Reaction coordinates, one-dimensional Smoluchowski equations, and a test for dynamical self-consistency

Peters, B.; Bolhuis, P.G.; Mullen, R.G.; Shea, J.E.

DOI
10.1063/1.4775807

Publication date
2013

Document Version
Final published version

Published in
Journal of Chemical Physics

Reaction coordinates, one-dimensional Smoluchowski equations, and a test for dynamical self-consistency

Baron Peters,[1,2] Peter G. Bolhuis,[3] Ryan G. Mullen,[1] and Joan-Emma Shea[2]
[1]*Department of Chemical Engineering, University of California, Santa Barbara, California 93106, USA*
[2]*Department of Chemistry and Biochemistry, University of California, Santa Barbara, California 93106, USA*
[3]*van 't Hoff Institute for Molecular Sciences, University of Amsterdam, PO Box 94157, 1090 GD Amsterdam, The Netherlands*

(Received 1 May 2012; accepted 27 December 2012; published online 1 February 2013)

We propose a method for identifying accurate reaction coordinates among a set of trial coordinates. The method applies to special cases where motion along the reaction coordinate follows a one-dimensional Smoluchowski equation. In these cases the reaction coordinate can predict its own short-time dynamical evolution, i.e., the dynamics projected from multiple dimensions onto the reaction coordinate depend only on the reaction coordinate itself. To test whether this property holds, we project an ensemble of short trajectory swarms onto trial coordinates and compare projections of individual swarms to projections of the ensemble of swarms. The comparison, quantified by the Kullback-Leibler divergence, is numerically performed for each isosurface of each trial coordinate. The ensemble of short dynamical trajectories is generated only once by sampling along an initial order parameter. The initial order parameter should separate the reactants and products with a free energy barrier, and distributions on isosurfaces of the initial parameter should be unimodal. The method is illustrated for three model free energy landscapes with anisotropic diffusion. Where exact coordinates can be obtained from Kramers-Langer-Berezhkovskii-Szabo theory, results from the new method agree with the exact results. We also examine characteristics of systems where the proposed method fails. We show how dynamical self-consistency is related (through the Chapman-Kolmogorov equation) to the earlier isocommittor criterion, which is based on longer paths. [http://dx.doi.org/10.1063/1.4775807]

INTRODUCTION

Accurate reaction coordinates are of paramount importance for simulations of activated processes, as they enable correct free energy calculations,[1,2] efficient calculations of true rate constants,[3] and perhaps most importantly, provide invaluable mechanistic insight.[1,4] Recent years have seen substantial advances in our understanding of reaction coordinates and in computational methods to obtain them.[5–27] Some of these advances[5,11–16,23] were enabled by transition path sampling[1,28] and related methods[29–32] that generate barrier crossing trajectories without biasing the natural dynamics along *a priori* chosen reaction coordinates. Several of these trajectory-based methods seek reaction coordinates that can accurately predict the committor at each configuration. The committor at a configuration $\mathbf{x}$, denoted $p_B(\mathbf{x})$, is the fraction of trajectories initiated with random momenta from $\mathbf{x}$ that reach the product state B before reaching the reactant state A.

One motivation for obtaining reaction coordinates is to reduce the dynamics of activated processes in high dimensional systems to a simple Smoluchowski model for motion along a one-dimensional reaction coordinate. Reduction to such a Smoluchowski model (or, equivalently, an overdamped Langevin equation) would provide valuable mechanistic insight. However, reduction to a low dimensional Smoluchowski equation is only possible for activated processes that obey special conditions: (1) the transition paths should follow a single channel (or transition tube[26]) representing the dynamical bottleneck in phase space, (2) dynamics projected along the reaction coordinate must be in the high friction (diffusive) limit,[29] and (3) there must be a clear separation of time scales[30] between the reaction time and all faster relaxation times within the metastable state to ensure that a reaction coordinate exists.[31] These restrictions may exclude many interesting applications. For example, extensively studied activated processes such as S_N2 reactions[32] and ion-pair dissociation in solution[33,34] have inertial dynamics that prevent reduction to a Smoluchowski model. Another example is protein folding, which involves a complex network of pathways between many intermediates.[10,35–37,83] A recent study accurately mapped the committor for protein folding intermediates,[10] but the existence of multiple channels with disparate mobilities prevents a simple reduction to a low dimensional Smoluchowski model. Nevertheless, there are activated processes that can be described by low dimensional Smoluchowski equations. For example, such a reduction has been successful for nucleation processes in lattice models[38] and atomistic simulations of simple systems.[39,40] Depending on the supersaturation, nucleation can involve high free energy barriers and therefore involves a clear time scale separation. Theoretical studies additionally suggest that nucleation trajectories pass through a single dynamical bottleneck and that a nucleus size variable largely controls the

committor probability.[41–44] Thus it should be possible to map atomistic or otherwise high dimensional models of some nucleation processes onto one-dimensional Smoluchowski models. In the context of nucleation, Smoluchowski equations and associated mean first passage time calculations provide an equivalently one-dimensional alternative to the sometimes problematic one-dimensional framework of classical nucleation theory.[39,40]

