Finite Temperature String Method for the Study of Rare Events†

Weinan E and Weiqing Ren

Department of Mathematics and PACM, Princeton University, Princeton, New Jersey 08544

Eric Vanden-Eijnden*

Courant Institute, New York University, New York, New York 10012

Received: September 30, 2004; In Final Form: December 15, 2004

A method is presented for the study of rare events such as conformational changes arising in activated processes whose reaction coordinate is not known beforehand and for which the assumptions of transition state theory are invalid. The method samples the energy landscape adaptively and determines the isoprobability surfaces for the transition: by definition the trajectories initiated anywhere on one of these surfaces has equal probability to reach first one metastable set rather than the other. Upon weighting these surfaces by the equilibrium probability distribution, one obtains an effective transition pathway, i.e., a tube in configuration space inside which conformational changes occur with high probability, and the associated rate. The method is first validated on a simple two-dimensional example; then it is applied to a model of solid−solid transformation of a condensed system.

1. Introduction

Activated processes such as nucleation events during phase transition, conformational changes of macromolecules, or chemical reactions usually occur on a time scale that is much larger than the microtime scale in the system. The reason is that these processes require an unusually large thermal fluctuation to drive the system over some energy barrier separating the conformations. Because of the wide separation of time scales, it is impossible to study activated processes by conventional molecular dynamics simulations.

David Chandler has been instrumental in the development of alternative techniques for determining transition pathways and rates in complex systems. In most of these techniques, one manages to introduces an appropriate bias on the dynamics which greatly enhances the probability of observing the portion of the trajectories during which they perform a transition from one metastable set to another. For instance, in the Bennett−Chandler two-step procedure,[1,2] which builds on transition state theory (TST),[21,22] the rate is computed by launching trajectories and running them both backward and forward in time from an appropriately chosen dividing surface between the sets rather than from the sets themselves. If the dividing surface is well chosen (which can be done when the mechanism of transition is known beforehand), the trajectories started from the surface have a relatively high probability to go to one set forward in time and to the other backward in time, meaning that the procedure allows one to observe true dynamical hopping trajectories between these sets.

More recently, other techniques using biased dynamics have been introduced which require no a priori knowledge of the mechanism of transition. These are based on Pratt's idea[20] of sampling dynamical trajectories conditional on them making a transition from metastable set to the other. For instance, the transition path sampling technique of Chandler and collaborators,[16,17] or the action based methods introduced by Elber and collaborators,[11] deal with the true dynamical trajectories, except that the methods only sample trajectories that hop from one metastable set to another. This way, these methods cleverly manage to observe these trajectories precisely during the windows of time when a transition happens.

On the other hand, very often the dynamical trajectories involved in the hopping events from one metastable set to another are quite complicated objects in complex systems, and the information they provide on the mechanism of the transition is very indirect. Arguably, the best reaction coordinate to describe this mechanism consists of the isoprobability surfaces, or committer surfaces.[16,17] By definition, on these surfaces, the probability that a trajectory reach one metastable set before the other is uniform. For instance, the committer $^1/_2$ surface defines the transition state since it is the surface such that a trajectory launched anywhere on the surface has probability half to go first to one set and half to go to the other. In addition, weighting these surfaces with the equilibrium Gibbs distribution allows one to determine where the transition trajectories are most likely to pass through the surface, since by definition there is no bias in the way switching and nonswitching trajectories hit these surfaces (more precisely: by definition, a trajectory hitting the surface anywhere has a uniform probability to either be a transition trajectory or be a trajectory that returns to the set it comes from).

Therefore, it is desirable to identify these isoprobability surfaces to describe the transition. However, this is not easy to do from the dynamical hopping trajectories. Indeed, these are by definition parametrized by time, which is not a good indicator of the advancement of the reaction. In general, two trajectories leaving a set at the same time will hit an isoprobability surface at different times. Even worse, a single hopping trajectory may hit an isoprobability surface many times during a transition.

The method we propose in this paper allows one to identify the isoprobability surfaces directly, i.e., without the intermediate step of determining the hopping trajectories first. The method

† Part of the special issue "David Chandler Festschrift".

