Prepared for publication in *J. Chem. Phys.*
15 June 2004

Accurate vibrational-rotational partition functions and standard-state free energy values for H_2O_2 from Monte Carlo path-integral calculations

Vanessa Audette Lynch,[a] Steven L. Mielke,[b*] and Donald G. Truhlar[a*]

[a]*Department of Chemistry and Supercomputer Institute, University of Minnesota, 207 Pleasant St. S.E., Minneapolis, Minnesota 55455-0431*

[b]*Department of Chemistry, Northwestern University, Evanston, Illinois 60208-3113*

Abstract

Accurate quantum mechanical partition functions and absolute free energies of H_2O_2 are determined using a realistic potential energy surface [J. Koput, S. Carter, and N. C. Handy, J. Phys. Chem. A **102**, 6325 (1998)] for temperatures ranging from 300 K to 2400 K by using Monte Carlo path integral calculations with new, efficient polyatomic importance sampling methods. The path centroids are sampled in Jacobi coordinates via a set of independent ziggurat schemes. The calculations employed enhanced-same-path extrapolation of trapezoidal Trotter Fourier path integrals, and the paths were constructed using fast Fourier sine transforms. Importance sampling was also used in Fourier coefficient space, and adaptively optimized stratified sampling was used in configuration space. The free energy values obtained from the path integral calculations are compared to separable-mode approximations, to the Pitzer–Gwinn approximation, and to values in thermodynamic tables. Our calculations support the recently proposed revisions to the JANAF tables.

*Email: slmielke@chem.northwestern.edu; truhlar@chem.umn.edu

I. INTRODUCTION

A key function of statistical mechanics is to provide methods for calculating thermodynamic properties, such as free energies and entropy, in terms of molecular properties. Rotational-vibrational partition functions are the generator for thermodynamic and kinetic quantities such as free energies, equilibrium constants, and reaction rates.[1-3] The standard method of calculating partition functions is to obtain or estimate the energy eigenvalues of a system and to sum their Boltzmann factors. To obtain an accurate result, it is necessary to know a very large number of rovibrational energies. This limits the applicability of the method to small systems unless one is willing to accept severe approximations such as the popular harmonic oscillator–rigid rotator approximation or the use of classical mechanics. For many systems of current interest, these approximations are not even qualitatively correct; in many other cases, such as molecules with torsions, they are not quantitatively accurate. Path integral methods[4-7] provide an alternative approach that can be used without such approximations, and we have used them to calculate accurate rovibrational partition functions for H_2O[8,9] and HCl.[10,11]

The vibrational-rotational, partition function of a molecule is defined as[1-3]

$$Q(T) = \sum_n e^{-E_n / k_B T} \tag{1}$$

where E_n is the energy of vibration-rotation state n, k_B is Boltzmann's constant, and T is the temperature. The standard method of calculating partition functions by summing their Boltzmann factors becomes very difficult as the dimensionality of the system increases, partly because the number of rovibrational eigenvalues required becomes very large and partly because the calculation of accurate eigenvalues for large systems is a

demanding task. Converged vibrational-rotational eigenvalue calculations have been successfully calculated this way for small systems such as H_2O[12] and CH_4;[13] however, it is very difficult to obtain enough accurate eigenvalues to calculate $Q(T)$ for more complex systems[14-24] unless additional approximations are made. One such approximation often made is the harmonic oscillator–rigid rotator approximation. Such approximations significantly reduce the computational cost; however, these methods are not accurate enough to treat many important systems quantitatively, for example, systems that include torsions, such as H_2O_2. Path integral methods offer the ability to calculate accurate values of the vibrational-rotational partition function for systems with torsions and large anharmonicity with less computational cost than calculating the necessary energy eigenvalues. The efficiency of Monte Carlo path integral calculations depends strongly on the sampling scheme employed. In the present article, we will present a new scheme for importance sampling in Jacobi coordinates that greatly increases the efficiency compared to our prior[25] approach.

In Sections II–V, we review the theory and computational methods for calculating the partition function for a given potential energy surface (PES). Section VI summarizes separable approximation methods, in particular the harmonic-oscillator approximation, the rigid-rotator approximation, and the Pitzer–Gwinn method. Section VII compares the accurate results from path integral calculations on H_2O_2 to results obtained by the approximate methods of Sect. VI, with both sets of calculations based on the potential energy surface (PES) of Koput et al.[24] The free energy values calculated from the path integral calculations are also compared to those available in thermodynamics tables, which are based, to the extent possible, on experiment.[26,27]

II. THEORY

II.A. Path integrals

If we neglect complications from nuclear spin statistics and dissociative states (both of which are unimportant for the present application), the exact expression for the quantum mechanical internal (vibrational-rotational) partition function of a molecule is[4,5]

$$Q(T) = \frac{1}{\sigma^{\text{sym}}} \int d\mathbf{x} \oint \mathbf{D}[\mathbf{x}(s)] \exp\left(-\frac{1}{\hbar} A[\mathbf{x}(s)] \right) \tag{2}$$

where $\hbar$ is Planck's constant divided by 2π, σ^{sym} is the symmetry number of the system,[28,29] $\oint \mathbf{D}[\mathbf{x}(s)]$ is the integral over all closed paths $\mathbf{x}(s)$ whose centroid position occurs at $\mathbf{x}$, and $A[\mathbf{x}(s)]$ is the action integral of path $\mathbf{x}(s)$ and is given by the expression

$$A[\mathbf{x}(s)] = \int_0^{\beta\hbar} ds \left(\frac{\mu}{2} \left(\frac{d\mathbf{x}}{dt} \right)^2 + V(\mathbf{x}) \right), \tag{3}$$

where β is $1/k_B T$, μ is the reduced mass of the system, $V(\mathbf{x})$ is the potential energy at point $\mathbf{x}$, and s is the distance along the path.

Fourier path integral (FPI) methods[4-11,25,30-41] represent the deviations of the paths from free-particle paths by a Fourier expansion. The Fourier representation transforms the path integral into an infinite-dimensional Riemann integral. The transformation from an integral over an infinite number of paths to an integral over Fourier coefficients has two advantages. First, it provides a compact and well-ordered representation of all closed paths that have the centroid position $\mathbf{x}$. Second, the use of an infinite Fourier series to represent each path allows for straightforward approximate treatments based on

truncating the number of Fourier coefficients from infinity to some finite value K or based on integrating the paths with a finite number of points.

The conventional approach to FPI calculations truncates the expansion at K terms, and the results converge as $O(1/K)$, i.e., in the limit that $K \to \infty$, $\dfrac{Q^{\infty}(T) - Q^{[K]}(T)}{1/K}$ is finite. A more efficient approach, introduced by Coalson,[33] that also uses a finite expansion but additionally uses a K-dependent rescaling of the Fourier coefficients, results in $P = K + 1$ evenly spaced points on the paths being distributed as they would be in an infinite-dimensional Fourier expansion. Coalson[33] showed that if quadratures over such paths are evaluated using a P-point trapezoidal rule, the method is isomorphic to the widely used trapezoidal Trotter discretized path integral scheme; hence, we refer to this as the trapezoidal Trotter FPI (TT-FPI) method. Partition function estimates obtained via TT-FPI converge as $O(1/K^2)$. Recently, Predescu et al.[39-40] have developed a scheme called re-weighted Weiner-FPI (RW-WFPI) by a different rescaling of the Fourier coefficients; formally this results in an $O(1/K^3)$ convergence for well behaved potentials, but the methods approach their asymptotic convergence rates more slowly than either the conventional or TT-FPI approaches do. We have recently shown[42] that using the TT-FPI approach allows one to construct lower-order path integral estimates at essentially no additional cost, and that appropriate sequences of these estimates converge monotonically on a path-by-path basis. We further showed that these estimates display regular convergence that permits highly accurate extrapolation to the infinite-P limit. This property makes the TT-FPI approach very attractive for practical computations.

In the TT-FPI method,[11,33,42] we only need to determine the path at $P = K + 1$ discretized points s_n that are evenly spaced in imaginary time. We expand the closed paths $\mathbf{x}(s)$ in a Fourier series, and the coordinate at point n is given by

$$\bar{x}_j(s_n) = x_j + \sum_{k=1}^{K} a_{jk} \sin\left(\frac{\pi(n-1)k}{K+1}\right) - x_j^{\mathrm{C}} \tag{4}$$

where a_{jk} is a coefficient in the Fourier series, the summation term is called the relative path, x_j^{C} is the centroid position of the relative path (which must be removed if the full path is to be centered on $\mathbf{x}$) and, as above, K is the length of the Fourier expansion. The fact that the path is closed is reflected by the fact that s_1 equals s_{P+1}. Since we only require evenly spaced points, the relative paths can be readily constructed using a fast Fourier sine transform.[11] The centroid of the relative path may be calculated by averaging the positions of the $P = K + 1$ discretized points which yields

$$x_j^{\mathrm{C}} = \frac{1}{K+1} \sum_{i=1}^{K+1} \sum_{k=1}^{K} a_{jk} \sin\left(\frac{\pi(i-1)k}{K+1}\right). \tag{5}$$

Since we use Eq. (4), the configuration space coordinates, $\mathbf{x}$, in equations like (2) and (3) refer to the centroid position rather than to a specific point on the path. We note that in several previous papers[8-11,25,38] we chose $\mathbf{x}$ to be the initial position of a path rather than its centroid. In our importance sampling scheme that will be discussed below, we will sample paths based on the value of the single point $\mathbf{x}$, and the centroid position of a path is a better way to characterize a path than any single point on the path.

Upon simplification, and using mass-scaled Jacobi coordinates with all masses scaled to μ, Eq. (2) can then be expressed as[11,33,42]

$$Q^{[P]}(T) = \frac{J(T)}{\sigma^{\mathrm{sym}}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left(\prod_{j=1}^{N} dx_j \right) \left(\prod_{j=1}^{N} \prod_{k=1}^{K} da_{jk} \right)$$

$$\times \exp\left[\left(- \sum_{j=1}^{N} \sum_{k=1}^{K} \frac{a_{jk}^2}{2\,\sigma_{k;K}^2} \right) - \frac{S(\mathbf{x}, \mathbf{a})}{\hbar} \right] \tag{6}$$

where $J(T)$ is the Jacobian[4] of the transformation from the integral over paths to the integral over Fourier coefficients given by

$$J(T) = \prod_{j=1}^{N} \sqrt{\frac{\mu}{2\pi\beta\hbar}} \sum_{k=1}^{K} \frac{1}{\sigma_{k;K}\sqrt{2\pi}} \tag{7}$$

where $\sigma_{k;K}$ is a fluctuation parameter and is given by the equation

$$\sigma_{k;K}^2 = \frac{2\beta\hbar^2}{\mu} \frac{1}{\left[2(K+1) \sin \dfrac{k\pi}{2(K+1)} \right]^2}, \tag{8}$$

and $S(\mathbf{x},\mathbf{a})$ is the contribution of the potential energy (with a P-point trapezoidal rule integration) to the action integral for a given path calculated by the expression

$$S^{\mathrm{TT-FPI}}(\mathbf{x}, \mathbf{a}) = \frac{\hbar\beta}{2P} \sum_{i=1}^{P} (V[\mathbf{x}(s_i)] + V[\mathbf{x}(s_{i+1})]) \tag{9}$$

where $V(\mathbf{x})$ is the potential energy.

