

On the calculation of reaction rate constants in the transition path ensemble

Christoph Dellago, Peter G. Bolhuis, and David Chandler

Department of Chemistry, University of California, Berkeley, California 94720

(Received 27 October 1998; accepted 29 December 1998)

We present improved formulas for the calculation of transition rate constants in the transition path ensemble. In this method transition paths between stable states are generated by sampling the distribution of paths with a Monte Carlo procedure. With the new expressions the computational cost for the calculation of transition rate constants can be reduced considerably compared to our original formulation. We demonstrate the method by studying the isomerization of a diatomic molecule immersed in a Weeks–Chandler–Andersen fluid. The paper is concluded by an efficiency analysis of the path sampling algorithm. © 1999 American Institute of Physics. [S0021-9606(99)50513-8]

I. INTRODUCTION

The calculation of rate constants for chemical reactions continues to be a computationally demanding problem for all but the simplest systems.^{1,2} In addition to complications caused by the quantum nature of important degrees of freedom,² the definition of an appropriate reaction coordinate is a major challenge. This problem is particularly significant for reactions occurring in solution, where thermal fluctuations of the condensed environment play a crucial role.³ Indeed, the free energy landscape determining the mechanism and the rate of a reaction can be drastically altered by the presence of a solvent. The definition of a good reaction coordinate requires knowledge of the reaction mechanism. Conventional methods for identifying reaction mechanisms, such as eigenvector following, are based on a complete enumeration of stationary points of the potential energy surface. Focusing on details of the energy landscape, these methods neglect entropic effects entirely. Therefore, exploration of the potential energy surface with conventional methods cannot lead to a satisfactory understanding of solution chemistry. It is clear that the complexity of chemical reactions occurring in the liquid phase at finite temperature requires a different approach.

In a series of recent papers^{4–6} we proposed to solve this problem by adopting a reaction-coordinate-free description of the reaction. In this method the notion of a single, well-defined reaction path, such as a zero kinetic energy path, is replaced by the concept of a large set of possible markedly different paths: the *transition path ensemble*. By definition, this ensemble contains all pathways starting in the reactant state *A* and arriving in the product state *B* within a maximum time *t*. Since these paths are true dynamical trajectories at finite temperature, potential energy surfaces dense in saddle points and rough at a scale of $k_B T$ can easily be treated using the method.

To sample the transition path ensemble we have developed an efficient Monte Carlo procedure, which generates a sequence of paths by small displacements. The basic idea of the algorithm is to select a point along an existing path and

“shoot off” a new path with slightly changed momenta in forward and backward directions. Due to the chaotic nature of the dynamics the new trajectory will quickly diverge from the old one. Since regions *A* and *B* are stable and attract trajectories started nearby, the probability that the new path still connects *A* and *B* is large. An appropriate acceptance criterion guarantees that paths are harvested according to their proper weight. As a result of a path sampling simulation one thus obtains a set of reactive trajectories from which the reaction mechanism (or mechanisms) can be extracted.

The transition path sampling method does not require the definition of a reaction coordinate capable of describing the course of the entire reaction from *A* to *B*. Instead, it suffices to specify the stable states *A* and *B*. Since regions *A* and *B* are stable they can be studied by straightforward molecular dynamics simulation. Hence, characterizing *A* and *B* is usually much simpler than defining a good reaction coordinate.

In contrast to reaction pathways obtained by eigenvector following or similar methods, trajectories generated in the transition path sampling method are true dynamical trajectories free of any bias. Hence, the ensemble of paths harvested in a transition path simulation can be used to calculate reaction rates. The present article describes the efficient calculation of reaction rate constants within the transition path ensemble. We develop improved expressions for the “left–right” time correlation function speeding up the computation of rate constants considerably compared to our original formulation.^{4–6} To demonstrate the method, we calculate a reaction rate constant for isomerizations of a diatomic molecule immersed in a WCA (Weeks–Chandler–Andersen) solvent. For low barrier heights the results of the transition path sampling simulations are compared with results obtained from a straightforward molecular dynamics simulation and excellent agreement is found. Finally, an efficiency analysis of the path sampling is used to determine an optimal acceptance probability of $\sim 40\%$.

This article is organized as follows. In Sec. II we review the theoretical basis of the transition path sampling method and derive a new expression for the reaction rate constant.

Algorithms for efficient sampling of the distribution of paths are discussed in Sec. III. The model system is described in Sec. IV, and numerical results are presented in Sec. V. In Sec. VI an efficiency analysis is carried out, and conclusive remarks are made in Sec. VII.

II. THEORY

A. Reaction rate constants

Consider a dynamical system with two stable states, A and B . These states are assumed to be stable in the sense that trajectories initiated in A or B will stay there for a time long compared to the characteristic time for molecular motions, τ_{mol} . Accordingly, transitions between A and B are rare and a well defined rate constant exists. This transition rate constant is related to the time correlation function

$$C(t) \equiv \frac{\langle h_A(x_0) h_B(x_t) \rangle}{\langle h_A \rangle}, \quad (1)$$

where brackets $\langle \cdots \rangle$ denote equilibrium ensemble averages and $x_t = \{q_t, p_t\}$ is the set of coordinates and momenta specifying the state of the system at time t . h_A and h_B are characteristic functions indicating if the system is in A or B , respectively:

$$h_{A,B}(x) = \begin{cases} 1 & \text{if } x \in A, B, \\ 0 & \text{if } x \notin A, B. \end{cases} \quad (2)$$

The phase space vector x_t evolves according to a set of deterministic equations of motion, such as those of Hamiltonian dynamics. Hence, $x_t = x_t(x_0)$ is completely determined by the initial condition x_0 at time $t=0$. The correlation function $C(t)$ is the conditional probability to find the system in region B at time t provided it started in A at time $t=0$.

Since transitions between the stable states A and B are rare, $C(t)$ approaches its asymptotic value exponentially⁷

$$C(t) \approx \langle h_B \rangle (1 - e^{-t/\tau_{\text{rxn}}}), \quad (3)$$

where $\tau_{\text{rxn}} = (k_{A \rightarrow B} + k_{B \rightarrow A})^{-1}$ is the characteristic reaction time of the system, and $k_{A \rightarrow B}$ and $k_{B \rightarrow A}$ are the forward and backward reaction rate constants, respectively. If the time scales are well separated, i.e., the reaction time τ_{rxn} is much larger than the molecular time τ_{mol} , there exists a time regime $\tau_{\text{mol}} < t \ll \tau_{\text{rxn}}$ in which $C(t)$ grows linearly:

$$C(t) \approx k_{A \rightarrow B} t. \quad (4)$$

Consequently, the reactive flux $k(t) \equiv dC(t)/dt$ displays a plateau whose value is the forward reaction rate constant $k_{A \rightarrow B}$.^{1,7}

One may, in principle, compute $C(t)$ from a single, very long molecular dynamics trajectory. Such an approach is clearly impractical for systems in which the reaction time τ_{rxn} is much longer than τ_{mol} . As shown in our recent work,⁴⁻⁶ this problem of disparate time scales can be avoided by using the *transition path sampling* method. In the following paragraphs we first motivate this approach and then explain the method.

