A skewed-momenta method to efficiently generate conformational-transition trajectories

James MacFadyen and Ioan Andricioaei[a)]
Department of Chemistry and The Program in Bioinformatics, University of Michigan, Ann Arbor, Michigan 48109

(Received 6 May 2005; accepted 21 June 2005; published online 24 August 2005)

We present a novel computational method, the skewed-momenta method (Skew'M), which applies a bias to the Maxwell distribution of initial momenta used to generate ensembles of trajectories. As a result, conformational transitions are accentuated and kinetic properties are calculated more effectively. The connection to the related puddle jumping method is discussed. A reweighting scheme permits the exact calculation of kinetic properties. Applications are presented for the rapid calculation of rate constants for molecular isomerization, and for the efficient reconstruction of free-energy profiles using a straightforward modification of the Jarzynski identity. © *2005 American Institute of Physics*. [DOI: 10.1063/1.2000242]

I. INTRODUCTION

There has been considerable recent interest in using approaches that generate ensembles of dynamical paths to calculate kinetic properties of conformational transitions[1,2] or to reconstruct entire free-energy profiles.[3] These approaches are expected to be useful, in particular, for large-dimensional systems (such as proteins or nucleic acids) because they need not calculate saddle points, which are exponentially many in the number of degrees of freedom. The fundamental object for these types of calculations is an average of an observable A over an infinite ensemble of trajectories that are initiated from the equilibrium distribution of phase-space vectors $\mathbf{\Gamma}=(\mathbf{q},\mathbf{p})$, and which evolve up to time t according to a specific dynamical flow, which may or may not conserve an equilibrium distribution. The observable A can be a function of the end points $\mathbf{\Gamma}_0$ and $\mathbf{\Gamma}_t$ or a functional of the entire trajectory leading to $\mathbf{\Gamma}_t$:

$$C(t)=\langle A[\mathbf{\Gamma}(t)]\rangle, \tag{1}$$

where $\langle\cdots\rangle$ indicates an average over the trajectory ensemble.

For the purposes of our method we wish to rewrite this trajectory average as an average over the initial distribution, which is assumed to be in equilibrium. If the system evolves according to deterministic (e.g., Hamiltonian) dynamics, each trajectory is uniquely determined by its initial point and Eq. (1) can be written without modification as an average over the canonical phase-space distribution.

On the other hand, if the system evolves according to some stochastic scheme, each initial point can lead to a multitude of trajectories. We note, however, that so long as each trajectory is initiated from an equilibrium distribution, Eq. (1) can still be rewritten as an average over the initial distribution:

$$C(t)=\langle Q(\mathbf{\Gamma}_0;t)\rangle_0, \tag{2}$$

where $Q(\mathbf{\Gamma}_0;t)$ is the average of the observable A over all realizations of the dynamics of duration t beginning from the initial phase-space point $\mathbf{\Gamma}_0$. We have replaced $A[\mathbf{\Gamma}(t)]$, a functional of the trajectory, with a function of the initial conditions. The notation $\langle\cdots\rangle_0$ indicates an average over the equilibrium distribution at $t=0$.

In the particular case when a reaction-rate constant is sought, the average in Eq. (1) takes the form of a correlation function representing an ensemble-averaged conditional probability,

$$C(t)\equiv\frac{\langle h_{\mathcal{R}}(\mathbf{\Gamma}_0)h_{\mathcal{P}}(\mathbf{\Gamma}_t)\rangle_0}{\langle h_{\mathcal{R}}(\mathbf{\Gamma}_0)\rangle_0} \tag{3}$$

with

$$h_{\mathcal{R},\mathcal{P}}(\mathbf{\Gamma})=\begin{cases}1 & \text{if } \mathbf{\Gamma}\in\mathcal{R},\mathcal{P}\\ 0 & \text{otherwise,}\end{cases} \tag{4}$$

where $\mathcal{R}$ and $\mathcal{P}$ stand for the reactant and product macrostates, respectively. The dynamics of the system are assumed to be deterministic so that $\mathbf{\Gamma}_t$ is uniquely determined by $\mathbf{\Gamma}_0$. $C(t)$ is properly a ratio of two separate averages as defined by Eq. (2); however, for the purposes of our method this distinction is immaterial.

The correlation function in Eq. (3) is in effect related to the exponential of the *reversible* work (i.e., the free energy) involved in confining the phase space trajectories initiated from R to region P at time t: $C(t)=\exp(-\Delta W_{\mathrm{RP}}^{\mathrm{rev}}(t)/k_BT)$.[1] For times smaller than the equilibration time of the system but greater than the time required for transient behavior to relax, $C(t)$ is a linear function and its time derivative corresponds to the forward rate constant $k_{\mathcal{R}\to\mathcal{P}}$.[4]

Another particular correlation function of recent interest involves averages over fast, irreversible trajectories with a nonconservative flow obtained by augmenting the dynamics with a time-dependent pulling potential. Based on the Jarzynski identity,[5] this strategy reconstructs free-energy pro-

[a)]Electronic mail: andricio@umich.edu

files along a given reaction coordinate $z(\Gamma(t))$ from single-molecule pulling experiments.[3] The pulling potential V is assumed to depend explicitly upon only the pulling coordinate z and the time t. This leads to the second particular form of Eq. (1) that we shall employ in the present paper:

$$C(t) \equiv \exp(-\beta F(z)) = \langle \delta(z - z(\mathbf{\Gamma}(t))) e^{-\beta(W_t - V(z(\mathbf{\Gamma}(t),t))} \rangle, \tag{5}$$

where, in contrast to Eq. (3), the average is over an ensemble of *trajectories* with initial points canonically distributed on the time-dependent potential at $t=0$. We note, however, that we can recast this average in the form of Eq. (2):

$$\exp(-\beta F(z)) = \langle Q_t(\mathbf{\Gamma}_0) \rangle_0, \tag{6}$$

where $Q_t(\mathbf{\Gamma}_0)$ is the average of $\delta(z-z_t)\exp(-\beta(W_t-V_t))$ over all trajectories of length t initiated from the initial point $\mathbf{\Gamma}_0$.

W_t is the *irreversible* work done along a particular trajectory in time t, defined by $W_t \equiv \int_0^t d\tau [\partial H((\mathbf{\Gamma}(\tau), \tau) / \partial \tau]$. Because the system's Hamiltonian evolves at a finite rate, the dynamical flow of the system no longer preserves an equilibrium distribution. Moreover, W_t here is a functional of the trajectory, rather than a function of the end points.