10.1021/jp0455430 CCC: $30.25
Published on Web 02/09/2005

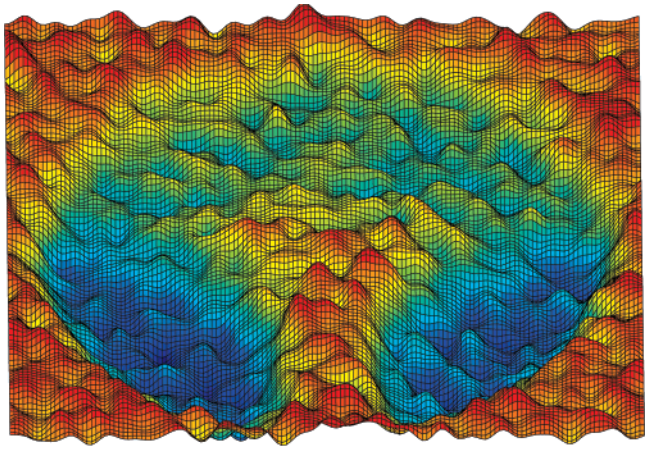

Figure 1. Multiscale two-dimensional potential used in the illustrative toy example. At a suitable temperature, the two dark blue regions are metastable states in which the process spends most of its time; yet transitions occur infrequently, mostly via the light blue/yellow channel—see Figure 2.

builds on standard results from stochastic process theory which characterizes the isoprobability surfaces. From an implementation viewpoint, the method is a direct extension of the zero-temperature string method,[6] except that the potential forces are replaced by some constrained thermodynamic averaged forces. The method performs a constrained sampling of the equilibrium distribution of the system in a collection of hyperplanes parametrized by a string which is updated self-consistently until it approximates locally the isoprobability surface. The region in these planes where the equilibrium distribution is concentrated determines a transition tube in configuration space in which transition takes place with high probability. The method only uses constrained simulations where every trajectory is statistically significant, and it remains efficient even in the absence of a dividing surface with low recrossing rate.

2. An Illustrative Example

The methodology we propose is best illustrated first on a toy example which, while simplistic, retains some of the nontrivial features displayed by more realistic systems. The example consists of a two-dimensional system whose dynamics is governed by the Langevin equation

$$\gamma \dot{q} = -\nabla V(q) + \xi(t) \tag{1}$$

Here γ is the friction coefficient, $\xi(t)$ is a white noise with $\langle \xi_j(t)\xi_k(0)\rangle = 2\gamma k_B T \delta_{jk}\delta(t)$, and $V(q)$ is a multiscale potential with a very large number of critical points (minima, saddle points, etc.) whose three-dimensional plot and isolines are shown respectively in Figure 1 and Figure 2. We assume that $\gamma = 1$ and take $k_B T = 1$, which is about the size of the small energy barriers in the potential, but lower than the big energy barriers (for instance, the yellow-blue channel is about 5 energy units above the two dark blue regions). At this temperature, we observe the following from the direct numerical simulations of (1):

1. The two regions within the green curves in Figure 2 (which are roughly the two dark-blue regions in Figures 1) are metastable sets, in the sense that the solution of (1) spends most of its time in these two regions. Notice that the details of the trajectory (two portions of which are shown in blue in Figure 2) within these sets is very complicated.

2. Transitions between these two metastable sets are rare but they do occur. While the details of the trajectories of these transition events are very complicated, they are very likely to go through the region between the two red curves in Figure 2 (which is roughly the yellow-blue channel in Figure 1); for obvious reasons, we shall refer to the region between the two red curves as the effective transition tube.

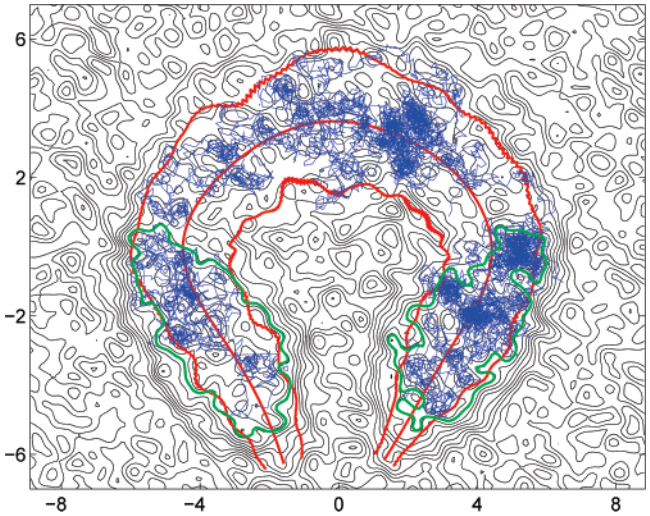

Figure 2. Metastable sets (regions within the green curves), the mean transition path φ (center red curve), and the boundaries of the effective switching tube (external red curves) shown on the isolines of the potential in Figure 1. Also shown in blue are two dynamical paths by which the system switches from the left metastable set to the right one and conversely. Here the thermal energy is on the order of the small-scale features in the potential, and the notion of minimal energy path (MEP) is insufficient; instead, the rather complicated dynamical switching paths lie within the much simpler switching tube, which is identified by the finite temperature string method.