B. $C(t)$ from a “free energy” difference

To determine reaction rate constants in the transition path ensemble we first write $C(t)$ in terms of the equilibrium phase space distribution $\rho(x_0)$,

$$C(t) = \frac{\int dx_0 \rho(x_0) h_A(x_0) h_B(x_t)}{\int dx_0 \rho(x_0) h_A(x_0)}. \quad (5)$$

Since in the above equation both integrands are manifestly positive, we can interpret $C(t)$ as a ratio of two partition functions. Consequently, one can use standard free energy estimation methods to calculate the associated “free energy” difference,^{6,8}

$$\Delta F(t) = -\ln C(t). \quad (6)$$

Thus, the calculation of $C(t)$ is equivalent to the calculation of the “reversible work” necessary to confine the endpoint of a dynamical path initiated in A at time 0 to region B at time t .

For a given time t , the free energy difference (6) may be computed by umbrella sampling.⁹ Imagine that region B is defined by the value of an order parameter $\lambda(x)$:

$$x \in B \quad \text{if} \quad \lambda_{\min} \leq \lambda(x) \leq \lambda_{\max}. \quad (7)$$

The distribution $P(\lambda, t)$ of the order parameter λ at time t for trajectories starting in A at time $t=0$ is given by

$$P(\lambda, t) \equiv \frac{\int dx_0 \rho(x_0) h_A(x_0) \delta[\lambda - \lambda(x_t)]}{\int dx_0 \rho(x_0) h_A(x_0)}, \quad (8)$$

where $\delta(x)$ is Dirac’s delta function. If the distribution $P(\lambda, t)$ is known, we can calculate $C(t)$ by integrating $P(\lambda, t)$ over all order parameters belonging to B :

$$C(t) = \int_{\lambda_{\min}}^{\lambda_{\max}} d\lambda P(\lambda, t). \quad (9)$$

Since transitions from A to B are rare, $P(\lambda, t)$, is very small in region B and it is impractical to calculate $P(\lambda, t)$ directly. To determine $P(\lambda, t)$ by umbrella sampling, we first define a sequence of $N+1$ overlapping regions $B[i]$ such that $B_0 = B$ and the union $\cup_{i=0}^N B[i]$ of all regions comprises the whole phase space. The regions $B[i]$ are defined through

$$x \in B[i] \Leftrightarrow \lambda_{\min}[i] \leq \lambda(x) \leq \lambda_{\max}[i]. \quad (10)$$

For $0 < i < N$ we require region $B[i]$ to overlap with the neighboring regions $B[i-1]$ and $B[i+1]$. Next, one calculates the distribution of the order parameter λ in each of the “windows” $B[i]$ separately:

$$P(\lambda, t; i) \equiv \frac{\int dx_0 \rho(x_0) h_A(x_0) h_{B[i]}(x_t) \delta[\lambda - \lambda(x_t)]}{\int dx_0 \rho(x_0) h_A(x_0) h_{B[i]}(x_t)}. \quad (11)$$

Comparison of Eqs. (8) and (11) shows that $P(\lambda, t; i)$ and $P(\lambda, t)$ are proportional in the window $B[i]$:

$$P(\lambda, t) \propto P(\lambda, t; i) \quad \text{for} \quad \lambda_{\min}[i] \leq \lambda \leq \lambda_{\max}[i]. \quad (12)$$

By matching the histograms $P(\lambda, t; i)$ in overlapping regions and normalizing the resulting distribution one obtains $P(\lambda, t)$. Then, according to Eq. (9), $C(t)$ is determined by integration. We have reduced the calculation of $C(t)$ to the

calculation of the histograms $P(\lambda, t; i)$ in the windows $B[i]$. In the following paragraphs we explain how to do so efficiently.

Equation (11) suggests that we write $P(\lambda, t; i)$ as an average over a distribution function $f_{AB[i]}(x_0, t)$:

$$P(\lambda, t; i) = \frac{\int dx_0 f_{AB[i]}(x_0, t) \delta[\lambda - \lambda(x_t)]}{\int dx_0 f_{AB[i]}(x_0, t)} \\ = \langle \delta[\lambda - \lambda(x_t)] \rangle_{AB[i]}, \quad (13)$$

where

$$f_{AB[i]}(x_0, t) \equiv \rho(x_0) h_A(x_0) h_{B[i]}(x_t). \quad (14)$$

$f_{AB[i]}(x_0, t)$ is the distribution function of all initial condition x_0 in region A leading to trajectories ending exactly in $B[i]$ at time t . $P(\lambda, t; i)$ can be calculated as a weighted average over all paths connecting A and $B[i]$ in time t . Accordingly, we call the quantity

$$\langle A(x_{t'}) \rangle_{AB[i]} = \frac{\int dx_0 f_{AB[i]}(x_0, t) A[x_{t'}(x_0)]}{\int dx_0 f_{AB[i]}(x_0, t)}, \quad (15)$$

a *path average*, and the distribution function $f_{AB[i]}(x_0, t)$, a *path ensemble*. Note that t' can be different from t . As shown in our previous work,⁴⁻⁶ path ensembles can be sampled using a Monte Carlo procedure. Efficient algorithms to do so are described in Sec. III.

C. An advantageous factorization of $C(t)$

Using the method described above, it is possible to calculate $C(t)$ for a number of different times. Then $k(t)$ is determined by numerical differentiation. If $k(t)$ displays a plateau, the rate constant is determined by the value of $k(t)$ in the plateau region. Since this procedure involves many “free energy” calculations, it is very time consuming for all but the simplest systems, making such an approach impractical. In our previous work^{5,6} we derived new expressions for the rate constant which circumvent this problem and require only a single “free energy” calculation. In the present paper we develop improved expressions leading to an increase of efficiency of almost an order of magnitude.

