Approximating Large-Basis Coupled-Cluster Theory Vibrational Frequencies using Focal-Point Approximations

Philip M. Nelson, Zachary L. Glick, and C. David Sherrill

Center for Computational Molecular Science and Technology, School of Chemistry and Biochemistry, and School of Computational Science and Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0400, USA

The focal-point approximation can be used to estimate a high-accuracy, slow quantum chemistry computation by combining several lower-accuracy, faster computations. We examine the performance of focal-point methods combining second-order Møller–Plesset perturbation theory (MP2) with coupled-cluster singles, doubles, and perturbative triples [CCSD(T)] for the calculation of harmonic frequencies, and of fundamental frequencies using second-order vibrational perturbation theory (VPT2). In contrast to standard CCSD(T), the focal-point CCSD(T) method approaches the complete basis set (CBS) limit with only triple-ζ basis sets for the coupled-cluster portion of the computation. The predicted harmonic and fundamental frequencies were compared to experimental values for a set of 20 molecules containing up to six atoms. The focal-point method combining CCSD(T)/aug-cc-pV(T+d)Z with CBS-extrapolated MP2 has mean absolute errors versus experiment of only 7.3 cm^{-1} for the fundamental frequencies, which is essentially the same as the mean absolute error for CCSD(T) extrapolated to the CBS limit using the aug-cc-pV(Q+d)Z and aug-cc-pV(5+d)Z basis sets. However, for H_2O, the focal-point procedure requires only 3% of the computation time as the extrapolated CCSD(T) result, and the cost savings will grow for larger molecules.

I. INTRODUCTION

Wavefunction-based quantum chemistry provides a hierarchy of methods with a fundamental trade-off between accuracy and computational cost. A well-established problem associated with this trade-off is the slow convergence of the correlation energy with increasing basis set size due to the electron cusp condition. This poor convergence can be addressed using either explicitly-correlated methods or basis set extrapolation techniques. Combining large basis sets with expensive correlated methods is undesirable, but is seemingly necessary for high levels of accuracy.

One promising solution to this problem is the focal-point approach,[1,2] which approximates a desired level of theory (method/basis) with a combination of calculations in smaller basis sets and calculations performed with less correlated methods. Focal-point type methods have been applied to the calculation of properties such as geometries,[3–5] interaction energies,[6–8] thermochemical properties,[9,10] as well as harmonic vibrational frequencies.[11–17]

At its core, the focal-point approach assumes that convergence of the correlation energy to the full configuration interaction limit and of the one-particle basis set to the complete basis set (CBS) limit are approximately independent and can therefore be accounted for separately. A popular application of the focal-point approach is the approximation of the "gold-standard" coupled-cluster singles, doubles, and perturbative triples [CCSD(T)][18] energy in the CBS limit by combining large-basis (lg) second-order Møller–Plesset perturbation theory (MP2) with CCSD(T) in some small (sm), incomplete basis set, as shown in Eq. 1.

$$\begin{aligned} E(\text{CCSD(T)/lg}) \approx & E(\text{MP2/lg}) + \\ & [(E(\text{CCSD(T)/sm}) - E(\text{MP2/sm})] \end{aligned} \tag{1}$$

The above scheme is written compactly as CCSD(T)/CBS(lg; δ:sm) throughout this work. The focal-point energy is written as a linear combination of energies; consequently, the focal-point estimation scheme readily extends to the extrapolation of any quantity that is a linear function of the energy. Of particular interest in this work are the derivatives of the energy up to fourth order, used in calculating the fundamental vibrational frequencies.

While the focal-point approach is mathematically simple, its routine use is limited in practice by the flexibility of quantum chemistry software. In geometry optimizations, for example, linear combinations of gradients are used to form a focal-point gradient at each iteration of the optimization, so the software must contain a focal-point driver or else support

user-defined gradients with a method-agnostic application programming interface (API). As meaningful calculation of vibrational frequencies requires a tightly optimized minimum of the potential energy surface, focal-point frequencies are only readily calculated using software that can automate focal-point methods. The Psi4 software program has been improved to automate such procedures. We recently used Psi4 to perform a systematic study of optimized CCSD(T) focal-point geometries, which we found to be in good agreement with standard CCSD(T) calculations performed in much larger basis sets.[5] In particular, we found that when CCSD(T) in an (aug-)cc-pVXZ basis set is combined with CBS-extrapolated MP2, the accuracy in optimized geometries is similar to that using standard CCSD(T) with a basis set one ζ level higher [i.e., (aug-)cc-pV(X+1)Z], while reducing the computational cost by an order of magnitude or more. This greatly improves the feasibility of obtaining benchmark-quality geometries for small molecules. Compared to explicitly-correlated methods, the focal-point scheme is much simpler to implement. We refer the reader to Ref. 5 for recommendations on specific combinations of basis sets for the CCSD(T) and MP2 calculations that optimize cost-accuracy trade-offs, which will be used in the present work. Here we examine the accuracy of focal-point CCSD(T) for calculating harmonic and fundamental vibrational frequencies.

Harmonic vibrational frequencies of a molecule are easily calculated using the Hessian matrix, but these cannot be easily compared to experiment, as the observed fundamental frequencies include anharmonic effects. A very simple and often-used approach is to scale the harmonic frequencies by some factor determined by benchmark studies at a given level of theory.[19] Such an approach assumes that the anharmonicity of any mode is always a fixed fraction of the harmonic frequency, and moreover it provides a limited ability to correct for systematic errors in the theoretical method used to compute the harmonic frequencies. One would not expect such a simple correction scheme to provide highly accurate fundamental frequencies, but recent work by Zapata Trujillo and McKemmish[20,21] indicates that explicit correlation estimates of the CCSD(T)/CBS limit harmonic frequencies can be scaled to fundamental frequencies with a median absolute deviation of only 15.9 cm^{-1} and a root mean squared error (RMSE) of 32 cm^{-1}. These values are reduced to 7.5 and 28 cm^{-1} if a different scale factor is used for high-, medium-, and low-freqency modes.

Quantum chemistry approaches to explicitly compute the effect of anharmoinicity include vibrational perturbation theory (VPT2),[22,23] vibrational self-consistent field (VSCF),[24,25]

and virtual-state configuration interaction (VCI),[26,27] among others. VPT2 requires calculation of the cubic and semi-diagonal quartic derivatives of the energy function, and these are typically obtained with a finite-difference procedure. The high computational cost of calculating the higher-order derivatives has limited the application of coupled-cluster anharmonic frequency calculations to very small systems. However, focal-point approximations may make CCSD(T) computations more feasible in this context, as we explore here.

VPT2 is based on non-degenerate perturbation theory and can encounter difficulties when dealing with nearly degenerate states, the so-called Fermi and Darling–Dennison resonances. There are many approaches to handle such resonance cases, and the choice can have a large impact on the calculated fundamental frequencies for some systems.[28] We assume that the accuracy of such schemes is independent from choices of basis set or electron correlation and that any improvements in treating Fermi resonances can be dealt with separately from improving basis set convergence.

Only a few studies to date examine the basis set convergence of CCSD(T) harmonic frequencies. Tew et al. have investigated the basis set convergence of the CCSD(T) harmonic frequencies of 17 molecules.[29] The convergence of frequencies to spectroscopic accuracy is very slow, requiring either extrapolations with very large basis sets or explicitly-correlated methods to converge results to within a few cm^{-1}.[30]

Focal-point methods have already found application in calculating harmonic vibrational frequencies, typically by direct extrapolation of the frequencies themselves.[11–16,31] In particular, combining focal-point coupled-cluster harmonic frequencies with a more approximate method for computing anharmonicities has been a popular choice in the literature to calculate fundamental frequencies for larger molecular systems such as uracil,[11] glycine,[13] glycine dipeptide,[14] pyruvic acid,[15] and serine.[16] Morgan et al. use a focal-point approach to approximate the CCSDT(Q)/CBS limit of anharmonic frequencies of formaldehyde, combining CBS extrapolated CCSD(T) with CCSDT and CCSDT(Q) computations in a smaller basis set.[32] In recent work, Varandas has examined the accuracy of CBS-extrapolated CCSD(T) for the calculation of harmonic frequencies and found that a focal-point approach combining MP2 with small basis CCSD(T) produced errors similar to CBS-extrapolated CCSD(T).[17] A systematic investigation of the accuracy of focal-point fundamental frequencies has not yet been performed.