A one-dimensional Smoluchowski description requires knowledge of the free energy as a function of the reaction coordinate and the diffusivity as a function of a reaction coordinate.[29] Given a reaction coordinate, there are established computational methods for computing these properties.[38,45–53] However, a key obstacle is often identifying a one-dimensional reaction coordinate. Transition path sampling data can be used to identify reaction coordinates systematically in some systems.[5,12–16,23,42,54,55] Unfortunately, there are nucleation problems that are not amenable to a transition path sampling approach. For example, in crystal nucleation from a supersaturated solution there is currently no proper way to control the supersaturation along a transition path without disrupting the dynamics for solute insertions or solvent-solute swaps.[56] In other cases, for example, in methane hydrate nucleation,[57] the transition paths are prohibitively long for transition path sampling simulations. Therefore, it would be useful to identify optimal reaction coordinates using only short time dynamics, and not full-length transition paths. In this paper we aim to develop such a method.

The new method is based on the fact that when reduction of an activated process to a Smoluchowski model is possible, the physical collective variable that predicts the committor must also accurately predict its own stochastic drift and diffusion dynamics along the transition pathway, while other collective variables cannot. Thus, when a physical reaction coordinate obeys a Smoluchowski equation, it might be identified as that collective variable that is sufficient to predict its own evolution toward reactants and products. This principle suggests a dynamical self-consistency criterion to screen candidate reaction coordinates using short trajectory data in some systems. We show here that the new criterion is closely related to the standard committor analysis for reaction coordinate optimization. However, unlike earlier isocommittor criteria,[25,58] the new criterion does not require information about the long time fate of trajectories. Our method for screening candidate reaction coordinates does not require a computation of committors[5] or committor realizations.[12] Instead, the proposed method needs as input an initial collective variable for harvesting an ensemble of initial configurations from which the short trajectory data are generated.

The paper is organized as follows. We start with some theoretical background on the one-dimensional Smoluchowski equation and its relation to the algorithm for our approach. We illustrate the method on model potentials with varying degrees of diffusion anisotropy. Two of the examples have a quadratic potential with a coordinate independent (but still anisotropic) diffusion tensor so that accurate reaction coordinates can be obtained from the Kramers-Langer-Berezhkovskii-Szabo[6,59,60] theory for comparison. The first example is pertinent to studies of polymorph selection during nucleation from a mixture of oppositely charged colloids.[40,61,62] A second example reveals analogies between anisotropic diffusion over a narrow saddle and the dynamical caging and frozen solvent regimes of the inertial Grote-Hynes theory.[41–43,63] Finally, a third example illustrates how the method can fail when certain requirements are not met by the initial order parameter. We end with concluding remarks.

THEORY

When projected onto suitable collective variables some activated processes can be modeled by diffusion on a free energy landscape. The mathematical description of such processes is an overdamped Langevin equation or equivalently a Smoluchowski equation.[29] In the most general case, the Smoluchowski equation can be written in terms a multidimensional free energy landscape, and a coordinate dependent diffusion tensor,

$$\frac{\partial \xi}{\partial t} = e^{+\beta F(\mathbf{q})} \frac{\partial}{\partial \mathbf{q}} \left\{ e^{-\beta F(\mathbf{q})} D(\mathbf{q}) \frac{\partial \xi}{\partial \mathbf{q}} \right\}, \tag{1}$$

where $\mathbf{q}$ denotes the multidimensional collective variables pertinent to the reaction coordinate, $F(\mathbf{q})$ is the free energy landscape, and $D(\mathbf{q})$ is the coordinate dependent diffusion tensor, and $\xi(\boldsymbol{q}, t) = p(\boldsymbol{q}, t)/\exp[-\beta F(\boldsymbol{q})]$ is the ratio of the time dependent probability density $p(\boldsymbol{q}, t)$ to the equilibrium distribution $\exp[-\beta F(\mathbf{q})]$, i.e., the Kramers crossover function.[29]

