

Quantum Mechanical Single Molecule Partition Function from Path Integral Monte Carlo Simulations

Shaji Chempath* and Alexis T. Bell†

Department of Chemical Engineering, University of California, Berkeley, California 94720

Cristian Predescu‡

*Department of Chemistry and Kenneth S. Pitzer Center for Theoretical Chemistry,
University of California, Berkeley, California 94720*

(Dated: March 24, 2006)

An algorithm for calculating the partition function of a molecule with the path integral Monte Carlo method is presented. Staged thermodynamic perturbation with respect to a reference harmonic potential is utilized to evaluate the ratio of partition functions. Parallel tempering and a new Monte Carlo estimator for the ratio of partition functions are implemented here to achieve well converged simulations that give an accuracy of 0.04 kcal/mol in the reported free energies. The method is applied to various test systems, including a catalytic system composed of 18 atoms. Absolute free energies calculated by this method lead to corrections as large as 2.6 kcal/mol at 300 K for some of the examples presented.

Keywords: path integral Monte Carlo, partition function, harmonic approximation, free energy, thermodynamic perturbation

I. INTRODUCTION

The thermodynamics of chemical reactions are in principle completely determined by the molecular partition functions of the reactants and products. If the energy states of a molecule are quantized, then the molecular partition function $Z(\beta)$ is defined by the equation

$$Z(\beta) = \sum_i \exp(-\beta E_i), \quad (1)$$

where E_i is the energy of a molecule in state i and at inverse temperature $\beta = 1/k_bT$. The values of the Born-Oppenheimer potential $V(\mathbf{x})$ for reactants and products can be determined from ab-initio quantum chemical calculations. In order to account for the quantum effects associated with the motion of nuclei, the molecules are often treated as rigid-rotor harmonic oscillators. In this approximation, the molecular potential energy surface (PES) is

approximated as a quadratic form around the minimum energy configuration \mathbf{x}_0 :

$$V(\mathbf{x}) = V(\mathbf{x}_0) + (\mathbf{x} - \mathbf{x}_0)^T \cdot \mathbf{K} \cdot (\mathbf{x} - \mathbf{x}_0) \quad (2)$$

where \mathbf{K} is the Hessian matrix evaluated at the local minimum \mathbf{x}_0 . The eigenvectors of \mathbf{K} correspond to the normal modes of vibrations and the eigenvalues of \mathbf{K} correspond to the square of the vibrational frequencies of the normal modes.

However, there are many cases where the harmonic approximation fails. For the partition function of a vibrational mode, the harmonic approximation produces:

$$Z(\beta) = \sum_{n=0}^{\infty} \exp[-(n + 1/2)\beta\hbar\omega] \quad (3)$$

where ω is the vibrational frequency. For large ω , say a vibrational mode with a frequency of around 3000 cm^{-1} , which roughly corresponds to a CH bond vibration, only the first term (ground state) in the summation is required to obtain convergence of Eq. 3. The distance covered by the first vibrational state can be estimated as $\sqrt{\hbar/\mu/\omega}=0.1\text{\AA}$, assuming a mass of $\mu=1$ amu, which corresponds to a hydrogen atom. In this case, the harmonic approximation gives the correct answer provided that the potential remains harmonic for displacements as large as 0.1\AA from the minimum energy geometry. For a small value of ω , for example, a vibrational frequency of 30 cm^{-1} , the first 20 terms in the summation must be considered in order for Eq. (3) to converge to within 95% of the actual value of $Z(\beta)$. The distance covered by the 20th vibrational state is estimated to be 7\AA . The harmonic approximation gives the correct answer only if the potential remains parabolic for a displacement of 7\AA from the minimum energy geometry, a condition that will not be met for molecular systems. Following these arguments it becomes evident that, for a given vibrational mode, the harmonic approximation will become increasingly inaccurate as the temperature increases.

The limitations of the harmonic approximation have been discussed previously¹⁻³ and several attempts have been made to correct for this limitation in calculation of partition functions. In the classical treatment, the summation in Eq. 1 is replaced by an integral in the phase space. Kolossvary² introduced a mode-integration algorithm (MINTA) for calculating the configuration integral over all degrees of freedom of a molecular system. He utilized the harmonic approximation for effective sampling of the phase space. Chang et al.¹ compared the use of different coordinate systems for MINTA and a mode-scanning algorithm and concluded that such algorithms work better with bond-angle-torsion coordinates than with Cartesian coordinates.

Lynch and coworkers⁴ have employed a path integral Monte Carlo technique to calculate the quantum mechanical partition function (Eq. 1) of hydrogen peroxide (H_2O_2). They performed importance sampling in the Fourier coefficient

space and utilized an adaptively-optimized stratified sampling. An interesting feature of the algorithm is the so-called “path-by-path” extrapolation technique devised by Mielke and Truhlar,⁵ which reduces the number of Trotter slices necessary for accurate path integration. These authors have recently⁶ applied their algorithm to isotopically substituted variants of H_2O_2 , obtained by exchanging ^{18}O and deuterium for oxygen and hydrogen in H_2O_2 . An analytical potential energy surface developed by Koput et al.⁷ was employed in these studies and Jacobi coordinates were utilized to facilitate convergence of the simulations.

Miller and Clary^{8,9} introduced a torsional path integral Monte Carlo approach to handle anharmonic PESs. Bond lengths and angles were kept fixed, and only the torsional degrees were allowed to vary. The authors utilized the harmonic approximation in the torsional coordinates as the reference state and then made recourse to thermodynamic perturbation in order to calculate the ratio of the partition functions between the reference state and the real state.

In the current work, we illustrate the advantages of a path integral Monte Carlo approach (PIMC) based on a recently developed fourth-order short-time approximation¹⁰ and the fast sampling algorithm.¹¹ The method presented here combines the efficiency of the PIMC technique with that of the parallel tempering algorithm and a new estimator for ratios of partition functions. The end result is a technique capable of providing accurate results despite our choice to perform the sampling in Cartesian coordinates, rather than to utilize torsional or other Jacobi coordinates. This feature makes the approach quite general and applicable to any system for which a potential energy surface is available. In the following sections, we demonstrate the capabilities of the method by applying it to systems which contain up to 18 atoms. As already mentioned, we use parallel tempering to ensure adequate sampling of the configuration space. In addition we also introduce a new estimator for evaluating the ratio of partition functions from parallel tempering simulations (see Appendix B). The potential energy of molecules is described using an analytical forcefield expression which is parameterized to match the Born-Oppenheimer surface. We use thermodynamic perturbation with respect to a harmonic potential to calculate the absolute free energy of the molecules. The method is described in Section II and the applications of the method to test cases are given in Section III. Some details of the implementation are given in the Appendices A-C.

II. METHODS

A. Path integral Monte Carlo

The path integral Monte Carlo method utilized has been described previously.^{11–14} This approach only requires energy evaluations and does not require derivatives. The code for path integral Monte Carlo and free energy evaluation is available from our website.¹⁵

The probability of finding the system at position x is described by the Feynman-Kac formula, which for the diagonal elements of a one dimensional system is given by

$$\rho(x, \beta) = \frac{1}{\sqrt{2\pi\sigma^2}} \mathbb{E} \exp \left[-\beta \int_0^1 V(x + \sigma B_u^0) du \right], \quad (4)$$

with \mathbb{E} denoting an average, $\sigma = (\hbar^2 \beta / m_0)^{1/2}$, and B_u^0 describing a Gaussian stochastic process known as a Brownian bridge. In short, the Brownian bridge is an ensemble of trajectories which, upon discretization of the one-dimensional integral in the exponent, transforms into an ensemble of points in the configuration space. The distribution of the ensemble of points is a certain correlated Gaussian distribution, which, in a Monte Carlo simulation, is further multiplied by a Boltzmann weight. The averaging implied by \mathbb{E} is performed over all possible closed paths that start and end at x .

From a numerical point of view, we are free to further approximate the Gaussian distribution and to choose the quadrature technique in an advantageous way, so as to minimize the computational cost. We treat this aspect of the problem by employing a fourth-order short-time approximation that has the advantage of requiring only calls to the potential function.¹⁰ Forces or higher order derivatives are not required. In order to cope with the aforementioned correlation of the beads, special sampling techniques must be utilized and our choice, the fast sampling algorithm, is among the most efficient.¹¹ The justification of the path integral technology employed is not within the scope of the present paper and can be found in the cited literature. However, some general features relevant to the present development are reviewed in Appendix A.

