered in the present paper, we find $\lambda_{\text{fund}}(\text{expt.}) = 0.9627(11)$, where the number in parentheses refers to the standard deviation in the fit parameter. The RMSD in the scaled frequencies is 24.9 cm⁻¹ (compared to 79.8 cm⁻¹ unscaled, i.e., “RMS anharmonicity” in this special case), the max + D and max − D (maximum positive and negative errors) being 56.9 and −59.2 cm⁻¹, respectively.

A very similar RMSD is obtained if the harmonic frequencies are replaced by CCSD(T)-F12c harmonic frequencies.
suggested early on by Radom and co-workers\textsuperscript{25} are close enough to our fitted values, the greatest difference being 0.003. However, with the 6-31G(d) basis set, the differences are 0.004 and 0.006 for the TPSS and MP2 methods, respectively.

The wave number linear scaling fit proposed in ref 139 was found not to yield better results.

**Approximate Harmonic Frequencies to Exact Anharmonic ZPVEs.** Detailed values are again given in the Supporting Information. The values for the diatomics are thus obtained as $\omega_0 = \omega_{\text{eq}} - \omega_{\text{anhar}}$ or eq 5 in ref 141), BH\textsubscript{2} and CC\textsubscript{2} (139) Some additional details are given as footnotes to the Supporting Information.

Scaling factors are given in Table 5, while RMSDs are given in Table 6. In light of the greatly reduced dependence (compared to the fundamentals) of the ZPVE_{\text{eq}} on the $X_p$, it is presumably not surprising that not only are scaling factors $\lambda_{\text{ZPVE}}$ (expt.) closer to unity, but the uncertainty on them is much smaller. For instance, fitting ZPVE_{\text{eq}} \approx \omega_{\text{eq}}\lambda_{\text{ZPVE}} (expt.) over our training set, we found $\lambda_{\text{ZPVE}} (\text{expt.}) = 0.9856 \pm 0.0053$, which for a RMD of only 0.04 kcal/mol (compared to 0.20 kcal/mol unscaled), a max $+D$ of 0.08 kcal/mol, and a max $-D$ of $-0.07$ kcal/mol. Considering eq 12, we can expect for the scaling factors that

\begin{align}
\lambda_{\text{harmonic}} > \lambda_{\text{ZPVE}} & > \lambda_{\text{harmonic}} + \frac{\lambda_{\text{linear}}}{2} > \lambda_{\text{harmonic}} + \frac{\lambda_{\text{fundamental}}}{2} > \lambda_{\text{linear}} + \frac{\lambda_{\text{fundamental}}}{2} \\
& > \lambda_{\text{fundamental}}
\end{align}

Obviously, for the experimental frequencies $\lambda_{\text{harmonic}} (\text{expt.}) \approx 1$ by definition, and we found $\lambda_{\text{fundamental}} (\text{expt.}) = 0.9627$ earlier; therefore, $\lambda_{\text{ZPVE}} (\text{expt.}) = 0.9856$ is, as we may expect, somewhat larger than $\lambda_{\text{harmonic}} + \lambda_{\text{fundamental}} / 2$. Our $\lambda_{\text{ZPVE}} (\text{expt.})$ is very close to the value of 0.9863 found by Perdew and co-workers\textsuperscript{159} for a
different set of mostly diatomic molecules; those authors proposed determining ZPVE$_{\text{valence}} \approx (5/8)\text{ZPVE}_{\text{full,calc}} + (3/8)\text{ZPVE}_{\text{full,exp}}$ which works out to be 0.9861 in this case.

Using yet another set of small molecules, Alecu et al.\textsuperscript{28} surveyed scaling factors for harmonic frequencies and ZPVEs for a very large number of levels of theory and proposed a "universal scaling factor ratio" of $x_{\text{full}}/x_{\text{ZPV}} \approx 1.014 \pm 0.002$, which would lead to 0.9862 $\pm$ 0.0019 for ZPVE$_{\text{valence}}$ (exp.)/ZPVE$_{\text{full}}$ (exp.), essentially the same as Perdew et al.\textsuperscript{12} and the present authors. (As an aside, Irikura et al.\textsuperscript{160} found 0.9949 $\pm$ 0.0124 for the same ratio, using a very different reference set that contains many second and third row diatomics, and pressimistic error bars on the reference data.)