Here we examine the quality of focal-point approximations to CCSD(T)/CBS harmonic

and fundamental frequencies for a number of small molecules where large-basis extrapolation estimates of the CCSD(T)/CBS frequencies are feasible, and we also compare to experimental values.

II. METHODOLOGY

CCSD(T) harmonic and fundamental vibrational frequencies were obtained for the twenty molecules in Table I and the results compared to experimental frequencies. Gas-phase experimental fundamental frequencies were available for all 20 molecules, for a total of 89 unique vibrational modes. Experimentally deduced harmonic frequencies are available for 14 of the molecules, for a total of 52 unique vibrational modes. All species considered here are neutral and closed-shell.

The correlation-consistent basis sets of Dunning[55,56] [cc-pV(X+d)Z, X ∈ {D,T,Q,5}] are used in all calculations, abbreviated as XZ throughout this paper. These basis sets can be additionally augmented with diffuse functions, denoted with the aug- prefix,[57] abbreviated as aXZ throughout this paper. For calculations including core-valence correlation, the aug-cc-pwC(X+d)Z basis sets[58] were used, abbreviated as awCXZ throughout this paper. The Dunning basis sets are designed to converge smoothly to the CBS limit.

The CBS limit of the correlation energy, which converges much more slowly than the self-consistent field (SCF) energy, can be approximated by extrapolating from multiple finite-basis computations. Helgaker and coworkers derived an inverse cubic relationship between basis set cardinal number X and the correlation energy,[59] given in Eq. 2. The CBS correlation energy can then be approximated with a two-point fit, as shown in Eq. 3.

$$E_{\rm corr}(X) = E_{\rm corr}({\rm CBS}) + AX^{-3} \tag{2}$$

$$E_{\rm corr}({\rm CBS}) = \frac{E_{\rm corr}(Y)Y^{-3} - E_{\rm corr}(X)X^{-3}}{Y^{-3} - X^{-3}} \tag{3}$$

The extrapolated correlation energy is then added to the SCF energy calculated with the largest basis. For compactness, such two-point CBS extrapolations are written here as (aug-)cc-pV[XY]Z or simply (a)[XY]Z.

Previous work on CBS extrapolated vibrational frequencies has typically proceeded by extrapolating the observed spectroscopic constants and structural parameters from separate vibrational calculations for each level of theory involved in the extrapolation,[11–17] as

	Unique	Refs.		This Work: CCSD(T)	
Molecule	Modes	ν	ω	/CBS(a[Q5]Z)	Focal-Point
BH	1	33	33	y	y
HF	1	34	34	y	y
CO	1	33	33	y	y
N_2	1	33	33	y	y
F_2	1	33	33	y	y
H_2O	3	35	36	y	y
HNC	3	37	None	y	y
HCN	3	38	39	y	y
HCP	3	40	41	y	y
HNO	3	42, 43	None	y	y
CO_2	3	44	45	y	y
SO_2	3	46	46	y	y
C_2H_2	5	47	47	-	y
HOOH	6	47	None	-	y
HNCO	6	48	None	-	y
H_2CO	6	49	50	-	y
CH_2NH	9	47	None	-	y
C_2H_2O	9	47	51	-	y
C_2H_4	12	52	53	-	y
$C_2H_2F_2$	10[a]	54	None	-	y
Total Modes	89		52	26	

[a] Two modes of $C_2H_2F_2$ were omitted due to lack of gas-phase experimental data for comparison.

TABLE I. List of molecules studied and references for experimental values for fundamental and harmonic vibrational frequencies.

described in Ref. 3. Our approach differs by directly calculating the derivatives of the focal-point energy function at its minimum geometry. This approach provides not only a more automated interface for the user, but a more natural way to calculate the frequencies.

High-quality optimized geometries and frequencies were calculated for the diatomic and triatomic molecules listed in Table I at the CCSD(T)/CBS(a[Q5]Z) level of theory. This includes a total of 26 unique modes, 20 of which have experimentally deduced harmonic values. Anharmonic VPT2 calculations with a[Q5]Z on the larger molecules (more than three atoms) listed in Table I were deemed too computationally expensive to be practicable. CCSD(T)/CBS(a[Q5]Z) has been found to be converged to within 1 cm^{-1} of the CBS limit for CCSD(T),[30] and these results are used as reference values for such. Single-basis CCSD(T) was performed on all molecules using the DZ, aDZ, TZ, and aTZ basis sets. Focal-point methods of the form CCSD(T)/CBS(large; δ:small) were used according to Eq. 1, where the δ:small term is the difference between MP2 and CCSD(T) correlation energy in the specified smaller basis set. For the larger basis, a CBS extrapolation is performed at the MP2 level following Eqs. 2 and 3. Focal-point CCSD(T) was performed on all molecules using the following basis set combinations: CCSD(T)/CBS([TQ]Z; δ:DZ), CCSD(T)/CBS(a[TQ]Z; δ:aDZ), CCSD(T)/CBS([Q5]Z; δ:TZ), and CCSD(T)/CBS(a[Q5]Z; δ:aTZ). These combinations of basis sets for the MP2 and CCSD(T) portions of the focal-point procedure have been identified in previous work as having ideal cost-accuracy trade-offs.[5] Additionally, calculations including core-valence correlation were performed using the CCSD(T)/CBS(awC[Q5]Z, δ :awCTZ) focal-point scheme.

All calculations were performed with version 1.4 of the open-source Psi4 software package.[60] All SCF, MP2 and CCSD(T) calculations performed in this work use the density-fitting approximation. SCF calculations used the (aug-)cc-pV(X+d)Z-JKFIT[61] auxiliary basis sets for calculating the Coulomb and exchange integral matrices. MP2 and CCSD(T) used the (aug-)cc-pV(X+d)Z-RI[62] auxiliary basis sets for density fitting.

All molecular geometries were tightly converged with maximum residual forces less than 2×10^{-6} Hartree au^{-1} and max residual displacements less than 6×10^{-6} au. All energies and densities were converged to 1×10^{-12} Hartree and 1×10^{-10} au^{-3}, respectively. All MP2 and CCSD(T) calculations used the frozen-core (fc) approximation except for those explicitly denoted as including core-valence correlation by the use of awCXZ basis sets.

Hessians for harmonic frequencies were obtained by finite-difference of analytic gradients

using a five-point stencil and a displacement size of 0.005 au. The cubic and semi-diagonal quartic derivatives needed for anharmonic frequencies were obtained by finite-difference of analytic gradients using displacements along the reduced normal mode coordinates. A displacement size of 0.1 reduced dimensionless units, scaled by the square-root of the harmonic frequencies,[63] was used, which was found to give numerically satisfactory results. For this study, one of us (P.M.N.) developed an open-source implementation of the VPT2 method described in Ref. 23, available on GitHub.[64] VPT2 is well known to suffer from issues in dealing with states involving Fermi resonances. Martin et al. proposed a method to identify resonant terms in the VPT2 equations, the eponymous Martin's test, and remove their contributions to the anharmonicity constants, effectively deperturbing the resonant interactions.[65] Identification of the Fermi resonant terms requires threshold values for energy separation between states and for the coupling between states, the Martin's test value. An energy separation of 200 cm^{-1} and a Martin's test value of 1 cm^{-1} were used in this work to identify Fermi resonant terms. The interaction between the depertubed states can then be reintroduced variationally by diagonalizing an effective Hamiltonian between the resonant states. Such an approach is known as GVPT2 or VPT2+F in the literature.[28,66]

III. RESULTS AND DISCUSSION

A. Defining target accuracy

Before using the focal-point approach to approximate CCSD(T)/CBS(a[Q5]Z) frequencies, it is necessary to define an acceptable degree of accuracy. One way to bound the focal-point 'target accuracy' is to consider errors in the reference CCSD(T)/CBS(a[Q5]Z) calculation. Although high quality, these frequencies still contain small errors from a variety of sources: neglected electron correlation, density-fitting, the frozen-core approximation, relativistic effects, Born–Oppenheimer effects, and remaining basis-set incompleteness. Therefore, a focal-point method that reproduces CCSD(T)/CBS(a[Q5]Z) frequencies within this average margin of error is very satisfactory.