Equation (6) is an infinite-dimensional integral over an infinite domain. To evaluate this by Monte Carlo methods, we define a finite region of space in which to sample. The contribution to the total partition function should be greatest near and around the equilibrium geometry. As the system gets further away from the equilibrium structure, the contribution to the partition function becomes negligible. We can therefore define a finite volume of configuration space where the contribution to the partition function needs to be calculated.

If we convert the integral in Eq. (6) to an average and employ Monte Carlo methods, we can reduce the evaluation of the integral to the problem of sampling a region of space and summing the contribution from each sample. We first restrict the N-dimensional configuration space to a finite domain D of volume V_D. Then we choose the number, P, of points to be used in the integration of the paths, and adopt Coalson's rescaling[11,33] which permits us to obtain these points on an infinite-dimensional path using only a finite Fourier expansion with $K = P - 1$ terms. This yields an approximate value, called $Q_{\text{int}}^{[P]}(T)$ or $Q_{\text{TT-FPI}}^{[P]}(T)$, for the internal partition function, where the P in the superscript indicates the number of quadrature points. Then, by multiplying and dividing Eq. (6) by $Q_{\text{fp}}^{[P]}(T)$, the P-point approximation to the free particle partition function, and by recognizing that $Q_{\text{fp}}^{[P]}(T)$ corresponds to the same N-dimensional space, we obtain

$$Q_{\text{int}}^{[P]}(T) = \frac{Q_{\text{fp}}^{[P]}(T)}{\sigma^{\text{sym}}} \frac{\int_D d\mathbf{x} \int_{-\infty}^{\infty} d\mathbf{a} \left(\exp\left(-\sum_{j=1}^{N} \sum_{k=1}^{K} \frac{a_{jk}^2}{2\sigma_{k;K}^2} \right) \exp\left[\frac{-S(\mathbf{x},\mathbf{a})}{\hbar} \right] \right)}{V_D Q_{\text{fp}}^{[P]}(T)}$$

$$= \frac{Q_{\text{fp}}^{[P]}(T)}{\sigma^{\text{sym}}} \frac{\int_D d\mathbf{x} \int_{-\infty}^{\infty} d\mathbf{a} \left(\exp\left(-\sum_{j=1}^{N} \sum_{k=1}^{K} \frac{a_{jk}^2}{2\sigma_{k;K}^2} \right) \exp\left[\frac{-S(\mathbf{x},\mathbf{a})}{\hbar} \right] \right)}{V_D \int_{-\infty}^{\infty} d\mathbf{a} \exp\left(-\sum_{j=1}^{N} \sum_{k=1}^{K} \frac{a_{jk}^2}{2\sigma_{k;K}^2} \right)} \tag{10}$$

which is in a form that can be readily used in Monte Carlo evaluation. Note that the free particle partition function—which is independent of P—is given by[4]

$$Q_{\mathrm{fp}}^{[P]}(T) = Q_{\mathrm{fp}}(T) = V_D\left(\frac{\mu}{2\pi\beta\hbar^2}\right)^{N/2} \tag{11}$$

where V_D is the volume of the domain D. It is worth noting that in prior work[11] we have used a slightly different notation where we placed K instead of P in brackets to emphasize the Fourier-like approach for sampling the infinite-dimensional paths via a finite, but rescaled, Fourier expansion, but we have changed the notation since the relevant approximation is a quadrature approximation and with our new extrapolation approach[42] it is more convenient to consider the results in this light.

II.B. Extrapolation

The extrapolation of the path integral results to the infinite-P limit in the enhanced same-path extrapolation (ESPE) approach[42] is done by fitting the $Q_{\mathrm{TT-FPI}}^{[P]}(T)$ for the highest three available values of P (in the present article we use P, $P/2$, and $P/3$) to the asymptotic behavior

$$Q_{\mathrm{TT-FPI}}^{[P]}(T) = Q^{\mathrm{ESPE}}(T) + \frac{C_2}{P^2} + \frac{C_3}{P^3}, \tag{12}$$

where C_2 and C_3 are fitting parameters. As we shall demonstrate below, when appropriate values of P are used the accuracy of the extrapolated result is comparable to that of the $Q_{\mathrm{TT-FPI}}^{[P]}(T)$ data, i.e., the extrapolation error is small compared to the statistical sampling errors.

III. ALGORITHM

The choice of mass-scaled Jacobi coordinates is different than was used in our prior[25] work. We label H-O-O-H as A-B-C-D. We remove the center-of-mass motion trivially, so the dimension, *N,* of the configuration space is 9, and this space is spanned by three 3-vectors. We choose a basis in which one vector connects B to A, one connects C to D, and one connects the centers of mass of these two fragments (CD to BA).[43] The magnitudes of these coordinates are the two OH bond distances and a distance that is similar to the OO bond distance; this identification with physical distances will be useful in the configuration space importance sampling introduced below. Importance sampling is used in both the Fourier coefficient space and the configuration space.

In order to evaluate Eq. (10) with Monte Carlo methods, we first define a density function $g(\mathbf{a})$. For an N-dimensional system, the function $g(\mathbf{a})$ is a normalized $(N \times K)$-dimensional Gaussian probability density function in Fourier coefficient ($\mathbf{a}$) space given by

$$g(\mathbf{a}) = \frac{\exp\left(- \sum_{j=1}^{N} \sum_{k=1}^{K} \frac{a_{jk}^2}{2\sigma_{k;K}^2} \right)}{\int_{-\infty}^{\infty} d\mathbf{a} \exp\left(- \sum_{j=1}^{N} \sum_{k=1}^{K} \frac{a_{jk}^2}{2\sigma_{k;K}^2} \right)}. \tag{13}$$

With this definition, our expression for the partition function becomes

$$Q_{\text{int}}^{[P]}(T) = \frac{Q_{\text{fp}}(T)}{\sigma^{\text{sym}}} \frac{\int_D d\mathbf{x} \int_{-\infty}^{\infty} d\mathbf{a} \exp\left[\frac{-S(\mathbf{x},\mathbf{a})}{\hbar} \right] g(\mathbf{a})}{V_D}. \tag{14}$$

Importance sampling provides an efficient method of sampling the Fourier coefficient degrees of freedom. In general, the importance function should reflect the shape of the

integrand or a chief contributor to that shape; in the present work the importance function for the Fourier space is taken as $g(\mathbf{a})$, defined in Eq. (13).

In addition to the importance sampling in Fourier space, importance sampling is used in the configuration space, with an importance function $f(\mathbf{x})$, to be explained below. Multiplying and dividing Eq. (14) by $f(\mathbf{x})$ and the average, $\langle f \rangle$, of $f(\mathbf{x})$, we obtain

$$Q_{\text{int}}^{[P]}(T) = \frac{Q_{\text{fp}}(T)}{\sigma^{\text{sym}}} \frac{\langle f \rangle}{\langle f \rangle} \frac{\int_D d\mathbf{x} \int_{-\infty}^{\infty} d\mathbf{a} \dfrac{\exp[-S(\mathbf{x},\mathbf{a})/\hbar]}{f(\mathbf{x})} g(\mathbf{a}) f(\mathbf{x})}{V_D} \tag{15}$$

where the average of $f(\mathbf{x})$ is defined as

$$\langle f \rangle = \frac{\int_D d\mathbf{x}\, f(\mathbf{x})}{\int_D d\mathbf{x}} = \frac{\int_D d\mathbf{x}\, f(\mathbf{x})}{V_D}. \tag{16}$$

If we substitute (16) in the denominator of Eq. (15) we obtain

$$Q_{\text{int}}^{[P]}(T) = \frac{Q_{\text{fp}}(T)\langle f \rangle}{\sigma^{\text{sym}}} \frac{\int_D d\mathbf{x} \int_{-\infty}^{\infty} d\mathbf{a} \dfrac{\exp[-S(\mathbf{x},\mathbf{a})/\hbar]}{f(\mathbf{x})} g(\mathbf{a}) f(\mathbf{x})}{\int_D d\mathbf{x} f(\mathbf{x})}. \tag{17}$$

We now define another density function $p(\mathbf{x})$ as

$$p(\mathbf{x}) = \frac{f(\mathbf{x})}{\int_D d\mathbf{x}\, f(\mathbf{x})} \tag{18}$$

so that our expression for the partition function becomes

$$Q_{\text{int}}^{[P]}(T) = \frac{Q_{\text{fp}}(T)\langle f \rangle}{\sigma^{\text{sym}}} \int_D d\mathbf{x} \int_{-\infty}^{\infty} d\mathbf{a} \frac{\exp[-S(\mathbf{x},\mathbf{a})/\hbar]}{f(\mathbf{x})} g(\mathbf{a}) p(\mathbf{x}). \tag{19}$$

The Monte Carlo estimate of Eq. (19) is obtained by choosing samples i in $(\mathbf{x},\mathbf{a})$ space where the components of $\mathbf{a}$ are sampled with relative weight $g(\mathbf{a})$, and the

components of the coordinate space, **x**, are sampled with relative weight $p(\mathbf{x})$. Then Eq. (19) becomes

$$Q_{\rm int}^{[P]}(T) = \frac{Q_{\rm fp}(T)\langle f\rangle}{\sigma^{\rm sym}} \frac{1}{N^{\rm samples}} \sum_{i=1}^{N^{\rm samples}} \frac{\exp[-S(\mathbf{x}_{(i)},\mathbf{a}_{(i)})/\hbar]}{f(\mathbf{x}_{(i)})} \tag{20}$$

where $N^{\rm samples}$ is the number of samples. Because we will use uncorrelated samples, the variance of the estimate of Eq. (20) is

$$w = \frac{Q_{\rm fp}^2(T)\langle f\rangle^2}{(\sigma^{\rm sym})^2 N^{\rm samples}[N^{\rm samples}-1]} \times \left[\sum_{i=1}^{N^{\rm samples}} \left(\frac{\exp[-S(\mathbf{x}_{(i)},\mathbf{a}_{(i)})/\hbar]}{f(\mathbf{x}_{(i)})} \right)^2 - \frac{1}{N^{\rm samples}} \left(\sum_{i=1}^{N^{\rm samples}} \frac{\exp[-S(\mathbf{x}_{(i)},\mathbf{a}_{(i)})/\hbar]}{f(\mathbf{x}_{(i)})} \right)^2 \right] \tag{21}$$

which is the square of the standard error of $Q_{\rm int}^{[P]}(T)$.

IV. SAMPLING STRATEGY

The details of the scheme for sampling **a** with relative weight $g(\mathbf{a})$ have been presented previously.[11] In the present work, we use a new scheme for sampling **x** with relative weight $f(\mathbf{x})$ defined later in this section. In sampling the 9-dimensional configuration space, we define the domain of interest as a direct product of three 3-dimensional annuli. It is useful to visualize a D-dimensional configuration space sample point (CSSP) as consisting of $D/3$ vectors (of length 3) corresponding to the Jacobi vectors[43] that define the coordinate system. Each of these vectors can be considered as a (generalized) Jacobi magnitude (the 2-norm of the vector) and a direction. The Jacobi directions define the angular degrees of freedom, and the magnitudes define the Jacobi magnitude degrees of freedom (which we also refer to as "radial" degrees of freedom).