To calculate $C(t)$ in the interval $[0, T]$ we factorize it as

$$C(t) = \frac{\langle h_A h_B(t) \rangle}{\langle h_A h_B(t') \rangle} \times \frac{\langle h_A h_B(t') \rangle}{\langle h_A \rangle} = \frac{\langle h_A h_B(t) \rangle}{\langle h_A h_B(t') \rangle} \times C(t'), \quad (16)$$

where both t and t' are in $[0, T]$. For notational convenience we have dropped the argument of h_A , which is always assumed to be x_0 . Also, we write $h_B(t)$ instead of $h_B(x_t)$. Next, we define the function

$$H_B(x_0, T) \equiv \max_{0 \leq t \leq T} h_B(x_t). \quad (17)$$

The indicator function $H_B(x_0, T)$ is unity if a trajectory starting from x_0 at time $t=0$ visits region B in the interval $[0, T]$ and vanishes otherwise. $H_B(x_0, T)$ is unity even if the system arrives in B at a time $t < T$ but lies outside of B at time T . Since $H_B(x_0, T)$ vanishes only if $h_B(x_t)$ vanishes for all $0 \leq t \leq T$ and $H_B(x_0, T)$ is unity otherwise, we can insert $H_B(x_0, T)$ into the averages of Eq. (16) obtaining

$$C(t) = \frac{\langle h_A h_B(t) H_B(T) \rangle}{\langle h_A h_B(t') \rangle} \times \frac{\langle h_A H_B(t) \rangle}{\langle h_A h_B(t') H_B(T) \rangle} \times C(t'), \quad (18)$$

where we have also multiplied and divided the right hand side of the equation by $\langle h_A H_B(T) \rangle$. For simplicity, we have dropped the argument x_0 of $H_B(x_0, T)$. Equation (18) is valid for all $t, t' \in [0, T]$. Here, the quantity

$$\langle h_B(t) \rangle_{AB} \equiv \frac{\langle h_A h_B(t) H_B(T) \rangle}{\langle h_A H_B(T) \rangle} \\ = \frac{\int dx_0 \rho(x_0) h_A(x_0) H_B(x_0, T) h_B(x_t)}{\int dx_0 \rho(x_0) h_A(x_0) H_B(x_0, T)} \quad (19)$$

can be interpreted as an average of $h_B(x_t)$ over the distribution

$$F_{AB}(x_0, T) \equiv \rho(x_0) h_A(x_0) H_B(x_0, T). \quad (20)$$

$F_{AB}(x_0, T)$ is the ensemble of all initial conditions x_0 in A leading to trajectories visiting B in the interval $[0, T]$. In contrast to $f_{AB}(x_0, T)$, $F_{AB}(x_0, T)$ contains also those trajectories reaching B but leaving it again before T . Using this notation Eq. (18) can be rewritten as

$$C(t) = \frac{\langle h_B(t) \rangle_{AB}}{\langle h_B(t') \rangle_{AB}} \times C(t'). \quad (21)$$

The distribution $F_{AB}(x_0, T)$ may also be sampled by the Monte Carlo techniques described in Sec. III, so that an efficient calculation of $\langle h_B(t) \rangle_{AB}$ is possible. Thus, the correlation function $C(t)$ in the time interval $[0, T]$ can be determined from a single path sampling simulation [to calculate $\langle h_B(t) \rangle_{AB}$] and a single “free energy” calculation [to calculate $C(t')$ at a certain time t']. The “free energy” calculation consists of a series of transition path simulations performed in overlapping windows. We note that t' can be chosen to be much smaller than T for an efficient calculation of $C(t')$.

D. Summary

We now summarize the procedure for the calculation of rate constants in the transition path ensemble as follows. First, one performs a path sampling simulation with the path ensemble $F_{AB}(x_0, T)$ from Eq. (20) to calculate $\langle h_B(t) \rangle_{AB}$ in the interval $[0, T]$. If the time derivative of $\langle h_B(t) \rangle_{AB}$ does not display a plateau, one has to repeat the path sampling simulation with a longer total time T . If $d\langle h_B(t) \rangle_{AB}/dt$ reaches a plateau, one chooses a time t' in the interval $[0, T]$ and computes $C(t')$ by employing the umbrella sampling technique described above. Since $\langle h_B(t') \rangle_{AB}$ is already known, one can now calculate $C(t)$ from Eq. (21) and extract the rate constant $k_{A \rightarrow B}$ from the plateau of

$$k(t) = \frac{dC(t)}{dt} = \frac{\langle \dot{h}_B(t) \rangle_{AB}}{\langle h_B(t') \rangle_{AB}} \times C(t'). \quad (22)$$

E. Reactive flux formalism

For adjacent regions A and B , i.e., $h_A(x) + h_B(x) = 1$, the transition path sampling formalism reduces to the reac-

tive flux formalism.⁷ In the latter, the transition-state-theory reaction rate constant, k_{TST} , is corrected by a time dependent dynamical transmission coefficient $\kappa(t)$:

$$k(t) = k_{\text{TST}}\kappa(t). \quad (23)$$

In general, the transmission coefficient κ is different from one due to recrossings of the surface separating A from B . It is usually determined by computing dynamical trajectories initiated at the dividing surface. The transition state theory estimate k_{TST} is a purely static quantity depending mainly on the probability of finding the system at the dividing surface. To demonstrate the equivalence of the reactive flux method and the transition path sampling we approximate both $C(t')$ and $\langle h_B(t') \rangle_{AB}$ linearly for small t' . Using Eq. (23), we then obtain:

$$k(t) = \langle \dot{h}_B(t) \rangle_{AB} \times \frac{k(0)t'}{\langle \dot{h}_B(0) \rangle_{AB}t'} = \frac{\langle \dot{h}_B(t) \rangle_{AB}}{\langle \dot{h}_B(0) \rangle_{AB}} \times k_{\text{TST}}. \quad (24)$$

Thus, $\langle \dot{h}_B(t) \rangle_{AB}$ normalized by its value at $t=0$ is identically the transmission coefficient of the reactive flux formalism.

III. SAMPLING THE TRANSITION PATH ENSEMBLE

As shown in Sec. II, the calculation of the time correlation function $C(t)$ involves the calculation of averages over the path ensembles $f_{AB}(x_0, T)$ and $F_{AB}(x_0, T)$ defined in Eqs. (14) and (20), respectively. While $f_{AB}(x_0, T)$ is the distribution function of all paths starting in A and ending exactly in B at time T , $F_{AB}(x_0, T)$ is the distribution of all paths starting in A and visiting B in the interval $[0, T]$. Averages over these path ensembles can be calculated by Monte Carlo sampling. A new path is generated from an old one, which is then accepted or rejected according to a detailed balance criterion.⁶ The efficiency of such a procedure depends on the ability to generate new paths which are sufficiently different from the old paths but are still accepted with a reasonable probability. We will discuss this point quantitatively in Sec. VI. In our previous work, we have devised algorithms (called “shooting” and “shifting”) which provide an efficient way to generate paths. We shortly describe these techniques and explain how they can be used to calculate rate constants. In addition to the shooting and shifting moves we introduce path reversal moves.