B. Potential energy surface

The energy of a molecular system can be evaluated accurately by many quantum chemistry software packages. Given the coordinates of the nuclei, such algorithms return the energy of the system within the Born-Oppenheimer approximation at a given level of theory. However, during a Monte Carlo simulation it is not economical to use a

quantum chemistry package at every iteration of the simulation. Instead, we use a force field approximation where the energy of the structure is written in terms of bond lengths, angles, and torsions:

$$V = \sum_{bonds} \frac{k_r}{2}(r - r_0)^2 + \sum_{angles} \frac{k_\theta}{2}(\theta - \theta_0)^2 + \sum_{torsions} \sum_{n=0}^5 a_n \cos(\phi - \phi_0)^n \quad (5)$$

Forcefield parameters used for all molecules studied here are given as supplementary information.¹⁶ The parameters of the force field are adjusted to match the Born Oppenheimer PES as closely as possible. B3LYP density functional theory is used in defining the potential energy surface of all the molecules. All QM calculations were done using the Gaussian quantum chemical package.¹⁷ The 6-311++G** basis set was used for ethanol, dimethyl ether, and glycine. In the case of palladium containing-complexes the LANL2DZ/6-31G* basis set was used. The equilibrium bond lengths and bond angles used in the forcefield expressions are taken to be the ones obtained from geometry optimization using the Gaussian software. The spring constants, namely k_r , k_θ , and a_n , in the forcefield expressions are adjusted to match the eigen values of the Hessian matrix at the minimum energy geometry. The Hessian matrix was obtained analytically by using the frequency evaluation module within the Gaussian software. In the case of H_2O_2 we used the potential energy surface defined by Koput et al.⁷

C. Thermodynamic perturbation

We will calculate the ratio of partition functions of the real system with a reference system. The real system is represented by a potential V_0 which is parameterized to match the Born Oppenheimer potential as explained previously. The reference potential, denoted by V_1 , is created by increasing the spring constants (k_r , k_θ , a_n) of all terms in the forcefield to a large value. The spring constants are made large enough to make all the vibrational modes of the molecule to have frequencies larger than 200 cm^{-1} . The harmonic approximation is valid for such a potential. Thus, we will use the harmonic approximation to calculate Z_1 , the partition function for potential V_1 . Then, from a path integral Monte Carlo simulation, the ratio of partition functions is evaluated as

$$\frac{Z_1}{Z_0} = \left\langle \frac{\rho_1(\mathbf{x})}{\rho_0(\mathbf{x})} \right\rangle_0 \quad (6)$$

where ρ_0 and ρ_1 refers to the probability distributions with the potentials V_1 and V_0 respectively. It should be noted that the reference potential V_1 refers to a rigid-rotor for which harmonic approximation is valid. The real potential V_0 does not assume rigidity, and thus the partition function Z_0 is calculated without assuming rotational-

vibrational decoupling. For all the molecules studied here the reference potential V_1 is also given in the supplementary information.¹⁶ The original potential V_0 may have more than one minimum on its potential energy surface. For example a rotating CH_3 group will give rise to 3 minima spaced 120 degree apart in the rotational angle, leading to a torsional term of the form $1 - 4 \cos^3(\phi) + 3 \cos(\phi)$. In the reference potential V_1 , all such terms are modified to $1 - \cos(\phi)$ so that there is only one stationary point on the reference PES. Note that if there are multiple minima on the PES the harmonic approximation can not be used to calculate the partition function for V_1 .

D. Replica exchange simulations

The technique of replica exchange^{18,19} (or parallel tempering) is a crucial aspect of the present development. A straight forward evaluation of the integral in Eq. 6 from a single simulation may not converge in a reasonable time. We use a set of parallel Monte Carlo simulations, performed at different potentials characterized by a parameter λ . $\lambda_0 = 0$ corresponds to V_0 and $\lambda_M = 1$ corresponds to V_1 . The word *replica* is used to refer to an individual simulation and there are a total of $M + 1$ replicas. For intermediate values of λ the potential energy is defined as

$$V(\lambda_i, \mathbf{x}) = V_0(\mathbf{x})(1 - \lambda_i) + V_1(\mathbf{x})\lambda_i \quad i = 0, 1, 2 \dots M \quad (7)$$

Thus the ratio between partition functions is calculated in stages as:

$$\frac{Z_1}{Z_0} = \prod_{i=0}^{M-1} \frac{Z(i+1)}{Z(i)} \quad (8)$$

$Z(i)$ refers to the partition function with the potential $V(\lambda_i)$. The technique of parallel tempering attempts to swap configurations among neighboring replicas according to the Metropolis algorithm, so that the Boltzmann distributions of the individual replicas are preserved. Nevertheless, the mixing involved has the effect of decreasing the correlation times: a walker moving in a highly confined environment will get a chance of escaping certain local minima by borrowing from the larger mobility of a neighboring walker.

To further improve the speed of Monte Carlo simulations we added an additional set of replicas with increasing temperatures, T_i . In the parallel tempering framework, these high temperature replicas jump over the barriers in phase space much faster than the low temperature replicas. As we have already mentioned, exchange of configurations between the parallel replicas prevents the system from being stuck at a local minimum. Thus, each replica is characterized by λ_i and T_i . In a typical calculation we use 32 parallel replicas running at different T and λ . The first 16 replicas are at constant T (usually 300 K) and varying λ . $\lambda = 1$ for the first simulation, and λ decreases

monotonically to $\lambda = 0$ for the 16th replica. The next 16 replicas are run with $\lambda = 0$ but T is varied from 300 K to a large value of T_{max} .

The replica simulations are run in parallel on a Linux cluster with MPI parallel programming software. Exchange of configurations is attempted between adjacent replicas at regular intervals during the Monte Carlo simulation. These exchanges are accepted with probability:

$$\min \left\{ 1, \frac{\rho_i(\mathbf{x}')\rho_j(\mathbf{x})}{\rho_i(\mathbf{x})\rho_j(\mathbf{x}')} \right\}. \quad (9)$$

where the i^{th} replica contains configuration \mathbf{x} , and j^{th} replica contains configuration \mathbf{x}'

The values of λ_i and T_i are adjusted to make sure that the acceptance rates for an exchange attempt is about 40%. The tuning algorithm is explained in the Appendix C.

As implied by Eqs. 6 and 8, the ratios of partition functions can be computed in the form (here, $j = i \pm 1$)

$$\frac{Z_j}{Z_i} = \frac{\int \rho_i(\mathbf{x})r_{ij}(\mathbf{x})d\mathbf{x}}{\int \rho_i(\mathbf{x})d\mathbf{x}} = \frac{\int d\mathbf{x} \int d\mathbf{x}' \rho_i(\mathbf{x})\rho_j(\mathbf{x}')r_{ij}(\mathbf{x})}{\int d\mathbf{x} \int d\mathbf{x}' \rho_i(\mathbf{x})\rho_j(\mathbf{x}')}, \quad (10)$$

where,

$$r_{ij}(\mathbf{x}) = \rho_j(\mathbf{x})/\rho_i(\mathbf{x}).$$

In the context of parallel replica simulations, where we simulate in parallel the distribution $\rho_j(\mathbf{x}')$ as well, we have observed that the following estimator is statistically more robust. Define the function

$$S_{ij}(\mathbf{x}, \mathbf{x}') = \frac{\rho_i(\mathbf{x}')\rho_j(\mathbf{x})}{\rho_i(\mathbf{x})\rho_j(\mathbf{x}')} \quad (11)$$

and let $\epsilon > 0$ be a small number. The new estimator is a function of both variables \mathbf{x} and \mathbf{x}' and reads

$$r_{ij}(\mathbf{x}, \mathbf{x}') = \frac{\rho_j(\mathbf{x})}{\rho_i(\mathbf{x})} \times \begin{cases} 2, & \text{if } S_{ij}(\mathbf{x}, \mathbf{x}') < 1 - \epsilon, \\ 0, & \text{if } S_{ij}(\mathbf{x}, \mathbf{x}') > (1 - \epsilon)^{-1}, \\ 1, & \text{otherwise.} \end{cases} \quad (12)$$

Notice that the function $S_{ij}(\mathbf{x}, \mathbf{x}')$ is also utilized for the acceptance test (Eq. 9). The number ϵ is introduced so that to deal with certain degenerate cases (e.g., if the distributions are equal). Although the estimator is correct for all values of ϵ , it is better behaved for small values, such as the square root of the floating precision. The derivation of the estimator is done in the Appendix B.

E. Simulation details

The code for conducting the path integral Monte Carlo is available for download from our website.¹⁵ The inputs to the code are the minimum energy structure and a file containing the force field parameters. The number of beads in the path integral, the number of parallel replicas, and the total number of MC iterations is specified at the beginning of the run. For data analysis, the total simulation time of each replica is divided into 200 *blocks*. The first 100 blocks are used for the equilibration of the simulation and the next 100 blocks constitute the production part of the run. During the equilibration period the MC displacements experienced by the atoms are adjusted to have an overall 40% acceptance ratio for the displacement attempts. The values of T_i , and λ_i are also adjusted during the equilibration to have 40% acceptance for the exchange of replicas between the parallel simulations. Thermodynamic averages and ratios of partition functions are accumulated during the production part of the run. The error bar, ϵ_A , in the estimate of a quantity A is defined as:

$$\epsilon_A = \frac{2\sigma_A}{\sqrt{N_b}}, \quad (13)$$

where N_b is the number of blocks used in the production part of the simulation (usually 100), and σ_A is the standard deviation in the block average values of A . Each block in the MC simulation consists of 1000-100,000 MC iterations depending on the system size. The largest simulation, which contains 18 atoms, took 3 days of CPU time on with 40 processes running in parallel on a Linux cluster operating with Intel(R) Xeon(TM) CPUs.