Equations (3), (5), and (6) are the central equations for the two correlation functions that constitute the focus of the numerical method we present herein. Although their calculations do not require foreknowledge of the energy landscape, a significant computational burden still remains because in complex systems one expects a multitude of reaction paths to contribute to the average.

In the case of Eq. (3), if too few of the sampled paths reach the product state, the calculation of $C(t)$ will not converge. This issue is common in systems with barriers that are large compared with $k_B T$, in which the fraction of trajectories started in the reactant macrostate which reach the product state is vanishingly small. This is in effect an enthalpic barrier problem. For many-dimensional systems, there also exists an entropic barrier problem,[6] i.e., the paths that reach the product region are not expected to cover the path space uniformly, and finding the right ones is expected to be exponentially hard in the dimensions of the path space.[7]

A similar sampling problem occurs in the case of Eq. (5); although low-work trajectories contribute most in the formula, they are rarely sampled. This issue is especially prevalent in the case of free-energy profiles found from computational pulling simulations, which we consider in this paper. In order to surmount the barriers typically found in such profiles, a guiding potential moving at speeds at which the trajectories are far from reversible typically performs significantly more work than a reversible trajectory, so that the average in Eq. (5) converges slowly. In most instances, a smaller number of slow-pulling trajectories will provide a more accurate estimate of the potential of mean force.[8] However, in numerical analyses, the time required for a specific estimate varies linearly with the speed of the pulling simulation but not with the number of trajectories, as easy parallelization allows for the evaluation of additional trajectories without increasing the simulation time. Therefore, the most efficient estimate of a free-energy surface may, in fact, be found from many fast trajectories rather than a few slower ones.

In order to surmount the difficulties presented above, we introduce a skewed-momenta (Skew'M) method, in which we calculate the quantities given in Eqs. (3) and (5) as weighted averages over initial phase-space distributions in which the momenta are "skewed" to bias the dynamics along certain directions. In the calculation of $C(t)$ in Eq. (3), we thereby produce trajectories that can surmount entropic and enthalpic barriers, increasing accuracy of the estimate; in the calculation of potentials of mean force from Eq. (5), we produce low-work trajectories that provide a more accurate estimation of the underlying free-energy surface.

The paper is organized as follows. We start with a background of existing methods that alter the initial distributions in the "reactant" basin focusing, in particular, on the puddle jumping method of Tully and coworkers,[9,10] which constitutes the inspiration for the Skew'M method, developed in Sec. III. We continue with a description of our method as applied to Eqs. (3) and (5), with numerical examples for each case. We end with a concluding discussion.

II. BACKGROUND

In a recent development, Corcelli *et al.*[10] have introduced a convenient bias function of general applicability and of promise to accelerate the convergence of rate calculations in systems with large enthalpic barriers. They apply a "puddle" potential (used previously by the same group to enhance thermodynamic averaging[9]) that changes the potential-energy surface from which the trajectories are initiated to become

$$V^*(\mathbf{q}) = \begin{cases} V(\mathbf{q}) & \text{if } V(\mathbf{q}) \geqslant V_{\text{pud}} \\ V_{\text{pud}} & \text{otherwise,} \end{cases} \tag{7}$$

and the correlation function in Eq. (3) can be written, without approximation,

$$C(t) = \frac{\int d\mathbf{\Gamma}_0 \rho^*(\mathbf{q}_0)\rho(\mathbf{p}_0)w(\mathbf{q}_0)h_R(\mathbf{\Gamma}_0)h_P(\mathbf{\Gamma}_t)}{\int d\mathbf{\Gamma}_0 \rho^*(\mathbf{q}_0)\rho(\mathbf{p}_0)w(\mathbf{q}_0)h_R(\mathbf{\Gamma}_0)} = \frac{\langle h_R(\mathbf{\Gamma}_0)h_P(\mathbf{\Gamma}_t)w(\mathbf{q}_0)\rangle^*}{\langle h_R(\mathbf{\Gamma}_0)w(\mathbf{q}_0)\rangle^*}, \tag{8}$$

where $w=\exp[\beta(V^*(\mathbf{q}_0)-V(\mathbf{q}_0))]$, $\rho^*(\mathbf{q}_0)$ is the equilibrium spatial distribution on V^*, $\rho(\mathbf{q}_0)$ is the equilibrium distribution of momenta, and $\langle \cdots \rangle^*$ indicates that the average is over the equilibrium distribution corresponding to $V^*(\mathbf{q})$.

The puddle potential removes the deep energy minima that would ordinarily dominate the initial distribution; trajectories from these deep minima have little chance of crossing a large barrier into the product region. This strategy bears resemblance to the hyperdynamics method of Voter,[11,12] in which the bottom of the potential wells are raised without affecting the barrier tops, and to the related accelerated dynamics method developed by Hamelberg *et al.*[13] In the same category of approaches are the methods of Laio and

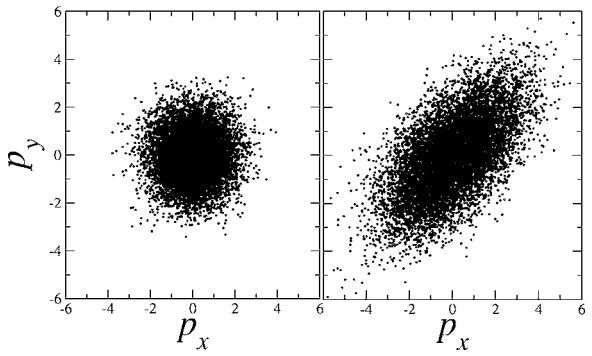

FIG. 1. A spherical and skewed-momentum distribution in two-dimensional momentum space.

Parrinello,[14] of Huber and Kim,[15] as well as Grubmuller's "conformational flooding" method.[16] In an even earlier reference, the strategy bears resemblance, to some extent, with that used by Carter *et al.*[17] to generate a constrained "blue-moon" ensemble from which one originates free trajectories to obtain rates from correlation functions. Most recently, Mella[18] extended the approach of Corcelli *et al.* to momentum space by drawing momenta from a Boltzmann-type initial distribution with an artificially high temperature.

Because the strategy of modifying the initial phase-space distribution proved to be effective in accelerating the convergence of rate constant calculations, we thought to explore further modifications that might improve the same calculation. To this end, we propose here a novel modification of the method of Corcelli *et al.*[10] Instead of applying a puddle potential or sampling momenta at a higher temperature, we skew the momentum distribution along certain directions in conformational space (see Fig. 1). These directions are to be chosen to correspond to the local slow manifold, which is the conformational subspace in which the natural dynamics of the systems evolves more slowly than in the rest of the space. By increasing the probability to sample initial momenta that have large-magnitude components in the slow manifold, the subsequent relaxation dynamics is accelerated relative to that of the equilibrium distribution. In a procedure similar to the Corcelli *et al.* method, the trajectories are reweighted and the appropriate correlation function is eventually recovered.