A. Shooting

Consider a path starting at x_0 in A and reaching region B after going over a high potential energy barrier. One could try to create a new path by applying a small displacement to the original old initial conditions x_0 to obtain a new path starting from $x_0 + \delta x_0$. In general, complex systems are chaotic and, more significantly, trajectories are very unstable in the vicinity of saddle points. Small changes in the initial conditions x_0 can therefore lead to large changes in the trajectory. Therefore δx_0 must be very small in order to generate paths which connect A and B , ensuring a reasonable

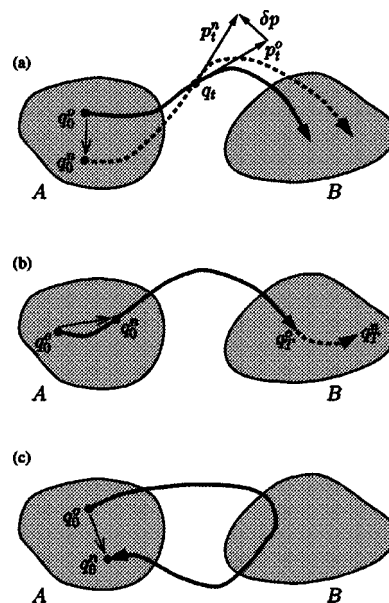


FIG. 1. Schematic representation of the shooting move (a), the shifting move (b), and the path reversal move (c).

acceptance probability. As a consequence, the initial conditions x_0 move slowly through the allowed region making the sampling of the path ensemble inefficient.

This problem can be remedied by changing the state of the system at time t instead of time 0, i.e., by changing x_t for some $t \in [0, T]$. Since the old path connects A and B , there must exist intermediate states x_t for which even a considerable displacement δx_t leads to trajectories still connecting A and B with a high probability. Of course, trajectories are still chaotic, but since regions A and B attract trajectories started nearby, a high acceptance probability can be expected for paths generated by changing an intermediate phase space point x_t . More specifically, in a shooting move one first selects a point $x_t = \{q_t, p_t\}$ at random along the current path. Then one changes the momenta p_t at time t at random by a small amount δp , keeping the coordinates q_t fixed. Integration of the equations of motion backwards to $t=0$ and forwards to $t=T$ from the new state $x_t^n = (q_t, p_t + \delta p)$ yields a new trajectory with initial conditions x_0^n . This procedure is illustrated in part (a) of Fig. 1.

In order to sample the path ensemble F_{AB} correctly, we accept a new path according to the Metropolis criterion

$$P_{\text{acc}}(o \rightarrow n) = \min \left[1, \frac{F_{AB}(x_0^n, T) P_{\text{gen}}(n \rightarrow o)}{F_{AB}(x_0^o, T) P_{\text{gen}}(o \rightarrow n)} \right]. \quad (25)$$

This choice satisfies detailed balance.⁶ Here, $P_{\text{acc}}(o \rightarrow n)$ is the probability for accepting the new path from the old one, $P_{\text{gen}}(o \rightarrow n)$ and $P_{\text{gen}}(n \rightarrow o)$ are the probabilities to generate the new path from the old one and vice versa. The generation probability $P_{\text{gen}}(o \rightarrow n)$ depends on how the point x_t is selected along the path and how the momentum change δp is chosen. For a symmetric generation probability, i.e., $P_{\text{gen}}(o \rightarrow n) = P_{\text{gen}}(n \rightarrow o)$, the acceptance probability reduces to

$$P_{\text{acc}}(o \rightarrow n) = \min \left[1, \frac{\rho(x_0^n)}{\rho(x_0^o)} h_A(x_0^n) H_B(x_0^n; T) \right]. \quad (26)$$

The acceptance probability of a shooting move in the ensemble $f_{AB}(x_0, T)$ is obtained simply by replacing $H_B(x_0^n; T)$ with $h_B(x_0^n)$ in Eq. (26). Since the acceptance probability of paths generated by shooting depends on the momentum displacement δp , one can optimize the efficiency of the algorithm by controlling δp .

B. Shifting

A computationally simple way to generate a new path from an old one is to translate the initial conditions in time. In other words, one obtains a new initial condition x_0^n by evolving the old initial condition x_0^o by a time t :

$$x_0^n = x_t^o(x_0^o). \quad (27)$$

The shifting time t can be positive as well as negative. This path move is schematically depicted in part (b) of Fig. 1. Provided one chooses t at random from a distribution symmetric about $t=0$, the acceptance probability for a shifting move is also given by Eq. (26). The combination of shifting and shooting moves provides an efficient and correct algorithm to sample the transition path ensemble.

Since shifting a path by a small time t requires the integration of the equations of motion for only a few time steps, a shifting move is much less expensive than a shooting move. It is therefore advantageous to attempt shifting moves much more frequently than shooting moves, and to store all the new trajectory segments calculated in subsequent shifting moves. Consider, for example, a sequence of N consecutive shifting moves, each of which translates the initial conditions by a small time Δt forward or backward. If the trajectory information is stored for each shifting move, later shifting moves can be performed with little computational cost. Thus, to do N shifting moves of time Δt , on average it is necessary to integrate the equations of motion only for a time $\sqrt{N}\Delta t$. Of course, the trajectory information is lost if the shifting moves are interrupted by a shooting move. Long sequences of shifting moves correspond to a random walk of the initial conditions in time. Even if the shifting moves do not change the path significantly in the sense that they lead to a new transition state, they are important for statistical refinement of the correlation function $C(t)$. Shifting moves help to extract as much information as possible from paths obtained using the shooting algorithm.

C. Path reversals

Another path move, which can be useful under certain circumstances, is path reversal. In this path move one obtains new initial conditions x_0^n by taking the final point of the path x_T and reversing the momenta:

$$x_0^n = \{q_0^n, p_0^n\} = \{q_T^o, -p_T^o\}. \quad (28)$$

This path move is schematically depicted in part (c) of Fig. 1. Since in the case of the path ensemble $f_{AB}(x_0, t)$ defined in Eq. (14), x_T^o is always in region B , path reversals cannot be employed to sample $f_{AB}(x_0, t)$. In the case of the en-

semble $F_{AB}(x_0, T)$ of paths visiting B in the time interval $[0, T]$, however, is it possible to find x_T in region A . For such paths a reversal is always accepted. Since no integration of the equations of motion is required for a path reversal, the computational cost of such moves is negligible and path reversal moves should be performed whenever possible. In fact, path reversals can facilitate ergodic sampling of the path space if qualitatively different pathways connecting A and B exist.