III. RESULTS AND DISCUSSION

A. Test Cases: harmonic oscillator and H₂O₂

For a harmonic oscillator the harmonic expression (Eq. 2) is exact and an analytical expression for Z can be obtained. However, we apply our methodology to calculate the partition function to make sure that our code is correct. We consider a hypothetical molecule with two atoms of mass 12.0 and 1.0 amu connected by a 1.0 Å bond having spring constant $k_r = 100$ kcal/Å². We define V_1 to be the potential where the spring constant is increased from 100 to 1000 kcal/Å². The ratio of partition functions, Z_1/Z_0 for this system can be calculated analytically to be 5.865. Our path integral simulation code returns a value of 5.853 ± 0.008 with the use of 6 parallel replicas and 40000 iterations per block. These values were obtained with 96 beads for each atom considered in the system. We also conducted a series of simulations where we varied N , the number of beads in the path integral, keeping everything

else constant. The calculated ratios of partition functions are plotted in Fig. 1. The case when $N=1$ corresponds to a classical simulation and the case of large N corresponds to the quantum mechanical limit. We see that for a value of $N=96$ a result very close to the quantum limit is obtained and there is no need to use a larger number of beads. In all subsequent simulations we have used 96 beads to represent the paths in the path integral, unless otherwise mentioned.

We applied our method to H_2O_2 molecule as a further test of our code. This system is anharmonic in the O-O bond rotation and absolute partition functions have been reported in the literature.⁴ The PES for hydrogen peroxide is well characterized and is available in the form of a down-loadable subroutine.^{7,20} Lynch et al.⁴ used this potential energy surface and calculated the rotational-vibrational partition function of H_2O_2 at 300K to be 3.00×10^{-9} . Using the same potential to define V_0 , our PIMC calculations gave a value of $2.95 \pm 0.03 \times 10^{-9}$ for the partition function. We used 100000 iterations per block in these simulations. Since the simulation contains 100 blocks the total number of MC iterations is 10^7 which is roughly of the same order as reported by Lynch et al.⁴ However we have used Cartesian coordinates for sampling, which makes our methodology more general and applicable to more complicated systems given below. To define the harmonic reference potential V_1 , we used the simple forcefield of the form given in Eq. 5. The parameters defining V_1 are given in the supplementary information.¹⁶

B. Ethanol and dimethyl ether

The change in enthalpy for converting ethanol to dimethyl ether, both of which have the generic formula $\text{C}_2\text{H}_6\text{O}$, is +12.2 kcal/mol based on the heats of formation given in the CRC handbook.²¹ A calculation using B3LYP density functional theory with 6-311++G** gives a value of +11.1 kcal/mol. This error arises from the inherent error in the electronic structure theory used. If a higher level theory and a larger basis set is used the electronic energy calculation can be made more accurate. For example, calculations with G3 theory²² predicts enthalpy difference to be +11.7 kcal/mol (using the G3B3 keyword in Gaussian¹⁷). However, in this paper, our focus is on the entropic contribution to free energy as given by the vibrational partition function. The experimentally measured change in entropy for converting ethanol to dimethyl ether at 300K is -3.6 cal/mol/K. However a harmonic approximation with the Hessian calculated at the B3LYP/6-311++G** level of theory predicts an entropy change of +0.09 cal/mol/K. We used our PIMC method to calculate the partition functions for ethanol and dimethyl ether more accurately. In order to do this, we defined a forcefield which is parameterized to match the B3LYP/6-311++G** potential energy surface. The parameterized forcefield is given in the supplementary information.¹⁶ The partition functions, free energies and entropies obtained by different methods are given in Table I. The difference between free energy estimated by harmonic

approximation and path integral Monte Carlo is significant (2.56 kcal/mol) in the case of ethanol. The change in entropy for converting ethanol to dimethyl ether as evaluated by PIMC is -3.7 cal/mol/K which is very close to the experimental value of -3.6 cal/mol/K. By contrast, the harmonic approximation predicted a change of +0.09 kcal/mol/K. Both ethanol and dimethyl ether have torsional angles which lead to multiple minima on their PES. Fig. 2 shows the distribution in C-O torsional angle of dimethyl ether as obtained from our path integral Monte Carlo with 96 beads. The fact that the distribution in the C-O torsion angle has three equal peaks when $\lambda = 0$ shows that the simulation is accessing all areas of phase space without getting stuck in the local minima. The potential energy corresponding to the C-O torsion angle ϕ in dimethyl ether is $V_0 = 1.32[1 - \cos(3\phi)]$. For the reference potential V_1 , this function changes to $V_1 = 200[1 - \cos(\phi)]$. Fig. 2 also shows the distribution in the C-O angle for $\lambda = 1$ as obtained from the simulations. When $\lambda = 0.01$ an intermediate distribution is obtained.

As mentioned previously, the error in determining partition functions using the harmonic approximation increases with increasing temperature. This is because at higher temperatures more vibrational levels are populated and the molecules explore a greater number of configurations far from the equilibrium geometry. Since our simulations contain parallel replicas which were run at higher temperatures it is straight forward to calculate the partition functions at higher temperatures. At 603 K, the free energies of ethanol, obtained using the path integral Monte Carlo and harmonic approximation differ by 7.4 kcal/mol.

C. Glycine conformers

Glycine is the smallest amino acid and it has many stable minima on its Born-Oppenheimer potential energy surface. Miller and Clary⁹ calculated the free energies of these different conformations using torsional path integral Monte Carlo. They separated torsional degrees of freedom from the rest and mapped the PES of glycine in terms of the three torsional angles. Path integral Monte Carlo was done in the reduced space and the contribution from the rest of the degrees of freedom were estimated using the regular harmonic approximation. This approach works well when torsional degrees of freedom are separable from the rest. Our aim was to check whether an all-atom path integral Monte Carlo will converge for such a system with 10 atoms. We consider two of the conformers shown in Fig. 3. Miller and Clary used a single forcefield in terms of torsional angles which is valid for the entire phase space of Glycine. They defined different conformers of glycine based on the value of torsional angles. For example the value of C-C torsion angle is between 100 and 260 degrees for conformer-I. In our calculations we define two different forcefields for these two different conformers. We treat the different conformers as different molecules and optimize

separate forcefields for each conformer. The geometries of conformer-I and conformer-III were optimized using the B3LYP/6-311++G** level of theory and the forcefield terms were optimized to match the Hessian calculated with this level of theory. The forcefield terms are given in the supplementary information.¹⁶ The results are summarized in Table II. These simulations used 100000 blocks to get an accuracy of 0.02 kcal/mol in the reported free energies.

Our calculations which allows all degrees of freedom to fluctuate gives values slightly different from Miller and Clary’s results. The entropy difference between the conformers obtained from our calculations is 3.7 cal/mol/K whereas Miller and Clary reports 1.1 cal/mol/K. A possible explanation is that Miller and Clary used a forcefield with 1729 coefficients which maps out the PES in the torsional space very accurately. The forcefield that we used (see Eq. 5) is an all atom forcefield, which allows bond angles and bond lengths to vary, but it contains many fewer coefficients. A more accurate forcefield expression would allow us to improve upon the current results. However, here we have demonstrated that the absolute free energies of glycine can be obtained with 0.02 kcal/mol statistical accuracy for a given PES.

D. CH₄ activation by palladium catalyst

Palladium dissolved in sulfuric acid is found to be an active catalyst for converting methane to acetic acid and the reaction pathway for this system has been studied recently.²³ Palladium exists in the form of Pd(HSO₄)₂ where the bisulfate anions form bidentate ligands as shown in Fig. 4. One of the principal steps is the methane activation step where a methane molecules loses a proton and forms a Pd-CH₃ bond. This reaction can be written as:



The structure of the resulting species is shown in Fig. 4. The free energy change for this reaction as obtained by harmonic approximation is 15.31 kcal/mol/K. This includes the change in electronic energy which is 4.90 kcal/mol. However a normal mode analysis of Pd(CH₃)(HSO₄)(H₂SO₄) shows that there are eleven vibrational modes with frequencies smaller than 200 cm⁻¹. The harmonic approximation is not valid along these vibrational modes. The entropic contribution to the free energy of Pd(CH₃)(HSO₄)(H₂SO₄) from these low frequency modes (as estimated with the harmonic approximation) is 12.74 kcal/mol. We used our path integral Monte Carlo method to calculate the partition functions and free energies more accurately. The results for Pd(HSO₄)₂ and Pd(CH₃)(HSO₄)(H₂SO₄) are shown in Table III. We also applied the path integral Monte Carlo method to methane. As expected, in the case of methane, the calculated partition functions obtained using the path integral Monte Carlo and the harmonic approxi-

mation agree well with each other. From the table it can be seen that the harmonic approximation overestimates the entropy in case of $\text{Pd}(\text{HSO}_4)_2$ and $\text{Pd}(\text{CH}_3)(\text{HSO}_4)(\text{H}_2\text{SO}_4)$. However these errors cancel out when the free energy change of the reaction is calculated. The overall change in free energy for the reaction is 15.21 kcal/mol determined using the path integral Monte Carlo method (after adding 4.9 kcal/mol electronic contribution). This is very close to the value of 15.31 kcal/mol obtained using the harmonic approximation

IV. CONCLUSIONS

We have demonstrated that, given a potential energy surface, absolute molecular free energies of molecules can be calculated with error bars of less than 0.04 kcal/mol. The method is designed to work with Cartesian coordinates and does not require the determination of Jacobi coordinates or the separation of torsional and non-torsional degrees of freedom. The advantage of using Cartesian coordinates is that it allows the handling of complex molecules (e.g., $\text{Pd}(\text{HSO}_4)_2$), for which it is not clear how to define Jacobi Coordinates.