If we were to use CCSD(T)-F12c/cc-pVQZ-F12 harmonic frequencies, which effectively represent the valence CCSD(T) basis set limit, we would have a fitted scale factor of 0.9859 (essentially the same as that for experimental frequencies) and an associated RMSD = 0.037 kcal/mol. Using conventional CCSD(T), we reach RMSD = 0.043 kcal/mol with the aug-cc-pVQZ(T+d)Z basis set, which increases to 0.059 kcal/mol with the cc-pVQZ(T+d)Z basis set.

For the lower-level methods, let us turn again to the def2-QZVPD basis set for performance near the basis set limit. Two of the double hybrids reach RMSDs as low as 0.049 (DSD-PBE86-D2) and 0.052 (B2GP-PLYP) kcal/mol, followed by 0.065 kcal/mol for B2LYP. Especially the former two numbers are gratifyingly close to the much more expensive values for CCSD(T); we note that B2GP-PLYP has a noticeably smaller scaling factor than the two other double hybrids, that is, there is more overestimation in the harmonic frequencies that needs to be compensated for by scaling down.

SCS-MP2 performs somewhat better than straight MP2, but both have twice the RMSD of the double hybrids. Still, with the def2-QZVPD basis set, many conventional DFT functionals outperform MP2 and yield performances similar to SCS-MP2; the exceptions are M06, M06-2X, PBE0, and oB97X-D.

Unsurprisingly perhaps, many of the smaller basis sets are wholly inadequate for MP2 and SCS-MP2; for the double hybrids, def2-TZVP appears to be a good compromise between RMSD and computational cost. Basis set sensitivity is weaker for the conventional DFT functionals, but still, one would like something better than 6-31G(d) or even N07D; again, def2-TZVP appears to be an attractive choice. The CBS-QB3 thermochemistry protocol\textsuperscript{13} specifies a B3LYP/CBSB7 ZPVE scaled by 0.9900, which is basically identical to the 0.9894 obtained in the present work. Likewise, the