Unlike any of the methods mentioned above, the presently proposed skewed-momenta (Skew'M) method involves accentuating the dynamics only along particularly relevant directions. A distinct feature of the present method is also that proper reweighting permits calculation of the exact underlying kinetics. Additionally, because the puddle method designs changes in the *potential* energy (i.e, modifies the coordinate distribution), while the proposed method alters the *kinetic* energy (i.e., modifies the momentum distribution), the latter method can be trivially combined with the former to enable a more efficient exploration of the entire phase-space dynamics.

Our Skew'M method has the additional advantage of being easily applicable to high-dimensional systems. The method of Corcelli *et al.* is efficient in such systems only if the puddle can be selectively applied across a few pertinent degrees of freedom. By contrast, the Skew'M method can be applied without modification to trajectories involving concerted changes in many degrees of freedom.

In addition to accelerating trajectories over enthalpic barriers, our method can also accelerate the convergence in high-dimensional systems with entropic barriers: the momentum distribution can be chosen to encourage motion along a particular direction while discouraging motion in directions that might lead the trajectories away from the states of interest.

As another application of our Skew' M method, we study the reconstruction of free-energy profiles using the Jarzynski identity. While at first sight this is a problem distinct from that of calculating correlations such as those in the central formula Eq. (3), Jarzynski's identity can be cast in terms of an equilibrium average, Eq. (6), as explained in the Introduction. We can then bias the dynamics to follow the motion of the pulling potential, enhancing our sampling of the low-work tail of the work distribution and thereby increasing the accuracy of our calculation.

III. THE SKEW'M METHOD

Our method relies upon the identification of a $3N$-dimensional vector in configuration space $\hat{\mathbf{e}}_s$, which points along a favored direction for the motion of the system. We then choose the initial momenta for our trajectory ensemble from a Gaussian distribution artificially extended in the direction of $\hat{\mathbf{e}}_s$, as illustrated in the right panel of Fig. 1. In the case of free-energy reconstructions from Eq. (6), we often wish to induce the motion along a predefined pulling direction, in which case $\hat{\mathbf{e}}_s$ can be found from inspection. Identifying $\hat{\mathbf{e}}_s$ is more difficult in the application of our method to the calculation of rate constants from Eq. (3); we postpone a more involved discussion until Sec. III D.

Because it is cumbersome to generate a Gaussian distribution oriented along an arbitrary axis in the natural coordinates of $\mathbf{p}$, we generate our momenta in a rotated coordinate system $\mathbf{p}'$ in which $\hat{\mathbf{e}}_s$ lies along the p_1' axis and then rotate the coordinate axes to place the generated momenta in the original frame. To this end, we seek a rotation matrix R that rotates $\hat{\mathbf{e}}_s$ onto the p_1 axis; this matrix will transform momenta generated in the $\mathbf{p}'$ system to the natural coordinates of the system. We present a detailed algorithm for the calculation of R in Sec. III B.

In the case of Eq. (6), the situation is once again simplified by prior knowledge of the pulling direction: $\hat{\mathbf{e}}_s$ can be taken without loss of generality to lie along one of the natural Cartesian axes of the system, so that the $\mathbf{p}$ and $\mathbf{p}'$ systems are equivalent and no axis rotations need be performed.

A. Skewed momenta and reweighting

At equilibrium, the components of the the momentum vector $\mathbf{p}$ are drawn from the distribution

$$\rho(\mathbf{p}) = C \exp(-\mathbf{p}^T A \mathbf{p}), \tag{9}$$

where C is a normalization constant and A the $3N \times 3N$ diagonal matrix $A = 1/2\beta M^{-1}$, in which $\beta = 1/k_B T$ and M is the mass matrix.

We choose instead from a biased distribution ρ_B defined in the $\mathbf{p}'$ coordinate system as

$$\rho_B(\mathbf{p}) = C' \exp(-\mathbf{p}'^T A' \mathbf{p}'), \tag{10}$$

where C' is another normalization constant and A' is a diagonal matrix with entries A'_i proportional to the variance of the Gaussian distribution along the ith axis of the rotated coordinate system. Since we wish to bias our dynamics along $\mathbf{p}_0$, A'_1 should be the longest axis, but there is in general no restriction on the diagonal entries of A' save that they be positive. The $\mathbf{p}'$ are related to $\mathbf{p}$ via $\mathbf{p}=R\mathbf{p}'$, where R is the rotation matrix introduced above.

We write the average in Eq. (2) in the exact factorized form

$$C(t) = \frac{\int d\mathbf{\Gamma}_0 \rho(\mathbf{q}_0)\rho_B(\mathbf{p}_0)w(\mathbf{p}_0)Q_t(\mathbf{\Gamma}_0)}{\int d\mathbf{\Gamma}_0 \rho(\mathbf{q}_0)\rho_B(\mathbf{p}_0)w(\mathbf{p}_0)}, \tag{11}$$

in which $\rho(\mathbf{q}_0)$ is the equilibrium spatial distribution, $\rho_B(\mathbf{p}_0)$ is defined in Eq. (10), and the weighting function $w(\mathbf{p}_0) = \exp(\mathbf{p}_0'^T A' \mathbf{p}'_0 - \mathbf{p}_0^T A \mathbf{p}_0)$. $Q_t(\mathbf{\Gamma}_0)$ is defined as before as the average of some observable $A[\mathbf{\Gamma}(t)]$ over all trajectories of length t initiated from the initial point $\mathbf{\Gamma}_0$.

$C(t)$ can then be calculated as a weighted average of $Q_t(\mathbf{\Gamma}_0)$ over the biased momentum distribution:

$$C(t) = \frac{\langle Q_t(\mathbf{\Gamma}_0)w(\mathbf{p})\rangle_B}{\langle w(\mathbf{p})\rangle_B}, \tag{12}$$

where the notation $\langle\cdots\rangle_B$ indicates that the momenta in the calculation are drawn from the biased distribution.