IV. MODEL

To illustrate the method described above, we calculate the transition rate constant for a diatomic molecule with two stable states immersed in a fluid of purely repulsive particles. A similar model has been studied by Straub, Borkovec, and Berne.¹⁰ The model consists of N two-dimensional particles of mass m interacting via the purely repulsive Weeks–Chandler–Andersen potential

$$V_{\text{WCA}}(r) = \begin{cases} 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] + \epsilon & \text{if } r \leq r_{\text{WCA}} \equiv 2^{1/6}\sigma, \\ 0 & \text{if } r > r_{\text{WCA}}, \end{cases} \quad (29)$$

where r is the interparticle distance and ϵ and σ are parameters specifying the strength and the interaction radius of the potential, respectively. In addition, two of the N particles interact via the double well potential

$$V_{\text{dw}}(r) = h \left[1 - \frac{(r - r_{\text{WCA}} - w)^2}{w^2} \right]^2. \quad (30)$$

Here, r is the distance of the two particles belonging to the diatomic molecule. The parameter h controls the height of the barrier between the stable states located at $r = r_{\text{WCA}}$ (the compact state) and $r = r_{\text{WCA}} + 2w$ (the extended state), respectively. The system evolves according to Hamilton's equations of motion in a simulation box with periodic boundary conditions.

If the barrier height h is large, transitions between the compact and the extended state are rare. In this case the diatomic oscillates in one of the stable states for a long time until, due to collisions with the solvent particles, sufficient energy to cross the barrier is deposited into the reaction coordinate. Then a quick transition occurs, and the energy of the diatomic is absorbed by the solvent, stabilizing the diatomic molecule again. In the following we calculate rate constants for such isomerizations.

V. NUMERICAL RESULTS

Since the system evolves at constant total energy E with a fixed center of mass, the appropriate distribution function of initial conditions x_0 is the microcanonical distribution with the additional constraint of a vanishing total momentum P ,

$$\rho(x_0) = \delta(\mathcal{H}(x_0) - E) \delta(P). \quad (31)$$

Accordingly, the momentum displacement δp used in the shooting algorithms must be chosen to conserve both the

total energy \mathcal{H} and the total momentum P of the system. In this case the acceptance probability for shooting moves simplifies to

$$P_{\text{acc}}(o \rightarrow n) = h_A(x_0^n) H_B(x_0^n, T). \quad (32)$$

In other words, any new path generated by a shooting move is accepted if it connects A and B . Since Hamilton's equation of motion conserves both the total energy \mathcal{H} and the total momentum P , the acceptance probability for a shifting move is also given by Eq. (32).

Our numerical results are presented in reduced units, i.e., lengths are measured in units of σ , energies in units of ϵ , masses in units of m , times in units of $\tau \equiv (m\sigma^2/\epsilon)^{1/2}$, and transition rate constants in units of τ^{-1} . In all our simulations we employ the velocity Verlet algorithm with a time step of $\Delta t = 0.002\tau$ to integrate the equations of motion.

For shooting moves all the components of the momentum displacement vector δp are chosen from a Gaussian distribution with a certain width. Next, components of δp corresponding to a nonvanishing total momentum are removed. Then, δp is added to the old momentum p_i^o yielding the new momentum $p_i^n = p_i^o + \delta p$ which is rescaled to conserve the total energy E . This procedure for changing p_i is symmetric in the sense that the generation probability for the move $p_i^o \rightarrow p_i^n$ equals the generation probability for the backwards move $p_i^n \rightarrow p_i^o$. Hence, such shooting moves obey the detailed balance condition and lead to a correct sampling of the transition path ensemble. The average magnitude of δp was

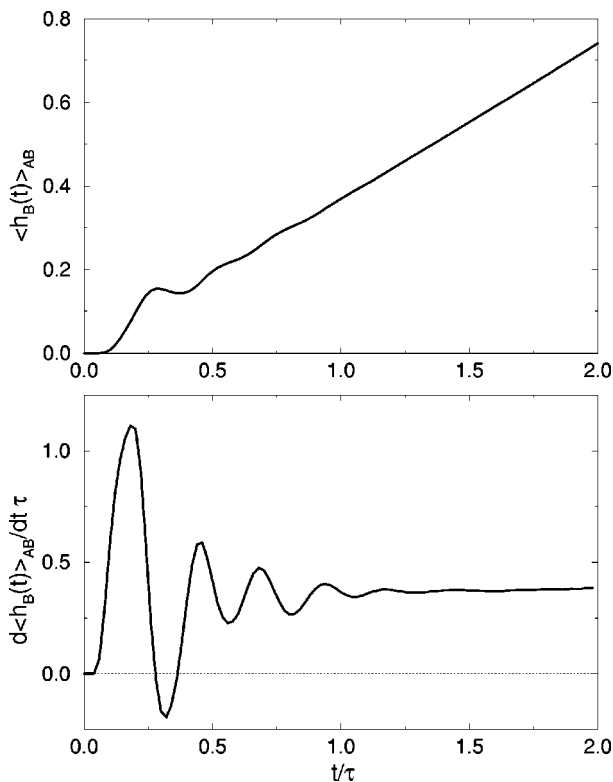


FIG. 2. Path average $\langle h_B(t) \rangle_{AB}$ (top) and its time derivative $d\langle h_B(t) \rangle_{AB}/dt$ (bottom) for the diatomic in the WCA fluid. The parameters of the simulation were $N=9$ (7 solvent particles and 2 particles belonging to the diatomic), $E=9\epsilon$, $h=6\epsilon$, $w=0.25\sigma$, $\rho=0.6\sigma^{-3}$, $R_A=1.30\sigma$, and $R_B=1.45\sigma$.

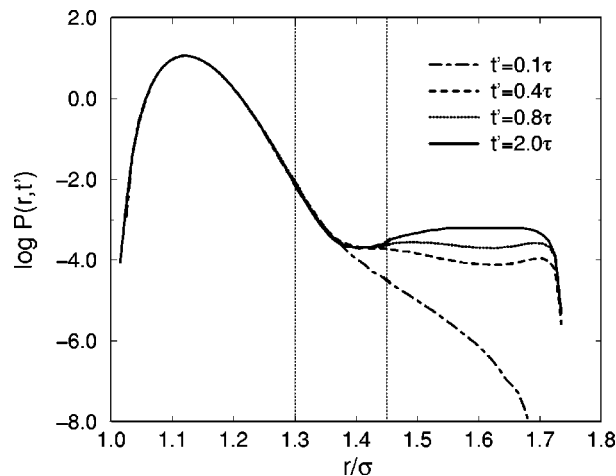


FIG. 3. Logarithm of the probability distributions $P(r, t')$ of the diatomic distance r the end of the path for different total path length $t' = 0.1\tau$, 0.4τ , 0.8τ , and 2.0τ for the same set of parameters used to obtain the results shown in Fig. 2. The graphs were obtained by matching the distributions of r obtained from 5 independent simulations where r was confined to windows defined by $0.20\sigma < r < 1.22\sigma$, $1.21\sigma < r < 1.26\sigma$, $1.25\sigma < r < 1.30\sigma$, $1.29\sigma < r < 1.46\sigma$, and $1.45\sigma < r < \infty$. The vertical dotted lines indicate the limits $R_A = 1.30\sigma$ and $R_B = 1.45\sigma$ of the initial and final region, respectively. Integration over the final region B yields the values of the correlation function $C(t')$ for the different values of t' : $C(0.1\tau) = 1.306 \times 10^{-6}$, $C(0.4\tau) = 3.003 \times 10^{-5}$, $C(0.8\tau) = 6.548 \times 10^{-5}$, $C(2.0\tau) = 1.456 \times 10^{-4}$.

chosen to optimize the sampling rate. A detailed discussion of this point appears in Sec. VI. After each shooting move, 100 shifting moves of 10 time steps Δt each are performed. Despite the large number of shiftings, the computational cost of the whole shifting sequence is only a fraction of the cost for a shooting move. We perform path reversal moves whenever it is possible.