Each of the Jacobi directions corresponds to a unit 3-vector, and sampling these directions is equivalent to performing a uniform sample of the surface of a 3-sphere. Thus, each D-dimensional CSSP can be rewritten as: $(x_1, x_2, \ldots . x_D) \equiv (R_1 \mathbf{V}_1, R_2 \mathbf{V}_2, \ldots, R_{D/3} \mathbf{V}_{D/3})$, where each R_i is a scalar corresponding to the magnitude of a Jacobi vector, and each $\mathbf{V}_i$ is a unit vector in 3-space corresponding to the direction of a Jacobi vector.

The finite sampling domain may be defined by specifying upper and lower bounds, $R_{\mathrm{l},i}$ and $R_{\mathrm{u},i}$, on each of the R_i values. The volume of the sampling domain is then given by

$$V_D = \prod_{i=1}^{D/3} \left(\frac{4\pi}{3} \left[R_{\mathrm{u},i}^3 - R_{\mathrm{l},i}^3 \right] \right). \tag{22}$$

This approach permits a large reduction in the volume of the sampling domain as compared to a minimum-volume hyperannulus (which is the sampling domain we used for 3- and 4-body problems in prior work[8,9,11,25,44]), and it also permits considerable simplification in the calculation of the average values of the importance function on the various strata. The values of the lower limits can all be taken as 0 with safety and with only a very slight increase in cost, provided that we are using a good importance-sampling scheme. Good values of the upper limits are easily identified by trial, and the performance penalty (if we are using a good importance sampling function) for only lightly optimizing these parameters is modest.

The method of stratified sampling can be utilized, along with importance sampling, in evaluating the coordinate space integral in the partition function in order to further reduce the sampling error. Stratification involves subdividing the configuration

space into disjoint strata. The effect of combining importance sampling with stratified sampling allows a stratum where the contribution to the partition function is large or has a large sampling variance to be sampled more frequently. In our calculations, we use an adaptively optimized stratified sampling (AOSS) scheme that has been presented previously;[25,38] this is a one-dimensional scheme (previously the coordinate chosen was the hyperradius; the choice of stratification coordinate for the present work is explained below) with N^{strata} strata. In this scheme, after distributing a number of initial samples (typically 10–20% of the total number of intended samples) in a stratum-blind manner, the optimal distribution of the remaining samples that should minimize the variance is estimated. As the calculation progresses, and more accurate information becomes available, these estimates are periodically updated. This updating occurs N^{sweep} times, where, for the present set of calculations, N^{sweep} is taken as 20.

Combining AOSS and importance sampling, the expression for the partition function becomes

$$Q_{\text{int}}^{[P]}(T) = \sum_{n=1}^{N^{\text{strata}}} \frac{Q_{\text{fp},n}(T)\langle f\rangle_n}{\sigma^{\text{sym}} N^{\text{samps},n}} \sum_{i=1}^{N^{\text{samps},n}} \frac{\exp[-S(\mathbf{x}_{(i)},\mathbf{a}_{(i)})/\hbar]}{f(\mathbf{x}_{(i)})} \tag{23}$$

where $N^{\text{samps},n}$ is the number of samples in stratum n. The total number of samples is

$$N^{\text{samps}} = \sum_{n=1}^{N^{\text{strata}}} N^{\text{samps},n} \tag{24}$$

and

$$\langle f\rangle_n = \frac{\int_{D_n} d\mathbf{x} f(\mathbf{x})}{V_n} \tag{25}$$

where D_n is the domain corresponding to stratum n, and V_n is its volume.

The variance of the estimate in stratum n is given by

$$w_n = \frac{Q_{\mathrm{fp},n}(T)^2 \langle f \rangle_n^{\,2}}{(\sigma^{\mathrm{sym}})^2 N^{\mathrm{samps},n}[N^{\mathrm{samps},n}-1]} \times \left[\sum_{i=1}^{N^{\mathrm{samps},n}} \left(\frac{\exp[-S(\mathbf{x}_{(i)},\mathbf{a}_{(i)})/\hbar]}{f(\mathbf{x}_{(i)})} \right)^2 - \frac{1}{N^{\mathrm{samps},n}} \left(\sum_{i=1}^{N^{\mathrm{samps},n}} \frac{\exp[-S(\mathbf{x}_{(i)},\mathbf{a}_{(i)})/\hbar]}{f(\mathbf{x}_{(i)})} \right)^2 \right] \quad (26)$$

and the overall variance is

$$w = \sum_{n=1}^{N^{\mathrm{strata}}} w_n^{\,2} \,. \quad (27)$$

The 2σ uncertainty in the partition function estimate is simply $2\sqrt{w}$.

In previous work,[9,11,25,44] the importance-sampling scheme that was employed involved an importance function expressed in internal coordinates; in particular, we used a product of Gaussians in the atom-atom distances. It is complicated to directly sample from such a distribution due to the nontrivial Jacobian between the internal and the Jacobi coordinates. In practice, one is forced to sample via rejection. For triatomic systems, this is affordable, but for 4-body and larger systems one must generate a very large number of candidate configurations to get one that will be accepted, and this makes the cost of selecting CSSPs dominate the overall cost; thus, the efficiency of the approach is much lower than the approach described next.

The importance function used for sampling the configuration space in the present work has the form

$$f(x_1, x_2, \ldots, x_D) = f_R(R_1, R_2, \ldots R_{D/3}) f_A(\theta_1, \theta_2, \ldots), \quad (28)$$

where f_R is of the form

$$f_R(R_1, R_2, \ldots R_{D/3}) = \prod_{i=1}^{D/3} f_{R,i} = \prod_{i=1}^{D/3} \exp(-(R_i - R_i^0)^2 / 2\Delta_{R_i}^2) . \quad (29)$$

and f_A is of the form

$$f_A(\theta_1, \theta_2, \ldots) = \prod_{i=1}^{N_\theta} f_{A,i}(\theta_i) = \prod_{i=1}^{N_\theta} \exp(-(\theta_i - \theta_i^0)^2 / 2\Delta_{\theta_i}^2) , \quad (30)$$

where N_θ and the angles θ_i are defined below. A key aspect is that these angles are defined in terms of the Jacobi direction vectors and are independent of the Jacobi magnitudes.

Certain details of our new importance sampling strategy will depend on the definition of the strata, and in the discussion that follows, we will assume that the strata boundaries are independent of the Jacobi magnitude degrees of freedom. We experimented with a number of different stratification coordinates, and we eventually chose the angle between the plane defined by the B→A and CD→BA vectors and the plane defined by the C→D and CD→BA vectors, where CD and BA denote the centers of mass of CD and BA, respectively; we shall refer to this as the Jacobi dihedral angle, φ. It can easily be calculated from the Jacobi coordinates.

Our prior importance sampling approach using general internal coordinates required selecting points uniformly from within the domain and then using rejection to reshape the distribution. For $D \geq 9$ this can require generating thousands of uniformly distributed CSSPs before obtaining one that is not rejected. In the present approach we write our importance sampling function in a way that allows us to sample certain subsets of the coordinates independently from others, and this greatly reduces the amount of computational effort that is lost during rejection sampling.

For a D-dimensional system, the $D/3$ Jacobi magnitudes can be selected independently of each other if we restrict ourselves to an importance function of the form of Eqs. (28)–(30). We select points from distributions of the form $R_i^2 \exp(-(R_i - R_i^0)^2 / 2\Delta_{R_i}^2)$ via a ziggurat scheme,[45,46] where the R^2 prefactor is part of the Jacobian of a 3-sphere (this Jacobian is hidden in $d\mathbf{x}$ in eq. (19)). The ziggurat sampling method[45,46] is a robust and efficient method for rejection sampling from a distribution; the method was introduced for sampling monotonic or symmetric unimodal distributions, but for the present work it is generalized to sample asymmetric distributions with a single maximum. The ziggurat scheme involves covering the desired distribution with a set (we use 256) of disjoint, equal volume, rectangles and then rejection sampling over that space; a detailed explanation of the method is given in an appendix. For most samples the computational effort consists mainly of the generation of a single uniform random number and a couple of comparisons. In a small percentage of cases, which for typical parameters is only ~2–3% of the time, one also needs to calculate a value of the sampling function and possibly reject the sample. The overall rejection rate of this approach is typically ~1–2%, and thus the cost of generating samples is comparable to the cost of generating uniform random numbers. The ziggurat approach requires the generation of a set of preliminary tables giving details of the set of covering rectangles, etc.; the cost to generate this information for each distribution is negligible since it needs to be done only once prior to the start of the Monte Carlo calculation.

Given a set of appropriate Jacobi magnitudes, we can complete the selection of the CSSP by randomly selecting $D/3$ Jacobi direction vectors. This is equivalent to uniformly selecting $D/3$ points on the surface of a 3-sphere, and we do this either by the

angular importance sampling scheme discussed below or, if we don't importance sample in the angular coordinates, by the scheme of Marsaglia.[47]

The Jacobi magnitude sampling function provides a fast and effective way of limiting CSSPs to physically reasonable choices. We can, however, increase our efficiency even further by importance sampling within the remaining $2D/3$ Jacobi degrees of freedom. Each individual Jacobi direction vector has no preferred orientation, but the angle between any two of the Jacobi directions can serve as an appropriate variable for further importance sampling. The appropriate choice and number of angles to use will strongly depend on the system being studied. We will limit further consideration to the special case of H_2O_2, and in particular to a scheme where we importance sample the angles between a particular H atom, the center of mass of the OH fragment that includes this H atom, and the center of mass of the other OH fragment (this is similar to, but not identical to, importance sampling in the HOO angles). Thus, we choose $N_\theta = 2$ and we choose these two angles from a $\sin(\theta_i)\exp(-(\theta_i - \theta_i^0)^2 / 2\Delta_{\theta_i}^2)$ distribution (where the sine term is the Jacobian for an angle between two vectors) via another ziggurat scheme. Next, we choose a stratum that still needs a sample (as determined by our stratified sampling scheme) and uniformly select a Jacobi dihedral angle, φ, from within the chosen stratum. We then use these six (R_1, R_2, R_3, θ_1, θ_2, φ) independently selected coordinates to define the configuration of the centroid position, except for an arbitrary overall rotation. We arbitrarily assign a direction for the CD→BA Jacobi vector (we assume that the PES is invariant to overall rotations and translations) and choose the other two Jacobi vectors to be consistent with θ_1, θ_2, and φ, where the angle between

B→A and CD→BA must be $\pi - \theta_1$, and the angle between C→D and CD→BA must be θ_2. A valid choice for the Jacobi coordinates of the centroid position is then simply given by [$R_1 \cos(\pi - \theta_1)$, $R_1 \sin(\pi - \theta_1)$, 0, $R_2 \cos(\theta_2)$, $R_2 \sin(\theta_2)\cos(\varphi)$, $R_2 \sin(\theta_2)\sin(\varphi)$, R_3, 0, 0]. The overall cost of this approach does not differ greatly from the cost of purely uniform sampling (i.e., the cost is dominated by the cost of generating six uniform random numbers and the six trigonometric functions needed above to completely specify the centroid position), but as we will see below, it is substantially more effective.

