Approximate inclusion of quantum effects in transition path sampling

Dimitri Antoniou[1] and Steven D. Schwartz[1,2,3,a)]
[1]*Department of Biophysics, Albert Einstein College of Medicine, 1300 Morris Park Ave., Bronx, New York 10461, USA*
[2]*Department of Biochemistry, Albert Einstein College of Medicine, 1300 Morris Park Ave., Bronx, New York 10461, USA*
[3]*Institut des Hautes Études Scientifiques, 91440 Bures-sur-Yvette, France*

(Received 3 September 2009; accepted 19 November 2009; published online 11 December 2009)

We propose a method for incorporating nuclear quantum effects in transition path sampling studies of systems that consist of a few degrees of freedom that must be treated quantum mechanically, while the rest are classical-like. We used the normal mode centroid method to describe the quantum subsystem, which is a method that is not CPU intensive but still reasonably accurate. We applied this mixed centroid/classical transition path sampling method to a model system that has nontrivial quantum behavior, and showed that it can capture the correct quantum dynamical features.
 [doi:10.1063/1.3272793]

I. INTRODUCTION

The identification of the reaction path in complex systems remains a challenge in theoretical chemistry. The only currently available methods for studying complex potential surfaces without *a priori* assumptions about the reaction coordinate are the transition path sampling (TPS)[1,2] and the string method.[3,4] We have recently used TPS in the analysis of two enzymatic systems,[5–7] for which there are strong indications that quantum effects are significant.

Because it is not possible to numerically study the quantum properties of systems, like proteins, that consist of thousands of degrees of freedom, approximate methods have been developed that incorporate quantum effects. Especially successful have been the potential energy quantum mechanical/molecular mechanical (QM/MM) models, which divide the system into a small quantum region (typically it includes atoms that directly affect the chemical step), and the rest of the system which is treated classically.

Quantum effects enter molecular dynamics (MD) simulations in two ways. First, the potential that governs the motions in the quantum region includes quantum contributions from their electronic structure. These contributions are calculated using semiempirical molecular orbital theory, density functional theory, or *ab initio* methods. A significant advantage of these methods is that in the QM/MM scheme, these quantum contributions to the potential enter only as a modification of the force field that governs the dynamics of the MD simulation. For example, this is the implementation that is used in simulation packages such as CHARMM and AMBER.

The second way that quantum effects enter a MD simulation is through the dynamics of the atoms of the quantum region. Several works have included such effects at various levels of theory. Hammes-Schiffer[8] uses a quantum/classical MD scheme and Truhlar and co-worker[9] uses a variational Transition State Theory/semiclassical multidimensional tunneling model. Also, centroid path integral methods have been used in calculating quantum corrections of the classical free energy.[10] Finally, centroid molecular dynamics (CMD) has been used to calculate the nuclear quantum effects in a proton transfer enzymatic reaction.[11]

In this paper we propose a simple method to incorporate into TPS the second type of quantum effects mentioned above, i.e., nuclear quantum effects. A TPS simulation using a package like CHARMM or AMBER can already include the first type (electronic structure contributions to the potential). Since TPS simulations of enzymatic systems are extremely CPU intensive, a requirement for including nuclear quantum effects is that the method should not have an excessive CPU overhead, but at the same time it should be relatively accurate. A method that satisfies these requirements is the centroid normal mode method, which is an established method in quantum simulations.[11–13]

In the TPS simulations reported in this paper, we use ordinary classical MD for the classical degrees of freedom and the normal mode centroid method for the quantum ones. This is similar to the work of Reichman,[14] who applied the nudged elastic band method to mixed quantum/classical systems. We will analyze a system that has nontrivial quantum behavior and we will examine whether a TPS method that uses centroids for the quantum subsystem is able to capture the correct qualitative quantum dynamical behavior.

II. TRANSITION PATH SAMPLING

TPS can simulate rare events without knowledge of a reaction coordinate or the transition state by performing a Monte Carlo search in the trajectory space. A brief description of the TPS algorithm follows.

Let us assume a transition between stable basins R and P (i.e., reactants and products). Since R and P are long-lived states, they can accommodate equilibrium fluctuations, and can be characterized by an "order parameter," which can be used as a criterion for deciding whether the system is localized in R or P. Let us further assume that we somehow know

a)Electronic mail: sschwartz@aecom.yu.edu.