Previous path integral Monte Carlo calculations by Lynch et al.⁴ have used a free-particles reference state which makes the perturbation between real and reference system quite large. We have used a harmonic potential as the reference state, which makes the real and reference systems quite similar. The major perturbations are only along the low frequency vibrational modes. Since the perturbations are smaller, the path integral Monte Carlo method converges more rapidly, making it possible to treat large systems (in the range of tens of atoms) with reasonable computational effort.

Parallel tempering is shown to provide efficient sampling of the configuration space. Also, the new estimator (Eq. 12) for ratios of partition functions proved to be more accurate than the standard estimator. Although it alleviates the overlap problem, the new estimator does not eliminate it completely. Our only choice to deal with the remaining numerical difficulties is to increase the overlap between the distributions of consecutive replicas, especially around those values of the parallel tempering parameters where a quasi phase-transition happens. If one also takes into account that the number of replicas increases with the dimensionality of the system, one can see that the technique presented becomes rather costly for large dimensional systems. Nevertheless, treating systems of up to 100 atoms appears to be a feasible task on the standard Linux clusters available nowadays to many groups.

In case of ethanol we find that the free energies calculated by the harmonic approximation leads to an error of 2.6 kcal/mol at 300K, and 7.4 kcal/mol at 603 K. The free energies calculated with path integral Monte Carlo correctly predicts the change in entropy for converting ethanol to dimethyl ether. In case of glycine we have demonstrated that

the partition functions can be calculated with an accuracy 0.02 kcal/mol for a given forcefield. However, our results with a simple forcefield which allow all degrees of freedom to fluctuate are different from the results in literature which uses a forcefield in the torsional space. Thus, there is a need for developing analytical potential energy surfaces which accurately describe the Born-Oppenheimer PES of molecules. In case of the methane activation by $\text{Pd}(\text{HSO}_4)_2$, our path integral Monte Carlo simulations lead to a correction of around 1 kcal/mol at 300K in the individual free energies of reactants and products. However these errors cancel out during the evaluation of overall free energy change in the reaction.

Acknowledgments

The authors greatly appreciate their discussions with Arup Chakraborty, who stimulated the initiation of this work and helped set its direction. SC and ATB acknowledges the Methane Conversion Cooperative funded by BP. CP acknowledges support in part from the National Science Foundation Grant No. CHE-0345280 and from the Director, Office of Science, Office of Basic Energy Sciences, Chemical Sciences, Geosciences, and Biosciences Division, U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

APPENDIX A: SHORT DESCRIPTION OF THE PATH INTEGRAL MONTE CARLO TECHNIQUE

In this section, we give a description of those aspects of a path integral Monte Carlo technique that are relevant for the issue of formulating estimators. To make our point, it is enough to consider a general random series representation of the Feynman-Kac formula for the diagonal elements of the density matrix.¹² We utilize a one-dimensional notation because the multi-dimensional generalization is rather trivial to spot. We let $\Omega = \mathbb{R}^{\mathbb{N}}$ be the space of all sequences $\bar{a} := \{a_1, a_2, \dots\}$ and define the infinite-dimensional standard normal distribution

$$dP[\bar{a}] = \prod_{k=1}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-a_k^2/2} da_k. \quad (\text{A1})$$

In other words, under the distribution $dP[\bar{a}]$, the random variables $\bar{a} := \{a_1, a_2, \dots\}$ are independent identically distributed standard normal variables. The Feynman-Kac formula for the diagonal elements of the density matrix reads^{12,24}

$$\begin{aligned} \rho(x; \lambda, \beta) &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\Omega} dP[\bar{a}] \\ &\times \exp \left\{ -\beta \int_0^1 V_{\lambda} \left[x + \sigma \sum_{k=1}^{\infty} a_k \Lambda_k(u) \right] du \right\}. \end{aligned} \quad (\text{A2})$$

For the equality expressed by Eq. (A2) to be true, the functions $\Lambda_k(u)$ must be generated according to the following Ito-Nisio prescription. One considers any arbitrary $L^2([0, 1])$ orthonormal basis $\{\lambda_k(\tau)\}_{k \geq 0}$ such that $\lambda_0(\tau) = 1$. Then the functions defined by

$$\Lambda_k(u) = \int_0^u \lambda_k(\tau) d\tau, \quad \text{for } k \geq 1,$$

are safe to use in the Feynman-Kac formula. The quantity σ is defined by the equation $\sigma = (\hbar^2 \beta / m_0)^{1/2}$, with $\beta = 1/(k_B T)$ being the inverse temperature and m_0 representing the mass of the particle. The interaction is described by a linear combination of “soft” and “hard” potentials

$$V_{\lambda}(x) = (1 - \lambda)V_0(x) + \lambda V_1(x). \quad (\text{A3})$$

For a d -dimensional system, one utilizes an independent random series for each additional degree of freedom. The prefactor $(2\pi\sigma^2)^{-1/2}$ must be replaced by $(2\pi\sigma^2)^{-d/2}$, assuming that all particles have equal masses. The last assumption carries no loss of generalization, provided that mass-scaled coordinates are employed.

The interesting property of the random series representation expressed by Eq. (A2) is that the Gaussian distribution $dP[\bar{a}]$ does not depend upon any of the parameters defining the physical system. As emphasized many times in the context of defining estimators for other thermodynamic quantities,^{12,25} this rescaling of the kinetic operator part of

the Feynman-Kac formula is crucial in obtaining well-behaved estimators that have statistical properties quite similar to their classical analogues. Thus, the ratio $Z(\lambda_j, \beta_j)/Z(\lambda_i, \beta_i)$ can be computed as the Monte Carlo average

$$\frac{\int_{\mathbb{R}} dx \int_{\Omega} dP[\bar{a}] e^{-\beta_i \int_0^1 V_{\lambda_i} [x + \sigma_i \sum_k a_k \Lambda_k(u)] du} r_{ij}(x, \bar{a})}{\int_{\mathbb{R}} dx \int_{\Omega} dP[\bar{a}] e^{-\beta_i \int_0^1 V_{\lambda_i} [x + \sigma_i \sum_k a_k \Lambda_k(u)] du}}, \quad (\text{A4})$$

where

$$r_{ij}(x, \bar{a}) = (\sigma_j / \sigma_i) e^{-\beta_j \int_0^1 V_{\lambda_j} [x + \sigma_j \sum_k a_k \Lambda_k(u)] du} \times e^{\beta_i \int_0^1 V_{\lambda_i} [x + \sigma_i \sum_k a_k \Lambda_k(u)] du} \quad (\text{A5})$$

is an estimating function that mimics almost entirely the statistical properties of the corresponding estimator for classical distributions. The advantage of such estimators consists in the fact that any possible violent fluctuations of the quantities depending on the replica exchange parameters are tempered by the degree to which the potential itself varies.

APPENDIX B: RATIOS OF PARTITION FUNCTIONS BY REPLICA EXCHANGE MONTE CARLO

The technique of replica exchange^{18,19} (or parallel tempering) is a crucial aspect of the present development. For the resulting Monte Carlo code to require minimal input from the user and minimal knowledge of the molecular structure, we do not attempt to isolate the vibrational, rotational, or torsional degrees of freedom or use them as guidance for the Metropolis sampler. Rather, the basic Metropolis sampler with individual particles subjected to random displacements is utilized. Due to possible mismatch between the strength of forces to which individual particles or groups of atoms are subjected, attaining ergodicity during the finite length of the simulation is no trivial task. An example is the ethane molecule, where the two methyl groups rotate with ease, yet are subjected to strong interactions along the vibrational degree of freedom. Treating the two degrees of freedom on an equal footing (by utilizing roughly the same maximal displacements regardless of direction) is, clearly, not an optimal sampling strategy. Indeed, a recent study of Lynch et al.⁴ demonstrates that proper use of Jacobi mass-scaled coordinates improves the quality of the sampling in a non-trivial way.