### Table 5. Optimal Scale Factors for ZPVEs

<table>
<thead>
<tr>
<th></th>
<th>def2-SVP</th>
<th>def2-SVPD</th>
<th>def2-TZVP</th>
<th>def2-TZVPD</th>
<th>def2-TZVPF</th>
<th>def2-TZVPFD</th>
<th>def2-QZVP</th>
<th>def2-QZVPD</th>
<th>6-31G(d)</th>
<th>N07D</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCS-MP2</td>
<td>0.9710</td>
<td>0.9798</td>
<td>0.9791</td>
<td>0.9818</td>
<td>0.9760</td>
<td>0.9755</td>
<td>0.9762</td>
<td>0.9766</td>
<td>0.9700</td>
<td>0.9703</td>
</tr>
<tr>
<td>MP2</td>
<td>0.9971</td>
<td>0.9971</td>
<td>0.9784</td>
<td>0.9812</td>
<td>0.9760</td>
<td>0.9755</td>
<td>0.9762</td>
<td>0.9766</td>
<td>0.9700</td>
<td>0.9701</td>
</tr>
<tr>
<td>B2GP-PLYP</td>
<td>0.9716</td>
<td>0.9773</td>
<td>0.9752</td>
<td>0.9766</td>
<td>0.9737</td>
<td>0.9746</td>
<td>0.9743</td>
<td>0.9746</td>
<td>0.9649</td>
<td>0.9707</td>
</tr>
<tr>
<td>B2LYP</td>
<td>0.9808</td>
<td>0.9859</td>
<td>0.9832</td>
<td>0.9844</td>
<td>0.9818</td>
<td>0.9827</td>
<td>0.9824</td>
<td>0.9827</td>
<td>0.9735</td>
<td>0.9794</td>
</tr>
<tr>
<td>DSD-PBE86-D2</td>
<td>0.9784</td>
<td>0.9848</td>
<td>0.9834</td>
<td>0.9852</td>
<td>0.9818</td>
<td>0.9827</td>
<td>0.9823</td>
<td>0.9827</td>
<td>0.9731</td>
<td>0.9799</td>
</tr>
<tr>
<td>M06-2X</td>
<td>0.9762</td>
<td>0.9754</td>
<td>0.9754</td>
<td>0.9760</td>
<td>0.9748</td>
<td>0.9751</td>
<td>0.9753</td>
<td>0.9754</td>
<td>0.9745</td>
<td>0.9734</td>
</tr>
<tr>
<td>M06</td>
<td>0.9916</td>
<td>0.9951</td>
<td>0.9876</td>
<td>0.9881</td>
<td>0.9860</td>
<td>0.9883</td>
<td>0.9856</td>
<td>0.9885</td>
<td>0.9829</td>
<td>0.9886</td>
</tr>
<tr>
<td>TPSS</td>
<td>0.9710</td>
<td>0.9741</td>
<td>0.9729</td>
<td>0.9733</td>
<td>0.9723</td>
<td>0.9725</td>
<td>0.9725</td>
<td>0.9727</td>
<td>0.9632</td>
<td>0.9706</td>
</tr>
<tr>
<td>B3LYP</td>
<td>0.9912</td>
<td>0.9935</td>
<td>0.9896</td>
<td>0.9898</td>
<td>0.9883</td>
<td>0.9885</td>
<td>0.9891</td>
<td>0.9892</td>
<td>0.9825</td>
<td>0.9883</td>
</tr>
<tr>
<td>PBE0</td>
<td>0.9817</td>
<td>0.9848</td>
<td>0.9830</td>
<td>0.9834</td>
<td>0.9821</td>
<td>0.9824</td>
<td>0.9828</td>
<td>0.9830</td>
<td>0.9733</td>
<td>0.9802</td>
</tr>
<tr>
<td>B97-1</td>
<td>0.9903</td>
<td>0.9929</td>
<td>0.9900</td>
<td>0.9903</td>
<td>0.9890</td>
<td>0.9892</td>
<td>0.9896</td>
<td>0.9897</td>
<td>0.9822</td>
<td>0.9884</td>
</tr>
<tr>
<td>TPSS</td>
<td>1.0191</td>
<td>1.0141</td>
<td>1.0100</td>
<td>1.0123</td>
<td>1.0103</td>
<td>1.0106</td>
<td>1.0105</td>
<td>1.0107</td>
<td>0.9941</td>
<td>1.0005</td>
</tr>
<tr>
<td>M06L</td>
<td>0.9858</td>
<td>0.9893</td>
<td>0.9824</td>
<td>0.9829</td>
<td>0.9817</td>
<td>0.9823</td>
<td>0.9813</td>
<td>0.9814</td>
<td>0.9812</td>
<td>0.9850</td>
</tr>
<tr>
<td>B8P6</td>
<td>1.0207</td>
<td>1.0232</td>
<td>1.0219</td>
<td>1.0216</td>
<td>1.0181</td>
<td>1.0193</td>
<td>1.0191</td>
<td>1.0193</td>
<td>1.0131</td>
<td>1.0181</td>
</tr>
<tr>
<td>PBE</td>
<td>1.0179</td>
<td>1.0203</td>
<td>1.0171</td>
<td>1.0175</td>
<td>1.0160</td>
<td>1.0163</td>
<td>1.0171</td>
<td>1.0173</td>
<td>1.0095</td>
<td>1.0149</td>
</tr>
<tr>
<td>oB97X-D</td>
<td>0.9788</td>
<td>0.9821</td>
<td>0.9791</td>
<td>0.9796</td>
<td>0.9779</td>
<td>0.9782</td>
<td>0.9788</td>
<td>0.9789</td>
<td>0.9709</td>
<td>0.9777</td>
</tr>
</tbody>
</table>
anharmonic corrections. Well-known pathological basis set dependence for the bending molecule had to be removed from the sample on account of its TZVP basis set and various methods (Table 7). The acetylene appropriate factor? We decided to investigate this for the def2-6-31G(d) advocated early on by Pople et al. Our study on the amino acids. The scale factor for MP2/6-31G(2df,p) basis set; our scaling factors agree to one unit on B97-1, BP86, TPSS, B3LYP, and MP2 methods with the we can compare with the Radom database close to our 6-31G(d), 6-31G(d), and 6-31G(d) levels of theory, we do see a signiﬁcant improvement over the meta-GGA functionals, the results display a surprising degree of grid sensitivity, as was found previously for vibrational anharmonicities. Consequently, we used the “superﬁne” pruned (150,974) grid combination throughout. Somewhat expectedly, the scaling factors λZPVE,anharm are generally found to be quite close to λharm typically slightly smaller. For semirigid molecules like (aside from internal rotations and the NH3 inversion) the molecules in the present set, the performance beneﬁt of anharmonic over scaled harmonic ZPVEs appears to be quite small. We may, however, expect the situation to be different for less rigid molecules, where the quartic force ﬁeld calculation may be preferable. Approximately Anharmonic ZPVE to Exact Anharmonic ZPVE. Would calculating actual anharmonic force ﬁelds yield materially better ZPVEs than simply scaling ZPVEharm by an appropriate factor? We decided to investigate this for the def2-TZVP basis set and various methods (Table 7). The acetylene molecule had to be removed from the sample on account of its well-known pathological basis set dependence for the bending anharmonic corrections. The Allen ZPVEharm expression, being free of resonance denominators, is independent of any operator decisions concerning which anharmonic resonances to include or exclude. Particularly for the meta-GGA functionals, the results display a surprising degree of grid sensitivity, as was found previously for vibrational anharmonicities. Consequently, we used the “superﬁne” pruned (150,974) grid combination throughout. Somewhat expectedly, the scaling factors λZPVE,anharm are generally found to be quite close to λharm typically slightly smaller. For semirigid molecules like (aside from internal rotations and the NH3 inversion) the molecules in the present set, the performance beneﬁt of anharmonic over scaled harmonic ZPVEs appears to be quite small. We may, however, expect the situation to be different for less rigid molecules, where the quartic force ﬁeld calculation may be preferable. Approximately Anharmonic ZPVE to Exact Anharmonic ZPVE. Would calculating actual anharmonic force ﬁelds yield materially better ZPVEs than simply scaling ZPVEharm by an appropriate factor? We decided to investigate this for the def2-TZVP basis set and various methods (Table 7). The acetylene molecule had to be removed from the sample on account of its well-known pathological basis set dependence for the bending anharmonic corrections. The Allen ZPVEharm expression, being free of resonance denominators, is independent of any operator decisions concerning which anharmonic resonances to include or exclude.