The simplicity of the metastable sets and the effective transition tube compared to the dynamical trajectories suggests that the system effectively experiences a thermally averaged energy landscape, much smoother than the one shown in Figures 1 and 2. Indeed, at the given temperature where we study (1), most energy barriers are smaller than or on the order of $k_B T$, and they are therefore mostly irrelevant for the conformational changes since the system can easily hop over these small barriers. Only a few barriers bigger than $k_B T$ actually confine the system for long times in separate regions in configuration space and thereby define metastable states. Similarly, the effective transition tube is also determined by large scale features of the energy landscape. The green and red curves in Figure 2 were produced by the finite temperature string method, which also gives the following estimates for the left−right and right−left mean transition times (i.e., the inverse of the rates) to switch between the regions inside the green curves in Figure 2:

$$t_{L\to R} = 2.7 \times 10^3, \quad t_{R\to L} = 4.4 \times 10^3 \tag{2}$$

In this simple example, these times can also be estimated directly by simulating (1); from observing about 10^6 left−right and right−left transitions, one obtains the following:

$$t_{L\to R} = 2.9 \times 10^3, \quad t_{R\to L} = 4.4 \times 10^3 \tag{3}$$

These are in very good agreement with the times in (2) since both (2) and (3) are tainted with statistical errors and the string method relies on the assumption that the isoprobability surfaces are planes, which is only approximately satisfied in the example. It shows that the string method is able to extract the essential features of the energy landscape which determine metastable

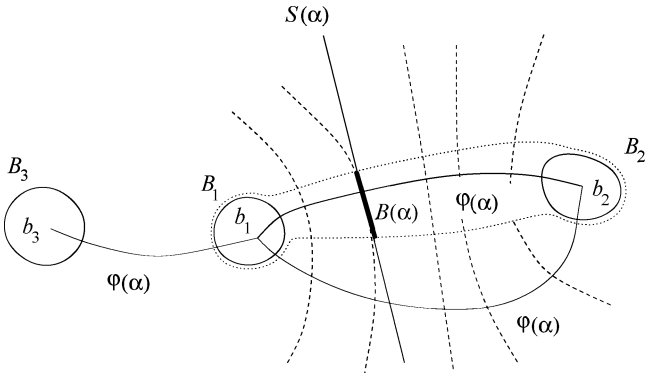

Figure 3. Schematic representation of a situation with three metastable sets, B_1, B_2, and B_3. The isoprobability surface for the transition between B_1 and B_2 are the dashed surfaces. The equilibrium distribution restricted to these surfaces is peaked in small subsets, the collection of which is the region inside the dotted curve: this is the effective switching tube, i.e., the tube inside which dynamical transition paths between B_1 and B_2 lie with high probability. The center of the tube is $\varphi(\alpha)$ and it is assumed that, inside the tube, the hypersurfaces associated with the reaction coordinate can be approximated by hyperplanes such as $S(\alpha)$. The equilibrium distribution restricted to the portion of $S(\alpha)$ relevant to this transition tube is peaked in $B(\alpha)$ (tick black segment). The portion of $S(\alpha)$ out of the tube is irrelevant; in practice, this external portion is not sampled anyway, because going out of $B(\alpha)$ by definition requires crossing barriers of a couple of k_BT at least. This guarantees that other switching tubes (like the one connecting B_1 and B_2 centered around the path $\varphi(\alpha)$ shown in gray) do not pollute sampling within the tube under consideration (since, if they did, the two tubes should really be considered as a single one at this temperature). For the situation as depicted, the method requires sampling successively the two effective switching tubes associated with the transition between B_1 and B_2, and the one associated with the transition between B_1 and B_3 (centered around the other path $\varphi(\alpha)$ shown in gray).

sets and effective transition tubes without having to resolve all the very complicated but mostly irrelevant details of the trajectories in the system.

3. The Method

We now introduce our method leaving the question of its theoretical justification to the next section. The method basically is a finite temperature generalization of the string method developed in refs 6 and 7. For simplicity of presentation, we assume here that there are only two metastable sets connected by a single effective transition tube. Extending this to more general situation with multiple metastable sets and tubes is rather straightforward and will be explained later (see also the caption in Figure 3).

Let $\varphi(\alpha)$ be a curve in configuration space parametrized by $\alpha = [0,1]$ whose end points, $\varphi(0)$ and $\varphi(1)$, belong to the two metastable sets (see the schematic representation in Figure 3). We wish to evolve $\varphi(\alpha)$ so that it converges toward the center of the effective transition tube. To this end we introduce an ensemble of realizations (or replicas) $\{\varphi^\omega(\alpha)\}$ whose mean is defined to be the string, i.e., $\langle\varphi^\omega(\alpha)\rangle \equiv \varphi(\alpha)$, and let these variables evolve by

$$\varphi_t^\omega = -(\nabla V(\varphi^\omega))^\perp + (\eta^\omega)^\perp + r\hat{t} \tag{4}$$