For vibrational frequencies, 'spectroscopic accuracy' is usually considered to mean that errors are within 1 cm^{-1}. Achieving this level of accuracy compared to experiment in calculated vibrational frequencies typically requires inclusion of effects beyond fc-CCSD(T).

Excluding such corrections, accuracy of around 5 cm^{-1} versus experimentally deduced harmonic frequencies can be expected. For example, Rauhut et al. report mean errors of 4.2 cm^{-1} for explicitly-correlated CCSD(T)-F12a/aTZ harmonic frequencies relative to experimentally deduced values.[47] CCSD(T)-F12a/aTZ can be expected to closely approach the basis set limit; these results provide a good estimate of what level of accuracy should be expected from CCSD(T)/CBS frequencies. We expect that the accuracy of CCSD(T)/CBS fundamental frequencies should be within nearly the same range, ignoring any errors from the treatment of anharmonic terms to obtain the fundamental frequencies, that is in the VPT2 treatment itself.

As noted above, in this work we compare theoretical values to both experimental fundamental vibrational frequencies and also to experimentally deduced harmonic vibrational frequencies. Unlike bond lengths, for which there is relatively plentiful experimental data against which to judge CCSD(T)/CBS values, experimentally deduced harmonic frequencies are difficult to acquire. Measured experimental frequencies incorporate the anharmonicity of vibrational modes, which must be accounted for to make a fair comparison. It is not trivial to separate the harmonic and anharmonic components of the frequency. According to Bartlett and Stanton, "the difficulty of obtaining harmonic frequencies experimentally is not as widely appreciated among the quantum chemical community as it should be."[67] Experimental values for harmonic frequencies are thus typically lacking for most polyatomic molecules.

For a selection of diatomic molecules and H_2O, experimental and calculated harmonic vibrational frequencies are compared in Table II. F_2 is a clear outlier; F has notoriously poor one-particle basis-set convergence.[68] Based on the results from Table II and the preceding discussion, we will assert that an ideal focal-point approximation to CCSD(T)/CBS should produce harmonic vibrational frequencies within approximately 5 cm^{-1} of the CCSD(T)/CBS value.

For the fundamental frequencies, the theoretical values can be directly compared to experiment. However, VPT2 is not an exact treatment of vibrational anharmonicity. Higher-order vibrational couplings not included in VPT2,[32] the aforementioned Fermi resonances, and numerical error in the finite-difference calculation of cubic and quartic derivatives are all likely sources of error. Nevertheless, the results in Table II show that the magnitude of errors versus experiment is generally consistent between the harmonic and VPT2 fundamental fre-

quencies. Thus, our target accuracy of 5 cm^{-1} for the focal point procedure, proposed above after consideration of the typical errors from harmonic frequencies at the CCSD(T)/CBS level, also seems appropriate for the computation of fundamental frequencies at this level.

TABLE II. Comparison of fundamental and harmonic frequencies (cm^{-1}) from experiment and theory [CCSD(T)/CBS(a[Q5]Z)]

Mol.	Fundamentals			Harmonics		
	Expt.	Theory	Error	Expt.	Theory	Error
BH	2269.2	2361.4	−6.7	2366.9	2361.4	−5.5
HF	3961.4	3962.9	+1.5	4138.3	4141.9	+3.6
CO	2143.2	2140.7	−2.5	2169.8	2166.9	−2.9
N_2	2329.9	2335.0	+5.1	2358.6	2363.0	+4.5
F_2	893.9	907.7	+13.8	916.6	930.7	+14.1
H_2O						
ω_1	1594.7	1596.3	+1.6	1649.3	1647.8	−1.5
ω_2	3657.1	3660.3	+3.2	3836.0	3831.9	−4.1
ω_3	3755.9	3758.1	+2.2	3947.3	3942.4	−4.9

B. Convergence to the CBS limit

As focal-point CCSD(T) is an approximation to complete basis set CCSD(T), we first establish its rate of convergence to the CBS limit and compare its convergence to that of standard CCSD(T) using single basis sets. Second, we analyze its performance in reproducing experimental vibrational frequencies, and we compare that performance to that of extrapolated CCSD(T)/CBS estimates.

Figure 1(a) displays the errors for fundamental and harmonic frequencies for the 5 diatomic and 7 triatomic molecules compared to CCSD(T)/CBS(a[Q5]Z) reference values. The corresponding error statistics are shown in Table III. While standard CCSD(T)/DZ results have large basis set incompleteness errors, increasing the basis set to TZ reduces the mean absolute errors by a factor of 3. The focal-point method using CCSD(T)/DZ decreases errors by a similar factor, but at a substantially reduced computational cost. The