We emphasize that the average represented by $Q_t(\mathbf{\Gamma}_0)$ is introduced in the interest of theoretical development only. In numerical simulations, $C(t)$ is found by averaging over the observable $A[\mathbf{\Gamma}(t)]$ directly. That is,

$$C(t) = \frac{\sum_i A_t^{(i)} w^{(i)}}{\sum_i w^{(i)}}, \tag{13}$$

where $A_t^{(i)}$ and $w^{(i)}$ are the values for the ith trajectory (with initial momenta drawn from the biased distribution) of the observable and the weighting factor, respectively.

A two-dimensional example of a spherical and a skewed-momentum distribution is displayed in Fig. 1; for this simple case,

$$\hat{\mathbf{e}}_s = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix},$$

and the variance along $\hat{\mathbf{e}}_s$ has been extended.

B. Rotations

A general rotation in $3N$ dimensions can be decomposed to into a series of main rotations around $3N-2$ invariant axes. Duffin *et al.*[19] present a general matrix for main rotations

$$R_{a,b}(\theta) = r_{ij} = \begin{cases} \cos(\theta) & \text{if } i=j=a \text{ or } i=j=b \\ \sin(\theta) & \text{if } i=a, j=b \\ -\sin(\theta) & \text{if } i=b, j=a \\ 1 & \text{if } i=j, j\neq a, j\neq b \\ 0 & \text{otherwise,} \end{cases} \tag{14}$$

which rotates the axis p_a in the direction of p_b by an angle θ. There are in general many series of main rotations which would place p'_1 along p_1. We choose, following Aguilera and Pérez-Aguila,[20] to write R as the product of $3N-1$ main rotation matrices such that the ith matrix rotates the vector component along the ith axis onto the i—1st axis, beginning with the $3N$th axis and moving iteratively through all axes:

$$R = R_{2,1}(\theta_1)R_{3,2}(\theta_2)\cdots R_{i,i-1}(\theta_i)\cdots R_{3N,3N-1}(\theta_{3N-1}). \tag{15}$$

The θ_i are identified by performing this series of rotations upon $\hat{\mathbf{e}}_s$, identifying the angle as $\arctan 2(e_i, e_{i-1})$, where e_i is the component of $\hat{\mathbf{e}}_s$ along the ith axis *after* the vector has been rotated for all axes greater than i.[19] The well-known function arctan 2 is defined as

$$\arctan 2(y,x) = \begin{cases} \arctan(y/x) & x>0 \\ \arctan(y/x)+\pi & x<0 \\ \pi/2 & x=0, y\geqslant 0 \\ -\pi/2 & x=0, y<0. \end{cases} \tag{16}$$

Note that the individual main rotation matrices differ only slightly from the identity matrix. Therefore, although the multiplication of square matrices ordinarily requires $O(n^3)$ operations, the product of a main rotation matrix and an arbitrary matrix of the same size requires only $O(n)$, making the expression for R in Eq. (15) $O(n^2)$, where n is the number of degrees of freedom in the system.

C. Choosing A'

To this point we have said nothing about the other entries of the matrix A', which we have defined implicitly in calculating the rotation matrix R. Although Eq. (12) is exact for any momentum distribution in the rotated reference frame, in practical implementations it will often be desirable to choose the entries of A' so that the momenta in the directions perpendicular to $\hat{\mathbf{e}}_s$ are similar to their equilibrium values, thereby minimizing their contribution to the scaling factor. However, in systems in which different degrees of freedom have different masses, the momentum-space shell of constant kinetic energy will be a high-dimensional ellipsoid whose axes may not align with the axes of the rotated reference frame. In such systems it is convenient to work in mass-weighted coordinates, in which the equilibrium shell is spherical. That is, define

$$\pi_i \equiv \frac{p_i}{\sqrt{m_i}}, \tag{17}$$

where m_i is the mass of the ith degree of freedom and p_i is the unweighted momentum. In these coordinates, the equilibrium matrix entries are simply $A_i=\beta/2$, and the equilibrium distribution can be reproduced in any rotated frame by choosing $A'_i=A_i=\beta/2$. The desired bias along $\hat{\mathbf{e}}_s$ can then be

obtained simply by choosing $A_1=\alpha$, for $\alpha<\beta/2$, and momenta in the natural coordinates of the system recovered subsequently by inverting Eq. (17).

D. The slow manifold

Complex molecules (such as proteins or nucleic acids) have a dynamical evolution in which fast oscillatory modes are coupled to slowly varying ones. While the *actual* time spent by an *ensemble* of molecules in the fast manifold is the same as that spent in the slow manifold, the *computer* time needed for convergence of properties in the slow manifold (when simulating a single molecule) is much larger than that spent in the fast manifold. Therefore, identifying the slow manifold and accentuating the motion along it is a good strategy for enhanced sampling of the overall configuration space. Moreover, in most cases, motion along the slow manifold includes the largest conformational changes (reactions), which are often of primary interest.[21,22] The objective of the method presented here is to develop a momentum distribution that will bias our path dynamics along the slow manifold, permitting the efficient calculation of kinetic properties of infrequent reactions.

Building on our previous work aimed at using slow manifold dynamics to enhance the calculation of thermodynamic properties,[22] we can identify the slow manifold by calculating $\mathbf{p}_s$, the average of the momentum $\mathbf{p}$, over t time units,

$$\hat{\mathbf{e}}_s=\frac{\mathbf{p}_s}{\|\mathbf{p}_s\|},\quad \text{with } \mathbf{p}_s=\frac{1}{t}\int_0^t \mathbf{p}(\tau)d\tau. \tag{18}$$

In order for $\hat{\mathbf{e}}_s$ to point along the slow manifold (i.e., for the components of momentum in the fast manifold to average to zero), one has to choose the averaging time t to be several times larger than the period of the fast modes but smaller than those of the slow modes.

Alternatively, $\hat{\mathbf{e}}_s$ can be found from either a normal-mode or a quasiharmonic-mode decomposition[24] by solving an eigenvalue-eigenvector problem,

$$\hat{\mathbf{e}}_s=\min_{\lambda}\{\hat{\mathbf{e}}|F\hat{\mathbf{e}}=\lambda\hat{\mathbf{e}}\}, \tag{19}$$

and choosing the lowest-eigenvalue eigenvector obtained from diagonalizing F, which is either the Hessian of the potential (in case of the harmonic modes) or the inverse of the covariance matrix of atomic fluctuations (in case of the quasiharmonic modes). The normal-mode decomposition can be performed for a minimized structure in the reactant well, and $\hat{\mathbf{e}}_s$ aligned along a low-frequency mode (or a linear combination of several such slow modes). The quasiharmonic calculation can be performed on the same trajectory that was used to generate the initial distribution of the starting conformations; again, a combination of slow quasiharmonic modes can be used.