Table 6. RMSD (kcal/mol) for ZPVEs after Scaling

<table>
<thead>
<tr>
<th>Method</th>
<th>aug-cc-pVQZ</th>
<th>aug-cc-pVTZ</th>
<th>cc-pVQZ</th>
<th>cc-pVTZ</th>
<th>aug-cc-pVQZ</th>
<th>aug-cc-pVTZ</th>
</tr>
</thead>
<tbody>
<tr>
<td>aug-cc-pVDZ</td>
<td>0.0374</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>def2-SVP</td>
<td>0.122</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>def2-SPVD</td>
<td>0.122</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>def2-TZVP</td>
<td>0.122</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>def2-TZVPD</td>
<td>0.122</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>def2-TZVPP</td>
<td>0.122</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>def2-TZVPPD</td>
<td>0.122</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>def2-QZVP</td>
<td>0.122</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>def2-QZVPD</td>
<td>0.122</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>def2-QZVPP</td>
<td>0.122</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>def2-QZVPPD</td>
<td>0.122</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-31G(d)</td>
<td>N07D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

0.9854 scale factor for B3LYP/6-31G(2df,p) specified in G4 and G4MP2 theory agrees almost perfectly with 0.9862. The 0.985 scale factor for B3LYP/cc-pV(T+d)Z in W1 theory, however, should be increased to about 0.989, as previously concluded in our study on the amino acids. The scale factor for MP2/6-31G(d) advocated early on by Pople et al. (0.9646) is also close to our ﬁtted value (0.9672). As an additional sanity check, we can compare with the Radom database for the PBE, PBE0, B97-1, BP86, TPSS, B3LYP, and MP2 methods with the 6-31G(2df,p) basis set; our scaling factors agree to one unit on the third decimal place, several of them to 1 decimal place.