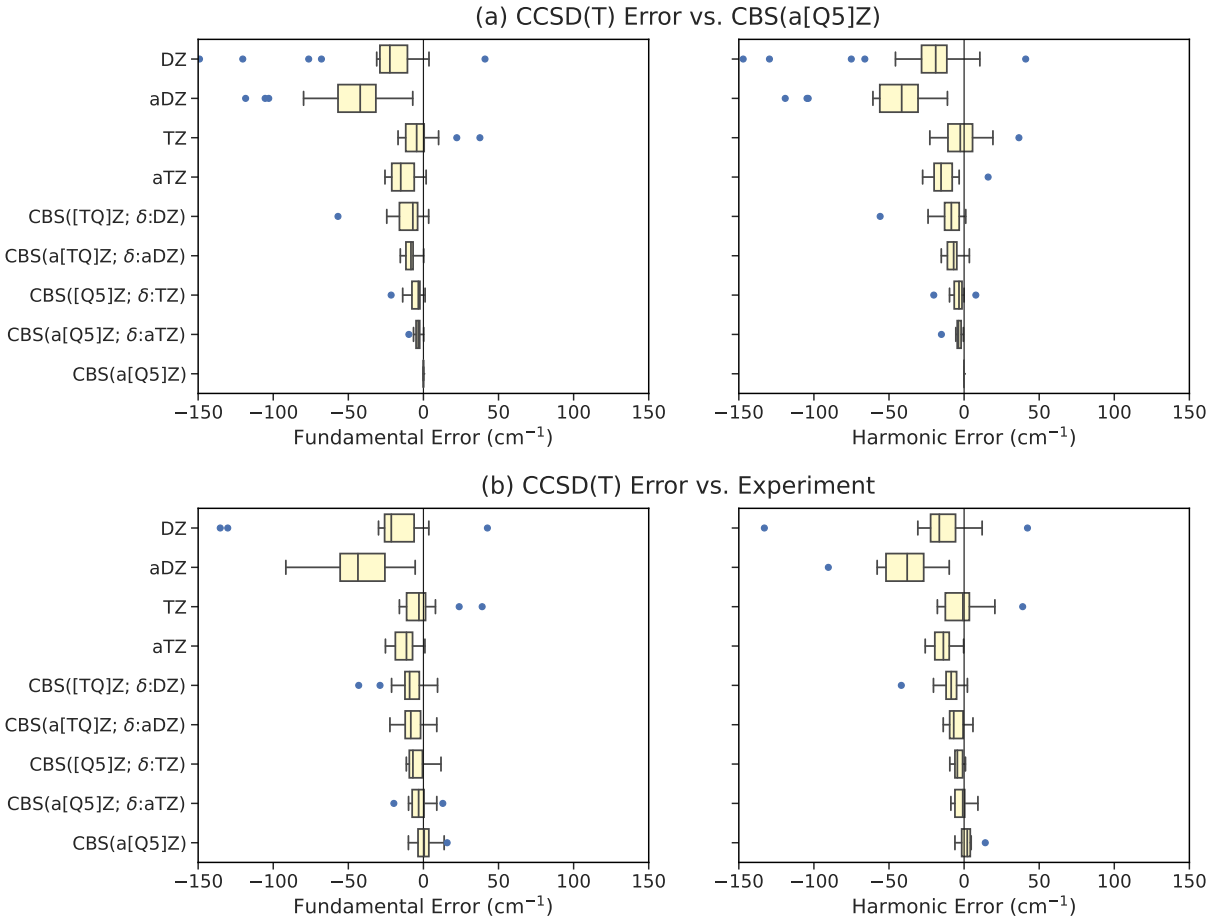

FIG. 1. (a) Comparison of error distributions for fundamental frequencies (left) and harmonic frequencies (right) versus CCSD(T)/CBS(a[Q5]Z) reference values. (b) Comparison of error distributions for fundamental frequencies (left) and harmonic frequencies (right) with respect to gas-phase experimental results. Data set includes 5 diatomic and 7 triatomic molecules, 26 modes total. Experimentally deduced harmonics include 20 modes. Boxes show the interquartile range, horizontal lines extend to the 1.5 interquartile range. Outlying data points shown as dots.

focal-point method using CCSD(T)/TZ further reduces errors by a factor of 2. Inclusion of diffuse functions in the basis set improves convergence to the CBS(a[Q5]Z) results for all focal-point methods.

The largest-basis focal-point procedure here, CCSD(T)/CBS(a[Q5]Z; δ:aTZ), reproduces the CCSD(T)/CBS(a[Q5]Z) harmonic frequencies with outstanding quantitative accuracy; the mean absolute error (MAE) is just 3.5 cm^{-1}. The fundamental frequencies for this focal-point procedure also have a MAE of 3.5 cm^{-1}.

TABLE III. Error of fundamental and harmonic frequencies with respect to CCSD(T)/CBS(a[Q5]Z) extrapolation for 5 diatomic and 7 triatomic molecules, 26 modes total. Included are median absolute deviation (MAD), mean absolute error (MAE), mean signed error (MSE), root mean squared error (RMSE), and largest absolute error (Max), all in cm^{-1}.

CCSD(T)	Fundamental Errors					Harmonic Errors				
Basis Set	MAD	MAE	MSE	RMSE	Max	MAD	MAE	MSE	RMSE	Max
DZ	23.6	31.8	−28.0	47.0	149.1	20.4	32.0	−27.1	47.5	147.1
aDZ	42.1	48.3	−48.3	55.4	118.4	41.5	46.8	−46.8	53.8	119.1
TZ	7.4	9.3	−3.0	12.2	37.6	10.4	10.6	−1.8	13.1	36.5
aTZ	15.1	13.7	−13.6	15.8	25.6	16.1	15.3	−14.1	16.8	27.6
estimated CBS										
[TQ]Z; δ:DZ	7.0	11.1	−10.8	15.8	56.9	8.5	10.4	−10.2	15.0	55.9
a[TQ]Z; δ:aDZ	8.3	8.8	−8.8	9.7	15.4	6.9	7.7	−7.4	8.7	15.2
[Q5]Z; δ:TZ	3.4	5.5	−5.4	7.4	21.5	3.6	4.8	−4.2	6.4	20.2
a[Q5]Z; δ:aTZ	3.5	3.7	−3.7	4.3	9.7	3.8	3.7	−3.7	4.6	15.1

Smaller-basis results tend to systematically underestimate the vibrational frequencies compared to the CBS(a[Q5]Z) results. As is typical, computations in the smaller basis sets result in bond distances that are too long and thus predict lower vibrational frequencies. As the basis set size is increased the bond lengths contract, driving the predicted frequencies up.

Figure 1(b) shows error distributions for fundamental and harmonic frequencies compared to experiment for the diatomic and triatomic molecules. The corresponding error statistics are shown in Table IV. The errors for finite-basis CCSD(T) methods and focal-point approximations are similar to but somewhat larger than those in 1(a), where the reference values were the two-point extrapolations CCSD(T)/CBS(a[Q5]Z). This indicates that the errors of CCSD(T)/CBS(a[Q5]Z) versus experiment are not large. Indeed, those errors are provided in Figure 1(b) and Table IV.

As both finite-basis CCSD(T) and focal-point CCSD(T) are approximations to the CCSD(T)/CBS limit, the accuracy of the latter is the limit of what one can expect for reproducing experimental frequencies, barring any fortuitous error cancellations. Higher

TABLE IV. Error of fundamental and harmonic frequencies with respect to experiment for 5 diatomic and 7 triatomic molecules, 26 fundamental and 20 harmonic modes total. Included are median absolute deviation (MAD), mean absolute error (MAE), mean signed error (MSE), root mean squared error (RMSE), and largest absolute error (Max), all in cm^{-1}.

CCSD(T)	Fundamental Errors					Harmonic Errors				
Basis Set	MAD	MAE	MSE	RMSE	Max	MAD	MAE	MSE	RMSE	Max
DZ	21.5	28.3	−24.0	44.2	135.3	19.1	24.9	−17.2	37.9	133.0
aDZ	43.6	42.8	−42.8	47.5	91.6	37.9	40.2	−40.2	44.3	90.4
TZ	6.1	9.2	−1.8	12.9	39.1	10.8	10.9	−1.5	14.3	39.0
aTZ	11.3	12.1	−12.0	14.1	25.3	13.8	14.1	−14.1	15.7	25.9
estimated CBS										
[TQ]Z; δ:DZ	9.4	11.6	−9.6	14.9	43.1	8.6	10.4	−10.0	14.0	41.8
a[TQ]Z; δ:aDZ	8.9	8.7	−7.7	10.4	22.3	6.9	6.5	−5.6	7.7	13.9
[Q5]Z; δ:TZ	7.2	6.4	−4.6	7.5	11.7	4.5	4.0	−3.8	4.8	9.5
a[Q5]Z; δ:aTZ	4.1	5.4	−2.9	7.2	19.7	1.5	3.4	−1.7	4.8	9.2
a[Q5]Z	4.0	4.8	0.7	6.3	15.8	3.7	4.0	1.5	4.9	14.1

excitations up to quadruples and pentuples (CCSDTQP) and core-correlation are known to have significant contributions to harmonic frequencies relative to fc-CCSD(T).[29] Harmonic frequencies accurate to within 1 cm^{-1} may require the inclusion of core-valence correlation, excitations up to quintuples, and scalar relativistic corrections. Corrections to the Born–Oppenheimer approximation may also be relevant in some systems. However, it appears that the corrections to harmonic frequencies for higher order excitations and for core-correlation tend to have opposite signs and are generally close in magnitude, resulting in good error cancellation for fc-CCSD(T).[68] It is worth noting that the maximum error in the CCSD(T)/CBS(a[Q5]Z) harmonic frequencies, around 14 cm^{-1}, corresponds to F_2, for which Tew et al. found that including explicit triples and quadruples provides a 12 cm^{-1} correction.[29] In contrast to other systems here, the core-correlation contributions do not cancel out this correction, so higher order excitations are necessary for accuracy approaching the experimental value for F_2.

Additionally, the density-fitting approximation used here is also known to incur some

very small errors in harmonic frequencies, typically less than 1 cm^{-1}.[69] Below, we explore the importance of including core-valence correlation in the focal-point procedure.