Here $\hat{t} = \varphi_\alpha/|\varphi_\alpha|$ is the unit tangent vector along φ and $a^\perp = a - (\hat{t}\cdot a)\hat{t}$ is the projection of the vector a in the hyperplane normal to the string $\varphi(\alpha)$, which we denote by $S(\alpha)$. η^ω is a white noise with covariance

$$\langle\eta_j^\omega(\alpha, t)\eta_k^\omega(\alpha', 0)\rangle = \begin{cases} 2k_BT\delta_{jk}\delta(t) & \text{if } \alpha = \alpha' \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

The scalar field $r \equiv r(\alpha, t)$ is a Lagrange multiplier term added in (4) to preserve some particular parametrization of the string φ chosen beforehand. For instance, a simple choice is the parametrization by normalized arclength. In this case, (4) must be supplemented by the constraint $(|\varphi_\alpha|)_\alpha = 0$ which determines r. Other parametrizations—for instance by energy-weighted arc length—can be straightforwardly implemented by modifying this constraint, as in the zero temperature string method.[6,7]

It can be checked that the equilibrium density function for (4) is given by

$$\rho(q, \alpha) = Z^{-1}(\alpha)e^{-\beta V(q)}\delta_{S(\alpha)}(q) \tag{6}$$

where $\delta_{S(\alpha)}(q)$ is the Dirac distribution concentrated on $S(\alpha)$, and $Z(\alpha) = \int_{S(\alpha)} e^{-\beta V(q)}d\sigma$ is a normalization constant. (6) allows us to characterize the effective transition tube, the free energy of the metastable states, and the associated transition rates.

1. Effective Transition Tube. By definition, the center of the transition tube is given by

$$\varphi(\alpha) = Z^{-1}(\alpha)\int_{S_{(\alpha)}} qe^{-\beta V(q)}\,d\sigma \tag{7}$$

Since $\rho(q, \alpha)$ is simply the equilibrium distribution of the system restricted to the plane $S(\alpha)$, one sees that the string $\varphi(\alpha)$ is the average position of the system within this plane. The width of the effective transition tube itself can be characterized by a few times the variance of q around $\varphi(\alpha)$; i.e., its local radius square can be defined as

$$R^2(\alpha) = \lambda Z^{-1}(\alpha)\int_{S(\alpha)} |q - \varphi(\alpha)|^2 e^{-\beta V(q)}\,d\sigma \tag{8}$$

where λ is a number of order unity (the effective transition tube in Figure 2 was defined with $\lambda = 3$). The integral of $\rho(q, \alpha)$ in the ball of radius $R(\alpha)$ centered around $\rho(\alpha)$ (denoted by $B(\alpha)$ in Figure 3) gives the probability that a dynamical trajectory involved in the transition event crosses the plane $S(\alpha)$ within this ball.

2. Free Energy. The planes $S(\alpha)$ (or more appropriately, their subsets $B(\alpha)$) are local approximations of the reaction coordinate for the conformational change. Therefore, the associated free energy is given by

$$F(\alpha) = -k_BT\ln\int_{S(\alpha)} e^{-\beta V(q)}\,d\sigma \tag{9}$$

The free energy necessarily has a minimum within each of the metastablesets B_1 and B_2 (because these are regions of high probability by definition), and at least one (sometimes more) maximum in between, whose value must be at least a couple of k_BT above the free energy at the metastable states (otherwise the transition would not be rare). Notice also that the free energy differences along $\varphi(\alpha)$ depend on the relative width of the tube, i.e., the variation in its radius $R(\alpha)$, which indicates that our method fully accounts for entropic effects − a high potential energy barrier around which the energy is rather flat may produce a lower free energy barrier.

In practice, (9) can be determined by standard thermodynamic integration.[5] Using the identity $\int F'(\alpha)\,d\alpha = F(\alpha)$, we obtain from (9) after integration by parts

$$F(\alpha_2) - F(\alpha_1) = \int_{\alpha_1}^{\alpha_2}\langle(\hat{t}\cdot\nabla V)\,((\hat{t}\cdot\varphi)_\alpha - \hat{t}_\alpha\cdot\varphi)\rangle\,d\alpha \tag{10}$$

This expression was first derived in ref 10. It is rather simple because we use the planes $S(\alpha)$; this eliminates the complicated Jacobian which arises when the sampling is performed in nonplanar surfaces.

3. Rates. In the high friction limit, the transition rates $k_{1,2}$ and $k_{2,1}$ from B_1 to B_2 and B_2 to B_1, respectively, are given by

$$k_{1,2} = (2\gamma\beta N_1\kappa)^{-1}, \quad k_{2,1} = (2\gamma\beta N_2\kappa)^{-1} \tag{11}$$

Here

$$N_1 = \int_{B_1} \mathrm{e}^{-\beta V(q)}\, \mathrm{d}q, \quad N_2 = Z^{-1} \int_{B_2} \mathrm{e}^{-\beta V(q)}\, \mathrm{d}q \tag{12}$$

with $Z = \int_{\mathbb{R}^d} e^{-\beta V(q)}\, \mathrm{d}q$, are the equilibrium probabilities within the sets B_1 and B_2, respectively, and

$$\kappa = \int_0^1 e^{\beta F(\alpha)} |\varphi_\alpha|\, \mathrm{d}\alpha \tag{13}$$

and (11) are further discussed below; these are the standard expressions for the rates of a one-dimensional diffusion process moving in the free energy potential $F(\alpha)$.