Using the momentum-averaging scheme, we showed that slow-mode directions are of promise for enhanced sampling and the exploration of large conformational changes, and have provided a hybrid Monte Carlo (MC) scheme to obtain exact thermal equilibrium properties.[23] For the case of the alternative choice of $\hat{\mathbf{e}}_s$ in Eq. (19), normal-mode analyses have provided considerable insight into the nature of collective motions in many proteins.[25–29] They have demonstrated that, in systems where an initial and a final structure are available, the first few low-frequency modes are often sufficient to describe the large-scale conformational changes involved in going from one structure to the other.[30,31] This strategy has worked well both for protein-DNA complexes[32] and for systems as large as the ribozome.[33]

In cases when the slow manifold is higher than one dimension (which is likely to be the case for complex biomolecular conformational changes), the guiding vector $\hat{\mathbf{e}}_s$ is to be calculated for each initial configuration or, if there is little variation in the slow direction for certain inital regions, a single $\hat{\mathbf{e}}_s$ can be applied to some or all of the initial points.

IV. APPLICATION TO KINETIC RATE CONSTANTS

As noted in the Introduction, Eq. (3) is actually a ratio of two expressions of the form given by Eq. (2). In practice, however, we generate the equilibrium spatial distribution $\rho(\mathbf{q})$ in the reactant region alone, so that the average over the container function h_R is unity, and Eq. (3) can be rewritten as an average over a biased distribution in the form of Eq. (12), with $Q_t=h_P(\mathbf{\Gamma}_t)$.

In our numerical examples, we generate the initial distribution using conventional Langevin dynamics. Direction $\hat{\mathbf{e}}_s$ is identified by averaging as in Eq. (18) and a dynamical trajectory is then initiated from each initial point using a momentum drawn from ρ_B. $C(t)$ is then calculated as sum of the weights of those trajectories that are in the product region at time t, divided by the sum of the weights of all initiated trajectories.

A. Model system examples

As a test of the efficiency of our Skew'M method, we apply it to the prototypical example of a particle moving in a potential in which a double well is coupled to n harmonic oscillators:

$$V_n(\mathbf{x},y)=\sum_{i=1}^{n} a\left(x_i-\frac{\gamma}{a}y\right)^2+y^2(y-b)^2. \tag{20}$$

We choose $a=500$ and $b=2$ so that the motion in the y direction is slow and the motion in the other directions is fast; for the coupling constant, we choose $\gamma=10$. $V(\mathbf{x},y)$ has stable minima when $y=0$ or $y=b$ and $x_i=\gamma y/a$ for all i. The reaction barrier is located at $y=1$, and has a minimum height 1 for $b=2$. We define $y<1$ to be the reactant region and all other y to be the product region. For this simple system, the theoretical rate constant can be easily calculated from transition state theory (TST). We choose $\beta=8$, corresponding to a barrier height of $8k_BT$, with a calculated rate of 1.47×10^{-4}. For this simple system, the TST result is close to exact. Beginning with an equilibrium spatial distribution of points in the reactant region generated with Langevin dynamics (also used to calculate $\hat{\mathbf{e}}_s$), dynamical trajectories were integrated over a time $t=2.0$ using a velocity Verlet algorithm with a time step of $\Delta t=0.01$. For a varying number

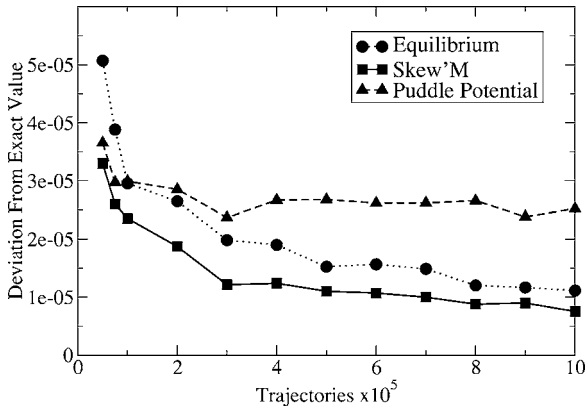

FIG. 2. Average deviation of calculated rate constant from TST value for 100 simulations on the potential $V(\mathbf{x})$ at $\beta=8$ with a varying number of initial points in each simulation.

of trajectories, $C(t)$ was calculated for an equilibrium momentum distribution in the regular fashion, for a Skew'M distribution with $A_1'=2.0$ and $A_i'=4.0$ (for all i), and for a puddle distribution on V with $V_{\text{pud}}=1.0$. After discarding the portion of the correlation function corresponding to the transient period, the slope of the correlation function was found using least-squares regression on the remaining data points. Figure 2 depicts the average deviation from the TST value over 100 independent repeats of the simulations for each of these three methods.

Additionally, using the same velocity Verlet integration, 100 simulation repeats were run on V for a variety of temperature values. Again, we compare the equilibrium momentum distribution (which varies with β), a puddle distribution with $V_{\text{pud}}=1.0$, and a momentum-weighted distribution with $A_i'=2.0$ and the other A_i' equal to their equilibrium values. Each simulation was composed of 5×10^5 initial points. Figure 3 depicts the average deviation over the 100 simulations from the TST value for each method at successive values of β.

At higher values of β or lower numbers of trajectories, the calculation from $C(t)$ using the equilibrium distribution becomes unreliable. At $\beta=12$, using 5×10^5 trajectories, 15% of our simulations showed no transitions at all: $C(t)=0$. Those simulations that do observe transitions from

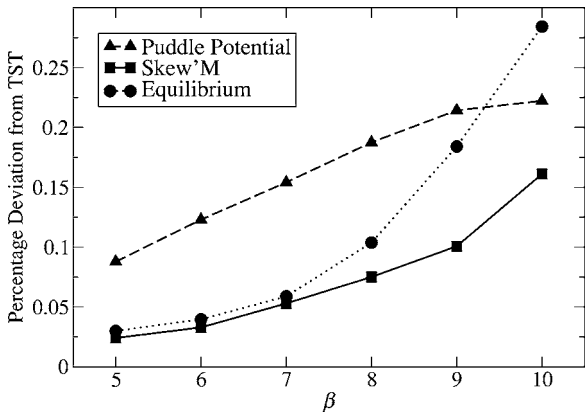

FIG. 3. Average deviation of calculated rate constant from TST value at varying β for 100 simulations on the potential $V(\mathbf{x})$ with 5×10^5 trajectories per simulation.