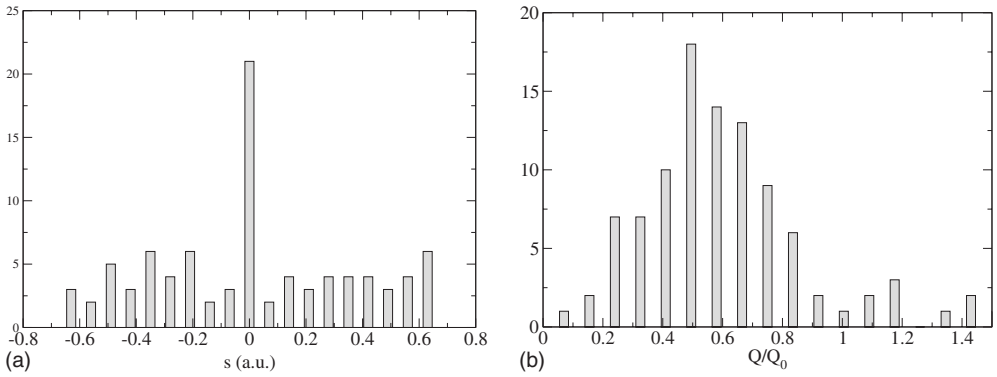

FIG. 1. Histograms for H tunneling, slow Q oscillation. Left, the histogram for s is peaked at the top of the symmetric double well. Right, the histogram for Q/Q_0 (Q_0 is the location of the classical transition state, so a classical-like dynamics corresponds to a histogram peaked at (1). The system corner cuts before it reaches the classical transition state.

one reactive trajectory that starts from R and ends in P. In the TPS algorithm, one randomly selects a time slice along this reactive trajectory, perturbs the momenta slightly at that time slice, and starting from that time slice and using the new momenta, the system is propagated forward and backward in time. The trajectory is then examined to determine whether the new trajectory is reactive or not. In usual Monte Carlo fashion, the new trajectory is accepted or rejected according to some probability distribution. Because of the ergodicity of classical dynamics of large systems, new trajectories are expected to quickly deviate from old ones, leading to a fast sampling of the trajectory space. Additionally, since the acceptance probability for trajectories is generated from a condition of detailed balance, if the initial distribution is stationary (in our case Boltzmann), then the resulting distribution of trajectories maintains the same stationary distribution.[15–17]

Once an ensemble of reactive trajectories has been generated, one can analyze it[2] to find the "separatrix," which is a surface with the property that trajectories that start from it with velocities that are assigned from a Boltzmann distribution have probability of 0.5 to end up in reactants. The reaction coordinate consists of the degrees of freedom that vary along a direction perpendicular to the separatrix. We briefly describe the steps of a typical analysis of the ensemble of reactive trajectories. (The results of this analysis for the system studied in this paper are shown in Figs. 1–3.) First, many new trajectories have to be shot from several slices of each reactive trajectory until one obtains structures with reactant committing probability of 0.5; second, to find the multidimensional reaction coordinate one has to guess for reaction coordinates and keep them fixed while performing a MD walk along the hypersurface defined by the remaining coordinates, then from structures along this MD walk, one has to shoot many trajectories and calculate the commitment probabilities; finally this procedure is iterated until one finds variables that when kept fixed, structures along the MD walk have commitment probability equal to 0.5.

III. CENTROID MOLECULAR DYNAMICS

The centroid concept was originally introduced in the path integral formulation of equilibrium quantum statistical mechanics. It is a mapping of a quantum particle to a polymer chain which has classical-like equations of motion, thus being appropriate as a starting point for semiclassical approximations. The centroid method was further developed in a series of papers by Cao and Voth.[18] They developed[19] an approximate method for computing the dynamics of quantum systems, called CMD. A recent alternative method that is related to CMD is the ring polymer MD.[20] We give a brief summary of CMD.

In CMD the motion of the centroid of the ith quantum particle is described by

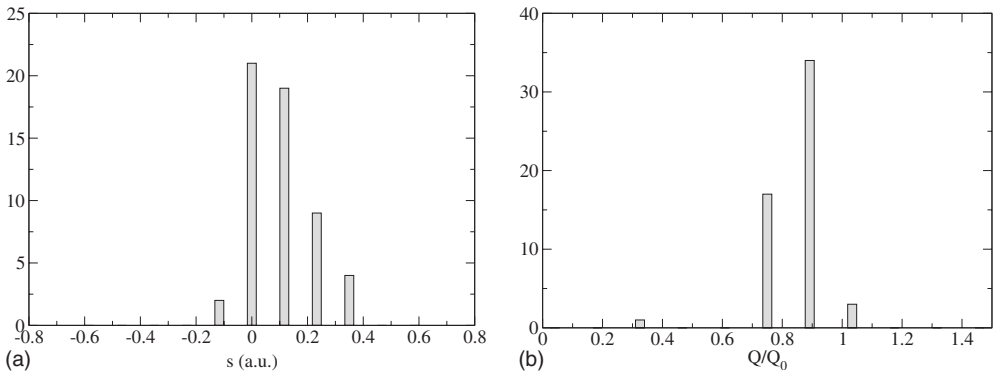

FIG. 2. Histograms for D tunneling, slow Q oscillation. The system corner cuts before it reaches the classical transition state, but it passes closer to it compared with the H tunneling case.