4. Theoretical Background

We now provide some theoretical background for the method proposed in the previous section. To this end, we characterize the effective transition tube in three ways, listed as C1, C2, and C3 below, which are equivalent but highlight different aspects of the method. We keep the discussion at a qualitative level here; the mathematical details have been presented elsewhere.[8]

C1. The curve at the center of the effective transition tube is the mean position of the system in the planes perpendicular to this curve. This follows from (7). Notice that C1 can be used as a definition for the minimum energy path (MEP) in a smooth energy landscape in the limit of zero temperature ($\beta \to \infty$); this can be easily seen by evaluating (7) by the Laplace method. Of course, we are not interested in smooth potentials and do not take the limit as $\beta \to \infty$, and C1 has to be thought of as the appropriate generalization of the MEP for dynamics over rough landscapes at finite temperature.

C2. The hyperplane $S(\alpha)$ is a local approximation around $\varphi(\alpha)$ of the isoprobability surfaces of the transition. Recall that if B_1 and B_2 are the metastable sets (corresponding, e.g., to the regions inside the green curves in Figure 2), the isoprobability surfaces of the transition are surfaces identified according to the following criterion: every trajectory initiated on one of these surfaces has the same fixed probability to reach B_1 before B_2. As already noted in refs 16 and 17, these isoprobability surfaces contain the correct information about the dynamics. They alone allow one to describe the dynamics of the rare transition events between B_1 and B_2, i.e., these surfaces constitute the right reaction coordinate for these transitions. Furthermore, because our method identifies the effective transition tube, i.e., the ball $B(\alpha)$ within $S(\alpha)$ which has a significant probability to be traversed by the trajectories involved in the transition, it automatically identifies the local portions of the isoprobability surface that are relevant because they have a high probability to be visited (recall that while the probability to reach B_1 before B_2 from any point on the surface is constant by definition, the probability to get to a point on this surface from either B_1 or B_2 is not uniform).

The effective transition tube allows us to identify the transition state region to be the region on the tube where the free energy is within k_BT of the maximum free energy along the tube. Indeed, by definition this region is also the one which contains the isoprobability surfaces with probability close to $^1/_2$. It is very important to notice that the transition state region may be quite wide, in which case the transition state theory[21,22] will not apply. For instance, in the example considered before, the entire light-blue/yellow channel is the transition state region. A wide transition state region is related to a high recrossing rate of any dividing surface.

C3. Among all possible choices, the path $\varphi(\alpha)$ satisfying (7) minimizes the transition rates. Equivalently, it maximizes κ given in (13). This characterization of $\varphi(\alpha)$ (or equivalently, of the plane $S(\alpha)$, i.e., the reaction coordinate) is not really surprising. A similar feature is observed in the context of TST,[2] where the dividing surface can actually be optimized by minimizing the TST rate it produces (this property was used in ref 10 to identify an optimal dividing plane). Of course, the present rate minimization principle is much more general than the one arising in TST because the expressions in (11) have a wider range of validity as we already made clear. This last characterization of $\varphi(\alpha)$ is not used explicitly in our method, but it may be taken as a starting point for generalizing our method if necessary (for instance by considering reaction coordinates which are not locally planar if this approximation were to break down).

5. Numerical Implementation

In practice, (4) is solved by considering M realizations of φ^ω, $\{\varphi^j\}_{j=1}^M$, and approximating the string as

$$\varphi = M^{-1} \sum_{j=1}^{M} \varphi^j \tag{14}$$

To integrate (4) we use a time-splitting scheme which takes advantage of the intrinsic description of the string. Each φ^j is discretized into a number, say N, of points which move according to the first two terms, $-(\nabla V(\varphi^\omega))^\perp + (\eta^\omega)^\perp$, at the right-hand-side of (4). After a number of steps determined by the accuracy requirement for the constraint $(|\varphi_\alpha|)_\alpha = 0$, a reparametrization step is applied to reinforce this constraint. This costs $O(N)$ operations. At the reparametrization step, it is also convenient to change N according to the accuracy requirement for the representation of the string. Once φ has converged to its steady-state value, no reparametrization step is necessary anymore and, using ergodicity, one can supplement the ensemble average by a time average using $\varphi = (MT)^{-1} \sum_{j=1}^M \int_0^T \varphi^j(t)\, \mathrm{d}t$ to obtain better statistics. Other averages such as (10) are evaluated similarly. Notice that once the string has converged and the hyperplanes do not need to be updated, the method becomes very similar to a standard sampling of blue moon type.