As described in Sec. II, to calculate a transition rate constant one first has to perform a path simulation to determine the path average $\langle h_B(t) \rangle_{AB}$. Such a simulation yields the correlation function $C(t)$ up to a multiplicative factor. To determine this multiplicative factor, a "free energy" calculation of $C(t')$ at an intermediate time t' is necessary. In the following paragraphs we present the results of a complete calculation of the transition rate constant for the diatomic in solution.

The interatomic distance r provides the natural order parameter for the definition of the stable regions A and B : we define regions A and B to contain all configurations with $r < R_A$ and $r > R_B$, respectively. Obviously, R_A and R_B should lie on different sides of the separating barrier and allow the stable regions to accommodate most of the equilibrium fluctuations around the potential energy minima. In all our simulations we use $R_A = 1.30\sigma$, $R_B = 1.45\sigma$, and a barrier width of $w = 0.25\sigma$. Consequently, the top of the barrier is at $r \sim 1.37\sigma$, and the minima of the bistable potential are at $r \sim 1.12\sigma$ and $r \sim 1.62\sigma$.

First, we determine the path average $\langle h_B(t) \rangle_{AB}$ in the interval $[0, T]$. Figure 2 shows $\langle h_B(t) \rangle_{AB}$ and its time derivative $d\langle h_B(t) \rangle_{AB}/dt$ as a function of t obtained from a path simulation consisting of 3.6×10^6 shots. The choice of parameters ($N=9$, $E=9\epsilon$, $h=6\epsilon$, $\rho=0.6\sigma^{-3}$) guarantees that the transition between A and B is rare and that a transition

TABLE I. Time correlation function $C(t')$ and rate constant $k_{A \rightarrow B}$ obtained for different times t' for the parameter set $N=9$, $E=9\epsilon$, $h=6\epsilon$, and $\rho=0.6\sigma^{-3}$.

t'	$C(t')$	$k_{A \rightarrow B}$
0.1τ	1.31×10^{-6}	$6.8 \times 10^{-5} \tau^{-1}$
0.4τ	3.00×10^{-5}	$7.9 \times 10^{-5} \tau^{-1}$
0.8τ	6.55×10^{-5}	$8.3 \times 10^{-5} \tau^{-1}$
2.0τ	1.46×10^{-4}	$7.5 \times 10^{-5} \tau^{-1}$

rate constant is well defined. As can be seen in the figures, $\langle h_B(t) \rangle_{AB}$ reaches a linear regime after $\sim 1.5\tau$. Accordingly, after a few oscillations, which arise from trajectories reaching B and then leaving, its time derivative displays a plateau.

Next, we calculate the correlation function $C(t')$ for an intermediate time t' for the same set of parameters. Figure 3 shows the distribution $P(r, t')$ of the order parameter r for different t' ranging from $t'=0.1\tau$ to $t'=2.0\tau$. $P(r, t')$ was obtained by using the overlapping windows $0.20\sigma < r < 1.22\sigma$, $1.21\sigma < r < 1.26\sigma$, $1.25\sigma < r < 1.30\sigma$, $1.29\sigma < r < 1.46\sigma$, and $1.45\sigma < r < \infty$. For each window we performed a path sampling simulation consisting of 5×10^5 shots in which the endpoint of the path was confined to the respective order parameter window.

As shown in Fig. 3, the probability to find the diatomic in region A (corresponding to small values of r) is high for all times t' . Since the transition from A to B is a rare event, the probability to be in the extended state B (corresponding to large values of r) is small. For longer times t' this probability increases leading to a small peak of $P(r, t')$ in region B . For times very large compared to the reaction time τ_{rxn} , $P(r, t')$ approaches the equilibrium distribution of the order parameter, which has two peaks of approximately the same height in regions A and B . By integrating $P(r, t')$ over region B , i.e., for $r > R_B$, we obtain the values for the correlation function $C(t')$ shown in column two of Table I. For

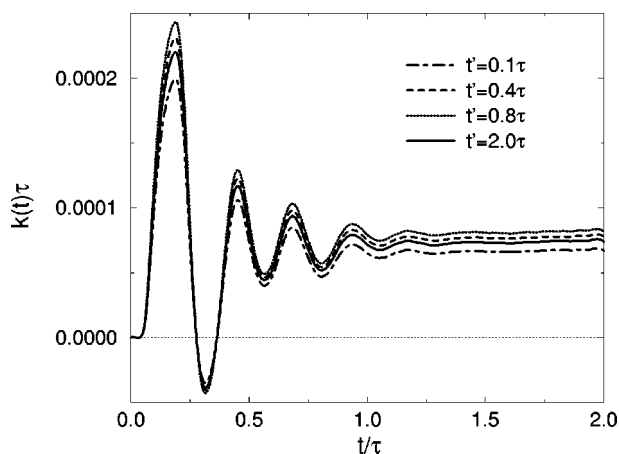


FIG. 4. Time dependent reaction rate $k(t)$ as a function of t obtained according to Eq. (21) by combining the $\langle h_B(t) \rangle_{AB}$ shown in Fig. 2 with the results shown in Fig. 3. Different graphs correspond to different values of t' . The reaction rate constants obtained from the plateau values of the curves are: $k(0.1\tau)=6.8 \times 10^{-5} \tau^{-1}$, $k(0.4\tau)=7.9 \times 10^{-5} \tau^{-1}$, $k(0.8\tau)=8.3 \times 10^{-5} \tau^{-1}$, and $k(2.0\tau)=7.5 \times 10^{-5} \tau^{-1}$. These differences are due to statistical errors in the “free energy” calculations.