Independently of how we do the importance sampling, we need to calculate the average value of the importance function over each of the strata since this appears in the numerator of eq. (23). (The need for these average values arises from the requirement that the distribution we sample be normalized.) Previously, we calculated[9,11,25,44] these normalization factors via a preliminary Monte Carlo sampling calculation that did not take advantage of any special variance reduction strategies. As the dimensionality increases, even these simple normalization factors can become expensive quantities to calculate. An additional benefit of importance sampling in functions of the separable form of Eq. (28) together with strata boundaries that are independent of the Jacobi magnitudes, is that we can use (for each stratum *j*):

$$\langle f \rangle_j = \langle f_A \rangle_j \prod_{i=1}^{D/3} \langle f_{R,i} \rangle_j \tag{31}$$

to simplify the formation of the average values of the importance function on the various strata. Reducing the dimensionality, via Eq. (31), over which the average values must be Monte Carlo sampled greatly reduces the cost of obtaining the required normalization

factors. The average of each $f_{R,i}$ can be calculated via a one-dimensional quadrature, but the average values of the angular parts of the importance function will usually need to be calculated via Monte Carlo. However, for the particular choice of angles and stratification boundaries chosen here, we can also obtain the angular averages via one-dimensional quadratures. In particular, the average values of the angular part of the importance function are then strata independent and can be calculated by

$$\left\langle f_A(\theta_1,\theta_2)\right\rangle_j = \left\langle f_{A,1}(\theta_1)\right\rangle_j \left\langle f_{A,2}(\theta_2)\right\rangle_j \tag{32}$$

where

$$\left\langle f_{A,i}(\theta_i)\right\rangle_j = \frac{1}{2}\int_0^\pi \sin\theta_i \exp\left[-\left(\theta_i-\theta_i^0\right)^2 \Big/ \left(2\bar{\Delta}_i^2\right)\right] \mathrm{d}\theta_i \,. \tag{33}$$

V. COMPUTATIONAL DETAILS

The PES used for the present set of calculations is the second of the two surfaces presented by Koput, Carter, and Handy (i.e., the fit presented in Table 3 of their paper,[24] which is the same surface that was used for their eigenvalue calculations). A FORTRAN version of this potential has been placed in the on-line POTLIB library.[48,49] The functional form used for this fit does not account for the full permutational symmetry of the molecule (it is symmetric with respect to exchange of the two OH groups but not with respect to exchanges of just the two H atoms nor with respect to exchanges of just the two O atoms); thus, the correct symmetry number to use is $\sigma^{\mathrm{sym}} = 2$.

It is also important to note that the functional form used for the PES does not yield an accurate global fit. Therefore, outside the domain of validity (such as in asymptotic regions corresponding to any bond dissociation) it can give energies that are

highly inaccurate. These regions of the surface do not contribute significantly to the partition function for the temperatures we consider here unless the fit returns spuriously low energies. To partially alleviate such complications (which should be rare for well chosen domain boundaries), we zero out the contributions of any path for which a configuration has an HH distance below 0.7 a_0, an OH distance below 1.0 a_0 or above 4.0 a_0, or an OO distance below 1.8 a_0 or above 4.0 a_0 (converged calculations were not especially sensitive to the presence or absence of these parameters.)

The masses of H and O are assumed to be 1.00782503 and 15.9949146 amu, respectively. The zero of energy is at the minimum of the potential, which occurs for the configuration with R_{OH} = 0.96265 a_0, R_{OO} = 1.45248 a_0, HOO angles of 99.906º, and a dihedral angle of 112.456º.

Table I gives the parameters used in calculating the FPI values for $Q_{\mathrm{TT-FPI}}^{[P]}(T)$. The chosen domain-boundary parameters are conservative values since tight optimization of these parameters yields negligible savings in computer time. The importance function width parameters were partially optimized in small initial trials. The importance function center parameters were chosen to be those of the global minimum of the PES; further optimization of these parameters typically yields very little increased efficiency. At each temperature an appropriate value of $P^{\max}$ was chosen that is divisible by 6, and a simultaneous set of results was obtained for $P = P^{\max}/3$, $P^{\max}/2$, and $P^{\max}$ points using the enhanced same path[42] (ESP) approach.

The classical (P = 1) results were obtained in separate calculations since for this special case the variance of the results obtained during an ESP calculation are much larger than those of other P values. This is predominately because the optimal width parameters

for the distribution of the centroid positions are quite narrow, and the individual points on a path with $P = P^{max}$ points—which are used in the ESP approach to construct the $P = 1$ result, can (at least for sufficiently low temperatures) be well removed from the optimal centroid position.

In past applications we have sometimes[9,11,25] sampled the relative paths (i.e., the **a** space) less frequently than we sampled the configuration space. For some applications and some path integral methods the path generation phase can dominate the overall computational effort, and reusing relative paths can provide the highest accuracy for a given computational cost, even though one then no longer has uncorrelated samples (which means that the confidence limits are no longer related in the usual way to the standard deviations); this is particularly true when unequally spaced quadrature nodes are required since then the paths must be generated using a matrix multiplication algorithm[25] which has a much steeper computational scaling than the fast Fourier sine transform scheme[25] that can be used to generate paths with equally spaced nodes (as in the case here). In the present application, we have chosen to use each path only once; thus, our sampling is uncorrelated and our uncertainty estimates are rigorous.

For the calculation at $T = 300$ K, approximately 70% of the computational cost was for evaluation of the PES, 29% for the generation of the Fourier paths, and 1% for the generation of the configuration space samples.

VI. APPROXIMATE METHODS

VI.A. Separable approximations

To obtain a simple expression for the vibrational partition function, the harmonic oscillator (HO) approximation can be made. The classical mechanical (CM) expression for the harmonic oscillator approximation to the vibrational partition function is[2]

$$Q_{\text{vib,CM}}^{\text{CHO}} = \prod_{m=1}^{F} \frac{1}{\hbar\beta\omega_m} \tag{34}$$

where ω_m is the vibrational frequency of mode m, and F is the number of vibrational modes (F = 6 for H_2O_2). The quantum mechanical expression for the harmonic oscillator approximation to the vibrational partition function is

$$Q_{\text{vib}}^{\text{HO}} = \prod_{m=1}^{F} \frac{e^{-\hbar\beta\omega_m/2}}{1-e^{-\hbar\beta\omega_m}}. \tag{35}$$

In the strictly harmonic approach, denoted HO, the vibrational frequencies are obtained by normal mode analysis based on the Hessian of the potential energy at the equilibrium geometry. Variations of Eq. (35) allow for improved accuracy. Three other approximations are obtained by rewriting Eq. (35) as

$$\begin{aligned} Q_{\text{vib}}^{\text{HO}} &= \tilde{Q}_{\text{vib}}^{\text{HO}} e^{-\beta E^{\text{G}}} \\ &= e^{-\beta E^{\text{G}}} \prod_{m=1}^{F} \frac{1}{1-e^{-\hbar\beta\omega_m}} \end{aligned} \tag{36}$$

where E^{G} is the ground state energy or zero point energy (ZPE). If the ZPE is calculated by

$$E^{\text{G}} = \frac{\hbar}{2} \sum_{m=1}^{F} \omega_m \tag{37}$$

we again obtain $Q_{\text{vib}}^{\text{HO}}$. If, however, we use the accurate ZPE value for the given PES, we denote the result as $Q_{\text{vib}}^{\text{HO}-\text{Z}}$ where HO-Z denotes the zero point corrected harmonic oscillator approximation. Instead of using the harmonic frequencies for ω_m in Eq. (36), we could substitute the accurate fundamentals for the given PES. It is interesting to test this approach because in some cases the fundamentals are the only data known; the resulting method is called $Q_{\text{vib}}^{\text{HO}-\text{ZF}}$ where the superscript HO-ZF represents the quasi-harmonic oscillator based on an accurate ZPE and accurate fundamentals for ω_m. The final approximation to be considered for the vibrational partition function involves using the accurate fundamentals (for the given PES) both for the frequencies in Eq. (36) and for the ZPE. Equivalently, one can just use these fundamentals in Eq. (35). The result is called $Q_{\text{vib}}^{\text{HO}-\text{F}}$ where the superscript HO-F denotes quasi-harmonic approximation based on fundamentals.

The frequencies and zero point energies[24,50-52] needed for all of these approximations are given in Tables II and III; the experimental values,[53-57] are also given for comparison. Table IV lists the sources of the ZPE and frequency estimates that are used in calculating the various approximate vibrational partition functions. Table II lists the values for the experimental[53-57] and calculated[24] fundamentals for H_2O_2 and the harmonic normal mode frequencies. Note that the frequencies in the tables are in spectroscopic units, and are given by

$$\nu = \omega / 2\pi c \,. \tag{38}$$

The calculated fundamentals were obtained by Koput et al. using the same H_2O_2 PES[24] as used here. The harmonic normal mode frequencies were calculated using POLYRATE.[50] Table III lists the ZPE estimated by a variety of approaches.

Table II shows a comparison of the experimentally observed fundamentals to the fundamentals calculated[24] from the PES. We can see that the symmetric (ν_1) and antisymmetric (ν_5) OH stretches differ from experiment by about 14 cm^{-1}. The symmetric (ν_2) and antisymmetric (ν_6) HOO bend differ by about 1 cm^{-1} and 3 cm^{-1}, respectively, and the symmetric OO stretch (ν_3) differs by about 8 cm^{-1}. Finally, the HOOH torsional vibrational mode (ν_4) differs by only 1 cm^{-1} from experiment. Table III shows a comparison of the zero point energy calculated from the experimentally observed fundamentals and one calculated with the fundamentals calculated from the PES. These two ZPE estimates differ by only 21 cm^{-1}, but both differ significantly from the accurate ZPE determined by variational calculations;[58] the differences are 146 and 167 cm^{-1}.