A major advantage of the method is that it uses as a subroutine anysolver for the original equations of motion in the system; the additional steps are to compute the tangent along the string and restrict the force field to the hyperplane perpendicular to string and from time to time to redistribute the images along the string to preserve parametrization. Therefore, the overall computational cost scales almost linearly with the number of images, $M \times N$, used to represent the string and the ensemble, and the computation can easily be parallelized. Notice also that the method can be used in an adaptive way where a small number of images is used at the beginning of the calculation, then progressively increase the number of images as the string converges toward its final steady value. This procedure saves computational cost at the early stages of the calculation where

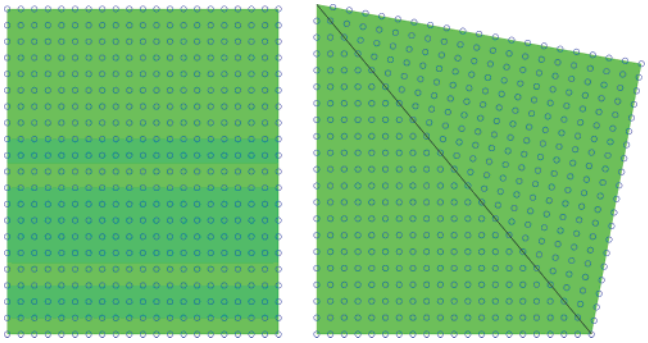

Figure 4. Initial (left) and final (right) states of the two-dimensional crystal. The diagonal line is the twin boundary which separates the two martensitic phases in the final state.

the string (and the associated hyperplanes) have not yet converged to the actual reaction coordinate of the transition.

For situations with more than one effective switching tubes between twometastable sets, each of these tubes must be obtained separately for a complete description of the transition events between these sets. In practice, this can be done by using an annealing procedure. The simulation is started at an artificially high temperature where the various effective switching tubes have clustered into a single one. The temperature is then slowly decreased, and the relevant tubes are identified as they become distinct. It is worth pointing out that, in the later stage of the computation where the temperature is back to the physical temperature, the presence of different switching tubes do not lead to practical difficulties as far as sampling is concerned, in the sense that they do not pollute each other. Indeed, along the reaction coordinate, these tubes are by definition separated by energy barriers of a couple of k_BT at least (since, otherwise, they should really be considered as a single tube); therefore, when one samples within the subset $B(\alpha)$ in the plane $S(\alpha)$ as defined earlier, one almost never observes a transition along $S(\alpha)$ to other regions where the plane may intersect another tube (see Figure 3), and in any case, these spurious transitions within $S(\alpha)$ can be rather easily disregarded by monitoring the simulation.

6. Example of Solid−Solid Transformation

This example illustrates the complex energy landscapes for systems in condensed phases. We consider a two-dimensional crystal of 20 × 20 arrayof atoms interacting via a pairwise potential which has three minima at $r_1 = 1$, $r_2 = 1.2$, and $r_3 = \sqrt{1+1.2^2}$. Assuming that the left and bottom sides of the crystal are fixed to some substrate, the two basic equilibrium states for this system are shown in Figure 4. These two states can be considered as the two variants of the martensitic phase, and we are interested in transformation from one variant to the other.

The finite temperature string method was used to analyze this transformation (see Figure 5). A movie of the transformation can been seen on.[23] From the atomic configurations along the string, i.e., the center of the effective switching tube, one sees that the crystal transforms by propagation of a twin boundary; the motion of this twin boundary by one atomic distance involve the motion of a twin dislocation associated with the phase transformation of individual atoms (upper panel in Figure 5). This transition path is preferred because the energy is concentrated for the most part at the dislocation (see the scale at the right of the upper panel in Figure 5). The energy landscape exhibits three scales (middle and lower panels in Figure 5). The