Curiously, basis sets with diffuse functions tend to notably degrade the agreement with experimental and CCSD(T)/CBS results for single-basis CCSD(T). In contrast, using diffuse functions in the focal-point procedure improves accuracy for both fundamental and harmonic results. This behavior is consistent with previous studies of focal-point geometries, where including diffuse functions in both the MP2 and CCSD(T) portions of the focal-point procedure consistently improved results.[5]

CCSD(T)/CBS(a[Q5]Z; δ:aTZ) is converged closely enough to the CBS limit so as to reproduce experimental values to the same level of accuracy, but at a substantially reduced computational expense (see below). Due to CCSD(T)/CBS(a[Q5]Z; δ:aTZ) systematically underestimating the frequencies at the CBS limit (see above), the method tends to slightly underestimate experimental frequencies while CCSD(T)/CBS(a[Q5]Z) tends to slightly overestimate them.

The focal-point extrapolated CCSD(T) results improve agreement with experiment significantly compared to the single-basis CCSD(T) results using the same basis set as used in the CCSD(T) portion of the focal-point procedure. For the (a)DZ results, the focal-point extrapolation reduces the MAEs versus experiment by up to a factor of 4. For the (a)TZ results, the MAEs are reduced by a factor of 2.

C. Comparison to experiment

In summary, at this point we have shown that focal-point methods combining MP2 in large basis sets with CCSD(T) in small or modest basis sets are very effective at approximating the CCSD(T)/CBS limit for both harmonic and fundamental vibrational frequencies. We have further shown that fc-CCSD(T)/CBS typically exhibits small errors versus experimental values (MAE of just a few cm^{-1}, although maximum errors can be larger in some cases). Given all this, we now turn to examining the accuracy versus experiment of focal-point methods [and single-basis CCSD(T)] for a larger set of molecules. We omit the highest-level CCSD(T)/CBS(a[Q5]Z) reference computations at this point due to the very high computational cost.

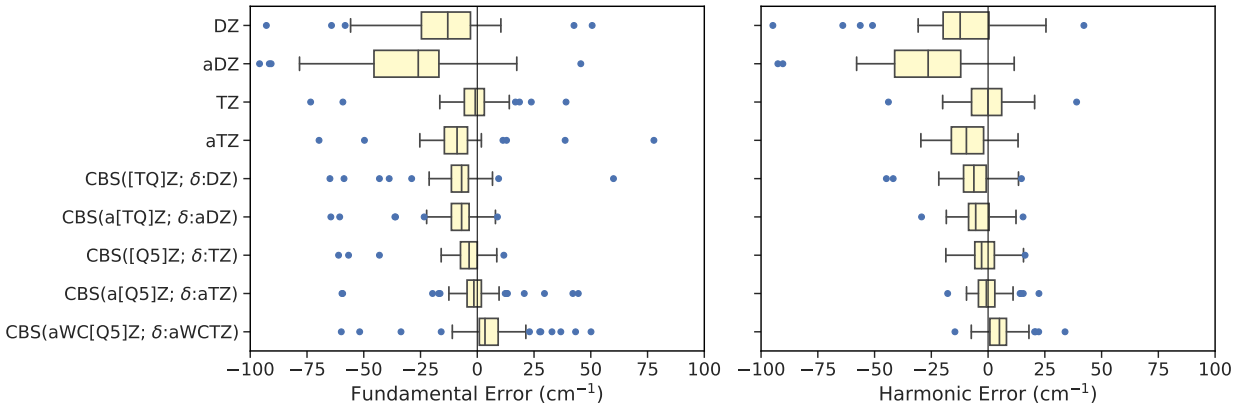

FIG. 2. Comparison of error distributions for fundamental frequencies (left) and harmonic frequencies (right) with respect to gas-phase experimental results. Fundamental data set includes 20 molecules, 89 modes total; harmonic data set includes 14 molecules, 52 modes total. Boxes show the interquartile range, horizontal lines extend to the 1.5 interquartile range. Outlying data points shown as dots.

TABLE V. Error of fundamental and harmonic frequencies with respect to experiment for full data set of 20 molecules. Included are median absolute deviation (MAD), mean absolute error (MAE), mean signed error (MSE), root mean squared error (RMSE), and largest absolute error (Max), all in cm^{-1}.

CCSD(T)	Fundamental Errors					Harmonic Errors				
Basis Set	MAD	MAE	MSE	RMSE	Max	MAD	MAE	MSE	RMSE	Max
DZ	13.4	21.7	−18.0	34.0	135.3	16.1	20.7	−13.3	31.1	133.0
aDZ	26.5	37.3	−27.8	55.1	359.5	26.4	33.7	−32.5	51.4	255.7
TZ	4.8	7.8	−2.2	13.4	73.4	6.6	8.6	−0.5	12.3	43.9
aTZ	9.8	12.3	−9.1	17.4	77.8	10.1	11.4	−9.5	13.9	29.6
estimated CBS										
[TQ]Z; δ:DZ	7.3	10.5	−8.0	15.8	64.9	7.8	9.8	−6.5	12.9	44.8
a[TQ]Z; δ:aDZ	7.5	9.6	−8.7	14.1	64.5	7.0	7.4	−3.9	9.1	29.3
[Q5]Z; δ:TZ	4.5	6.7	−4.9	11.7	61.1	4.6	5.5	−0.8	6.9	18.6
a[Q5]Z; δ:aTZ	4.0	7.4	−1.4	13.2	59.6	4.1	5.2	0.1	7.1	22.4
awC[Q5]Z; δ:awCTZ	4.7	9.4	4.8	15.4	59.9	5.8	7.0	5.7	9.6	33.9

Figure 2 shows the error distribution of fundamental and harmonic frequencies compared to experiment for all 20 molecules (14 with experimental harmonic values) listed in Table I. The corresponding error statistics are shown in Table V.

There is a rather extreme outlier that has been omitted from Fig. 2 for legibility, corresponding to the C-H bending mode of C_2H_2, with an experimental fundamental frequency observed at 612.6 cm^{-1} and an experimentally deduced harmonic frequency at 621.5 cm^{-1}. CCSD(T)/aDZ produces large errors of 359.5 cm^{-1} and -255.7 cm^{-1} for the fundamental and harmonic frequencies, respectively. This mode is known to be problematic[47] and the error stems from a poor description of carbon-carbon bonding in the out-of-plane bending modes when using correlated methods.[70] While the other levels of theory produce rather large errors for this mode, e.g., 44.5 cm^{-1} for CCSD(T)/CBS(a[Q5]Z; δ:aTZ), none have such catastrophic errors as CCSD(T)/aDZ. This is consistent with the observation by Fortenberry et al. that the poor description of the out-of-plane bending modes is particularly exacerbated by "saturation of the spd one-particle basis without inclusion of higher angular momentum functions."[70]

The impact of including core-valence effects is demonstrated by the focal-point method CCSD(T)/CBS(awC[Q5]Z; δ:awCTZ). The inclusion of core-valence correlation here does not provide systematic improvement to the predicted fundamental or harmonic vibrational frequencies. The overall accuracy remains nearly the same as the frozen-core focal-point approach with corresponding basis sets. Considering that the core-valence effects can add quite substantially to the computational cost, the frozen-core approximation to CCSD(T) would appear to be justified for the systems considered here, and improved accuracy would require inclusion of other effects, particularly higher excitations in the coupled-cluster treatment. Previous studies have demonstrated that higher-order electron correlation effects beyond CCSD(T) can affect vibrational frequencies by several cm^{-1}.[71] Additional improvements may also be afforded by increasing the basis sets in the CCSD(T) portion of the focal-point methods to quadruple-ζ, but the maximum errors in the harmonic frequencies for the focal-point methods based on valence triple-ζ basis sets, around 20 cm^{-1}, are not much larger than the maximum error for CCSD(T)/CBS(a[Q5]Z) in Table IV, around 14 cm^{-1}.

Note the presence of a few larger error modes (around $\pm$60 cm^{-1}) in the fundamental frequencies, even for the largest basis set methods presented here. In contrast, the harmonic modes have much tighter error distributions and smaller mean errors. The largest errors in

the fundamental frequencies correspond to the previously discussed C-H bending mode of C_2H_2 (experimental fundamental value of 612 cm^{-1}) and the torsional mode of HOOH (371 cm^{-1}).