For an asymmetric top ($I_A \neq I_B \neq I_C$), where I_X is the moment of inertia about principal axis X, the classical mechanical rigid-rotator approximation to the rotational partition function is[2]

$$Q_{\mathrm{rot,CM}}^{\mathrm{asym,e}} = \frac{\pi^{1/2}}{\sigma^{\mathrm{sym}}}\left(\frac{8\pi^2 I_A^e kT}{h^2}\right)^{1/2}\left(\frac{8\pi^2 I_B^e kT}{h^2}\right)^{1/2}\left(\frac{8\pi^2 I_C^e kT}{h^2}\right)^{1/2} \qquad (39)$$

where I_A^e, I_B^e, and I_C^e are the three principal moments of inertia evaluated at the equilibrium geometry, and *e* denotes that the geometry used to evaluate the moments of inertia is the equilibrium geometry. Since two of the moments of inertia (I_B^e and I_C^e) for H_2O_2 are of similar magnitude, we can make a symmetric top approximation. The

symmetric top approximation takes an average, $I_{\mathrm{M}}^{\mathrm{e}}$, of the two moments of inertia, $I_{\mathrm{B}}^{\mathrm{e}}$ and $I_{\mathrm{C}}^{\mathrm{e}}$, and then the classical rotational partition function expression for a symmetric top can be used

$$Q_{\mathrm{rot,CM}}^{\mathrm{sym,e}} = \frac{\pi^{1/2}}{\sigma^{\mathrm{sym}}}\left(\frac{8\pi^2 I_{\mathrm{M}}^{\mathrm{e}} kT}{h^2}\right)\left(\frac{8\pi^2 I_{\mathrm{A}}^{\mathrm{e}} kT}{h^2}\right)^{1/2} . \tag{40}$$

In order to incorporate the effect of averaging the geometry over the vibrational wave functions, we can evaluate effective moments of inertia, I_{M}^{0} and I_{A}^{0}, from the accurate relative energies of the $J = K = 0$, $J = 1$, $K = 0$, and $J = 1$, $K = 1$ rotational levels of the ground vibrational level, where J is the rotational quantum number, and K is the projection quantum number for a body-fixed axis. The calculated spectroscopic constants can be used to determine the value of the effective moments of inertia. (These accurate energies and spectroscopic constants are available in Koput et al.[24]) To determine the effective moments of inertia using the calculated spectroscopic constants we use the following expressions[2]

$$I_{\mathrm{A}}^{0} = \frac{h}{8\pi^2 cA}, \; I_{\mathrm{M}}^{0} = \frac{1}{2}\left(\frac{h}{8\pi^2 cB} + \frac{h}{8\pi^2 cC}\right) \tag{41}$$

where A, B, and C are the rotational spectroscopic constants, c is the speed of light, and h is Plank's constant. Using these moments of inertia yields

$$Q_{\mathrm{rot,CM}}^{\mathrm{sym,0}} = \frac{\pi^{1/2}}{\sigma^{\mathrm{sym}}}\left(\frac{8\pi^2 I_{\mathrm{M}}^{0} kT}{h^2}\right)\left(\frac{8\pi^2 I_{\mathrm{A}}^{0} kT}{h^2}\right)^{1/2} \tag{42}$$

where 0 denotes the ground vibrational level.

Alternatively, the expression for the quantum mechanical rotational partition function for a symmetric top, which is well known,[2] can be used. This yields

$$Q_{\text{rot}}^{\text{sym}} = \frac{1}{\sigma^{\text{sym}}} \sum_{J=0}^{\infty} \sum_{K=-J}^{J} (2J+1) \exp\left(-\alpha_{\text{M}} J(J+1) - (\alpha_{\text{A}} - \alpha_{\text{M}}) K^2\right) \tag{43}$$

where

$$\alpha_{\text{X}} = \frac{\hbar^2}{2 I_X kT}; \text{X} = \text{A or M}. \tag{44}$$

As for the classical expression for the symmetric top rotational partition function, the principal moments of inertia in the quantum mechanical partition function can be evaluated at the equilibrium geometry or from the lowest rotational levels of the ground vibrational level, denoted e and 0, respectively. Note that the approximations involving I_{X}^{e} may also be called rigid rotator (RR) approximations, and those involving I_{X}^{0} may be called vibrating rotator (VR) approximations.

The values of the moments of inertia needed for all these formulas are given in Table V.

The most basic separable approximation is to combine the HO approximation for vibration with the RR approximation. This yields the classical harmonic-oscillator rigid-rotor (CHO-RR) result:

$$Q^{\text{CHO}-\text{RR}} = Q_{\text{vib}}^{\text{CHO}} Q_{\text{rot,CM}}^{\text{asym,e}} \tag{45}$$

and the quantum result:

$$Q^{\text{HO}-\text{RR}} = Q_{\text{vib}}^{\text{HO}} Q_{\text{rot}}^{\text{sym,e}}. \tag{46}$$

We can also try to identify the best separable approximation to the vibrational-rotational partition function. To attempt to identify this, we will consider the product of Eq. (36)

and Eq. (43), where the principal moments of inertia are evaluated in the VR approximation. These products are denoted HO-Z-VR, HO-ZF-VR, and HO-F-VR.

VI.B. Non-separable approximations

The Pitzer–Gwinn[23,59-62] approximation can be used to approximate the value of the anharmonic quantum mechanical vibrational partition function

$$Q^{\mathrm{PG}} = e^{-\beta E^{G}} Q_{\mathrm{CM}}^{\mathrm{anhar}} \left(\frac{\widetilde{Q}^{\mathrm{HO}}}{Q^{\mathrm{CHO}}} \right). \tag{47}$$

We can apply this by calculating the anharmonic classical vibrational-rotational classical partition function, $Q_{\mathrm{CM}}^{\mathrm{anhar}}$, by setting $P = 1$ in the path integral algorithm. However, the PG approximation also depends on the choice of the ZPE estimate. We will test two approximations, one employing the *accurate* ZPE, to be called PG-AZPE, and one employing the *harmonic* ZPE, to be called PG-HZPE.

VII. RESULTS AND DISCUSSION

VII.A. FPI Results

Table VI contains for several temperatures the value of $P^{\max}$ used, the classical limit of the path integral calculation, $Q_{\mathrm{TT-FPI}}^{[P=1]}(T)$, the values of $Q_{\mathrm{TT-FPI}}^{[P]}(T)$ for $P = P^{\max}/2$, $P^{\max}/3$, and $P^{\max}$, the extrapolated value, $Q^{\mathrm{ESPE}}(T)$, and the 2σ statistical uncertainties for each of the partition function values, where σ denotes standard deviation. The small magnitudes of the 2σ statistical uncertainties given in Table VI result partly from the highly effective new importance-sampling scheme and partly because the calculations used

a large number (1×10^7–2×10^8) of Monte Carlo samples. The uncertainties for the ESPE calculations were estimated by assuming these partition functions had the same relative uncertainty (i.e., the ratio of the absolute uncertainty and the partition function) as in the calculations with $P = P^{\max}$. Additional uncertainties due to extrapolation error were assumed to be negligible; the validity of this assumption is considered further below.

It is interesting to note that in the high-temperature limit the relative uncertainty of the TT-FPI calculations is very similar to the relative uncertainty of the classical results, but that as the temperature decreases the relative uncertainties of the quantum calculations grow much more rapidly than that of the classical results. The ratios of the quantal to classical uncertainties are 1.1, 3.3, 13.7, and 38.8 at $T = 2400$, 600, 400, and 300 K respectively. Thus, at 300 K the quantal calculation requires ~1500 times as many samples (and each of these requires 90 times as many PES evaluations) to achieve the same relative accuracy as a classical calculation. This indicates that at moderate to high temperatures the present scheme of importance sampling the path centroids is sufficient to ensure efficient sampling of the full paths, but that at low temperatures—where the paths meander more widely over configuration space—more sophisticated importance-sampling schemes are desirable.

We monitored the effect of dissociated species on the calculated partition function by separately tabulating contributions from paths whose initial position was above the sum of the energy of the lowest energy dissociation channel[63] and the accurate ZPE[52] of the molecule (~0.1038 E_h). At the highest temperature considered here (2400 K) these amounted to less than 0.014% of the total partition function, which is negligible for the present purposes.

The maximum correction obtained via extrapolation was about 3.6% which, given the very high accuracy that the extrapolations typically achieve,[9,11,42] is sufficiently small that errors introduced from the extrapolation are expected to be negligible (i.e., small compared to the statistical uncertainties). Since in many contexts extrapolation is notoriously inaccurate, and our typical statistical uncertainties of ~0.1% are miniscule compared to the maximum extrapolation correction, it is worthwhile to provide a demonstration to support the plausibility of this expectation. The largest extrapolation correction (~3.6%) in Table VI occurs at $T = 600$ K where we used a $P^{\max}$ of 30; we repeated this calculation resetting $P^{\max}$ to 60 and using a different seed for the random number generator. The partition function estimates obtained from this new calculation were $0.014\ 648 \pm 0.000\ 012$, $0.014\ 003 \pm 0.000\ 011$, and $0.013\ 620 \pm 0.000\ 011$ at $P = 20$, 30, and 60 respectively. The extrapolated result is $0.013\ 493$, with an extrapolation correction of only ~0.93% (a reduction of about a factor of 4 from the $P^{\max} = 30$ calculation); we note that this value agrees with the extrapolated value presented in Table VI to within ~0.06%, which is within the limits of the statistical uncertainties and thus suggests that the extrapolation error is comparable to or smaller than this. If, instead of extrapolation, we were to rely only on the use of a single large value of P, we would require P to be at least 6 to 7 times larger than the value ($P^{\max} = 30$) used for the calculation in Table VI to reduce the truncation error to below that of the statistical uncertainty.

The present calculations are intended as benchmarks—and thus warrant high precision—but for many applications a final uncertainty of several percent (or even larger) is perfectly acceptable, so one might be tempted to use significantly lower values of $P^{\max}$ and rely on extrapolation for much larger corrections than those represented here. This

might well be acceptable in some situations, but a word of caution is warranted. The extrapolation correction of the ESPE approach[42] is extremely accurate in the regime where the partition function has approached the asymptotic convergence rate indicated in Eq. (12). For sufficiently low values of P, this expression will *not* hold and dramatically poorer results will be achieved. If, for instance, one chooses a $P^{\max}$ of 12 for the calculation at T = 600 K, the partition functions at P = 4, 6, and 12 are $0.052\,613 \pm 0.000\,040$, $0.028\,142 \pm 0.000\,023$, and $0.016\,753 \pm 0.000\,015$ respectively. Extrapolating these values yields a result of $0.013\,955$**,** which corresponds to an extrapolation correction of about 16.7% but results in an error of about 3.3% compared to the statistical uncertainty of only 0.09%. The partition function for P = 4 is larger than the converged result by a factor of about 3.9, and thus P = 4 is not sufficiently near the asymptotic limit to permit accurate extrapolation.

Table S-I (in supplementary material[64]) presents an extended table of extrapolated partition functions; only those values that are also present in Table VI will be considered further in the following sections.

Table VII compares the relative efficiencies of various subsets of importance sampling and stratification options for a calculation at T = 1000 K using the parameters presented in Tables I and VI. The overall increase in efficiency of a factor of ~1600 compared to purely uniform sampling is considerable. Importance sampling in the three Jacobi magnitudes provides the largest single improvement, but importance sampling in two Jacobi angles also provides a very significant enhancement. Stratification in the single Jacobi dihedral angle is observed to provide only a small benefit (factors of 1.4–1.7 depending on the type of importance sampling used), but for the present application this is

the floppiest coordinate, and thus this is not entirely unexpected. Interestingly, Table VII shows that importance sampling simultaneously in both the angular and radial coordinates improves the efficiency by more than the product of the efficiency gains for employing either of these techniques separately.

VII. B. Comparison to results from the separable approximation

We have compared the FPI calculations with the separable mode and Pitzer–Gwinn approximations which use the accurate fundamentals and harmonic frequencies of the Koput, Carter, and Handy surface.[24]

Table VIII lists the classical and quantum mechanical values for the vibrational partition function obtained using Eq. (34) and (36) at various temperatures. It shows that there is a large improvement at low temperatures by replacing $Q_{\mathrm{vib}}^{\mathrm{HO}}$ with $Q_{\mathrm{vib}}^{\mathrm{HO\text{-}Z}}$. The use of $Q_{\mathrm{vib}}^{\mathrm{HO-ZF}}$ might naively be expected to yield the most accurate harmonic value of the vibrational partition function since it uses both the accurate ZPE and the accurate fundamentals. This approximation predicts the lowest few eigenvalues quite accurately (and thus yields excellent results at very low temperatures), but it tends to systematically overestimate the energies of the higher eigenvalues, and thus it systematically underestimates the true partition function.