Approximate Anharmonic ZPVE to Exact Anharmonic ZPVE. Would calculating actual anharmonic force ﬁelds yield materially better ZPVEs than simply scaling ZPVEharm by an appropriate factor? We decided to investigate this for the def2-TZVP basis set and various methods (Table 7). The acetylene molecule had to be removed from the sample on account of its well-known pathological basis set dependence for the bending anharmonic corrections. The Allen ZPVEharm expression, being free of resonance denominators, is independent of any operator decisions concerning which anharmonic resonances to include or exclude. Particularly for the meta-GGA functionals, the results display a surprising degree of grid sensitivity, as was found previously for vibrational anharmonicities. Consequently, we used the “superﬁne” pruned (150,974) grid combination throughout. Somewhat expectedly, the scaling factors λZPVE,anharm are generally found to be quite close to λharm typically slightly smaller. For semirigid molecules like (aside from internal rotations and the NH3 inversion) the molecules in the present set, the performance beneﬁt of anharmonic over scaled harmonic ZPVEs appears to be quite small. We may, however, expect the situation to be different for less rigid molecules, where the quartic force ﬁeld calculation may be preferable. Approximately Anharmonic ZPVE to Exact Anharmonic ZPVE. Would calculating actual anharmonic force ﬁelds yield materially better ZPVEs than simply scaling ZPVEharm by an appropriate factor? We decided to investigate this for the def2-TZVP basis set and various methods (Table 7). The acetylene molecule had to be removed from the sample on account of its well-known pathological basis set dependence for the bending anharmonic corrections. The Allen ZPVEharm expression, being free of resonance denominators, is independent of any operator decisions concerning which anharmonic resonances to include or exclude. Particularly for the meta-GGA functionals, the results display a surprising degree of grid sensitivity, as was found previously for vibrational anharmonicities. Consequently, we used the “superﬁne” pruned (150,974) grid combination throughout. Somewhat expectedly, the scaling factors λZPVE,anharm are generally found to be quite close to λharm typically slightly smaller. For semirigid molecules like (aside from internal rotations and the NH3 inversion) the molecules in the present set, the performance beneﬁt of anharmonic over scaled harmonic ZPVEs appears to be quite small. We may, however, expect the situation to be different for less rigid molecules, where the quartic force ﬁeld calculation may be preferable. Approximately Anharmonic ZPVE to Exact Anharmonic ZPVE. Would calculating actual anharmonic force ﬁelds yield materially better ZPVEs than simply scaling ZPVEharm by an appropriate factor? We decided to investigate this for the def2-TZVP basis set and various methods (Table 7). The acetylene molecule had to be removed from the sample on account of its well-known pathological basis set dependence for the bending anharmonic corrections. The Allen ZPVEharm expression, being free of resonance denominators, is independent of any operator decisions concerning which anharmonic resonances to include or exclude. Particularly for the meta-GGA functionals, the results display a surprising degree of grid sensitivity, as was found previously for vibrational anharmonicities. Consequently, we used the “superﬁne” pruned (150,974) grid combination throughout. Somewhat expectedly, the scaling factors λZPVE,anharm are generally found to be quite close to λharm typically slightly smaller. For semirigid molecules like (aside from internal rotations and the NH3 inversion) the molecules in the present set, the performance beneﬁt of anharmonic over scaled harmonic ZPVEs appears to be quite small. We may, however, expect the situation to be different for less rigid molecules, where the quartic force ﬁeld calculation may be preferable. Approximately Anharmonic ZPVE to Exact Anharmonic ZPVE. Would calculating actual anharmonic force ﬁelds yield materially better ZPVEs than simply scaling ZPVEharm by an appropriate factor? We decided to investigate this for the def2-TZVP basis set and various methods (Table 7). The acetylene molecule had to be removed from the sample on account of its well-known pathological basis set dependence for the bending anharmonic corrections. The Allen ZPVEharm expression, being free of resonance denominators, is independent of any operator decisions concerning which anharmonic resonances to include or exclude. Particularly for the meta-GGA functionals, the results display a surprising degree of grid sensitivity, as was found previously for vibrational anharmonicities. Consequently, we used the “superﬁne” pruned (150,974) grid combination throughout. Somewhat expectedly, the scaling factors λZPVE,anharm are generally found to be quite close to λharm typically slightly smaller. For semirigid molecules like (aside from internal rotations and the NH3 inversion) the molecules in the present set, the performance beneﬁt of anharmonic over scaled harmonic ZPVEs appears to be quite small. We may, however, expect the situation to be different for less rigid molecules, where the quartic force ﬁeld calculation may be preferable. Approximately Anharmonic ZPVE to Exact Anharmonic ZPVE. Would calculating actual anharmonic force ﬁelds yield materially better ZPVEs than simply scaling ZPVEharm by an appropriate factor? We decided to investigate this for the def2-TZVP basis set and various methods (Table 7). The acetylene molecule had to be removed from the sample on account of its well-known pathological basis set dependence for the bending anharmonic corrections. The Allen ZPVEharm expression, being free of resonance denominators, is independent of any operator decisions concerning which anharmonic resonances to include or exclude.
DSD-PBE86 it goes from 26 to 15 cm$^{-1}$. For functionals like B3LYP and PBE0, no great improvement is seen, but that can be attributed to deficiencies in the harmonic frequencies. If CCSD(T)-F12c/cc-pVQZ-F12 harmonic frequencies are combined with the B3LYP anharmonicities, RMSD drops to just 12.5 cm$^{-1}$; if DSD-PBE86 anharmonicities are used instead, RMSD drops to 9 cm$^{-1}$.