The parallel tempering technique is capable of compensating for this loss in efficiency by performing exchanges between two statistically independent replicas of slightly different temperatures. The two replicas may differ through other parameters, as for instance the λ parameter utilized in the present study to construct intermediate Hamiltonians. A ladder of temperatures is created, having the form $T_{\min} = T_0 < T_1 < \dots < T_{n-1} = T_{\max}$. By successive

exchanges between neighboring replicas, a configuration may experience any of the temperatures between T_{\min} and T_{\max} . Because the maximal displacements increase with the temperature, the configuration is randomly subjected to maximal displacements of various sizes, without incurring a penalty in the acceptance rate. In turn, this increases the likelihood that, for example, the molecule of ethane rotates with sufficient frequency despite the fact that the maximal displacements for each given temperature are basically controlled by the vibrational degree of freedom. The same exchange mechanism helps explain the higher efficiency of the parallel tempering technique in surmounting potential barriers.¹⁹

Because the basic mechanism through which ratios of partition functions are computed is independent of the actual form of the distribution functions, we shall utilize some generic distributions in the present section, which are denoted by $\rho(\lambda, T, \mathbf{x})$ and are defined on a high-dimensional space \mathbb{R}^d . The distributions may depend upon several parameters that are used to differentiate the various parallel tempering replicas. In our case, the parameters are the temperature T and the scaling factor λ utilized for the definition of intermediate Hamiltonians. The probability distributions are assumed to be integrable for all $T > 0$ and $\lambda \geq 0$. In addition, we assume that, by making either T or λ large enough, the acceptance probabilities for swaps decrease to zero (in other words, the overlap of two distributions can be made arbitrarily small by increasing any of the aforementioned parameters for one of the replicas). We have divided the replicas utilized in the simulation in two groups. The distributions from the first group of replicas differ through the value of the spring constants of the force fields (that is, through λ). However, they have the same temperature $T = T_{\min}$. On the other hand, the replicas in the second group differ through the value of the temperature, but they have the same λ , namely $\lambda_{\min} = 0$, which corresponds to the normal values of the spring constants. A typical plot denoting the replicas in the space (λ, T) is shown in Fig. 5

For definiteness, let

$$\begin{pmatrix} \lambda_q & \lambda_{q-1} & \cdots & \lambda_1 & \lambda_0 & \lambda_0 & \cdots & \lambda_0 \\ T_0 & T_0 & \cdots & T_0 & T_0 & T_1 & \cdots & T_p \end{pmatrix} \quad (\text{B1})$$

be the schedule of parallel tempering parameters. There are $n = p + q + 1$ replicas, which can be uniquely identified according to their position in the list by an index i ranging from $-q$ to p . For ease of notation, we let $\rho_i(\mathbf{x})$ denote the distribution probability $\rho(\lambda, T, \mathbf{x})$ for the parameters (λ, T) corresponding to the index i . In the parallel tempering algorithm, swaps involving two neighboring indexes, say i and j with $j = i \pm 1$, are attempted periodically and accepted with the conditional probability^{18,19}

$$\min \left\{ 1, \frac{\rho_i(\mathbf{x}')\rho_j(\mathbf{x})}{\rho_i(\mathbf{x})\rho_j(\mathbf{x}')} \right\}. \quad (\text{B2})$$

The replica of index $i = 0$ is characterized by the parameters $T = T_0$ and $\lambda = \lambda_0$, parameters that set the temperature and the spring constants of the system of interest. This replica is involved in alternative swaps either with the replica characterized by (λ_1, T_0) or with the replica characterized by (λ_0, T_1) . It serves as a pivot linking the two groups of parameters. Because the replicas with large spring constants are unlikely to equilibrate by themselves during the finite time of the simulation, it is important that the acceptance probabilities for replicas of neighboring parameters are carefully tuned to some optimal value. The initialization and tuning of the schedule of parameters is discussed in the Appendix C. We mention that the parameters for the replica of index $i = 0$ are kept unchanged during the tuning process.

From Eq. (B2), one quickly infers that all quantities necessary for the evaluation of ratios of partition functions are, in fact, already computed for the purpose of implementing the parallel tempering technique. Let

$$Z_i = \int_{\mathbb{R}^d} \rho_i(\mathbf{x}) d\mathbf{x} \quad (\text{B3})$$

denote the partition function and let

$$\langle f(\mathbf{x}) \rangle_i = \frac{\int_{\mathbb{R}^d} \rho_i(\mathbf{x}) f(\mathbf{x}) d\mathbf{x}}{\int_{\mathbb{R}^d} \rho_i(\mathbf{x}) d\mathbf{x}}. \quad (\text{B4})$$

We may compute both

$$\frac{Z_j}{Z_i} = \left\langle \frac{\rho_j(\mathbf{x})}{\rho_i(\mathbf{x})} \right\rangle_i \quad \text{and} \quad \frac{Z_i}{Z_j} = \left\langle \frac{\rho_i(\mathbf{x})}{\rho_j(\mathbf{x})} \right\rangle_j, \quad (\text{B5})$$

by accumulating the respective ratios of probability densities before each swapping attempt involving the replicas i and $j = i \pm 1$. Because the correlation times for realistic simulations are larger than the number of sweeps after which a swap of configurations is attempted, there is little statistical penalty associated with the fact that we collect averages before each swapping attempt, rather than after each complete sweep or individual Monte Carlo step. The more relevant problem is the fact that the ratios described by Eq. (B5) suffer from the so-called overlap problem if the distributions $\rho_i(\mathbf{x})$ and $\rho_j(\mathbf{x})$ are not sufficiently close to each other. That is, due to poor statistics in the region of maximal overlap, the rate at which the estimators attain the normal distribution predicted by the central limit theorem might be very low. As such, the statistical error bars might not be representative of the actual errors.

Let us focus our attention on the computation of the ratio Z_j/Z_i as the average shown by Eq. (B5) and ask the question of whether or not it is possible to construct estimators that are better behaved than the implicit ones. The answer is affirmative and the remainder of the section is concerned with developing an example. We start with some

preparatory observations. Clearly,

$$\frac{Z_j}{Z_i} = \frac{\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \rho_j(\mathbf{x}) \rho_j(\mathbf{x}') d\mathbf{x} d\mathbf{x}'}{\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \rho_i(\mathbf{x}) \rho_j(\mathbf{x}') d\mathbf{x} d\mathbf{x}'}. \quad (\text{B6})$$

Define the set

$$D = \left\{ (\mathbf{x}, \mathbf{x}') \in \mathbb{R}^d \times \mathbb{R}^d : \frac{\rho_j(\mathbf{x})}{\rho_i(\mathbf{x})} \leq \frac{\rho_j(\mathbf{x}')}{\rho_i(\mathbf{x}')} \right\}, \quad (\text{B7})$$

let D^c be its complement, and let $I_D(\mathbf{x}, \mathbf{x}')$ be the indicator function of the set D (the function that takes the value 1 for all the points in D and 0 elsewhere). Notice that

$$I_D(\mathbf{x}', \mathbf{x}) = I_{D^c}(\mathbf{x}, \mathbf{x}'), \quad (\text{B8})$$

perhaps with the exception of the points on the frontier of D , which, nonetheless, carry zero statistical weight.

Now, let $\pi(\mathbf{x}, \mathbf{x}')$ be some integrable and symmetric probability distribution. By interchanging the coordinates \mathbf{x} and \mathbf{x}' and by using the symmetry of $\pi(\mathbf{x}, \mathbf{x}')$ as well as Eq. (B8), we obtain

$$\begin{aligned} & \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \pi(\mathbf{x}, \mathbf{x}') I_D(\mathbf{x}, \mathbf{x}') d\mathbf{x} d\mathbf{x}' \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \pi(\mathbf{x}, \mathbf{x}') I_{D^c}(\mathbf{x}, \mathbf{x}') d\mathbf{x} d\mathbf{x}'. \end{aligned} \quad (\text{B9})$$

On the other hand, since

$$I_D(\mathbf{x}', \mathbf{x}) + I_{D^c}(\mathbf{x}, \mathbf{x}') = 1,$$

it follows that

$$\begin{aligned} & \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \pi(\mathbf{x}, \mathbf{x}') d\mathbf{x} d\mathbf{x}' \\ &= 2 \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \pi(\mathbf{x}, \mathbf{x}') I_D(\mathbf{x}, \mathbf{x}') d\mathbf{x} d\mathbf{x}' \\ &= 2 \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \pi(\mathbf{x}, \mathbf{x}') I_{D^c}(\mathbf{x}, \mathbf{x}') d\mathbf{x} d\mathbf{x}'. \end{aligned} \quad (\text{B10})$$

We are now ready to construct two *uncorrelated* estimating functions for the quantity Z_j/Z_i , to be averaged against the distribution $\rho_i(\mathbf{x})\rho_j(\mathbf{x}')$, as prescribed by Eq. (B6). They are defined by the equations

$$r_{ij}(\mathbf{x}, \mathbf{x}') = 2 \frac{\rho_j(\mathbf{x})}{\rho_i(\mathbf{x})} I_D(\mathbf{x}, \mathbf{x}') \quad (\text{B11})$$

and

$$r_{ij}^c(\mathbf{x}, \mathbf{x}') = 2 \frac{\rho_j(\mathbf{x})}{\rho_i(\mathbf{x})} I_{D^c}(\mathbf{x}, \mathbf{x}'), \quad (\text{B12})$$

respectively. Their correctness follows from the symmetry of the product $\rho_j(\mathbf{x})\rho_j(\mathbf{x}')$ together with Eqs. (B10) and (B6). Their lack of correlation follows from the trivial equality $I_D(\mathbf{x}, \mathbf{x}') \cdot I_{D^c}(\mathbf{x}, \mathbf{x}') = 0$.