As mentioned above, we do not believe that further basis set improvements beyond CCSD(T)/CBS(a[Q5]Z; δ:aTZ) are likely to provide a systematic improvement in the agreement with experimental values. Accuracy beyond this level would likely need to include off-diagonal quartic terms (or higher) in the vibrational treatment,[32] explicit triples and approximate quadruples in the coupled-cluster computation, and potentially Born-Oppenheimer diagonal corrections. We do not expect relativistic effects to have a significant impact on systems containing first row elements.[71]

Finally, we may compare the accuracy (vs experiment) of the VPT2 focal-point CCSD(T)/CBS fundamental frequencies to that of the CCSD(T)/CBS fundamental frequencies obtained by simple scaling of explicitly correlated CCSD(T) harmonic frequencies, explored by Zapata Trujillo and McKemmish.[20] As mentioned in the Introduction, simple scaling with one fitted scale factor gives a median absolute deviation of 15.9 cm^{-1} and a root mean squared error of 32 cm^{-1}, and these values decrease to 7.5 and 28 cm^{-1} upon dividing the frequencies into three regions and using a separate fitted scale factor for each. According to Table V, the focal-point VPT2 approach using valence basis sets gives median absolute deviations in the range of 4-8 cm^{-1} depending on the particular focal-point method used; those performing the CCSD(T) portion of the computation with a triple-ζ quality basis are in the 4-5 cm^{-1} range. The RMSE values range from 11.7–15.8 cm^{-1} (11.7–13.2 cm^{-1} for focal-point approaches using triple-ζ basis sets for the CCSD(T) portion). For the interested reader, the Supplemental Material presents additional graphs and tables of the absolute errors in fundamental and harmonic frequencies, and their quartiles. Our interquartile ranges for the valence triple-ζ CCSD(T) based focal point approaches are 5.8 and 7.1 cm^{-1}, compared to 14.4 cm^{-1} using scaling with one scale factor, or 12.1 cm^{-1} using three scale factors. Thus, the focal-point VPT2 results are more reliable than the scaling approaches, as one would expect, but they come at a dramatically increased computational expense compared to computing only the harmonic terms and scaling them.

TABLE VI. Timings for H_2O VPT2 routine for various levels of theory.

CCSD(T) Basis Set	Wall Time (minutes)
DZ	8.0
aDZ	8.8
TZ	10.3
aTZ	19.8
estimated CBS	
[TQ]Z; δ:DZ	12.5
a[TQ]Z; δ:aDZ	15.6
[Q5]Z; δ:TZ	23.4
a[Q5]Z; δ:aTZ	44.9
awC[Q5]Z; δ:awCTZ	67.8
a[Q5]Z	1334.2

IV. TIMINGS

Table VI shows the total wall time required for the anharmonic VPT2 calculation on H_2O, using an (older) Intel Core i7-3930K central processing unit (CPU) @ 3.20GHz with 32GB of random-access memory (RAM) and a redundant array of independent disks (RAID 0) volume for scratch files. Inclusion of the MP2 calculation for the focal-procedure does increase the computational cost significantly beyond the time required for the CCSD(T) portion for this small molecule, but the cost-accuracy trade-off is very favorable: CCSD(T)/CBS(a[TQ]Z; δ:aDZ) for example achieves greater accuracy than CCSD(T)/aTZ at a lower cost. We note that the steeper scaling of CCSD(T) relative to MP2 means the proportion of time spent on MP2 in the focal-point calculation will steadily decrease for systems larger than H_2O. CCSD(T)/CBS(a[Q5]Z; δ:aTZ), which predicts harmonic and fundamental frequencies nearly equivalent to CCSD(T)/CBS(a[Q5]Z) results, requires just a fraction of the wall time, about 3%, for H_2O. The accuracy of this procedure at a comparatively low cost provides an option for the calculation of reliable vibrational properties of small molecules with multiple heavy atoms.

V. CONCLUSIONS

Focal-point approaches for extrapolating the costly gold-standard CCSD(T) are popular for estimating high quality energies at a reduced cost. However, focal-point methods have not yet been widely adopted for anharmonic vibrational frequency analysis. Focal-point approximations to large-basis CCSD(T) gradients have already been shown to work well in geometry optimizations;[5] this work examines their performance in calculating harmonic and fundamental vibrational frequencies.

We have examined focal-point approximations to CCSD(T)/CBS vibrational frequencies, and how they converge towards the CBS limit as the basis sets are increased, for a set of 12 molecules. We have also assessed their performance in reproducing experimental vibrational frequencies in a larger set of 20 molecules. Focal-point CCSD(T) using a triple-ζ basis set for the CCSD(T) portion of the computation is nearly at the CBS limit for harmonic and fundamental frequencies, and any further increase in the basis set is not expected to improve agreement with experiment. The best performing focal-point method, CCSD(T)/CBS(a[Q5]Z; δ:aTZ) has MAEs of 7.3 and 5.3 cm^{-1} in the fundamental and harmonic frequencies, respectively. In contrast, we find that single-basis CCSD(T) does not achieve comparable accuracy for harmonic or fundamental vibrational frequencies with double- or triple-ζ basis sets. Large basis set extrapolations, focal-point approximations, or explicitly-correlated methods are evidently necessary for benchmark levels of accuracy. The focal-point procedure requires only a very small fraction of the computational cost of a large basis set extrapolation like CCSD(T)/CBS(a[Q5]Z), 3% for H_2O, and has the same level of accuracy versus experiment for harmonic and fundamental frequencies. We recommend the use of focal-point coupled-cluster methods for the calculation of vibrational frequencies of benchmark accuracy as opposed to large basis set extrapolations, or as an alternative to explicitly-correlated methods. Of course, the computational cost of VPT2 CCSD(T)/CBS fundamental frequencies would be reduced further if the anharmonic terms only were computed at a lower level of theory, as has been done in some applications.[11,13–16]

SUPPLEMENTARY MATERIAL

The Supplementary Material contains Cartesian coordinates of all optimized geometries used in the present study (.xyz format), and all computed vibrational frequencies (.csv format). A PDF file provides quartile data for errors vs experimental fundamental and harmonic frequencies and an analog of Figure 2 showing unsigned errors.

ACKNOWLEDGMENTS

The authors thank Thomas Regan for performing some preliminary investigations for harmonic frequencies. We gratefully acknowledge support from the U.S. National Science Foundation (Grant No. CHE-1955940) and the Exascale Computing Project (Grant No. 17-SC-20-SC), a collaborative effort of the U.S. Department of Energy (DOE) Office of Science and the National Nuclear Security Administration.

DATA AVAILABILITY

The data that supports the findings of this study are available with the article and its supplementary material. The PyVPT2 code used to perform anharmonic frequency analysis is freely available as open-source software (Ref. 64).

REFERENCES

[1]A. L. L. East and W. D. Allen, "The heat of formation of NCO," J. Chem. Phys. **99**, 4638–4650 (1993).

[2]A. G. Császár, W. D. Allen, and H. F. Schaefer, "In pursuit of the ab initio limit for conformational energy prototypes," J. Chem. Phys. **108**, 9751–9764 (1998).

[3]C. Puzzarini, "Extrapolation to the complete basis set limit of structural parameters: Comparison of different approaches," J. Phys. Chem. A **113**, 14530–14535 (2009).

[4]J. Černý, M. Pitoňák, K. E. Riley, and P. Hobza, "Complete basis set extrapolation and hybrid schemes for geometry gradients of noncovalent complexes," J. Chem. Theory Comput. **7**, 3924–3934 (2011).

[5]C. E. Warden, D. G. A. Smith, L. A. Burns, U. Bozkaya, and C. D. Sherrill, "Efficient and automated computation of accurate molecular geometries using focal-point approximations to large-basis coupled-cluster theory," J. Chem. Phys **152**, 124109 (2020).