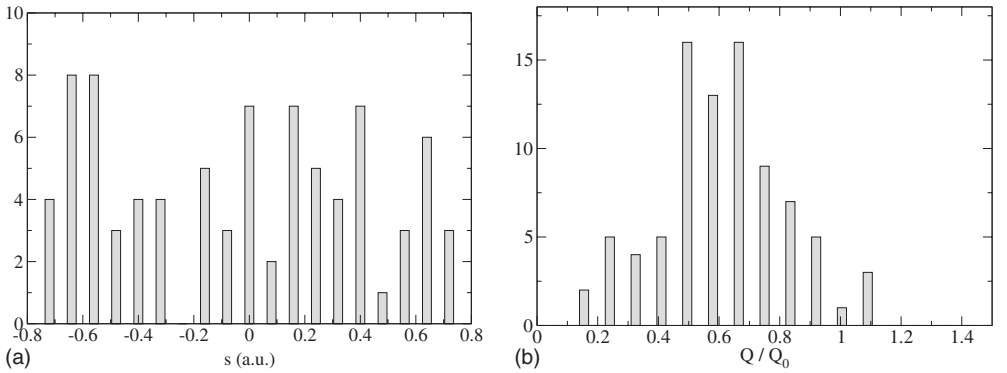

FIG. 3. H tunneling, fast Q oscillation. The period of the Q oscillation is four times slower than the time it takes for the s centroid to tunnel. As a result, many barrier heights are visible to the system for a value of s.

$$m_i\ddot{\boldsymbol{R}}_i = \langle \boldsymbol{F}_i(\boldsymbol{R}_1, \ldots, \boldsymbol{R}_N)\rangle_c, \tag{1}$$

where m_i is the mass of the particle and $\boldsymbol{R}_i$ is its centroid position, and the subscript c signifies a path integral average. The centroid variable is defined as the center of mass of a polymer chain of B beads (in the following, Latin subscripts refer to particles and Greek subscripts refer to beads),

$$\boldsymbol{R}_i^c = \frac{1}{B}\sum_{\beta=1}^{B} \boldsymbol{r}_i^\beta, \tag{2}$$

where each bead has mass m_i/B, and the centroid force $\boldsymbol{F}_i$ is the sum of forces on all beads,

$$\boldsymbol{F}_i(\boldsymbol{R}_1^c, \ldots, \boldsymbol{R}_N^c) = -\frac{1}{B}\sum_{\beta=1}^{B} \frac{\partial V(\boldsymbol{r}_i^\beta)}{\partial \boldsymbol{r}_i^\beta}. \tag{3}$$

The right-hand side of Eq. (1) is the path integral average of the centroid force $\boldsymbol{F}_i$, weighted by $\exp(-\beta V_{\text{eff}})$, where V_{eff} is the discretized potential of the beads,

$$V_{\text{eff}}(\boldsymbol{r}_i^\beta) = \sum_i \sum_\beta \left[\frac{1}{2}k_i(\boldsymbol{r}_i^\beta - \boldsymbol{r}_i^{\beta+1})^2 + \frac{1}{B}V(\boldsymbol{r}_i^\beta)\right], \tag{4}$$

with the spring constants between neighboring beads equal to $k_i = m_i B(k_B T)^2$.

To propagate, according to Eq. (1), we need to average the forces with respect to the bead coordinates. One way to accomplish this efficiently is with the normal mode transformation method.

A. Normal mode transformation

The advantage of the normal mode transformation[19] is that it separates the slower centroid motion from the faster bead motions, which allows a significant simplification in the calculation of the centroid force.

To obtain the normal mode coordinates $\boldsymbol{q}$, we make a unitary transformation of the bead coordinates $\boldsymbol{r}$,

$$\boldsymbol{q}_\alpha^i = \frac{1}{\sqrt{B}}\sum_{\beta=1}^{B} U_{\alpha\beta}\boldsymbol{r}_i^\beta \quad \boldsymbol{r}_i^\beta = \sqrt{B}\sum_{\alpha=1}^{B} U_{\beta\alpha}\boldsymbol{q}_\alpha^i,$$
$$U_{\alpha\beta} = \frac{1}{B}\left(\cos\frac{2\pi\beta\alpha}{B} - \sin\frac{2\pi\beta\alpha}{B}\right). \tag{5}$$

The unitary transformation U approximately diagonalizes the effective bead problem V_{eff} with eigenvalues

$$\omega_\alpha^i = \sqrt{\frac{2k_i}{m_\alpha^i}\left(1 - \cos\frac{2\pi\alpha}{B}\right)}, \quad \alpha = 1, \ldots, B. \tag{6}$$

The Bth mode has zero frequency, $\omega_B^i = 0$, and corresponds to the overall translation of the beads, i.e., the corresponding normal mode is the centroid coordinate, $\boldsymbol{q}_B^i = \boldsymbol{R}_i^c$.

The transformed effective bead potential (with the centroid term removed) is

$$V_{\text{eff}}(\boldsymbol{R}_i;\boldsymbol{q}_\alpha^i) = \sum_{i=1}^{N}\sum_{\beta=1}^{B-1}\left[\frac{1}{2}m_\alpha^i(\omega_\alpha^i\boldsymbol{q}_\alpha^i)^2 + \frac{1}{B}V(\boldsymbol{R}_i;\boldsymbol{q}_\alpha^i)\right]. \tag{7}$$

If we define $\boldsymbol{f}_\beta$ to be the force on a bead due to the potential V (i.e., excluding the fictional harmonic bead potential), the transformation between forces on the normal modes and on beads is

$$\tilde{\boldsymbol{f}}_\alpha^i = \frac{1}{\sqrt{B}}\sum_{\beta=1}^{B} U_{\alpha\beta}\boldsymbol{f}_i^\beta. \tag{8}$$