The first estimating function has a smaller variance than the second one. Since the expected values of the estimators are the same, it is enough to prove the inequality for their second order moments. We have

$$\begin{aligned} & \frac{\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \rho_i(\mathbf{x})\rho_j(\mathbf{x}')r_{ij}(\mathbf{x}, \mathbf{x}')^2 d\mathbf{x}d\mathbf{x}'}{\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \rho_i(\mathbf{x})\rho_j(\mathbf{x}')d\mathbf{x}d\mathbf{x}'} \\ &= 4 \frac{\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \rho_j(\mathbf{x})\rho_j(\mathbf{x}')[\rho_j(\mathbf{x})/\rho_i(\mathbf{x})]I_D(\mathbf{x}, \mathbf{x}')d\mathbf{x}d\mathbf{x}'}{\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \rho_i(\mathbf{x})\rho_j(\mathbf{x}')d\mathbf{x}d\mathbf{x}'} \\ &\leq 4 \frac{\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \rho_j(\mathbf{x})\rho_j(\mathbf{x}')[\rho_j(\mathbf{x}')/\rho_i(\mathbf{x}')]I_D(\mathbf{x}, \mathbf{x}')d\mathbf{x}d\mathbf{x}'}{\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \rho_i(\mathbf{x})\rho_j(\mathbf{x}')d\mathbf{x}d\mathbf{x}'}, \end{aligned}$$

where we have utilized that $\rho_j(\mathbf{x})/\rho_i(\mathbf{x}) \leq \rho_j(\mathbf{x}')/\rho_i(\mathbf{x}')$ on the set D . However, again by interchanging the variables \mathbf{x} and \mathbf{x}' and utilizing Eq. (B8), we obtain that the last term in the preceding equation is

$$\begin{aligned} & 4 \frac{\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \rho_j(\mathbf{x})\rho_j(\mathbf{x}')[\rho_j(\mathbf{x})/\rho_i(\mathbf{x})]I_{D^c}(\mathbf{x}, \mathbf{x}')d\mathbf{x}d\mathbf{x}'}{\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \rho_i(\mathbf{x})\rho_j(\mathbf{x}')d\mathbf{x}d\mathbf{x}'} \\ &= \frac{\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \rho_i(\mathbf{x})\rho_j(\mathbf{x}')r_{ij}^c(\mathbf{x}, \mathbf{x}')^2 d\mathbf{x}d\mathbf{x}'}{\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \rho_i(\mathbf{x})\rho_j(\mathbf{x}')d\mathbf{x}d\mathbf{x}'}, \end{aligned}$$

thus proving our claim.

Certain numerical issues appear if the frontier of the set D does not have zero statistical weight (for example, if the distributions are equal). Taking into consideration this issue as well, one ends up with the estimator given by Eq. (12) in the text, where a parameter $\epsilon > 0$ is introduced. We leave it for the reader to show that the estimator is correct for any value of $\epsilon > 0$.

To summarize in our simpler picture, a replica exchange simulation is conducted. Before any swapping attempt involving replicas i and $j = i \pm 1$, one verifies whether or not

$$\frac{\rho_j(\mathbf{x})}{\rho_i(\mathbf{x})} \leq \frac{\rho_j(\mathbf{x}')}{\rho_i(\mathbf{x}')}.$$

If the answer is positive, one accumulates the quantity $2\rho_j(\mathbf{x})/\rho_i(\mathbf{x})$, which constitutes an estimator, to be denoted by $\overline{Z_j/Z_i}$, for the ratio Z_j/Z_i . If the answer is negative, one accumulates $2\rho_i(\mathbf{x}')/\rho_j(\mathbf{x}')$, which constitutes the estimator $\overline{Z_i/Z_j}$ for the inverse ratio Z_i/Z_j . For each ratio, one also computes the relative statistical error bars, defined as twice the standard deviation divided by the value of the ratio. The ratio between the partition functions for the distributions characterized by the parameters (λ_q, T_0) and (λ_0, T_0) , corresponding to indexes $i = -q$ and $i = 0$ respectively, is evaluated as the product

$$\overline{Z_0/Z_{-q}} = \prod_{i=-q}^1 \overline{Z_{i+1}/Z_i}, \quad (\text{B13})$$

whereas the inverse ratio is given by

$$\overline{Z_{-q}/Z_0} = \prod_{i=-q}^1 \overline{Z_i/Z_{i+1}}. \quad (\text{B14})$$

Because the ratios $\overline{Z_{i+1}/Z_i}$ for $i = -q, -q + 1, \dots, 1$ are statistically independent, the relative error bars for $\overline{Z_0/Z_{-q}}$ are obtained by computing the square root of the sum of the square of the individual relative error bars. The relative error bars for the inverse ratio $\overline{Z_{-q}/Z_0}$ are evaluated in a similar fashion. These relative errors become absolute errors for the free energies, if divided by $\beta_0 = 1/(k_B T_0)$.

The agreement between $\overline{Z_0/Z_{-q}}$ and the inverse of $\overline{Z_{-q}/Z_0}$ within the computed relative error bars is a measure of the degree to which the overlap problem has been overcome. In principle, there are two ways of converging the simulation if agreement is not obtained. The first one is the brute-force approach, whereby the simulation is run until agreement is obtained. The second one is to insert more replicas between the end points $i = -q$ and $i = 0$. Research of Lu and Kofke²⁶ has demonstrated that increasing the number of intermediate states between the two end distributions, so that to improve the overlap between successive states, is the computationally more efficient approach. We have utilized this second technique and have increased the acceptance probability for swaps between the replicas $i = -q, -q + 1, \dots, 0$ until a good agreement between the direct and the inverse estimators has been obtained. We notice that, although the acceptance probability for swaps is also a measure of the degree of overlap between two distributions, in general, ensuring a constant acceptance probability does not automatically ensure uniform quality for partition functions. An adaptive algorithm that attempts to improve the overlap with a minimalist number of intermediate replicas can, in principle, be designed. As such, one could verify the values of $\overline{Z_{i+1}/Z_i}$ and the inverse of $\overline{Z_{i+1}/Z_i}$ for each pairs of replicas involved in direct exchanges and adjust the replica exchange parameters until agreement is obtained.

APPENDIX C: TUNING THE SCHEDULE OF REPLICA EXCHANGE PARAMETERS

An important problem in the implementation of the replica exchange technique is the computation of a schedule of parameters that ensures optimal efficiency for swaps. Thus, the parameters have to be modified so that the acceptance probabilities lie in a tight interval about some value that optimizes the number of exchanged replicas per unit of computational effort. This optimal value depends on many factors related to the specifics of the simulation. Nevertheless, a recent study²⁷ regarding the efficiency of exchange for parallel tempering in a classical canonical ensemble has revealed that the optimal acceptance probability is about 39% for many systems with continuous distributions.

It has been pointed out that the efficiency of swaps decreases quite slowly with the acceptance probability. More precisely, the efficiency remains within one half of the optimal value for all acceptance probabilities situated between 7% and 82%. It appears then reasonable to infer that values of the acceptance probabilities between 35% and 45% will maintain a high efficiency of exchange even for quantum simulations. In the present Appendix, we shall describe a simple algorithm for tuning the replica exchange parameters so that the acceptance probabilities lie in the desired interval $[0.35, 0.45]$.

For the sake of an example, assume that the tuning parameter is the temperature T and that n replicas are utilized. The n replicas are periodically involved in alternative swaps with the replicas of higher and lower temperatures, respectively. As an initial guess, one starts the simulation with the temperatures arranged in geometric progression. Thus, for $k = 0, 1, \dots, n - 1$, the initial temperatures are

$$T_k = T_{\min} \prod_{i=0}^k R_i, \quad \text{with} \quad R_i = \sqrt[n-1]{T_{\max}/T_{\min}}. \quad (\text{C1})$$

We shall assume that the minimal temperature T_{\min} is a temperature “of interest” and must be kept fixed. However, T_{\max} is allowed to change. Also, we shall assume that the equilibration part of the Monte Carlo simulation is divided in, say, 50 blocks and that acceptance probabilities are computed for each individual block and each pair of replicas involved in direct exchanges.

At the following step, one verifies whether or not the acceptance probabilities lie in the desired interval $[0.35, 0.45]$. Let Ac_i stand for the acceptance probability between replicas i and $i + 1$. If $Ac_i > 0.45$, then the ratio R_i is increased to the new value R'_i , according to the law

$$R'_i = 1 + (R_i - 1)\alpha_i, \quad (\text{C2})$$

where $\alpha_i > 1$ is some tuning factor. If $Ac_i < 0.35$, then R_i is decreased according to the equation

$$R'_i = 1 + (R_i - 1)/\alpha_i. \quad (\text{C3})$$

Once the new temperature ratios $R_i = T_{i+1}/T_i$ are determined, one may compute the temperature for the replica k by utilizing the fact that T_{\min} is kept fixed. We thus have

$$T_k = T_{\min} \prod_{i=0}^k R_i, \quad \text{for } k = 0, 1, \dots, n - 1. \quad (\text{C4})$$

The procedure is then iterated until the schedule of temperatures no longer changes or the number of equilibration blocks is exhausted.