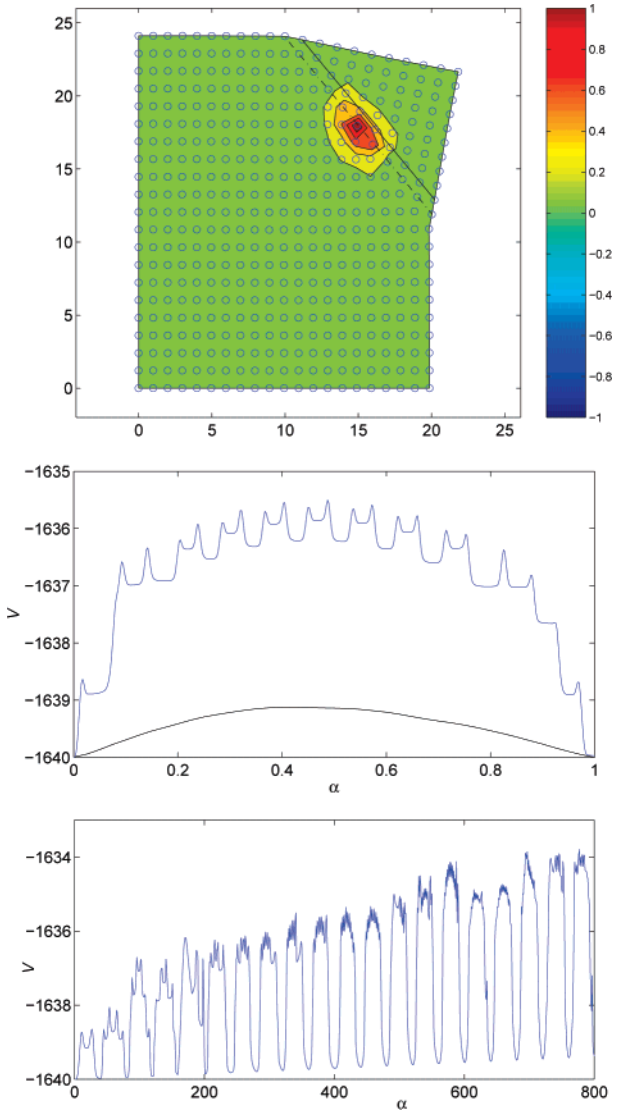

Figure 5. Upper panel: Snapshot of the crystal during transformation. The color bar shows the scale of the local energy of each atom; the region with higher energy corresponds to the twin dislocation. Middle panel: The free energy experienced by the crystal at finite temperature when the twin boundary moves by one atomic distance along the diagonal (black curve). Also shown for comparison is the potential energy along one particular MEP (blue curve, at zero temperature) which shows the features of the potential associated with the transformation of individual atoms along the twin boundary. Notice how much lower is the free energy barrier compared to the potential energy barrier, which can be attributed to the importance of entropic effects in this model. Lower panel: The energy landscape after the crystal is half transformed. Notice the appearance of three scales on the energy landscape.

largest scale associated with the position of the twin boundary, an intermediate scale associated with the propagation of the twin boundary by one atomic distance, and a small scale associated with the twin dislocation propagation. We choose k_BT to be between the energy barrier for twin dislocation propagation and twin boundary propagation. At this temperature, the barriers associated with the motion of the twin dislocation are irrelevant, and the metastable states along the way toward full transformation are the 20 states corresponding to the possible positions of the twin boundary, and no dislocation along it; the transition state between two successive metastable states where the twin boundary has advanced by one atomic distance is a wide region in phase space, where the twin boundary presents a dislocation along itself (the main energy barrier corresponding to creation

of such a dislocation). Notice that this region is not only wide along the string, but also in the direction transverse to it (i.e., the effective transition tube is wide); this can be deduced from the fact that the free energy barrier is so much lower than the energy barrier between two successive positions of the twin boundary; i.e., relative entropic effects are important (see the middle panel in Figure 5).

To obtain these results, the simulation was initiated with a string which interpolates linearly between the initial and final states shown in Figure 4. The calculation is started with few images, then refined along the way. At the end, 101 realizations were used, and 40 discretization points per motion of the twin boundary by one atomic distance were necessary. For comparison, we also computed one MEP associated with this transition using the zero temperature string method; this required 400 discretization points permotion of the twin boundary by one atomic distance. This example therefore makes apparent several advantages of the method. It only requires sampling the equilibrium distribution of the system in appropriate subsets in configuration space where, by definition, there is not any free energy barrier much larger than k_BT. This sampling was performed here using the Langevin dynamics in (4), but it could just as well have been done using standard techniques of constrained molecular dynamics, and the absence of significant barrier in each of the sampling windows guarantees that the sampling converges fast. Notice that this allows one to bypass completely any unconstrained dynamical simulation which is necessarily slow in systems with characteristics similar to our example, not only because the time between transitions is very large but also because the transitions themselves involve rather slow and complicated trajectories. In the present example, the trajectory of the twin dislocation during the propagation of the twin boundary by one atomic distance is actually very complicated, with the dislocation moving back and forth along the twin boundary in an essentially diffusive way (which is reminiscent of the behavior of the trajectories in the toy example; see the blue curves in Figure 2).

It is also important to stress that the computational effort of the methodis not significantly different from that of other umbrella or constrained sampling methods, such as blue moon sampling. The only additional cost comes from the transient stage of the simulation when the hyperplanes are updated so as to converge to local approximation of the isoprobability surface.