B. Implementation of CMD propagation

The time propagation of the centroid is computationally expensive because of the need to calculate the path integral average of the centroid force. However, the reformulation of the problem in terms of normal mode coordinates allows some approximations that drastically simplify the calculation.[12]

First, the masses m_α^i of the normal coordinates are taken to be small, so the path fluctuations around the centroid will be fast (compared with the motion of the centroid coordinate), leading to quick generation of normal mode configurations, so the average that leads to the centroid force can be computed efficiently. One can exploit this separation of time scales with the use of multistep MD propagation schemes.[21]

In addition, if the masses of the normal modes are taken equal to $m_\alpha^i = \mu_i(\omega_\alpha^i)^2$ (where μ_i is a proportionality constant), then all normal modes oscillate with the same period (equal to $2\pi\sqrt{\mu_i}$), so the time scale of the propagation is not limited by the time scale of the fastest normal mode.[19]

Second, a Nosé–Hoover thermostat is attached to *each* normal mode so that the normal modes that fluctuate around the centroid path sample their canonical distribution. Because of the well-know ergodicity difficulties of the Nosé–Hoover thermostat for harmonic degrees of freedom,[22] one attaches a Nosé–Hoover chain to each mode.

Third, because of time scale separation, instead of using the path integral average of the force in Eq. (1), one can use the instantaneous value of the force to propagate the centroid variable,

$$m_i \ddot{\boldsymbol{R}}_i \simeq \boldsymbol{F}_i(\boldsymbol{R}_1, \ldots, \boldsymbol{R}_N; \boldsymbol{q}_\alpha^i). \tag{9}$$

The path integral average of the force will be approximated well by CMD during a centroid variable time step, if the normal mode coordinates sample their phase space sufficiently during this interval. This scheme resembles Car–Parinello: instead of fast oscillations around the adiabatic Born–Oppenheimer surface, here we have fast oscillations around the centroid path. The total force on the ith centroid particle is

$$\boldsymbol{F}_i = \sum_\beta \boldsymbol{f}_i^\beta = \sqrt{B} \sum_{\alpha,\beta} U_{\beta\alpha} \tilde{\boldsymbol{f}}_\alpha^i. \tag{10}$$

The terms $\partial V / \partial \boldsymbol{r}_i^\beta$ have to be expressed in terms of the normal modes $\boldsymbol{q}_\alpha^i$ using Eq. (6). Generally the external potential mixes the normal modes, except when it is harmonic.

IV. SIMULATION DETAILS

As a proof of principle, we studied a system consisting of two quantum particles coupled to a classical bath. The quantum two-particle subsystem is described by the potential[23]

$$as^4 - bs^2 + \tfrac{1}{2}M\Omega^2(Q - Q_0)^2 + cs^2Q, \tag{11}$$

where $Q_0 = cs_0^2/(M\Omega^2)$. This potential describes a transferred particle s in a symmetric double well $as^4 - bs^2$, whose minima are in $\mp s_0$, which is coupled to a harmonic oscillator Q of frequency Ω. The coupling cs^2Q is symmetric for the transferred particle; this form is the so-called "rate-promoting vibration."[23] This potential has two minima at $(s,Q) = (\pm s_0, 0)$ and one saddle point at $(s,Q) = (0, Q_0)$.

This two-dimensional Hamiltonian has been studied by Benderskii and co-workers[24] using instanton techniques. At low temperature, this two-dimensional quantum system has very interesting quantum dynamic properties:[23] when the frequency of the rate-promoting vibration Q is smaller than half the inverted barrier frequency, the transferred particle s "cuts the corner" and tunnels toward products without passing near the saddle point (i.e., the classical transition state). For high-frequency vibrations Q, or large mass transferred particle s, the transferred particle follows the adiabatic path passing near the saddle point.

The bath consists of 400 classical oscillators, coupled to s through bilinear coupling (which is equivalent to a classical Langevin equation for the dynamics of s).

A. Quantum transferred particle

The quantum particle whose potential is the double well had mass equal to a H mass, and the double well potential parameters were $a = 0.001\ 25$ a.u. and $b = 0.01$ a.u. These correspond to a barrier height $V = 0.020$ a.u. (12.5 kcal/mol) and potential minima at ∓ 2 a.u. The inverted barrier frequency ω_b was 725 cm^{-1} and the well frequency was 1025 cm^{-1}.

B. Quantum oscillator

The quantum oscillator that was coupled to the quantum double well degree of freedom had mass equal to 4 C masses and frequency 0.0005 a.u. (110 cm^{-1}). The coupling constant for the symmetric term cs^2Q was 0.005 a.u. With these parameters, the barrier height at the saddle point was 3.75 kcal/mol and the dimensionless "corner-cutting criterion," i.e., frequency of Q divided by $\omega_b/2$, was 0.3 (values smaller than 1 are predicted by instanton methods to lead to corner cutting).

When we compared H tunneling with D tunneling, all the parameters other than the mass of s were the same. When we compared with a high-frequency Q oscillation, we used a frequency four times larger than the above value; to get a meaningful comparison and have the same barrier height at the saddle point, in the high-frequency Q oscillation case, we need to multiply the coupling constant between Q and s by the same factor of 4.