Estimating the ZPE using the fundamental frequencies tends to overestimate the effects of anharmonicity, resulting in predicted values that are too low. For H_2O_2, this estimate is too low by 146 cm^{-1}, whereas the harmonic approximation is too high by 112 cm^{-1}. This error causes the $Q_{\mathrm{vib}}^{\mathrm{HO-F}}$ values to be higher than those of the HO-ZF

approximation, and a fortuitous cancellation of errors often results in the HO-F values being more accurate (as is the case here). This same kind of cancellation has been noted previously for other systems.[60]

Table IX shows the classical and quantum mechanical values for the rotational partition function according to Eqs. (37)–(41). (A longer version of this table is in supplementary material.[64]) We can see that even though the B and C rotational constants differ by about 2.6%, the symmetric top approximation agrees extremely well with the asymmetric result since I_{M} is larger than $\sqrt{I_{\mathrm{B}} I_{\mathrm{C}}}$ by only about 0.008%. The classical and quantum mechanical symmetric partition functions agree within about 0.14%, and the difference between the vibrating rotator and rigid rotator approximations is about 1.9%. Thus, $Q(T)$ will be relatively insensitive to the choice of methods used to treat the rotational part of the partition function.

Table X gives values for separable approximations to the vibrational-rotational partition function and compares them to the TT-FPI-ESPE result and the accurate classical mechanical result. (A longer version of this table is in supplementary material.[64]) At 300 K, we find that the classical harmonic (CHO-RR) result differs from the accurate result by six orders of magnitude, while at 2400 K, the error between the CHO-RR result and the accurate result is 35%. A similar trend is observed when comparing the accurate classical results to the accurate QM results, with the classical result being nearly seven orders of magnitude too large at 300 K and 55% too large at 2400 K. The quantal HO-RR result show a dramatic improvement compared to the CHO-RR results. The deviation of the HO-RR partition function from the FPI result ranges from 73% at 300 K to 60% at 2400 K. Thus the HO-RR results still leave much room for improvement. We first improve upon the

HO-RR result by pulling out the harmonic ZPE estimate in the vibrational partition function according to Eq. (36) and replacing it with the accurate value. We also replace the rigid-rotator rotational partition function with that of the vibrating rotator approximation. A large improvement is observed in proceeding from the HO-RR method to the HO-Z-VR method due to these small changes. The errors in the partition function for the HO-Z-VR result (compared to the accurate FPI result) range from 53% at 300 K to 57% at 2400 K. The HO-Z-VR results show a small improvement compared to the HO-ZF-VR results due to the use of the calculated fundamentals instead of the harmonic frequencies in the vibrational partition function. The errors in the partition function for the HO-ZF-VR results range from 52% at 300 K to 51% at 2400 K. The best overall results are obtained from the HO-F-VR method; however, this is partly due to a fortuitous cancellation of errors as explained above. The errors in the partition function for the HO-F-VR result range from 4% at 300 K to 46% at 2400 K.

VII. C. Comparison to the Pitzer–Gwinn Approximation

Table XI shows values for the Pitzer–Gwinn approximation. The error in the PG-HZPE value of the partition function as compared to the FPI result is 38% at 300 K and 5% at 2400 K. The error in the PG-AZPE value of the partition function as compared to the FPI result is only 7% at 300 K and 1.7% at 2400 K.

VII. D. Free energies

Tables VI, X, and XI give the path integral results for the internal partition function $Q_{\rm int}$. From these we can calculate the standard-state Gibbs free energy at temperature *T*, G_T°, by

$$G_T^\circ = -RT\ln\frac{Q_{\rm trans}^\circ(T)Q_{\rm int}(T)}{N_A} \tag{48}$$

where *R* is the universal gas constant, $Q_{\rm trans}^\circ$ is the translational partition function evaluated under standard-state conditions (taken here as an ideal gas at a pressure of 0.1 MPa), and N_A is Avogadro's number. However, we have taken the zero of energy at the minimum of the potential surface, whereas the values in the thermodynamic tables, $\widetilde{G}_T^\circ$, are consistent with putting the zero of energy at the ground state. We denote this shift in the zero of energy by a tilde. Rewriting Eq. (48) with this shift gives

$$\widetilde{G}_T^\circ = -RT\ln\frac{Q_{\rm trans}^\circ(T)\widetilde{Q}_{\rm int}(T)}{N_A} \tag{49}$$

where

$$\widetilde{Q}_{\rm int} = Q_{\rm int}(T)e^{\beta E^{\rm G}} . \tag{50}$$

We use the accurate value (for this PES) of 5726.1 $\rm cm^{-1}$ for $E^{\rm G}$. The last column of Table XII gives the resulting path integral values of $\widetilde{G}_T^\circ$. Table XII contains the standard-state Gibbs free energy values for the corresponding partition functions in Table X and compares them to the accurate calculations and the accurate classical mechanical calculation. The errors in free energy calculated by exact classical mechanics, as compared to the FPI results,

which represent exact quantum mechanics for the given PES, are 39 kJ/mol at 300 K and 13.9 kJ/mol at 1500 K.

Table XI shows the standard-state Gibbs free energy values for various separable methods. We see that the errors in the CHO-RR free energy value range from 37 kJ/mol at 300 K to 8.6 kJ/mol at 2400 K. We find that the errors in the standard-state free energy calculated by the quantum mechanical HO-RR approximation, for the chosen PES, range from 3.3 kJ/mol at 300 K to 18.4 kJ/mol at 2400 K. The errors decrease greatly at lower temperatures where the separable approximation is more appropriate. The errors in the standard-state free energy values calculated by the HO-ZF-VR approximation range from 1.9 kJ/mol at 300 K to 16.6 kJ/mol at 2400 K. The errors in standard-state free energy values calculated by the HO-F-VR approximation range from 0.1 kJ/mol at 300 K to 12.3 kJ/mol at 2400 K.

Table XIII gives the standard-state Gibbs free energy values for the Pitzer–Gwinn approximations and compares them to the FPI results and to the data available from thermodynamic tables.[26,27] The Pitzer–Gwinn approximations provides a dramatic improvement from the exact classical anharmonic values. The errors in the PG-HZPE result range from 1.2 kJ/mol at 300 K to 1.0 kJ/mol at 2400 K. The errors in the PG-AZPE result range from 0.16 kJ/mol at 300 K to 0.39 kJ/mol at 2400 K.

VII. E. Comparison to experiment

Although the first goal of this project was to obtain accurate partition functions and standard-state free energies for a known realistic PES, the PES used in the calculations is good enough that we can also compare the results to experiment.

The quantities used in the thermodynamic tables[26,27] to which we will make comparison are, in the notation used here, $\frac{-[\widetilde{G}^\circ_{\mathrm{T}} - H^\circ_{298}]}{T}$ and $H^\circ_0 - H^\circ_{298}$, where H°_T is the standard-state enthalpy at temperature *T*. Using these values, one can calculate the Gibbs free energy function[65,66] defined as

$$\mathrm{gef}(T) = -\frac{[\widetilde{G}^\circ_{\mathrm{T}} - H^\circ_0]}{T}. \tag{51}$$

At 0 K, the standard state enthalpy, H°_0, is zero. At any temperature, the standard state Gibbs free energy, $\widetilde{G}^\circ_{\mathrm{T}}$, with the zero of energy at the ground state is obtained by multiplying the Gibbs free energy function by the temperature.

Table XII shows that the present calculations of the standard-state free energy agree with prior tabulations[26,27] based partly on experimental data within 1.2 kcal/mol for 300–2400 K, whereas the harmonic oscillator-rigid rotator approximation has errors of 3–12 kcal/mol. Our calculations show systematically better agreement with the tabulations of Dorofeeva et al.[27] than the earlier JANAF[26] values.

VIII. CONCLUDING REMARKS

This paper presents the first example of converged accurate quantal rovibrational partition functions for a polyatomic molecule with a torsion. The calculations are carried out by the TT-FPI-ESPE path integral method, and the statistical errors in the free energies are about 0.1% from 400 K to 2400 K; the calculations are most difficult at the lowest temperature (300 K) where the statistical error is 0.4%. Our results for the standard-state

free energies show excellent agreement with the recent data of Dorofeeva et al.,[27] much better than with the earlier[26] JANAF tabulations.

The availability of well-converged results allows us, for the first time, to test the effect of anharmonicity and vibrational-rotational coupling for a molecule with a torsion. We find that the errors in the standard-state free energy calculated by the quantum mechanical harmonic oscillator-rigid rotator approximation range from 3.3 kJ/mol at 300 K to 18.4 kJ/mol at 2400 K, and for the "best" separable rotation approximation (i.e., the HO-ZF-VR approximation) the errors range from 1.9 kJ/mol at 300 K to 14.0 kJ/mol at 2400 K. The classical harmonic oscillator-rigid rotator approximation to the standard-state free energy yields errors ranging from 37 kJ/mol at 300 K to 8.6 kJ/mol at 2400 K. The standard-state free energy calculated from the anharmonic quantum mechanical partition function obtained using the Pitzer–Gwinn approximation gives a slight overestimation for the PG-AZPE method and a slight underestimation for the PG-HZPE method. The errors in the PG-HZPE result range from 1.2 kJ/mol at 300 K to 1.0 kJ/mol at 2400 K. The errors in the PG-AZPE result range from 0.2 kJ/mol at 300 K to 0.4 kJ/mol at 2400 K.

We presented a very efficient new scheme for uncorrelated importance sampling in the configuration space when using Jacobi coordinates. This approach permits Monte Carlo samples of the configuration space to be chosen with minimal losses from rejection. The cost of choosing such samples is only slightly greater than the cost of uniform sampling (and requires only a very small percentage of the overall computational cost), but the efficiency can be about three orders of magnitude greater than uniform sampling. We expect that this approach will permit calculations of significantly larger systems than the one presented here.

ACKNOWLEDGMENTS

We thank Jacek Koput for providing us with a copy of the potential parameters and unpublished data on the zero-point energy. This work was supported in part by the National Science Foundation under grant no. CHE00-92019.

APPENDIX

In this article we have made frequent use of a slight extension of the ziggurat sampling scheme of Marsaglia and Tsang.[45,46] For completeness we will give here a brief statement of the method. We restrict attention to a one-dimensional, non-negative function $y = f(x)$ having a single finite maximum and a finite domain $[0, x_{\max}]$; the method can also be applied to functions with a finite area and an infinite domain as discussed elsewhere.[45,46] The original ziggurat method[45,46] was restricted to monotonic functions but it is easily and efficiently extended to functions with a single relative maximum.

The method works by covering the curve $y = f(x)$ with a set of equal-volume disjoint rectangles (we use a set of 256). These rectangles are oriented in a vertical stack with a single rectangle at any given height, y. The vertices of rectangle i are (x_i^l, y_i), (x_i^l, y_{i+1}), (x_i^h, y_i), and (x_i^h, y_{i+1}), where the values y_i are constructed via

$$y_1 = 0 \tag{A1}$$

$$y_{i+1} = y_i + V/(x_i^h - x_i^l) \tag{A2}$$

where V is the volume of the individual rectangles, and x_i^l and x_i^h are, respectively, the lowest and highest values of x such that $f(x) \geq y_i$. The vertices of the rectangles are

uniquely defined by the above equations and the value of V. We choose V to be the smallest value such that 256 rectangles entirely covers the area under the curve.