The parameters $\alpha_i > 1$ should be chosen low enough that the acceptance probabilities do not jump from values smaller than 35% to ones larger than 45%, or conversely. Provided that such oscillations do not occur, the values α_i must be kept as large as possible, to speed up the optimization of the schedule. Nevertheless, the maximal values of α_i get closer to 1 for larger dimensional systems, as can be deduced from the incomplete beta function law.^{28–30} One way to cope with the problem of having values of α_i that are too large or too small is to let the α_i 's be random variables uniformly distributed in the interval $(1, \alpha_{\max})$. That is, if ξ_i is a random number in the interval $(0, 1)$, then

$$\alpha_i = 1 + (\alpha_{\max} - 1)\xi_i. \quad (\text{C5})$$

In the present paper, we have utilized this first approach, with $\alpha_{\max} = 1.15$. Another way is to decrease the difference $\alpha_i - 1$ by another fixed factor $f > 1$, whenever oscillations outside the interval $[0.35, 0.45]$ are noticed. The first approach seems slightly less efficient, but it is easier to implement. Both algorithms are always convergent. Indeed, after several steps, the value of T_1 remains unchanged because T_0 is kept fixed. Once T_1 is fixed, again after several steps, the value of T_2 is determined and then kept unchanged. Provided that the number of equilibration blocks is large enough, the whole list of temperatures will eventually be exhausted. We see that, in the worst case scenario, the optimization of the schedule will be finished in a time roughly proportional to the number of replicas. In actual applications, the algorithm is faster.

For those replicas that are characterized by the same temperature $T = T_{\min}$ but by different Hamiltonians according to the different values λ , the same procedure is applied. Here, the initial guess is an equally spaced grid of mesh $\Delta_i = (\lambda_{\max} - \lambda_{\min})/(n - 1)$, with $\lambda_{\min} = 0$ and $\lambda_{\max} = 1$. Thus, we start with the guess

$$\lambda_k = \lambda_{\min} + \sum_{i=0}^k \Delta_i. \quad (\text{C6})$$

Then, the values of the displacements Δ_i are decreased or increased by some random factors $\alpha_i > 1$ between 1 and $\alpha_{\max} = 1.15$, according to whether the acceptance probabilities A_{c_i} are greater than 0.45 or smaller than 0.35. The remainder of the optimization follows closely the discussion for the case where the parallel tempering parameter is the temperature. We point out that the construction of the intermediate Hamiltonians must be performed in such a way that the system is integrable for all $\lambda \geq 0$. Also, if at the end of the simulation some acceptance probabilities have values larger than 45%, nothing “wrong” happens with the quality of the simulation (which is actually improved, provided that the largest temperature and value of λ are still sufficiently large). It is only the efficiency that suffers. The more worrisome scenario with respect to the quality of the simulation is when some acceptance probabilities become less than 35%. To prevent this, the values T_{\max} and λ_{\max} should be chosen small enough that the initial

acceptance probabilities are larger than the optimal values.

The stability of the algorithm also depends upon the precision with which the acceptance probabilities for swaps are computed. To minimize the statistical errors, it is worth computing the acceptance probabilities with the help of the formula

$$Ac_i = \langle ac_i \rangle \equiv \frac{\int d\mathbf{x} \int d\mathbf{x}' \rho_i(\mathbf{x}) \rho_{i+1}(\mathbf{x}') ac_i(\mathbf{x}, \mathbf{x}')}{\int d\mathbf{x} \int d\mathbf{x}' \rho_i(\mathbf{x}) \rho_{i+1}(\mathbf{x}')}, \quad (\text{C7})$$

where

$$ac_i(\mathbf{x}, \mathbf{x}') = \min \left\{ 1, \frac{\rho_i(\mathbf{x}') \rho_{i+1}(\mathbf{x})}{\rho_i(\mathbf{x}) \rho_{i+1}(\mathbf{x}')} \right\}$$

is the conditional acceptance probability. In other words, before each swapping attempt, we accumulate the averages of the conditional acceptance probability, a quantity that is computed anyway for the purpose of acceptance/rejection testing. Due to detailed balance, Ac_i can also be computed as the ratio between the numbers of accepted and attempted swaps, but the resulting estimator has a larger variance. Indeed, in the second case, we accumulate 1, if the swap is accepted, or 0, if it is not. By the idempotence of the estimator, the variance is $Ac_i - Ac_i^2$. On the other hand, the variance for the estimator appearing in Eq. (C7) is $\langle ac_i^2 \rangle - Ac_i^2$. The inequality $ac_i(\mathbf{x}, \mathbf{x}') \leq 1$ implies $ac_i(\mathbf{x}, \mathbf{x}')^2 \leq ac_i(\mathbf{x}, \mathbf{x}')$. Therefore, $\langle ac_i^2 \rangle \leq \langle ac_i \rangle = Ac_i$ and the statistical variance of the estimator $ac_i(\mathbf{x}, \mathbf{x}')$ is smaller.

-
- * URL: <http://zeolites.cqe.northwestern.edu/shaji>
- † Electronic address: bell@ccchem.berkeley.edu
- ‡ Electronic address: cpredescu@comcast.net
- ¹ C. E. Chang, M. J. Potter, and M. K. Gilson, *J. Phys. Chem. B* **107**, 1048 (2003).
- ² I. Kolossvary, *J. Phys. Chem. A* **101**, 9900 (1997).
- ³ K. Glaesemann and L. Fried, *J. Chem. Phys.* **123**, 034103 (2005).
- ⁴ V. A. Lynch, S. L. Mielke, and D. G. Truhlar, *J. Chem. Phys.* **121**, 5148 (2004).
- ⁵ S. L. Mielke and D. G. Truhlar, *Chem. Phys. Lett.* **378**, 317 (2003).
- ⁶ V. A. Lynch, S. L. Mielke, and D. G. Truhlar, *J. Phys. Chem. A* **109**, 10092 (2005).
- ⁷ J. Koput, S. Carter, and N. C. Handy, *J. Phys. Chem. A* **102**, 6325 (1998).
- ⁸ T. F. Miller and D. C. Clary, *J. Chem. Phys.* **119**, 68 (2003).
- ⁹ T. F. Miller and D. C. Clary, *Phys. Chem. Chem. Phys.* **6**, 2563 (2004).
- ¹⁰ C. Predescu, *J. Phys. Chem. B* **110**, 667 (2006).
- ¹¹ C. Predescu, *Phys. Rev. E* **71**, 10092 (2005).
- ¹² C. Predescu and J. D. Doll, *J. Chem Phys.* **117**, 7448 (2002).
- ¹³ C. Predescu and J. D. Doll, *Phys. Rev. E* **67**, 4119 (2003).
- ¹⁴ C. Predescu, *Phys. Rev. E* **69**, 5148 (2004).
- ¹⁵ <http://zeolites.cqe.northwestern.edu/shaji>.
- ¹⁶ See EPAPS Document No. — for all the forcefield parameters used here. This document can be reached through a direct link in the online article’s html reference section or via the EPAPS homepage (<http://www.aip.org/pubservs/epaps.html>).
- ¹⁷ M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, *Gaussian 03*, Revision C.02, Gaussian, Inc., Wallingford, CT, 2004.

- ¹⁸ C. J. Geyer, in E. M. Keramigas, editor, *Computing Science and Statistics* (Interface foundation, Fairfax, 1991), Proceedings of the 23rd Symposium on the Interface, 156–163.
- ¹⁹ K. Hukushima and K. Nemoto, *J. Phys. Soc. Japan* **65**, 1604 (1996).
- ²⁰ 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-Garca, , and J. C. Corchado, Potlib-online, <http://comp.chem.umn.edu/potlib>.
- ²¹ D. R. Lide, editor, *CRC Handbook of Chemistry and Physics* (CRC Press, 2005), chapter 5.
- ²² L. A. Curtiss, K. Raghavachari, P. C. Redfern, and J. A. Pople, *J. Chem. Phys.* **112**, 7374 (2000).
- ²³ S. Chempath and A. T. Bell, *J. Am. Chem. Soc.* (In Press).
- ²⁴ D. L. Freeman and J. D. Doll, *J. Chem. Phys.* **80**, 5709 (1984).
- ²⁵ C. Predescu, *J. Chem. Phys.* **123**, 217102 (2005).
- ²⁶ N. D. Lu and D. A. Kofke, *J. Chem. Phys.* **114**, 7303 (2001).
- ²⁷ C. Predescu, M. Predescu, and C. V. Ciobanu, *J. Phys. Chem. B* **109**, 4189 (2005).
- ²⁸ C. Predescu, M. Predescu, and C. V. Ciobanu, *J. Chem. Phys.* **120**, 4119 (2004).
- ²⁹ D. A. Kofke, *J. Chem. Phys.* **117**, 6911 (2002).
- ³⁰ D. A. Kofke, *J. Chem. Phys.* **121**, 1167 (2004).