C. Classical particles

Our bath consisted of 400 classical oscillators, each with mass equal to a C mass. They are all coupled bilinearly to s. The coupling constants c_i of these bilinear couplings, and the frequencies of the classical oscillators, were determined as follows.

We assumed a uniform distribution of frequencies for the classical oscillators with a discrete step $d\omega = 1/(2\omega_c)$ with $\omega_c = 725$ cm^{-1}. Then the coupling constants c_i were chosen so that they lead to an ohmic with exponential cutoff spectral density, $J(\omega) = \eta \exp(-\omega/\omega_c)$. The friction η was chosen so that $\eta/(m_s\omega_b) = 1$, i.e., an intermediate value of friction.

D. Centroid molecular dynamics

Each of the two quantum variables was described by a centroid with 32 beads. The time step for the centroid propagation was 5 fs, and ten normal mode MD short time steps were done for each long CMD time step. The temperature was $T = 300$ K and Nosé–Hoover thermostat chains of length 3 were attached to the centroids and to each normal mode.

One needs to choose the thermostat parameters carefully, so that there is an appropriate time scale separation.[12] The periods of all normal mode thermostats were 0.24 fs, while the period of the centroid thermostat was 7.6 fs (the inverted barrier frequency corresponds to a period of 46.1 fs). This

choice follows the careful analysis given by Klein and coworkers in the reference above. We again note that Chandler has proved that the implementation of the shooting algorithm in TPS yields a Boltzmann distribution of paths. The effect of the addition of centroid dynamics is to run the trajectories on a modified potential surface that is the centroid averaged potential. The convergence to a Boltzmann distribution still obtains—simply that on an altered potential. We also note that for the time spent in the barrier region (very short in all TPS simulations we have published, on the order of a few tens of femtosecond), the centroid approach should work well. Its well known failure in cases of both deep tunneling and long time scales in highly quantum regions is certainly avoided in the biological simulations we envision.

V. RESULTS

The system described above consists of 402 degrees of freedom, and because of its size one cannot obtain an exact quantum mechanical solution. However, one can guess the qualitative character of the quantum solution. We can view the full system as the two-dimensional problem equation (11) feeling an effect of friction due to its coupling to the 400 classical-like particles. We do not expect the friction to alter qualitatively the features of the quantum solution of the problem equation (11): a light transferred particle (like H) will corner cut before it reaches the classical transition state, while a heavier transferred particle (like D) or a light particle coupled to a high-frequency Q oscillation will both cross to the products nearer the classical transition state.

We will now examine if the proposed quantum TPS method is able to capture these trends. For the TPS "shooting" moves, we perturbed the velocities by 10% of their thermal kinetic value $\sqrt{k_B T/(2m)}$. The order parameters were defined as s values of <-0.2 a.u. or >0.2 a.u. for reactants/products, and the transition was considered reactive if the trajectory remained within one of the above order parameters for at least the last 500 time steps. After we generated a harvest of TPS reactive trajectories of length 375 fs, we constrained the s,Q degrees of freedom and did a MD for the rest of the coordinates to generate 30 configurations. Then we shot 50 new trajectories from each of these configurations to find the separatrix. Finally we plotted the histograms of the s,Q positions for the points on the separatrix, i.e., points that are members of the transition state ensemble.

In Fig. 1 we plot the results for H tunneling and the slow Q vibration with frequency of 110 cm^{-1} (we found 99 points of the separatrix). The horizontal axis for the Q histogram has been normalized with respect to the position Q_0 of the classical transition state. We see that the s histogram is sharply peaked at the top of the one-dimensional barrier, and that the Q histogram is peaked at a value half of the location of the classical transition state. So, the quantum TPS has captured the correct qualitative quantum feature, i.e., that the H corner cuts before it reaches the classical transition state.

In Fig. 2 we plot the results for D tunneling and a slow Q vibration with frequency of 110 cm^{-1} (we found 55 points of the separatrix). Compared with H tunneling, the Q value at which most tunneling takes place has shifted toward the classical transition state value.

Finally, in Fig. 3 we plot the results for H tunneling and a fast Q vibration with frequency of 440 cm^{-1} (we found 86 points of the separatrix). The reason that the s histogram is not peaked is the following. The Q vibration corresponds to a period of 75 fs and the centroid coordinate for s takes about 250–300 fs to tunnel over to the product side. Therefore the Q vibration makes four oscillations during the time it takes for s to tunnel, which means that during tunneling, s sees both high and low barriers.

In conclusion, we have proposed a scheme for including quantum effects in trajectories that are harvested by TPS. We showed that a TPS analysis based on this scheme is able to reproduce the correct dynamics for a system that has nontrivial quantum behavior. This scheme could be useful for including nuclear quantum effects in QM/MM MD simulations. We are currently working in implementing this method in the CHARMM package and eventually applying it to enzymatic systems.

ACKNOWLEDGMENTS

We acknowledge the Chemistry Division of the National Science Foundation for support of this work through Grant No. CHE-0714118 and the National Institutes of Health through Grant No. GM068036.