Sampling proceeds by choosing a random rectangle, i, and then a random distance, x, between x_i^l and x_i^h. If x also lies between x_{i+1}^l and x_{i+1}^h, then clearly every point in this rectangle with that value of x lies below the curve $f(x)$ and x can be accepted as a valid sample. Otherwise, we choose a random value, y_{trial}, between y_i and y_{i+1} and if y_{trial} is less than $f(x)$ we also accept x as a valid sample. If, instead, y_{trial} is greater than $f(x)$ then we reject this value of x and start from scratch.

REFERENCES

[1] D. F. Eggers, Jr., N. W. Gregory, G. D. Halsey, Jr., and B. S. Rabinovitch, *Physical Chemistry* (Wiley, New York, 1964).

[2] D. A. McQuarrie, *Statistical Mechanics* (Harper Collins, New York, 1976).

[3] T. L. Hill, *An Introduction to Statistical Thermodynamics* (Dover, New York, 1986).

[4] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).

[5] R. P. Feynman, *Statistical Mechanics: A Series of Lectures* (Benjamin, Reading, MA, 1972).

[6] L. S. Schulman, *Techniques and Applications of Path Integration* (Wiley, New York, 1981).

[7] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, and Polymer Physics*, 2nd ed (World Scientific, Singapore, 1995).

[8] R. Q. Topper, Q. Zhang, Y.-P. Liu, and D. G. Truhlar, J. Chem. Phys. **98**, 4991 (1993).

[9] S. L. Mielke, J. Srinivasan, and D. G. Truhlar, J. Chem. Phys. **112**, 8758 (2000).

[10] R. Q. Topper, G. J. Tawa, and D. G. Truhlar, J. Chem. Phys. **97**, 3668 (1992), **113**, 3930E (2000).

[11] S. L. Mielke and D. G. Truhlar, J. Chem. Phys. **114**, 621 (2001).

[12] H. Partridge and D. W. Schwenke, J. Chem. Phys. **106**, 4618 (1997).

[13] A. Chakraborty, D. G. Truhlar, J. M. Bowman, and S. Carter, J. Chem. Phys. (submitted).

[14] D. L. Gray and A. G. Robiette, Mol. Phys. **37**, 1901 (1979).

[15] W. T. Raynes, P. Lazzeretti, R. Zanasi, A. J. Sadlej, and P. W. Fowler, Mol. Phys. **60**, 509 (1987).

[16] X. G. Zhao, A. Gonzalez-Lafont, D. G. Truhlar, and R. Steckler, J. Chem. Phys. **94**, 5544 (1991).

[17] M. J. T. Jordan and R. G. Gilbert, J. Chem. Phys. **102**, 5669 (1995).

[18] J. Koput, Chem. Phys. Lett. **236**, 516 (1995).

[19] T. J. Lee, J. M. L. Martin, and P. R. Taylor, J. Chem. Phys. **102**, 254 (1995).

[20] S. Carter, H. M. Shnider, and J. M. Bowman, J. Chem. Phys. **110**, 8417 (1999).

[21] B. Kuhn, T. R. Rizzo, D. Luckhaus, M. Quack, and M. A. Suhm, J. Chem. Phys. **111**, 2565 (1999).

[22] N. J. Wright and R. B. Gerber, J. Chem. Phys. **112**, 2598 (2000).

[23] A. D. Isaacson and D. G. Truhlar, J. Chem. Phys. **75**, 4090 (1981).

[24] J. Koput, S. Carter, and N. C. Handy, J. Phys. Chem. A **102**, 6325 (1998).

[25] J. Srinivasan, Y. L. Volobuev, S. L. Mielke, and D. G. Truhlar, Comput. Phys. Commun. **128**, 446 (2000).

[26] *NIST-JANAF Thermochemical Tables*, edited by M. W. Chase (American Institute of Physics, Woodbury, NY, 1998).

[27] O. V. Dorofeeva, V. S. Iorish, V. P. Novikov, and D. B. Neumann, J. Phys. Chem. Ref. Data **32**, 879 (2003).

[28] E. B. Wilson, Jr., Chem. Rev. **27**, 17 (1940).

[29] G. Herzberg, *Infrared and Raman Spectra* (D. Van Nostrand, Princeton, 1945), p. 508.

[30] W. H. Miller, J. Chem. Phys. **63**, 1166 (1975).

[31] J. D. Doll and D. L. Freeman, J. Chem. Phys. **80**, 2239 (1980).

[32] R. D. Coalson, D. L. Freeman, and J. D. Doll, J. Chem. Phys. **85**, 4567 (1986).

[33] R. D. Coalson, J. Chem. Phys. **85**, 926 (1986).

[34] J. D. Doll, R. D. Coalson, and D. L. Freeman, Phys. Rev. Lett. **55**, 1 (1986).

[35] D. L. Freeman and J. D. Doll, Adv. Chem. Phys. **70B**, 139 (1988).

[36] R. D. Coalson, D. L. Freeman, and J. D. Doll, J. Chem. Phys. **91**, 4242 (1989).

[37] J. D. Doll, D. L. Freeman, and T. Beck, Adv. Chem. Phys. **78**, 61 (1990).

[38] R. Q. Topper and D. G. Truhlar, J. Chem. Phys. **97**, 3647 (1992).

[39] C. Predescu and J. D. Doll, J. Chem. Phys. **117**, 7448 (2002).

[40] C. Predescu and J. D. Doll, Phys. Rev. E **67**, 026124 (2003).

[41] C. Predescu, D. Sabo, and J. D. Doll, J. Chem. Phys. **119**, 4641 (2003).

[42] S. L. Mielke and D. G. Truhlar, Chem. Phys. Lett. **378**, 317 (2003).

[43] J. O. Hirschfelder and J. S. Dahler, J. Chem. Phys. **24**, 1258 (1956).

[44] S. L. Mielke and D. G. Truhlar, J. Chem. Phys. **115**, 652 (2001).

[45] G. Marsaglia and W. W. Tsang, J. Sci. Stat. Comput. **5**, 349 (1984).

[46] G. Marsaglia and W. W. Tsang, J. Stat. Softw. **5** (2000).

[47] G. Marsaglia, Ann. Math. Stat. **43**, 645 (1972).

[48] R. J. Duchovic, Y. L. Volobuev, G. C. Lynch, D. G. Truhlar, T. C. Allison, A. F. Wagner, B. C. Garrett, and J. C. Corchado, Comput. Phys. Commun. **144**, 169 (2002), **156**, 319E (2004).

[49] R. J. Duchovic, Y. L. Volobuev, G. C. Lynch, A. W. Jasper, D. G. Truhlar, T. C. Allison, A. F. Wagner, B. C. Garrett, J. Espinosa-García, and J. C. Corchado, POTLIB-online, http://comp.chem.umn.edu/potlib.

[50] J. C. Corchado, Y.-Y. Chuang, P. L. Fast, J. Villà, W.-P. Hu, Y.-P. Liu, G. C. Lynch, K. A. Nguyen, C. F. Jackels, V. S. Melissas, B. J. Lynch, I. Rossi, E. L. Coitiño, A. Fernandez-Ramos, J. Pu, T. V. Albu, R. Steckler, B. C. Garrett, A. D. Isaacson, and D. G. Truhlar, POLYRATE-version 9.1 (University of Minnesota, Minneapolis, 2002).

[51] J. Koput, S. Carter, and N. C. Handy, J. Chem. Phys. **115**, 8345 (2001).

[52] J. Koput, private communication.

[53] W. B. Cook, R. H. Hunt, W. N. Shelton, and F. A. Flaherty, J. Mol. Spectrosc. **155**, 84 (1992).

[54] W. B. Olson, R. H. Hunt, B. W. Young, A. G. Maki, and J. W. Brault, J. Mol. Spectrosc. **127**, 12 (1988).

[55] A. Perrin, A. Valentin, J. M. Flaud, C. Camy-peyret, L. Schriver, A. Schriver, and P. Arcas, J. Mol. Spectrosc. **171**, 385 (1995).

[56] C. Camy-Peyret, J. M. Flaud, J. W. C. Johns, and M. Noel, J. Mol. Spectrosc. **155**, 84 (1992).

[57] J. M. Flaud, C. Camy-Peyret, J. W. C. Johns, and B. Carli, J. Chem. Phys. **91**, 1504 (1989).

[58] S. Carter and N. C. Handy, J. Chem. Phys. **113**, 987 (2000).

[59] K. S. Pitzer and W. D. Gwinn, J. Chem. Phys. **10**, 428 (1942).

[60] D. G. Truhlar and A. D. Isaacson, J. Chem. Phys. **94**, 357 (1991).

[61] A. D. Isaacson and D. G. Truhlar, J. Chem. Phys. **80**, 2888 (1984).

[62] A. D. Isaacson, D. G. Truhlar, K. Scanlon, and J. Overend, J. Chem. Phys. **75**, 3017 (1981).

[63] X. Luo, P. R. Fleming, and T. R. Rizzo, J. Chem. Phys. **96,** 5659 (1992).

[64] See EPAPS Document No. E-JCPXXX-XXX-XXXXXX for electronic files that contain (a) an extended table of extrapolated partition functions, (b) an extended table of vibrational partition functions at 10 temperatures ranging from 298 K - 2400 K, and (c) an extended table of rotational partition functions at 10 temperatures ranging from 298 K - 2400 K. This document my be retrieved via the EPAPS homepage (http://www.aip.org/pubservs/epaps.html) or from ftp.aip.org in the directory /epaps/. See the EPAPS homepage for more information.

[65] G. Lewis, M. Randall, K. S. Pitzer, and L. Brewer, *Thermodynamics* (McGraw-Hill, New York, 1961).

[66] M. L. Martin, J. P. Francois, and R. Gijbels, J. Chem. Phys. **96**, 7633 (1992).

TABLE I. Parameters used in calculating $Q^{[P]}(T)$ at various temperatures.

Parameter	300 K	400 K	600–800 K	1000 K	1500 K	2400 K
N^{samples} [a]	2×10^8	2×10^8	5×10^7	1×10^7	1×10^7	1×10^7
$R_{\text{OH}}^{\text{min}}$ [b]	60	60	60	60	50	50
$R_{\text{OH}}^{\text{max}}$ [b]	100	100	100	100	105	105
$R_{\text{OH-OH}}^{\text{min}}$ [c]	290	290	290	290	270	270
$R_{\text{OH-OH}}^{\text{max}}$ [c]	450	450	450	450	480	480
$\Delta_{R_{\text{OH}}}$ [d]	0.05	0.055	0.08	0.09	0.12	0.18
$\Delta_{R_{\text{OH–OH}}}$ [e]	0.08	0.09	0.13	0.15	0.20	0.27
Δ_{θ_i} [f]	5	5	7	9	10	16

[a] The number of Monte Carlo samples; 10% are distributed in a strata-blind manner in the initial phase, the remaining samples are adaptively distributed among 20 equal-volume strata during 20 subsequent sampling phases.