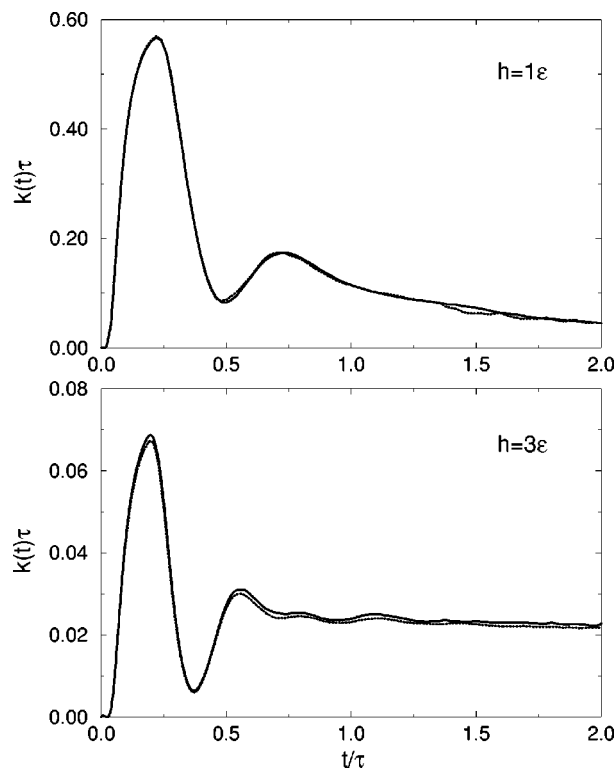


FIG. 5. Comparison of transition path sampling results with the results obtained by straightforward molecular dynamics (MD) simulation for $h=1\epsilon$ (top) and $h=3\epsilon$ (bottom). All other parameters are identical to the parameters used to obtain the results shown in Fig. 2. The solid line indicates $k(t)$ computed by path sampling whereas the dotted line indicates the MD result. The MD trajectories consisted of 200×10^6 steps of length $\Delta t=0.002\tau$.

the purpose of calculating a reaction rate constant it is necessary to compute $C(t')$ for a single value of t' . Here, we have computed $C(t')$ at a variety of t' for comparison.

By inserting $\langle h_B(t) \rangle_{AB}$ and $C(t')$ into Eq. (21) we obtain $C(t)$ and, by numerical differentiation, its time derivative $k(t)$. Figure (4) shows $k(t)$ calculated for different values of t' . The corresponding transition rate constants are shown in column three of Table I. All $k(t)$ curves lead to approximately the same plateau value and hence to the same transition rate constant. Differences in the curves are due to statistical errors in the calculation of $C(t')$ and $\langle h_B(t) \rangle_{AB}$. The computational cost for the calculation of $C(t')$ at $t'=0.1\tau$ is smaller by a factor of 20 than at $t'=2.0$. Thus, it is feasible to reduce the computational cost of transition rate calculations considerably by performing the “free energy” estimation of $C(t')$ at short times t' .

For low values of the barrier height h the results of the transition path sampling simulations can be compared with the results of straightforward molecular dynamics simulations. Figure 5 shows $k(t)$ as a function of t for the barrier heights $h=1\epsilon$ (top) and $h=3\epsilon$ (bottom). In the case $h=1\epsilon$, transitions between the compact and the extended state occur very frequently and there is no separation of time scales. Consequently, $k(t)$ does not reach a plateau but rather decays exponentially after the molecular time τ_{mol} . For $h=3\epsilon$ the transitions are rare [but still frequent enough for an MD (Molecular Dynamics) simulation] so that $k(t)$ reaches a

plateau and a transition rate constant is well defined. In both cases the result of the transition path sampling simulation (solid line) and the MD simulation (dotted line) agree very well.

It is interesting to compare the efficiency of straightforward molecular dynamics simulation with the efficiency of the path sampling method. In a straightforward molecular dynamics trajectory the average time needed to obtain a reaction event is precisely the reaction time τ_{rxn} . For barrier heights $h=1\epsilon$, $h=3\epsilon$, and $h=6\epsilon$, the average waiting times for a transition are approximately 1τ , 40τ , and $12\,500\tau$, respectively. By contrast, the time needed to generate statistically independent trajectories with the path sampling algorithm does not depend on the barrier height. As can be inferred from Figs. 2 and 5 trajectories of length 2τ are long enough to capture typical transitions. As shown in the next paragraph, trajectories are statistically independent if they are separated by $n_c \sim 5$ Monte Carlo cycles. The generation of a statistically independent pathway therefore requires the integration of the equations of motion for $\sim 5 \times 2\tau = 10\tau$. Including the cost of the “free energy” calculation we obtain an integration time of $\sim 20\text{--}40\tau$ per independent trajectory. Hence, for $h=1\epsilon$ direct MD simulation is more efficient than path sampling, but already for $h=3\epsilon$, path sampling exceeds in efficiency. At $h=6\epsilon$ path sampling is more efficient than straightforward simulation by a factor of ~ 500 . Clearly, with increasing separation between molecular time and reaction time scales the path sampling scheme becomes more and more efficient compared with straightforward dynamics. While for typical molecular systems straightforward molecular dynamics simulation is typically limited to time scales of nanoseconds, reactions with characteristic times many orders of magnitude larger can be treated with the transition path sampling method.

VI. EFFICIENCY

In general, the efficiency of a Monte Carlo simulation depends on how quickly the available space is sampled. Strong correlations between consecutive states lead to slow sampling and hence to a poor efficiency.¹¹ The efficiency of a path sampling simulation can therefore be assessed by calculating the autocorrelation function $c(n)$ of the fluctuation of an appropriate quantity $A(x)$ as a function of the number n of simulations cycles:

$$c(n) \equiv \frac{\langle \delta A(0) \delta A(n) \rangle_{AB}}{\langle \delta A^2 \rangle_{AB}}, \quad (33)$$

where $\delta A(n) \equiv A(n) - \langle A(n) \rangle_{AB}$. Fast decay of $c(n)$ indicates fast sampling and hence high efficiency. Figure 6 shows the correlation function $c(n)$ as a function of the number of simulation cycles for three different quantities $A(x)$: the indicator function $h_B(x_{T/2})$ at the midpoint of the path (solid line), the potential energy V_{top} of the system on top of the barrier (dotted line), and the transition time t_{tr} (broken line). The transition time t_{tr} is the time needed to cross the gap between A and B. All three curves decay quickly at short times but decorrelate more slowly for longer times. One can estimate the number n_c of correlated cycles by de-