APPENDIX: PROPAGATION ALGORITHMS

We summarize the propagation algorithms used in this paper. Details can be found in the original literature.

1. Definition of Liouville propagators

We split the total Liouville operator into the standard decomposition of derivatives with respect to position and momentum,[25]

$$iL = iL_{v,\mathrm{fluc}} + iL_r + iL_v + iL_{\mathrm{NH;nm}} + iL_{\mathrm{NH;c}}, \tag{A1}$$

where $L_{\mathrm{NH;nm}}, L_{\mathrm{NH;c}}$ are the Nosé–Hoover for the normal modes and centroids, and

$$iL_r = \sum_{n=1}^{B-1}\sum_{i=1}^{N} \dot{\boldsymbol{q}}_n^i \cdot \nabla_{\boldsymbol{q}_n^i} + \sum_{i=1}^{N} \dot{\boldsymbol{R}}_i \cdot \nabla_{\boldsymbol{R}_i},$$

$$iL_v = \sum_{n=1}^{B-1}\sum_{i=1}^{N} \frac{\tilde{\boldsymbol{f}}_n^i}{m_n^i} \cdot \nabla_{\dot{\boldsymbol{q}}_n^i} + \sum_{i=1}^{N} \frac{\mathbf{F}_i}{M_i} \cdot \nabla_{\dot{\boldsymbol{R}}_i}, \tag{A2}$$

$$iL_v^{\mathrm{fluc}} = \sum_{n=1}^{B-1}\sum_{i=1}^{N} -\frac{(\omega_n^i)^2 \boldsymbol{q}_n^i}{m_n^i} \cdot \nabla_{\dot{\boldsymbol{q}}_n^i}.$$

In addition to the usual MD terms L_r and L_v, there is a term $L_{v,\mathrm{fluc}}$ that describes fluctuations of the normal modes around the centroid.

The sum of the first terms in Eq. (A2) is the Liouville operator for the normal modes, $L^{\mathrm{nm}}=L_v^{\mathrm{nm}}+L_r^{\mathrm{nm}}+L_v^{\mathrm{fluc}}$, and the sum of the second terms is the Liouville operator for the centroid variable, $L^{\mathrm{c}}=L_v^{\mathrm{c}}+L_r^{\mathrm{c}}$.

2. Propagation of centroids

The centroid variable propagates one long time step Δt with the Liouville operator

$$e^{iL_c\Delta t} = e^{iL_{\mathrm{NH;c}}\Delta t/2} e^{iL_v^c \Delta t/2} e^{iL_r^c \Delta t} e^{iL_v^c \Delta t/2} e^{iL_{\mathrm{NH;c}}\Delta t/2}, \tag{A3}$$

which is the ordinary velocity Verlet with Nosé–Hoover.

3. Multistep propagation algorithms

Since the normal modes fluctuate in a time scale much shorter than the time scale of centroid motion, for efficient propagation of dynamics we need a time-reversible MD algorithm that can handle multiple time scales.[21] This category of algorithms is called reference system propagation algorithms (RESPAs). There are two classes of RESPAs: extended system outside reference (XO) and extended system inside reference (XI). XO-RESPAs are used for systems that have fast vibrations compared with the bath motions, while XI-RESPAs are used in the opposite limit.

In path integral calculations it is numerically efficient to have the thermostat evolved in similar time scale as the vibrations of the path integral polymer. Therefore XI-RESPA is appropriate for path integral MD.

The slow motion is integrated with time step Δt and the fast motion with time step $\delta t = \Delta t / N_{\mathrm{ms}}$. That is, one time step Δt for the slow motion is followed by N_{ms} time steps δt for the fast motions.

4. Propagation of normal modes

The propagation of the normal modes must be done with XI-RESPA. We want to do N_{ms} short steps δt, while the long time step is Δt,[12]

$$\begin{aligned} e^{-iL^{\mathrm{nm}}\Delta t} &= e^{iL_{\mathrm{NH}}^{\mathrm{nm}}\delta t/2} e^{iL_v^{\mathrm{nm}}\Delta t/2} e^{-iL_{\mathrm{NH}}^{\mathrm{nm}}\delta t/2} \\ &\quad \times \left[e^{iL_{\mathrm{NH}}^{\mathrm{nm}}\delta t/2} e^{iL_{\mathrm{F}}^{\mathrm{nm}}\delta t} e^{iL_{\mathrm{NH}}^{\mathrm{nm}}\delta t/2} \right]^{N_{\mathrm{ms}}} \\ &\quad \times e^{-iL_{\mathrm{NH}}^{\mathrm{nm}}\delta t/2} e^{iL_v^{\mathrm{nm}}\Delta t/2} e^{iL_{\mathrm{NH}}^{\mathrm{nm}}\delta t/2} \\ &= e^{iL_{\mathrm{NH}}^{\mathrm{nm}}\delta t/2} e^{iL_v^{\mathrm{nm}}\Delta t/2} e^{iL_{\mathrm{F}}^{\mathrm{nm}}\delta t} e^{iL_{\mathrm{NH}}^{\mathrm{nm}}\delta t/2} \\ &\quad \times \left[e^{iL_{\mathrm{NH}}^{\mathrm{nm}}\delta t/2} e^{iL_{\mathrm{F}}^{\mathrm{nm}}\delta t} e^{iL_{\mathrm{NH}}^{\mathrm{nm}}\delta t/2} \right]^{N_{\mathrm{ms}}-2} \\ &\quad \times e^{iL_{\mathrm{NH}}^{\mathrm{nm}}\delta t/2} e^{iL_{\mathrm{F}}^{\mathrm{nm}}\delta t} e^{iL_v^{\mathrm{nm}}\Delta t/2} e^{iL_{\mathrm{NH}}^{\mathrm{nm}}\delta t/2}. \end{aligned} \tag{A4}$$