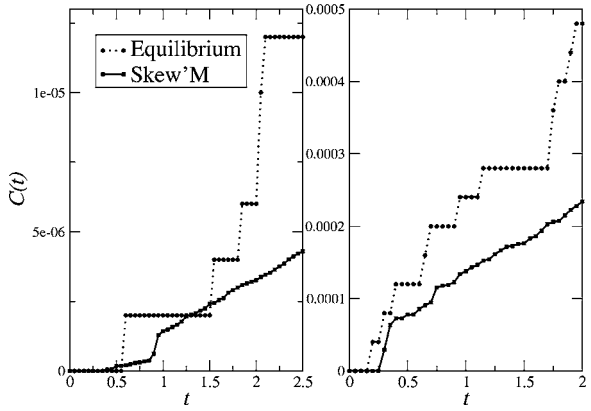

FIG. 4. Typical $C(t)$ at extremes of low-temperature (left panel) and low number of trajectories (right panel) for weighted and unweighted momentum distributions. The low-temperature and few-trajectory data was found from simulations with $\beta=12$ and 5×10^5 trajectories and $\beta=8$ and 2.5×10^4 trajectories, respectively.

the reactant to product regions were unconverged, showing no linear region from which a rate constant could be extracted. By contrast, weighted simulations with $A_1'=2.0$ demonstrate convergence in both low-trajectory and low-temperature simulations; a typical $C(t)$ for each case is depicted in Fig. 4.

B. Higher-dimensional model systems and dimensional scaling

Notably, in both Figs. 2 and 3 the puddle method shows diminished accuracy when compared with both the momentum-weighted distribution and the equilibrium distribution. These tests illustrate one drawback of the puddle method; although it is accurate when applied to one or two separable degrees of freedom, it is not efficient when applied across several separate terms in a potential, as the system tends to spread the distribution along directions that are not of importance in the reaction. Of course, for our simple system, we could apply a puddle only to the $y^2(y-b)^2$ term, and thereby attain accurate results. In general, however, for transitions involving the concerted motion of many degrees of freedom, the potential cannot be separated along an arbitrary direction and so the puddle method must be applied blindly across many degrees of freedom. Thus, in larger systems, we expect to find similar difficulties as we do in V. In order to investigate this behavior as a function of the number n of fast degrees of freedom, we consider simulations on $V_n(\mathbf{x})$ for a variety of n.

For V_2 through V_9, 100 simulations were initiated at $\beta=8$ with an equilibrium distribution of 5×10^5 initial points per simulation, and the average deviation of the computed rate from the TST value was compared with 100 simulations of a weighted-momentum simulation with $A_1'=2.0$ and all other $A_i'=4.0$. The results are depicted in Fig. 5. The accuracy of the rate constant decreases with increasing dimension, as the puddle potential tends to drive the initial states against the perpendicular harmonic oscillators. By contrast, the accuracy of the weighted-momentum simulation decreases only slightly as the number of dimensions increases.

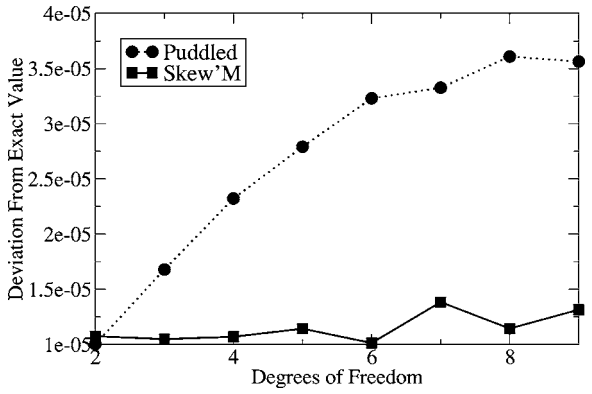

FIG. 5. Average deviation of calculated rate constant from TST value using a "puddle" of 1.0 and a weighted momentum distribution with $A_1'=2$ and all other $A_i'=4.0$ for 100 simulations on the potential $V_n(\mathbf{x})$ for a varying number of dimensions n at $\beta=8$ with 5×10^5 trajectories per simulation.

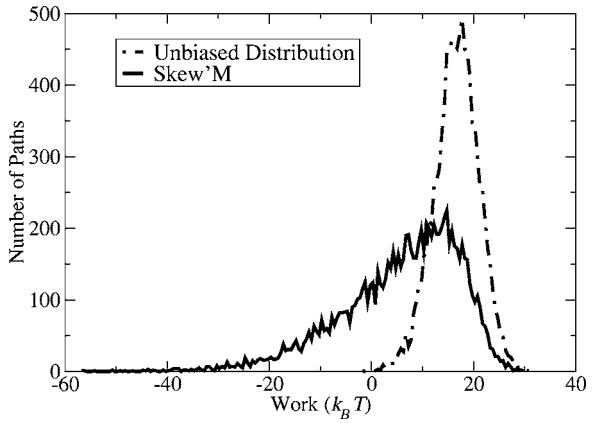

FIG. 6. Histogram of work values for 10 000 trajectories on $V_{\text{tot}}(x(t),t)$ for both an equilibrium distribution with variance of 1.0 and a biased distribution with variance of 16.0.

V. APPLICATION TO JARZYNSKI'S EQUALITY

We found it of interest to explore the use of the present method to the generation of the fast-pulling trajectories used in the Jarzynski equality to derive equilibrium free energy from nonequilibrium trajectories.[5] Jarzynski's equality states that

$$\exp(-\beta\Delta F)=\langle\exp(-\beta W_t)\rangle, \tag{21}$$

where ΔF is the free-energy change, W_t is the work performed on the system during each nonequilibrium trajectory of length t, and $\langle\cdots\rangle$ indicates an average over an infinite number of trajectories. Equation (6), which we presented in the Introduction, represents a specific application of Eq. (21) to free-energy reconstructions from pulling experiments. While Jarzynski's identity is exact, it suffers from the problem that the trajectories which count most (i.e., the low-work trajectories) are statistically rare. In computational simulations, which can only sample a finite number of trajectories, it is important to look for lower-work distributions.

As Jarzynski's equality is in the form of Eq. (1), we can apply to it our skewed-momenta method simply by setting $A[\mathbf{\Gamma}(t)]=\exp(-\beta W_t)$. However, we anticipate that our method will be most useful in the particular case of calculating free-energy profiles from pulling experiments, for which Hummer and Szabo have provided a modified form of Jarzynski's expression.[3]

The average defining the potential of mean force, Eq. (6), can be written as an average over a skewed distribution of initial momenta as described by Eq. (12). We anticipate that skewed trajectories have lower work, as the momenta can be biased so that important degrees of freedom tend to move in the same direction as the pulling potential. Specifically, the instantaneous contribution to the work of a given trajectory is found from an infinitesimal movement of the time-dependent guiding potential at constant position and momentum, $dW=(\partial V/\partial t)(\mathbf{\Gamma}(t))dt$. The dW contribution will tend to be smaller, or even negative, when the system follows the motion of the guiding potential naturally rather than being "pulled along."