TABLE I: Comparison of thermochemical properties calculated from path integral Monte Carlo(PIMC) and harmonic approximation (HAR). The partition function Z includes the rotational and translational contributions calculated using rigid rotor approximation. The standard free energy G and the entropy S are evaluated at 300 K

dimethyl ether		
	HAR	PIMC
$\ln(Z)$	-56.57	-54.18
G_{300}^0 , kcal/mol	33.72	32.30
S_{300}^0 , cal/mol/K	62.59	67.33
ethanol		
	HAR	PIMC
$\ln(Z)$	-56.90	-52.60
G_{300}^0 , kcal/mol	33.92	31.36
S_{300}^0 , cal/mol/K	62.50	71.03

TABLE II: Comparison of thermochemical properties of glycine conformers calculated from path integral Monte Carlo (PIMC) and harmonic approximation (HAR). The partition function Z includes the rotational and translational contributions calculated using rigid rotor approximation. The standard free energy G and the entropy S are evaluated at 300 K

Conformer-I		
	HAR	PIMC
$\ln(Z)$	-51.05	-50.24
G_{300}^0 , kcal/mol	30.43	29.95
S_{300}^0 , cal/mol/K	76.22	77.83

Conformer-III		
	HAR	PIMC
$\ln(Z)$	-50.59	-49.02
G_{300}^0 , kcal/mol	30.16	29.22
S_{300}^0 , cal/mol/K	78.34	81.45

TABLE III: Thermochemical properties of the Pd/methane system calculated from path integral Monte Carlo (PIMC) and harmonic approximation (HAR). The partition function Z includes the rotational and translational contributions calculated using rigid rotor approximation. The standard free energy G and the entropy S are evaluated at 300 K

<hr/> <hr/> Pd(HSO ₄) ₂ <hr/>		
	HAR	PIMC
$\ln(Z)$	-10.87	-12.48
G_{300}^0 , kcal/mol	6.48	7.44
S_{300}^0 , cal/mol/K	124.67	121.46
<hr/> <hr/>		
<hr/> CH ₄ <hr/>		
	HAR	PIMC
$\ln(Z)$	-26.89	-26.83
G_{300}^0 , kcal/mol	16.03	15.99
S_{300}^0 , cal/mol/K	49.18	49.31
<hr/> <hr/>		
<hr/> Pd(CH ₃)(HSO ₄)(H ₂ SO ₄) <hr/>		
	HAR	PIMC
$\ln(Z)$	-55.22	-56.60
G_{300}^0 , kcal/mol	32.92	33.74
S_{300}^0 , cal/mol/K	143.27	140.53

FIGURE CAPTIONS

FIG.1 Ratio of partition functions plotted against N , the number of beads used in the path integral Monte Carlo for harmonic oscillator at 300K.

FIG.2 Normalized probability distribution of the C-O torsion angle of dimethyl ether. $\lambda = 0$ corresponds to the real potential, $\lambda = 1$ corresponds to the reference potential, and $\lambda = 0.01$ corresponds to an intermediate potential. y-error bars for the histogram values are less than 0.0004 and, hence, not shown here.

FIG.3 (Color online) Two conformers of glycine which differ by a 180 degree rotation around the C-C bond. Color Key: Nitrogen(blue), Carbon(gray), Oxygen(red), Hydrogen(white).

FIG.4 (Color online) Structures of the catalyst and the activated methane species. Color Key: Sulfur(yellow), Carbon(gray), Oxygen(red), Hydrogen(white) Palladium(blue)

FIG.5 Typical schedule of parallel tempering parameters in the (λ, T) space. The replica (λ_0, T_0) corresponds to the physical system of interest. The double-headed arrows denote exchange of configurations between the replicas they connect.

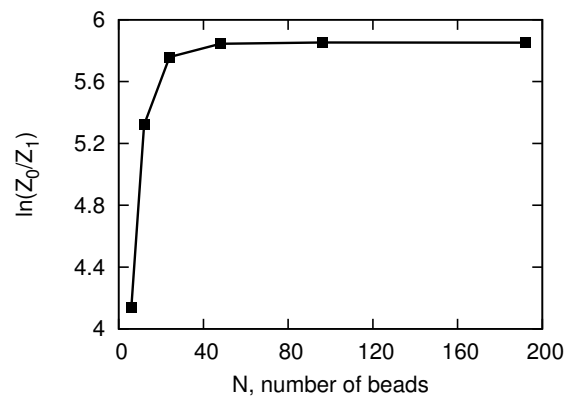


FIG. 1: Ratio of partition functions plotted against N , the number of beads used in the path integral Monte Carlo for harmonic oscillator at 300K.

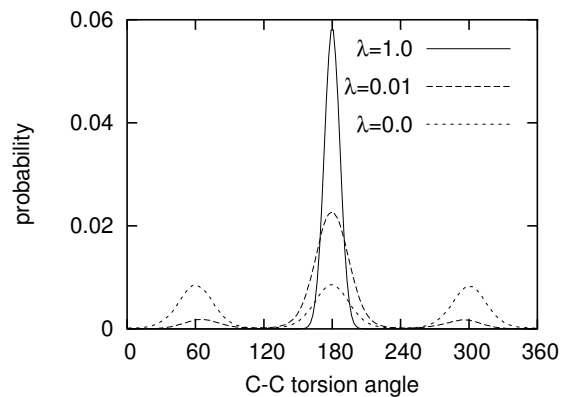


FIG. 2: Normalized probability distribution of the C-O torsion angle of dimethyl ether. $\lambda = 0$ corresponds to the real potential, $\lambda = 1$ corresponds to the reference potential, and $\lambda = 0.01$ corresponds to an intermediate potential. y-error bars for the histogram values are less than 0.0004 and, hence, not shown here.

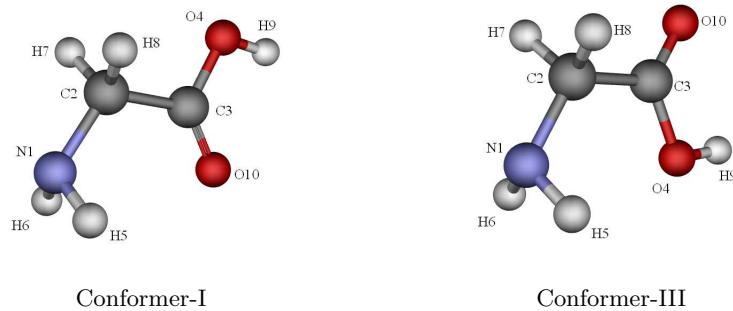


FIG. 3: (Color online) Two conformers of glycine which differ by a 180 degree rotation around the C-C bond. Color Key: Nitrogen(blue), Carbon(gray), Oxygen(red), Hydrogen(white).

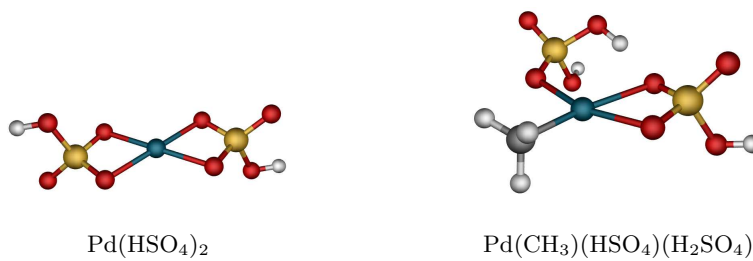


FIG. 4: (Color online) Structures of the catalyst and the activated methane species. Color Key: Sulfur(yellow), Carbon(gray), Oxygen(red), Hydrogen(white) Palladium(blue)

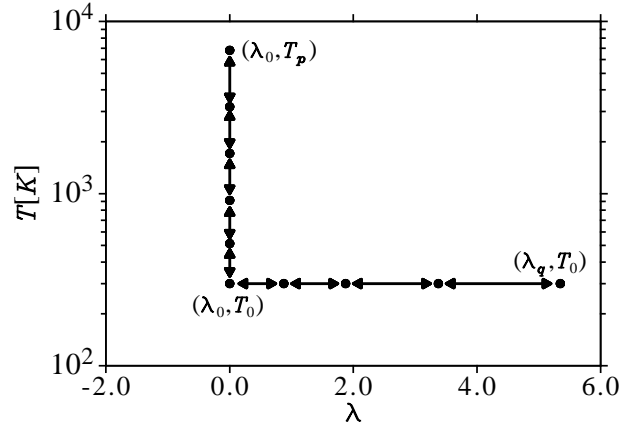


FIG. 5: Typical schedule of parallel tempering parameters in the (λ, T) space. The replica (λ_0, T_0) corresponds to the physical system of interest. The double-headed arrows denote exchange of configurations between the replicas they connect.