The operator

$$e^{iL_{\mathrm{F}}^{\mathrm{nm}}\delta t} = e^{iL_v^{\mathrm{fluc}}\delta t/2} e^{iL_r^{\mathrm{nm}}\delta t} e^{iL_v^{\mathrm{fluc}}\delta t/2} \tag{A5}$$

propagates the normal modes according to the fluctuating potential [last line in Eq. (A2), *not* the real potential].

The N_{ms} middle steps are done with this propagator. The Nosé–Hoover operator appears in every short time step because we are in the regime where XI-RESPA applies.

Before and after these N_{ms} short steps, we propagate the velocities with the *long* step $\Delta t/2$ using the real force propagator $e^{iL_v^{\mathrm{nm}}\Delta t/2}$.

5. Nosé–Hoover propagation

In normal mode the CMD one cannot use plain Nosé–Hoover because it fails to sample the phase space of harmonic oscillators, so it would not work for the normal modes. This is easily fixed by using a chain of thermostats, in fact as few as three thermostats overcome the harmonic oscillator ergodicity issue.[21]

We couple the particles $\boldsymbol{R}_i$ with a chain of $M=3$ thermostats ξ_i with masses M_{ξ_i}. The equations of motion are

$$\begin{aligned} M_i \ddot{\boldsymbol{R}}_i &= \boldsymbol{F}_i - M_i \dot{\boldsymbol{R}}_i \dot{\xi}_1, \\ M_{\xi_1} \ddot{\xi}_1 &= \left(\sum_{i=1}^{N} M_i \dot{\boldsymbol{R}}_i^2 - N_{\mathrm{f}} k_{\mathrm{B}} T \right) - M_{\xi_1} \dot{\xi}_1 \dot{\xi}_2, \\ M_{\xi_i} \ddot{\xi}_i &= M_{\xi_{i-1}} \dot{\xi}_{i-1}^2 - k_{\mathrm{B}} T - M_{\xi_i} \dot{\xi}_i \dot{\xi}_{i+1}, \quad i = 2, \ldots, M-1, \\ M_{\xi_M} \ddot{\xi}_M &= M_{\xi_{M-1}} \dot{\xi}_{M-1}^2 - k_{\mathrm{B}} T. \end{aligned} \tag{A6}$$

The Nosé–Hoover Liouville operator is the sum

$$\begin{aligned} iL_{\mathrm{NH}} &= \left(\sum_{i=1}^{M} iL_{\xi_i} + iL_v \right) + \left[\sum_{i=1}^{M-1} iL_{G\dot{\xi}_i} + iL_{v_{\dot{\xi}}\dot{\xi}_1} \right] + iL_{\dot{\xi}_M} \\ &= \left(\sum_{i=1}^{M} \dot{\xi}_i \frac{\partial}{\partial \xi_k} - \dot{\xi}_1 \dot{\boldsymbol{R}} \cdot \nabla_{\dot{\boldsymbol{R}}} \right) \\ &\quad + \left[\sum_{i=1}^{M-1} G_i \frac{\partial}{\partial \dot{\xi}_i} - \dot{\xi}_i \dot{\xi}_{i+1} \frac{\partial}{\partial \dot{\xi}_i} \right] + G_M \frac{\partial}{\partial \dot{\xi}_M}, \end{aligned} \tag{A7}$$

where

$$\begin{aligned} G_1 &= \frac{1}{M_{\xi_1}} \left(\sum M \dot{R}^2 - N_{\mathrm{f}} k_{\mathrm{B}} T \right), \\ G_i &= \frac{1}{M_{\xi_i}} \left(M_{\xi_{i-1}} \dot{\xi}_{i-1}^2 - k_{\mathrm{B}} T \right). \end{aligned} \tag{A8}$$

The factorization used for the Nosé–Hoover propagation is [the terms in parentheses and brackets in Eq. (A9) correspond to those similarly marked in the first line of Eq. (A7)]

$$\begin{aligned} e^{iL_{\mathrm{NH}}\delta t/2} &= e^{iL_{\dot{\xi}_M}\delta t/4} \left[\prod_{i=1}^{M-1} e^{iL_{G\dot{\xi}_i}\delta t/8} e^{iL_{v_{\dot{\xi}}\dot{\xi}_i}\delta t/4} e^{iL_{G\dot{\xi}_i}\delta t/8} \right] \\ &\quad \times \left(\prod_{i=1}^{M} e^{iL_{\dot{\xi}_i}\delta t/2} e^{iL_v \delta t/2} \right) \\ &\quad \times \left[\prod_{i=1}^{M-1} e^{iL_{G\dot{\xi}_i}\delta t/8} e^{iL_{v_{\dot{\xi}}\dot{\xi}_i}\delta t/4} e^{iL_{G\dot{\xi}_i}\delta t/8} \right] e^{iL_{\dot{\xi}_M}\delta t/4}. \end{aligned} \tag{A9}$$