[b] The minimum and maximum allowed OH distance in mass-scaled bohr. The coordintates are scaled to a reduced mass of 1 m_e (i.e., the mass of an electron).

[c] The minimum and maximum allowed OH–OH distance in mass-scaled bohr.

[d] The OH importance function width parameter (in bohr) in the importance function; the center is always taken as 1.8192 a_0.

[e] The OH–OH importance function width parameter (in bohr) in the importance function; the center is always taken as 2.77611 a_0.

[f] The importance function width parameter (in degrees) for sampling either of the two Jacobi angles; the center is always taken as 102.934º.

TABLE II. Calculated and experimentally observed fundamentals and harmonic normal mode frequencies (in cm^{-1}).

Mode[a]	Harmonic[b]	Fundamentals[c]	Experimental
ν_1	3807.7	3624.3	3609.80[e]
ν_2	1437.2	1396.9	1395.88[f]
ν_3	910.9	874.4	865.94[g]
ν_4	381.9	371.6	370.89[h]
ν_5	3809.2	3625.1	3610.66[d]
ν_6	1329.8	1267.8	1264.58[f]

[a] The vibrational modes are symmetric (ν_1) and antisymmetric (ν_5) OH stretches, symmetric (ν_2) and antisymmetric (ν_6) HOO bends, symmetric OO stretch (ν_3), and the HOOH torsion (ν_4).
[b] The harmonic normal mode frequencies as calculated using the POLYRATE v9.1 program.
[c] The calculated fundamentals as obtained by Koput, Carter, and Handy using their H_2O_2 potential energy surface.[24]
[d] Ref. 53
[e] Ref. 54
[f] Ref. 55
[g] Ref. 56
[h] Ref. 57

TABLE III. Comparison of zero-point energies of H_2O_2 from various sources.

Source	ZPE (cm^{-1})
From calculated fundamentals	5580.1
From harmonic frequencies	5838.4
Accurate for this PES[a]	5726.1

[a] the variational result of Ref. 52.

TABLE IV. Source of zero point energies and frequency values used for various approximate partition functions

Method	ZPE	frequencies
$Q_{\mathrm{vib}}^{\mathrm{HO}}$	harmonic	harmonic
$Q_{\mathrm{vib}}^{\mathrm{HO\text{-}Z}}$	accurate	harmonic
$Q_{\mathrm{vib}}^{\mathrm{HO-ZF}}$	accurate	fundamentals
$Q_{\mathrm{vib}}^{\mathrm{HO-F}}$	from fundamentals	fundamentals

TABLE V. Values of the moments of inertia used in the calculations (in kg m^2).[a]

$I_{\mathrm{A}}^{\mathrm{e}}$	2.76 (-47)
$I_{\mathrm{B}}^{\mathrm{e}}$	3.18 (-46)
$I_{\mathrm{C}}^{\mathrm{e}}$	3.27 (-46)
$I_{\mathrm{M}}^{\mathrm{e}}$	3.23 (-46)
I_{A}^{0}	2.78 (-47)
I_{M}^{0}	3.27 (-46)

[a] Powers of ten are in parentheses.

TABLE VI. TT-FPI partition functions and the 2σ statistical uncertainties for various values of P and various temperatures.

	300 K	400 K	600 K	800 K
$P^{\max}$	90	60	30	30
$Q^{[1]}$	$(1.90264\pm0.00021)\times10^{-2}$	$1.66816\pm0.00014)\times10^{-1}$	3.50548 ± 0.00090	30.2048 ± 0.0051
$Q^{[P^{\max}/3]}$	$(3.984\pm0.018)\times10^{-9}$	$(6.8927\pm0.0073)\times10^{-6}$	$(1.8298\pm0.0014)\times10^{-2}$	1.08642 ± 0.00047
$Q^{[P^{\max}/2]}$	$(3.409\pm0.015)\times10^{-9}$	$(5.9612\pm0.0065)\times10^{-6}$	$(1.5567\pm0.0013)\times10^{-2}$	1.01020 ± 0.00044
$Q^{[P^{\max}]}$	$(3.092\pm0.013)\times10^{-9}$	$(5.4381\pm0.0061)\times10^{-6}$	$(1.4001\pm0.0012)\times10^{-2}$	0.96477 ± 0.00043
Q^{ESPE}	$(2.995\pm0.013)\times10^{-9}$	$(5.2746\pm0.0059)\times10^{-6}$	$(1.3501\pm0.0011)\times10^{-2}$	0.94972 ± 0.00042

	1000 K	1500 K	2400 K
$P^{\max}$	30	18	18
$Q^{[1]}$	160.499 ± 0.071	3358.6 ± 1.6	116785 ± 81
$Q^{[P^{\max}/3]}$	16.481 ± 0.013	1174.18 ± 0.72	76478 ± 59
$Q^{[P^{\max}/2]}$	15.856 ± 0.013	1136.06 ± 0.69	75744 ± 58
$Q^{[P^{\max}]}$	15.479 ± 0.012	1112.91 ± 0.66	75301 ± 58
Q^{ESPE}	15.354 ± 0.012	1105.11 ± 0.66	75152 ± 58

TABLE VII. Relative efficiencies of various subsets of configuration space importance sampling and stratification options for a calculation at T = 1000 K using the parameters in Tables I and VI.

CSIS scheme	stratification	relative efficiency[a]
A[b] and R[c]	yes	1600
A only	yes	27
R only	yes	56
none	yes	1.4
A and R	no	930
A only	no	19
R only	no	39
none	no	1.0

[a] The efficiency is relative to a calculation without stratification or either type (radial, R, or angular, A) of configuration space importance sampling (CSIS). All calculations included importance sampling in the Fourier coefficient space.
[b] Denotes importance sampling in the two Jacobi angles as further discussed in the text.
[c] Denotes importance sampling in the three Jacobi magnitudes as further discussed in the text

TABLE VIII. Classical mechanical and quantum mechanical values of Q_{vib} at various temperatures[a]

T(K)	Q_{vib}^{CHO}	Q_{vib}^{HO}	Q_{vib}^{HO-Z}	Q_{vib}^{HO-ZF}	Q_{vib}^{HO-F}
300	8.52(−6)	8.36(−13)	1.43(−12)	1.45(−12)	2.92(−12)
800	3.06(−3)	8.21(−5)	1.00(−4)	1.06(−4)	1.38(−4)
2400	2.23	1.37	1.47	1.67	1.82

[a] Powers of ten are in parentheses.

TABLE IX. Classical mechanical and quantum mechanical values of $Q_{\mathrm{rot}}(T)$ at various temperatures

T(K)	$Q_{\mathrm{rot,CM}}^{\mathrm{asym,e}}$	$Q_{\mathrm{rot,CM}}^{\mathrm{sym,e}}$	$Q_{\mathrm{rot}}^{\mathrm{sym,e}}$	$Q_{\mathrm{rot,CM}}^{\mathrm{sym,0}}$	$Q_{\mathrm{rot}}^{\mathrm{sym,0}}$
300	965.3	965.3	966.6	983.6	985.0
800	4203.3	4203.7	4205.8	4283.4	4285.6
2400	21841.2	21843.1	21846.8	22257.4	22261.1

TABLE X. Vibrational-rotational partition functions at various temperatures[a]

T(K)	CHO-RR	HO-RR	HO-Z-VR	HO-ZF-VR	CM, anhar.[b]	HO-F-VR	ESPE
300	8.23(−3)	8.08(−10)	1.41(−9)	1.43(−9)	1.90(−2)	2.88(−9)	3.00(−9)
400	7.12(−2)	1.59(−6)	2.43(−6)	2.48(−6)	1.67(−1)	4.20(−6)	5.27(−6)
600	1.49	4.60(−3)	6.13(−3)	6.38(−3)	3.51	9.05(−3)	1.35(−2)
800	1.29(+1)	3.45(−1)	4.30(−1)	4.54(−1)	3.02(+1)	5.91(−1)	9.50(−1)
1000	6.87(+1)	5.80	6.94	7.43	1.60(+2)	9.17	1.54(+1)
1500	1.44(+3)	4.37(+2)	4.96(+2)	5.46(+2)	3.36(+3)	6.28(+2)	1.11(+3)
2400	4.88(+4)	2.99(+4)	3.26(+4)	3.72(+4)	1.17(+5)	4.06(+4)	7.52(+4)

[a] Powers of ten in parentheses.

[b] The accurate classical result obtained by a TT-FPI calculation with $P = 1$.

TABLE XI. Partition functions obtained by the Pitzer–Gwinn approximation compared to the accurate path integral results at various temperatures[a]

T(K)	PG-HZPE	PG-AZPE	$Q^{\mathrm{ESPE}}(T)$
300	1.87(–9)	3.20(–9)	3.00(–9)
400	3.73(–6)	5.58(–6)	5.27(–6)
600	1.08(–2)	1.42(–2)	1.35(–2)
800	8.09(–1)	9.90(–1)	9.50(–1)
1000	1.35(+1)	1.59(+1)	1.54(+1)
1500	1.02(+3)	1.14(+3)	1.11(+3)
2400	7.16(+4)	7.66(+4)	7.52(+4)

[a] Powers of ten in parentheses.

TABLE XII. Standard-state Gibbs free energy values, $\widetilde{G}_T^\circ$, (in kJ/mol) at various temperatures.

T(K)	CHO-RR	HO-RR	HO-Z-VR	HO-ZF-VR
300	–96.18	–55.93	–57.32	–57.35
400	–114.97	–79.36	–80.77	–80.84
600	–158.44	–129.60	–131.04	–131.23
800	–207.55	–183.48	–184.94	–185.31
1000	–260.87	–240.31	–241.81	–242.38
1500	–407.62	–392.77	–394.34	–395.55
2400	–704.88	–695.14	–696.86	–699.48

T(K)	HO-F-VR	CM, anhar.[a]	ESPE[b]
300	–59.10	–98.27	–59.20
400	–82.59	–117.81	–83.35
600	–132.98	–162.71	–134.98
800	–187.05	–213.22	–190.21
1000	–244.13	–267.93	–248.41
1500	–397.29	–418.21	–404.35
2400	–701.23	–722.30	–713.50

[a] The accurate classical result obtained by a TT-FPI calculation with $P = 1$.
[b] The enhanced same path extrapolated TT-FPI results

TABLE XIII. Standard-state Gibbs free energy values, $\widetilde{G}_T^\circ$, (in kJ/mol) at various temperatures

T(K)	PG-HZPE	PG-AZPE	ESPE[a]	JANAF[b]	JPCRD[c]
300	–58.02	–59.36	–59.20	–59.04	–59.20
400	–82.19	–83.53	–83.35	–83.05	–83.34
600	–133.87	–135.21	–134.98	–134.54	–134.97
800	–189.14	–190.48	–190.21	–189.78	–190.19
1000	–247.37	–248.71	–248.41	–248.03	–248.36
1500	–403.35	–404.69	–404.35	–404.03	–404.09
2400	–712.55	–713.89	–713.50	N/A	–712.29

[a] The enhanced same path extrapolated TT-FPI results
[b]Ref. 26
[c]Ref. 27