7. Concluding Remarks

The method we have presented is a powerful tool for determining effective transition pathways, free energy profiles, and transition rates in complex systems with multiscale energy landscapes. The main assumption the method relies on is that the equilibrium distribution be localized on the isoprobability surfaces used as reaction coordinate and that these isoprobability surfaces can be locally approximated by hyperplanes. When this assumption is satisfied, the method allows us to determine self-consistently these hyperplanes by evolving a smooth curve which evolves toward the center of a tube by which transitions occur with high probability. The method improves upon the sampling techniques of the blue moon type since it does not require specifying a reaction coordinate beforehand. It is very similar to blue moon sampling at the later stage of the computation when the hyperplanes approximating the reaction coordinate have been determined. For this reason, it carries almost the same computational cost as constrained sampling technique of the blue moon type.

Finally, we must note that in certain situations, the assumption that the reaction coordinates are locally planar will be violated—for instance if the effective transition tube is highly curved in configuration space. For such situations, the finite temperature string method will have to be modified. A possibility is to use C3 above as a starting point for determining self-consistently nonplanar reaction coordinates, but this will definitely constitute a formidable numerical challenge. On the other hand, less ambitious solutions may be found for this problem on a case by case basis. It may amount for instance, to appropriately redefining the problem in curvilinear coordinates in which the reaction coordinate is locally planar again, or applying the method to a smaller sets of coarse-grained variables suitably defined. The latter solution may also render the method practical in larger systems where a full computation with the original degrees of freedom is too expensive.

Acknowledgment. We thank David Chandler for his suggestions to improve the presentation of this paper. We are also grateful to Giovanni Ciccotti and Bob Kohn for helpful discussions. The work of W.E is supported in part by NSF Grant DMS01-30107. The work of E.V.-E. is supported in part by NSF Grant DMS01-01439, DMS02-09959 and DMS02-39625. The work of W.R. is supported in part by NSF Grant DMS97-29992.

Supporting Information Available: A movie of the martensitic transformation (gif file). This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Bennett, C. H. **1975** In *Diffusion in Solids: Recent Developments*; Nowick, A. S., Burton, J. J., Eds.; Academic Press: New York, 1975.

(2) Chandler, D. *J. Chem. Phys.* **1978**, *68*, 2959−2970.

(3) Carter, E. A.; Ciccotti, G.; Hynes, J. T.; Kapral, R. *Chem. Phys. Lett.* **1989**, *156*, 472−477.

(4) Ciccotti, G. In *Proceedings of the NATO ASI on Computer Simulation in Materials Science*; Meyer, M., Pontikis, V., Eds.; Kluwer: Dordrecht, The Netherlands, 1991.

(5) Frenkel, D.; Smit, B. **1996** *Understanding Molecular Simulation*; Academic: San Diego, CA, 1996.

(6) E, W.; Ren, W.; Vanden-Eijnden, E. *Phys. Rev. B.* **2002**, *66*, 052301.

(7) E, W.; Ren, W.; Vanden-Eijnden, E. *J. Appl. Phys.* **2003**, *93*, 2275−2282.

(8) E, W.; Vanden-Eijnden, E. In *Multiscale Modelling and Simulation*; Attinger, S., Koumoutsakos, P., Eds.; LNCSE 39; Springer: Berlin, 2004.

(9) Straub, J. E. In *Computational Biochemistry and Biophysics*; Becker, O. M., et al., Eds.; Marcel Dekker: New York, 2001.

(10) Schenter, G. K.; Mills, G.; Jónsson, H. *J. Chem. Phys.* **1994,** *101*, 8964.

(11) Olender, R.; Elber, R. *J. Chem. Phys.* **1996,** *105*, 9299−9315.

(12) Ruiz-Montero, M. J.; Frenkel, D.; Brey, J. *J. Mol. Phys.* **1997**, *90*, 925−941.

(13) Voter, A. F. *Phys. Rev. Lett.*. **1997**, *78*, 3908−3911.

(14) Henkelman, G.; Jóhannesson, G.; Jónsson, H. In *Progress on Theoretical Chemistry and Physics*; Schwartz, S. D., Ed.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 2000.

(15) Sorensen, M. R.; Voter, A. F. *J. Chem. Phys.* **2000**, *112*, 9599−9606.

(16) Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. *Annu. Rev. Phys. Chem.* **2002**, *59*, 291.

(17) Dellago, C.; Bolhuis, P. G.; Geissler, P. L. *Adv. Chem. Phys.* **2002**, 123.

(18) Passerone, P.; Parrinello, M. *Phys. Rev. Lett.* **2001**, *87*, 108302−108305.

(19) Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12562−12566.

(20) Pratt, L. R. *J. Chem. Phys.* **1986**, *85*, 5045−5048.

(21) Eyring, H. *J. Chem. Phys.* **1935**, *3*, 107.

(22) Wigner, E. *Trans. Faraday Soc.* **1938**, *34*, 29.

(23) Movies of the martensitic transformation are available on:http://www.cims.nyu.edu/_weiqing/.