This propagation scheme is implemented in MD as follows:

$$\dot{\xi}_M \to \dot{\xi}_M + G_M \frac{\delta t}{4}, \tag{A10}$$

$$\dot{\xi}_i \to \dot{\xi}_i e^{-\dot{\xi}_{i+1}\delta t/8},$$

$$\dot{\xi}_i \to \dot{\xi}_i + G_i \frac{\delta t}{4}, \tag{A11}$$

$$\dot{\xi}_i \to \dot{\xi}_i e^{-\dot{\xi}_{i+1}\delta t/8}, \quad i = 1, \ldots, M-1,$$

$$\dot{\boldsymbol{R}} \to \dot{\boldsymbol{R}} e^{-\dot{\xi}_1 \delta t/2},$$

$$\mathrm{KE} \to \mathrm{KE} e^{-\dot{\xi}_1 \delta t}, \tag{A12}$$

$$\xi_i \to \xi_i + \dot{\xi}_i \frac{\delta t}{2}, \tag{A13}$$

the three steps of Eq. A11, (A14)

the step of Eq. A10. (A15)

In Eq. (A14) the quantity G_1 uses the updated value of the centroid variable kinetic energy, calculated in Eq. (A12).

[1] C. Dellago and D. Chandler, in *Bridging the Time Scales: Molecular Simulations for the Next Decade*, Lecture Notes in Physics Vol. 605, edited by P. Nielaba, M. Mareschal, and G. Ciccotti (Springer-Verlag, Berlin, 2003).
[2] C. Dellago and P. Bolhuis, Adv. Polym. Sci. **221**, 167 (2009).
[3] W. E, W. Ren, and E. Vanden-Eijnden, Chem. Phys. Lett. **413**, 242 (2005).
[4] P. Metzner, C. Schutte, and E. Vanden-Eijnden, J. Chem. Phys. **125**, 084110 (2006).
[5] J. E. Basner and S. D. Schwartz, J. Am. Chem. Soc. **127**, 13822 (2005).
[6] S. Saen-oon, V. Schramm, and S. Schwartz, Z. Phys. Chem. **222**, 1359 (2008).
[7] S. Saen-oon, S. Quaytman, V. Schramm, and S. Schwartz, Proc. Natl. Acad. Sci. U.S.A. **105**, 16543 (2008).
[8] S. R. Billeter, S. P. Webb, T. Iordanov, P. K. Agarwal, and S. Hammes-Schiffer, J. Chem. Phys. **114**, 6925 (2001).
[9] J. Gao and D. Truhlar, Annu. Rev. Phys. Chem. **53**, 467 (2002).
[10] A. Thomas, D. Jourand, C. Bret, P. Amara, and M. J. Field, J. Am. Chem. Soc. **121**, 9693 (1999).
[11] M. Wang, Z. Lu, and W. Yang, J. Chem. Phys. **124**, 124516 (2006).
[12] K. Kinugawa, P. More, and M. Klein, J. Chem. Phys. **106**, 1154 (1997).
[13] A. Witt, S. Ivanov, M. Shiga, H. Forbert, and D. Marx, J. Chem. Phys. **130**, 194510 (2009).
[14] Y. Brumer, A. Golosov, Z. D. Chen, and D. Reichman, J. Chem. Phys. **116**, 8376 (2002).
[15] P. Bolhuis, C. Dellago, and D. Chandler, Faraday Discuss. **110**, 421 (1998).
[16] C. Dellago, P. Bolhuis, F. Csajka, and D. Chandler, J. Chem. Phys. **108**, 1964 (1998).
[17] C. Dellago, P. Bolhuis, and D. Chandler, J. Chem. Phys. **108**, 9236 (1998).
[18] G. A. Voth, in *Theoretical Methods in Condensed Phase Chemistry*, edited by S. D. Schwartz (Kluwer, Dordrecht, 2000), p. 43.
[19] J. Cao and G. Voth, J. Chem. Phys. **101**, 6168 (1994).
[20] I. Craig and D. Manolopoulos, J. Chem. Phys. **123**, 034102 (2005).
[21] G. Martyna, M. Tuckerman, D. Tobias, and M. Klein, Mol. Phys. **87**, 1117 (1996).
[22] G. Martyna, M. Klein, and M. Tuckerman, J. Chem. Phys. **97**, 2635 (1992).
[23] D. Antoniou and S. D. Schwartz, J. Chem. Phys. **108**, 3620 (1998).
[24] V. Benderskii, D. Makarov, and C. Wight, Adv. Chem. Phys. **88**, 1 (1994).
[25] D. Frenkel and B. Smit, *Understanding Molecular Simulation*, 2nd ed. (Academic, New York, 2002).