[6]M. S. Marshall, L. A. Burns, and C. D. Sherrill, "Basis set convergence of the coupled-cluster correction, $\delta_{MP2}^{CCSD(T)}$: Best practices for benchmarking non-covalent interactions and the attendant revision of the S22, NBC10, HBC6, and HSG databases," J. Chem. Phys. **135**, 194102 (2011).

[7]L. A. Burns, M. S. Marshall, and C. D. Sherrill, "Comparing counterpoise-corrected, uncorrected, and averaged binding energies for benchmarking noncovalent interactions," J. Chem. Theory Comput. **10**, 49–57 (2014).

[8]S. Alessandrini, V. Barone, and C. Puzzarini, "Extension of the "cheap" composite approach to noncovalent interactions: The jun-ChS scheme," J. Chem. Theory Comput. **16**, 988–1006 (2020).

[9]A. Karton, E. Rabinovich, J. M. L. Martin, and B. Ruscic, "W4 theory for computational thermochemistry: In pursuit of confident sub-kJ/mol predictions," J. Chem. Phys. **125**, 144108 (2006).

[10]L. A. Curtiss, P. C. Redfern, and K. Raghavachari, "Gaussian-4 theory," J. Chem. Phys. **126**, 084108 (2007).

[11]C. Puzzarini, M. Biczysko, and V. Barone, "Accurate anharmonic vibrational frequencies for uracil: The performance of composite schemes and hybrid CC/DFT model," J. Chem. Theory Comput. **7**, 3702–3710 (2011).

[12]R. Kalescky, E. Kraka, and D. Cremer, "Accurate determination of the binding energy of the formic acid dimer: The importance of geometry relaxation," J. Chem. Phys. **140**, 084315 (2014).

[13]V. Barone, M. Biczysko, J. Bloino, and C. Puzzarini, "Accurate structure, thermodynamic and spectroscopic parameters from CC and CC/DFT schemes: the challenge of the conformational equilibrium in glycine," Phys. Chem. Chem. Phys. **15**, 10094–10111 (2013).

[14]C. Puzzarini, M. Biczysko, V. Barone, L. Largo, I. Peña, C. Cabezas, and J. L. Alonso, "Accurate characterization of the peptide linkage in the gas phase: A joint quantum-chemical and rotational spectroscopy study of the glycine dipeptide analogue," J. Phys. Chem. Lett. **5**, 534–540 (2014).

[15]V. Barone, M. Biczysko, J. Bloino, P. Cimino, E. Penocchio, and C. Puzzarini, "CC/DFT route toward accurate structures and spectroscopic features for observed and elusive conformers of flexible molecules: Pyruvic acid as a case study," J. Chem. Theory Comput. **11**, 4342–4363 (2015).

[16]M. Sheng, F. Silvestrini, M. Biczysko, and C. Puzzarini, "Structural and vibrational properties of amino acids from composite schemes and double-hybrid DFT: Hydrogen bonding in serine as a test case," J. Phys. Chem. A **125**, 9099–9114 (2021).

[17]A. J. C. Varandas, "Scale-free-modeling (harmonic) vibrational frequencies: Assessing accuracy and cost-effectiveness by CBS extrapolation," J. Chem. Phys. **157**, 174110 (2022).

[18]K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, "A fifth-order perturbation comparison of electron correlation theories," Chem. Phys. Lett. **157**, 479–483 (1989).

[19]J. P. Merrick, D. Moran, and L. Radom, "An evaluation of harmonic vibrational frequency scale factors," J. Phys. Chem. A **111**, 11683–11700 (2007).

[20]J. C. Zapata Trujillo and L. K. McKemmish, "VIBFREQ1295: A new database for vibrational frequency calculations," J. Phys. Chem. A **126**, 4100–4122 (2022).

[21]J. C. Zapata Trujillo and L. K. McKemmish, "Model chemistry recommendations for scaled harmonic frequency calculations: A benchmark study," J. Phys. Chem. A **127**, 1715–1735 (2023).

[22]H. H. Nielsen, "The vibration-rotation energies of molecules," Rev. Mod. Phys. **23**, 90–136 (1951).

[23]V. Barone, "Anharmonic vibrational properties by a fully automated second-order perturbative approach," J. Chem. Phys **122**, 014108 (2004).

[24]J. M. Bowman, "Self-consistent field energies and wavefunctions for coupled oscillators," J. Chem. Phys. **68**, 608–610 (1978).

[25]J. M. Bowman, "The self-consistent-field approach to polyatomic vibrations," Acc. Chem. Res. **19**, 202–208 (1986).

[26]J. M. Bowman, K. Christoffel, and F. Tobin, "Application of SCF-SI theory to vibrational motion in polyatomic molecules," J. Phys. Chem. **83**, 905–912 (1979).

[27]K. M. Christoffel and J. M. Bowman, "Investigations of self-consistent field, SCF CI and virtual state configuration interaction vibrational energies for a model three-mode system," Chem. Phys. Lett. **85**, 220–224 (1982).

[28]Q. Yang, M. Mendolicchio, V. Barone, and J. Bloino, "Accuracy and reliability in the simulation of vibrational spectra: A comprehensive benchmark of energies and intensities issuing from generalized vibrational perturbation theory to second order (GVPT2)," Front. Astron. Space Sci. **8**, 77 (2021).

[29]D. P. Tew, W. Klopper, M. Heckert, and J. Gauss, "Basis set limit CCSD(T) harmonic vibrational frequencies," J. Phys. Chem. A **111**, 11242–11248 (2007).

[30]J. M. L. Martin and M. K. Kesharwani, "Assessment of CCSD(T)-F12 approximations and basis sets for harmonic vibrational frequencies," J. Chem. Theory Comput. **10**, 2085–2090 (2014).

[31]D. Agbaglo and R. C. Fortenberry, "The performance of explicitly correlated wavefunctions [CCSD(T)-F12b] in the computation of anharmonic vibrational frequencies," Chem. Phys. Lett. **734**, 136720 (2019).

[32]W. J. Morgan, D. A. Matthews, M. Ringholm, J. Agarwal, J. Z. Gong, K. Ruud, W. D. Allen, J. F. Stanton, and H. F. I. Schaefer, "Geometric energy derivatives at the complete basis set limit: Application to the equilibrium structure and molecular force field of formaldehyde," J. Chem. Theory Comput. **14**, 1333–1350 (2018).

[33]K. K. Irikura, "Experimental vibrational zero-point energies: Diatomic molecules," J. Phys. Chem. Ref. Data **36**, 389–397 (2007).

[34]R. J. Le Roy, "Improved parameterization for combined isotopomer analysis of diatomic spectra and its application to HF and DF," J. Mol. Spectrosc. **194**, 189–196 (1999).

[35]J. Tennyson, N. F. Zobov, R. Williamson, O. L. Polyansky, and P. F. Bernath, "Experimental energy levels of the water molecule," J. Phys. Chem. Ref. Data **30**, 735–831 (2001).

[36]G. Strey, "Extremal properties of force constants: Part I," J. Mol. Spectrosc. **24**, 87–99 (1967).

[37]A. Maki and G. Mellau, "High-temperature infrared emission measurements on HNC," J. Mol. Spectrosc. **206**, 47–52 (2001).

[38]I. Suzuki, M. A. Pariseau, and J. Overend, "General quartic force field of HCN," J. Chem. Phys. **44**, 3561–3567 (1966).

[39]G. Strey and I. M. Mills, "The anharmonic force field and equilibrium structure of HCN and HCP," Mol. Phys. **26**, 129–138 (1973).

[40]A. Cabana, Y. Doucet, J.-M. Garneau, C. Pépin, and P. Puget, "The vibration-rotation spectrum of methinophosphide: The overtone bands $2\nu_1$ and $2\nu_3$, the summation bands $\nu_1 + \nu_2$ and $\nu_2 + \nu_3$, and the difference band ν_1 - ν_2," J. Mol. Spectrosc. **96**, 342–350 (1982).