Most of the meta-GGAs turn out to be quite problematic for anharmonicities owing to excessive grid sensitivity. Fitted scaling factors, once again, are quite close to the harmonic-to-harmonic ones.

### CONCLUSIONS

We have obtained a collection of uniform frequency scaling factors $\lambda_{\text{harm}}$ (approximate to true harmonic frequencies), $\lambda_{\text{anharm}}$ (ZPVE$_{\text{anharm}}$ to ZPVE$_{\text{true}}$), and $\lambda_{\text{fund}}$ (calculated harmonics to true fundamentals) for the Weigend–Ahlrichs and other selected basis sets for MP2, SCS-MP2, and a variety of DFT functionals including double hybrids. In addition, we have obtained, for selected levels of theory, scaling factors for true anharmonic ZPVEs and fundamental frequencies obtained from quartic force fields. Where comparison is possible, our scaling factors generally agree well with those previously obtained by the Radom and Truhlar groups.

For “harmonic to harmonic” scaling, the double hybrids B2LYP, B2GP-PLYP, and DSD-PBE86 clearly yield the best performance at RMSDs of around 10–12 cm$^{-1}$ for sufficiently large basis sets. For comparison, the valence CCSD(T) basis set limit represents a RMSD of 4.6 cm$^{-1}$.

For “harmonic to ZPVE$_{\text{anharm}}$” scaling, again the double hybrids are the best performers (reaching a RMSD of 0.05 kcal/mol with large basis sets, compared to 0.04 kcal/mol for valence CCSD(T) at the basis set limit), but functionals like B3LYP and B97-1 can still reach RMSDs in the 0.10 and 0.08 kcal/mol range, respectively. The use of explicit anharmonic ZPVEs from quartic force fields yields only a fairly marginal further improvement.

For “harmonic to fundamental” scaling, simple uniform scaling factors leave something to be desired in terms of performance; here, explicit calculation of anharmonicities does offer considerably better RMSDs. This becomes especially true if, in addition, the harmonic frequencies are replaced by basis set limit CCSD(T) values.

Among the various basis sets considered, def2-TZVP appears to offer the best compromise between quality and computational cost.

### ASSOCIATED CONTENT

#### Supporting Information

Unabridged refs 36, 74, and 107; (quasi-)experimental harmonic frequencies, fundamentals, and ZPVEs for the test molecules, and computed harmonic frequencies, ZPVE$_{\text{anharm}}$ anharmonic frequencies, and ZPVE$_{\text{anharm}}$ for the studied molecules at the levels of theory considered in this article. Tables containing unscaled RMSDs, standard deviations, and sum of the max+ and
max— deviations are also given. This material is available free of charge via the Internet at http://pubs.acs.org.

## AUTHOR INFORMATION

### Corresponding Author
E-mail: gershom@weizmann.ac.il. Fax: +972 8 9344142.

### Author Contributions
The article was written through contributions of all authors. All authors have given approval to the final version of the article.

### Notes
The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

The authors would like to thank Profs. Martin Suhm (U. of Göttingen, Germany) and Amir Karton (U. of Western Australia, Perth) for helpful discussions. This research was supported in part by the Lise Meitner-Minerva Center for Computational Quantum Chemistry and by the Helen and Martin Kimmel Center for Molecular Design.

## REFERENCES


36. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.;
Virial Theorems to the Theoretical Calculation of Molecular Potential
Orbital Studies of Vibrational Frequencies.
Functional Derived Vibrational Force Fields.
Formaldehyde, and Ethylene.
Mechanical (QM) Force Fields for Glyoxal, Acrolein, Butadiene,
Combination of Theoretical Ab Initio and Experimental Information to

Potential Energy for CO.

Factors.
Interaction, Density Functional Theory, and Semiempirical Scale
Evaluation of Hartree
The Journal of Physical Chemistry A
Harmonic Vibrational Frequencies: Scaling Factors for HF, B3LYP, and
Higher-Order Force Constants at Nonstationary Reference Geometries.

(49) Sadlej, A. J. Medium-Size Polarized Basis Sets for High-Level-