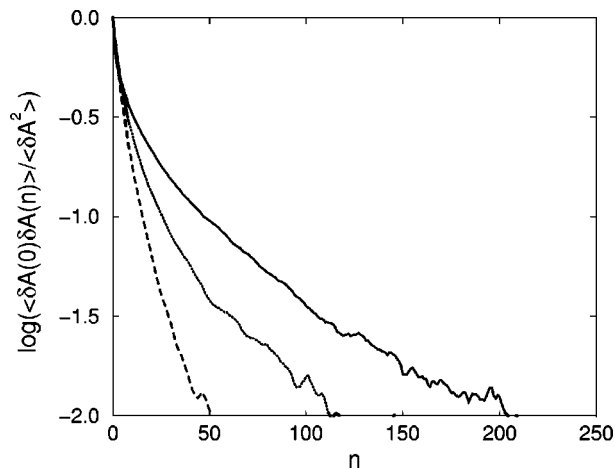


FIG. 6. Logarithm of the correlation functions $c(n)$ as a function of the number of cycles for the quantities $h_B(x_{T/2})$ (solid line), V_{tr} (dotted line), and t_{tr} (broken line) for the same parameters used to obtain the results shown in Fig. 2. The path sampling simulation consisted of 10^6 shots and 100 shifts of length 0.02τ after each shot. Paths generated by shooting were accepted with a probability of 40%.

fining a threshold $c^*(n)$ below which paths are regarded to be uncorrelated. Alternatively, n_c can be defined as the integral of $c(n)$, i.e., $n_c = \sum_n c(n)$. In our example, the number of correlated cycles obtained with a threshold of 0.5 is smaller by a factor of 2–3 than the n_c obtained from the integration of $c(n)$.

Since an exhaustive correlation analysis is computationally expensive, it is useful to examine the dependence of the number of correlated cycles, n_c , upon the acceptance probability P_{acc} . The number of correlated cycles n_c (for a threshold of 0.5) is shown as a function of the acceptance probability P_{acc} in Fig. 7 for $h_B(x_{T/2})$, V_{top} , and t_{tr} . Different acceptance probabilities were obtained by employing different magnitudes of the momentum displacement δp . A

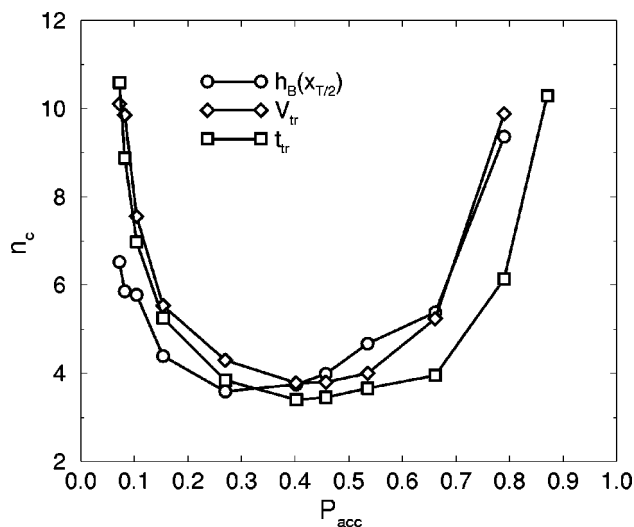


FIG. 7. Number n_c of correlated cycles as a function of the acceptance probability P_{acc} for $h_B(x_{T/2})$ (circles), V_{tr} (diamonds), and t_{tr} (squares). Each simulation consisted of 10^5 shots and 100 shifts of length 0.02τ after each shot. The number n_c of correlated cycles was obtained by determining where the corresponding correlation function falls below a threshold of 0.5.

small δp leads to a high acceptance probability but to a low efficiency because subsequent paths are highly correlated. On the other hand, a large momentum displacement causes subsequent paths to differ widely. However, such paths are only rarely accepted resulting again in low efficiency. Between these extremes n_c displays a rather wide minimum corresponding to the optimum efficiency. These results suggest that, as a rule of thumb, the magnitude of δp should be chosen to obtain an acceptance probability of about 40%. Taking into account that the computational cost of a rejected move is lower than for an accepted move, even smaller acceptance probabilities might be advantageous.

VII. CONCLUSION

The new formulas derived in this paper further reduce the cost of transition path sampling calculations of rate constants. Although we demonstrate the accuracy and practicality of the new formulas by studying a simple model, transition path sampling is not restricted to low-dimensional systems. Indeed, transition path sampling is practical for any system where straightforward molecular dynamics is feasible. Of course, the latter produces few if any transitions between long lived stable states while the former focuses entirely on such transitions.

The general feasibility of transition path sampling is the result of linear scaling. Namely, the computational cost for harvesting N statistically independent paths, each L time steps long, is the same as that for a straightforward trajectory that is $m \times N \times L$ steps long. The proportionality constant, m , is the number of simulation cycles required to harvest a statistically independent trajectory. This number is generally not large. For example, from the efficiency analysis presented in Sec. VI one finds $m \approx 5$ for the system studied herein. A similar analysis⁶ for structural transitions in a two-dimensional seven particle Lennard–Jones cluster yields m

≈ 4 . Furthermore, recent transition path sampling studies of hydrogen bond breaking¹² and ionic dissociation¹³ in bulk water indicate that $m < 10$ even in systems of high dimensionality. Thus, possible applications of the transition path sampling method include chemical reactions occurring in solution and the dynamics of complex biopolymers.

ACKNOWLEDGMENTS

This work was initiated with support from the National Science Foundation and completed with support from the Department of Energy through the Chemical Sciences Division of Lawrence Berkeley National Laboratory. We thank Phillip Geissler for a critical reading of the manuscript. C. Dellago gratefully acknowledges support from the Austrian Fonds zur Förderung der wissenschaftlichen Forschung grants J1302-PHY and J1548-PHY.

¹P. Hänggi, P. Talkner, and M. Borkovec, Rev. Mod. Phys. **62**, 251 (1990).

²W. H. Miller, J. Phys. Chem. A **102**, 793 (1998).

³M. Sprik, Faraday Discuss. **110**, 437 (1998).

⁴C. Dellago, P. G. Bolhuis, F. S. Csajka, and D. Chandler, J. Chem. Phys. **108**, 1964 (1998).

⁵C. Dellago, P. G. Bolhuis, and D. Chandler, J. Chem. Phys. **108**, 9236 (1998).

⁶P. G. Bolhuis, C. Dellago, and D. Chandler, Faraday Discuss. **110**, 421 (1998).

⁷D. Chandler, J. Chem. Phys. **68**, 2959 (1978).

⁸D. Chandler, Lecture 3 in the Proceedings of the Enrico Fermi Summer School on the "Computer Simulation of rare events and the dynamics of classical and quantum condensed matter systems," Lerici, Villa Mari gnola, Italy, 7–18 July 1997.

⁹D. Chandler, *Introduction to Modern Statistical Mechanics* (Oxford University, New York, 1987).

¹⁰J. E. Straub, M. Borkovec, and B. J. Berne, J. Chem. Phys. **89**, 4833 (1988).

¹¹M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Clarendon, Oxford, 1987).

¹²F. S. Csajka and D. Chandler, J. Chem. Phys. **109**, 1125 (1998).

¹³P. L. Geissler, C. Dellago, and D. Chandler, J. Phys. Chem. **103** (in press, 1999).