[41]J. Koput and S. Carter, "The potential energy surface and vibrational-rotational energy levels of HCP," Spectrochim. Acta A Mol. Biomol. Spectrosc. **53**, 1091–1100 (1997), ab Initio and Ab Initio Derived Force Fields: State of the Science.

[42]J. Petersen and M. Vervloet, "Infrared emission spectrum of HNO: The ν_1 band," Chem. Phys. Lett. **141**, 499–502 (1987).

[43]J. W. C. Johns and A. R. W. McKellar, "Laser Stark spectroscopy of the fundamental bands of HNO (ν_2 and ν_3) and DNO (ν_1 and ν_2)," J. Chem. Phys. **66**, 1217–1224 (1977).

[44]A. Requena, A. Bastida, and J. Zúñiga, "Curvilinear Jacobi and Radau normal coordinates for linear triatomic molecules. Application to CO_2," Chem. Phys. **175**, 255–264 (1993).

[45]A. Gershikov and V. Spiridonov, "Anharmonic force field of CO_2 as determined by a gas-phase electron diffraction study," J. Mol. Struct. **96**, 141–149 (1983).

[46]W. B. Person and G. Zerbi, *Vibrational Intensities in Infrared and Raman Spectroscopy* (Elsevier, Amsterdam, 1982).

[47]G. Rauhut, G. Knizia, and H.-J. Werner, "Accurate calculation of vibrational frequencies using explicitly correlated coupled-cluster theory," J. Chem. Phys **130**, 054105 (2009).

[48]M. J. Coffey, H. L. Berghout, E. Woods, and F. F. Crim, "Vibrational spectroscopy and intramolecular energy transfer in isocyanic acid (HNCO)," J. Chem. Phys. **110**, 10850–10862 (1999).

[49]D. E. Reisner, R. W. Field, J. L. Kinsey, and H. Dai, "Stimulated emission spectroscopy: A complete set of vibrational constants for $\tilde{X}^1A_1$ formaldehyde," J. Chem. Phys **80**, 5968–5978 (1984).

[50]D. C. Burleigh, A. B. McCoy, and E. L. Sibert, "An accurate quartic force field for formaldehyde," J. Chem. Phys. **104**, 480–487 (1996).

[51]A. L. L. East, W. D. Allen, and S. J. Klippenstein, "The anharmonic force field and equilibrium molecular structure of ketene," J. Chem. Phys. **102**, 8506–8532 (1995).

[52]D. Van Lerberghe, I. J. Wright, and J. L. Duncan, "High-resolution infrared spectrum and rotational constants of ethylene-H_4," J. Mol. Spectrosc. **42**, 251–273 (1972).

[53]D. A. Clabo, W. D. Allen, R. B. Remington, Y. Yamaguchi, and H. F. Schaefer, "A systematic study of molecular vibrational anharmonicity and vibration—rotation interaction by self-consistent-field higher-derivative methods. asymmetric top molecules," Chem. Phys. **123**, 187–239 (1988).

[54]N. C. Craig, K. L. Petersen, and D. C. McKean, "Vibrational and quantum chemical studies of 1,2-difluoroethylenes: Spectra of 1,2-$^{13}C_2H_2F_2$ species, scaled force fields, and dipole derivatives," J. Phys. Chem. A **106**, 6358–6369 (2002).

[55]T. H. Dunning, "Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen," J. Chem. Phys. **90**, 1007–1023 (1989).

[56]T. H. Dunning, K. A. Peterson, and A. K. Wilson, "Gaussian basis sets for use in correlated molecular calculations. X. The atoms aluminum through argon revisited," J. Chem. Phys. **114**, 9244–9253 (2001).

[57]R. A. Kendall, T. H. Dunning, and R. J. Harrison, "Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions," J. Chem. Phys. **96**, 6796–6806 (1992).

[58]K. A. Peterson and T. H. Dunning, "Accurate correlation consistent basis sets for molecular core–valence correlation effects: The second row atoms Al–Ar, and the first row atoms B–Ne revisited," J. Chem. Phys. **117**, 10548–10560 (2002).

[59]A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen, and A. K. Wilson, "Basis-set convergence in correlated calculations on Ne, N_2, and H_2O," Chem. Phys. Lett. **286**, 243–252 (1998).

[60]D. G. A. Smith, L. A. Burns, A. C. Simmonett, R. M. Parrish, M. C. Schieber, R. Galvelis, P. Kraus, H. Kruse, R. Di Remigio, A. Alenaizan, A. M. James, S. Lehtola, J. P. Misiewicz,

M. Scheurer, R. A. Shaw, J. B. Schriber, Y. Xie, Z. L. Glick, D. A. Sirianni, J. S. O'Brien, J. M. Waldrop, A. Kumar, E. G. Hohenstein, B. P. Pritchard, B. R. Brooks, H. F. Schaefer, A. Y. Sokolov, K. Patkowski, A. E. DePrince, U. Bozkaya, R. A. King, F. A. Evangelista, J. M. Turney, T. D. Crawford, and C. D. Sherrill, "Psi4 1.4: Open-source software for high-throughput quantum chemistry," J. Chem. Phys. **152**, 184108 (2020).

[61]F. Weigend, "A fully direct RI-HF algorithm: Implementation, optimised auxiliary basis sets, demonstration of accuracy and efficiency," Phys. Chem. Chem. Phys. **4**, 4285–4291 (2002).

[62]F. Weigend, A. Köhn, and C. Hättig, "Efficient use of the correlation consistent basis sets in resolution of the identity MP2 calculations," J. Chem. Phys **116**, 3175–3183 (2002).

[63]K. Yagi, K. Hirao, T. Taketsugu, M. W. Schmidt, and M. S. Gordon, "Ab initio vibrational state calculations with a quartic force field: Applications to H_2CO, C_2H_4, CH_3OH, CH_3CCH, and C_6H_6," J. Chem. Phys. **121**, 1383–1389 (2004).

[64]P. M. Nelson, PyVPT2, `https://github.com/philipmnel/pyvpt2`.

[65]J. M. L. Martin, T. J. Lee, P. R. Taylor, and J. François, "The anharmonic force field of ethylene, C_2H_4, by means of accurate ab initio calculations," J. Chem. Phys **103**, 2589–2602 (1995).

[66]P. R. Franke, J. F. Stanton, and G. E. Douberly, "How to VPT2: Accurate and intuitive simulations of CH stretching infrared spectra using VPT2+K with large effective hamiltonian resonance treatments," J. Phys. Chem. A **125**, 1301–1324 (2021).

[67]R. J. Bartlett and J. F. Stanton, "Applications of post-Hartree–Fock methods: A tutorial," in *Reviews in Computational Chemistry*, Vol. 5, edited by K. B. Lipkowitz and D. B. Boyd (VCH Publishers, New York, 1994) Chap. 2, pp. 65–170.

[68]T. A. Ruden, T. Helgaker, P. Jørgensen, and J. Olsen, "Coupled-cluster connected quadruples and quintuples corrections to the harmonic vibrational frequencies and equilibrium bond distances of HF, N_2, F_2, and CO," J. Chem. Phys. **121**, 5874–5884 (2004).

[69]U. Bozkaya and C. D. Sherrill, "Analytic energy gradients for the coupled-cluster singles and doubles with perturbative triples method with the density-fitting approimation," J. Chem. Phys. **147**, 044104 (2017).

[70]R. C. Fortenberry, T. J. Lee, and J. P. Layfield, "Communication: The failure of correlation to describe carbon=carbon bonding in out-of-plane bends," J. Chem. Phys **147**, 221101 (2017).

[71]B. Temelso, E. F. Valeev, and C. D. Sherrill, "A comparison of one-particle basis set completeness, higher-order electron correlation, relativistic effects, and adiabatic corrections for spectroscopic constants of BH, CH^+, and NH," J. Phys. Chem. A **108**, 3068–3075 (2004).