As a simple test of this approach, we apply Jarzynski's equality to a one-dimensional double well potential of the form

$$V(x)=x^2(x-a)^2+x \tag{22}$$

with a harmonic guiding potential

$$V_{\text{pull}}(x,t)=\frac{k(x-\upsilon t)^2}{2}, \tag{23}$$

where k is the spring constant and υ the pulling velocity. By artificially increasing the variance of the momentum distribution, we can alter the work distribution to include more low-work trajectories. To illustrate this, we run Langevin dynamics for a particle of unit mass on the potential $V_{\text{tot}}(x(t),t)=V(x(t))+V_{\text{pull}}(x(t),t)$ with $k_BT=1$, $k=100$, step size $\Delta t=0.001$, and friction constant $\gamma=0.2$ (in arbitrary units). We choose $\upsilon=4$ and $a=4$, so that the barrier height is many times k_BT and the pulling speed far from the reversible regime. The trajectories were run for a duration $t=1000$. Figure 6 displays a histogram of work values for 10 000 trajectories for both a equilibrium initial momentum distribution, drawn from a Gaussian distribution with zero average and unit variance, and a biased distribution with zero average and a variance of 16.0.

Because an individual trajectory's contribution to the average in Eq. (21) depends exponentially on the work performed on the system during the trajectory, we expect that the increased sampling of low-work trajectories will improve the accuracy of the free-energy calculation.

As a test of the practical efficiency of our method, we recreate the function $V(x)$ using Eqs. (6) and (12) from a number of high-speed pulling simulations using a technique analogous to the weighted histogram method:[3]

$$\exp(-\beta G_0(z))=\frac{\sum_t \frac{\langle\delta[z-z(t)]\mathcal{W}\rangle}{\langle\mathcal{W}\rangle}}{\sum_t \frac{\exp(-\beta V(z,t))}{\langle\mathcal{W}\rangle}}. \tag{24}$$

A detailed numerical implementation of this method is discussed in Ref. 3. $\mathcal{W}$ is the statistical weight of a trajectory,

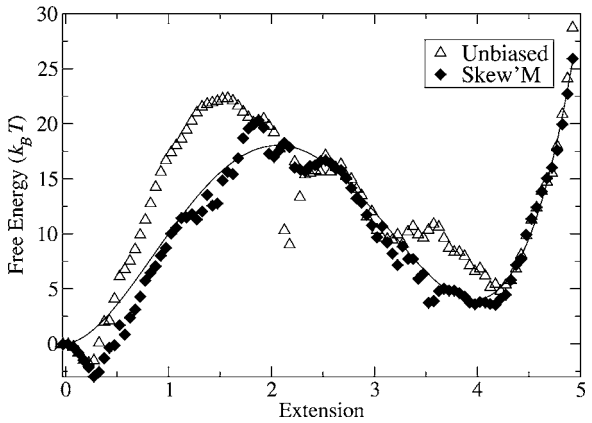

FIG. 7. Reconstruction of the potential $V(x)$ (solid line) from 10 000 pulling simulations with $a=4$, $v=4$, and $k=100$ for both an equilibrium momentum distribution and a biased momentum distribution with zero mean and variance of 16.0. The reconstruction from the equilibrium trajectories shows significant error, especially in the barrier region, while the biased reconstruction is reasonably accurate, save for a deviation near $x=0$ that is shared by the unbiased reconstruction and which is a consequence of poor sampling near the beginning of the pulling run.

and the averages are taken over the ensemble of trajectories. In the unbiased case, $\mathcal{W}=\exp(-\beta W_t)$, while in the biased case an additional factor must be included to account for the skewed-momentum distribution: $\mathcal{W}=\exp(-\beta W_t)w(\mathbf{p})$.

In our simple system, $z[\mathbf{\Gamma}(t)]=x(t)$, and we choose the dynamical constants Δt, γ, and k_BT to be the same as their values in our previous example. We initiate two pulling simulations, one with an equilibrium momentum distribution and one with a variance of 16, of 1000 trajectories each for $a=4$, $v=4$, $k=100$, and total time $t=5$, and reconstruct the free-energy profile using Eq. (24). The results of these simulations, displayed in Fig. 7, show an increased accuracy in the reconstruction using the Skew'M method.

In future work, it should be of interest to explore the combination of the Skew'M method to alter the momentum distribution with novel methods that either apply periodic loading[34] or Monte Carlo sampling of nonequilibrium trajectories from a work-weighted ensemble.[35] A quantum analog of the Jarzynski method has been recently proposed as a mean to describe the dephasing of quantum coherences.[36] It would also be of interest to explore whether biasing the flow of adiabatic states in the corresponding master equation in a numerical realization of a nonequilibrium line-shape measurement would yield a faster convergence of spectroscopic properties.

VI. CONCLUDING DISCUSSION

Although calculating the kinetic properties from a trajectory ensemble avoids the often difficult task of identifying transition states and reaction coordinates, in systems with high energetic or entropic barriers its accuracy will suffer, as few paths in the ensemble will make it to the product region. We have presented a method that mitigates this problem by drawing momenta for molecular dynamics from a distribution that is artificially enhanced along the system's slow degrees of freedom, which we find by averaging the momenta over a short dynamics run. Additionally, we have shown that our "skewed momenta" addresses a problem endemic to the reconstruction of free-energy profiles from fast-pulling experiments using Jarzynski's identity, namely, that although low-work trajectories have the largest statistical weight, they are rarely sampled, especially when the pulling is fast. Our method generates more low-work trajectories by biasing the relevant degrees of freedom so that they tend to move with the pulling potential, thereby lowering the work done on the system and increasing the accuracy of the calculated potential of mean force. This fact is in accord with recent theoretical work analyzing the Jarzynski equality for an ideal gas, for which Lua and Grosberg[37] have noted that the trajectories in the far tails of the Maxwell distribution are the key ones in determining the accuracy of the Jarzynski method when the system's volume is altered rapidly. While an ideal gas is a crude "flat landscape" approximation, it is important to realize that the far tails of the Boltzmann distribution are preferentially sampled, in our method, only along important degrees of freedom along which major conformational changes are expected.

The method is likely to be useful for the numerical calculation of other correlation functions of importance to complex molecules. An example is the orientation correlation functions of interest in nuclear magnetic resonance (NMR)-derived dynamical estimates for proteins and nucleic acids.[38] Such correlations are difficult to converge numerically when multiple conformations separated by large free-energy barriers contribute to their measurement.

Future improvements of the method presented here should take into account that the direction of the slow manifold in general can change (i.e., direction of importance can curve around conformational space). For the method to be effective in such instances, one would have to bias not only the initial momentum distribution, but, using a periodically updated $\hat{\mathbf{e}}_s$, also the actual trajectories that lead to relaxation out of the product well. In this case, one would also need to reweight the trajectory itself, not only the initial points. If the propagation uses Langevin dynamics, the formalism of stochastic path integrals[39] leads to the proper weight described by the exponential of a Onsager-Machlup action[40–42] that would have to be calculated along each trajectory according to the formula

$$S[\mathbf{\Gamma}(t)] = \int_{t_1}^{t_2} dt(M\ddot{\mathbf{q}}(t) + \gamma\mathbf{p}(t) + \nabla V[\mathbf{q}(t)])^2. \tag{25}$$

ACKNOWLEDGMENT

This work was supported by funds from the University of Michigan.

[1] C. Dellago, P. G. Bolhuis, F. S. Csajka, and D. Chandler, J. Chem. Phys. **108**, 1964 (1998).
[2] C. Dellago, P. G. Bolhuis, and D. Chandler, J. Chem. Phys. **110**, 6617 (1999).
[3] G. Hummer and A. Szabo, Proc. Natl. Acad. Sci. U.S.A. **98**, 3659 (2001).
[4] D. Chandler, *Introduction to Modern Statistical Mechanics* (Oxford University Press, New York, 1987).
[5] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).
[6] D. J. Bicout and A. Szabo, Protein Sci. **9**, 452 (2000).

[7] This is akin to the "golf course" problem often cited in minimization literature pertinent to protein folding (Ref. 43) having to do with finding a deep hole on a large, flat (i.e., enthalpically barrierless) surface.
[8] H. Oberhofer, C. Dellago, and P. L. Geissler, J. Phys. Chem. B **109**, 6902 (2005).
[9] J. A. Rahman and J. C. Tully, Chem. Phys. **285**, 277 (2002).
[10] S. A. Corcelli, J. A. Rahman, and J. C. Tully, J. Chem. Phys. **118**, 1085 (2003).
[11] A. F. Voter, Phys. Rev. Lett. **78**, 3908 (1997).
[12] A. F. Voter, J. Chem. Phys. **106**, 4665 (1997).
[13] D. Hamelberg, J. Mongan, and J. A. McCammon, J. Chem. Phys. **120**, 11919 (2004).
[14] A. Laio and M. Parrinello, Proc. Natl. Acad. Sci. U.S.A. **99**, 12562 (2002).
[15] G. A. Huber and S. Kim, Biophys. J. **70**, 97 (1996).
[16] H. Grubmuller, Phys. Rev. E **52**, 2893 (1995).
[17] E. Carter, G. Ciccotti, J. Haynes, and R. Kapral, Chem. Phys. Lett. **156**, 472 (1989).
[18] M. Mella, J. Chem. Phys. **122**, 204106 (2005).
[19] K. Duffin and W. Barnett, *Proceedings of the IEEE Conference on Scientific Visualization, Washington, DC, 1994* (IEEE Computer Society Press, Los Alamitos, CA, 1994), pp. 205–211.
[20] A. Aguilera and R. Pérez-Aguila, in *Short Communication Papers—Proceedings of the 12th International Conference on Computer Graphics, Visualization and Computer Vision, University of West Bohemia, Plzen, Czech. Republic*, 2–6 Feb. (Union Agency-Science, 2004), pp. 1–8.
[21] A. Amadei, A. Linssen, and H. Berendsen, Proteins **17**, 412 (1993).
[22] M. Balsera, W. Wriggers, Y. Oono, and K. Schulten, J. Phys. Chem. **100**, 2567 (1996).
[23] I. Andricioaei, A. R. Dinner, and M. Karplus, J. Chem. Phys. **118**, 1074 (2003).
[24] B. R. Brooks, D. Janezic, and M. Karplus, J. Comput. Chem. **16**, 1522 (1995).
[25] N. Go, T. Noguti, and T. Nishikawa, Proc. Natl. Acad. Sci. U.S.A. **80**, 3696 (1983).
[26] M. Levitt, C. Sander, and P. S. Stern, J. Mol. Biol. **181**, 423 (1985).
[27] B. Brooks and M. Karplus, Proc. Natl. Acad. Sci. U.S.A. **82**, 4995 (1985).
[28] J. P. Ma and M. Karplus, J. Mol. Biol. **274**, 114 (1997).
[29] Q. Cui, G. H. Li, J. P. Ma, and M. Karplus, J. Mol. Biol. **340**, 345 (2004).
[30] F. Tama and Y. H. Sanejouand, Protein Eng. **14**, 1 (2001).
[31] W. G. Krebs, V. Alexandrov, C. A. Wilson, N. Echols, H. Y. Yu, and M. Gerstein, Proteins **48**, 682 (2002).
[32] M. Delarue and Y. H. Sanejouand, J. Mol. Biol. **320**, 1011 (2002).
[33] F. Tama, M. Valle, J. Frank, and C. L. Brooks, Proc. Natl. Acad. Sci. U.S.A. **100**, 9319 (2003).
[34] O. Braun, A. Hanke, and U. Seifert, Phys. Rev. Lett. **93**, 158105 (2004).
[35] S. X. Sun, J. Chem. Phys. **118**, 5769 (2003).
[36] S. Mukamel, Phys. Rev. Lett. **90**, 170604 (2003).
[37] R. C. Lua and A. Y. Grosberg, J. Phys. Chem. B **109**, 6805 (2005).
[38] A. G. Palmer, Chem. Rev. (Washington, D.C.) **104**, 3623 (2004).
[39] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, Polymer Physics and Financial Markets*, 3rd ed. (World Scientific, Singapore 2004).
[40] L. Onsager and S. Machlup, Phys. Rev. **91**, 1505 (1953).
[41] R. Elber, J. Meller, and R. Olender, J. Phys. Chem. B **103**, 899 (1999).
[42] D. M. Zuckerman and T. B. Woolf, Phys. Rev. E **63**, 016702 (2001).
[43] J. D. Bryngelson and P. G. Wolynes, J. Phys. Chem. **93**, 6902 (1989).