The Smoluchowski equation is most easily constructed for a single coordinate q:[48,56] the free energy landscape $F(q)$ requires only a one-dimensional potential of mean force calculation[45,46] and the coordinate dependent diffusion coefficient $D(q)$ is scalar. The resulting Smoluchowski equation is

$$\frac{\partial \xi}{\partial t} = e^{+\beta F(q)} \frac{\partial}{\partial q} \left\{ e^{-\beta F(q)} D(q) \frac{\partial \xi}{\partial q} \right\}. \tag{2}$$

Established methods exist to compute the diffusivity[38,49–53,64] and free energy profiles[45,46] once an accurate scalar reaction coordinate q is known. The mean first passage time, the rate, the committor, and the committor isosurfaces are then also easy to obtain. For example, solving the backward Kolmogorov equation with boundary conditions $p_B(q_B) = 1$ and $p_B(q_A) = 0$ gives the committor along the scalar coordinate q:[7,14,18,65]

$$p_B(q_0) = \frac{\int_{q_A}^{q_0} \exp[\beta F(q) - \ln D(q)] dq}{\int_{q_A}^{q_B} \exp[\beta F(q) - \ln D(q)] dq}. \tag{3}$$

The integration limits q_A and q_B correspond to the boundaries of the stable states A and B, respectively. q_A and q_B should be carefully specified so that A and B include the typical fluctuations from their respective equilibrium positions but A and B should not encroach on the barrier region. For systems with a high activation barrier along the transition pathway, contributions to the integrals in Eq. (3) come primarily from regions near the free energy maximum. If the boundaries q_A and q_B are chosen on opposite sides of, and many k_BT below the free energy maximum, then the committor p_B is not sensitive to the precise location of q_A and q_B. For activation barriers much

higher than k_BT, a mean first passage time calculation gives the rate constant from state A to B as[66]

$$k_{A\to B} = \left[\int_{\cap} e^{\beta F(q) - \ln D(q)} dq \int_A e^{-\beta F(q)} dq\right]^{-1}. \quad (4)$$

The integral subscripts $\cap$ and A denote integration over the barrier region and stable basin of state A, respectively.[6,59] For high activation barriers, nearly all contributions to the integral with $e^{+\beta F(q)}$ come from regions near the barrier maximum. Likewise nearly all contributions to the integral with $e^{-\beta F(q)}$ come from regions near the reactant minimum.

Equations (3) and (4) accurately describe the committor and the interconversion kinetics for a one-dimensional Smoluchowski equation (2) between reactant (A) and product (B) states. However, $D(q)$ and $F(q)$ are obtained by projections of the multidimensional diffusion tensor and energy landscape. Unless the local diffusivity $D(q)$ and local mean force $-dF/dq$ are perfectly independent of all coordinates other than q, the one-dimensional Smoluchowski model (and all results obtained from it) must be regarded as an approximation. In practice, one-dimensional and low-dimensional Smoluchowski models are usually inexact, but conceptually useful, approximations to the dynamics of multidimensional systems.

Several strategies for constructing Smoluchowski models for high dimensional systems have recently emerged. The first strategy assumes the knowledge of the important collective variables, often two or more, and then examines the behavior of short trajectory "*swarms,*" i.e., trajectories launched with random Maxwell-Boltzmann momenta from specific configurations on the free energy landscape.[67] The drift and time-dependent covariance of the swarms reveal the local potential of mean force and the local coordinate dependent diffusion tensor.[40,50,67,68] These properties can be statistically combined to map the entire free energy landscape and the coordinate dependent diffusion tensor.[64] The accuracy of Smoluchowski models constructed in this way depends critically on whether the *a priori* chosen collective variables are pertinent reaction coordinates. For example, Berezhkovskii and Szabo have shown that errors in the reaction coordinate will artificially magnify the reaction rate as computed from a mean first passage time.[6]

Kramers-Langer-Berezhkovskii-Szabo (KLBS) theory[6] shows, in principle, how the optimal one-dimensional reaction coordinate can be obtained from a multidimensional space. KLBS theory considers a multidimensional free energy $F(\mathbf{q})$ and diffusion tensor $\mathbf{D}(\mathbf{q})$ where the components of $\mathbf{q}(\mathbf{x})$ are collective variables that depend on the more detailed (perhaps atomistic) coordinates $\mathbf{x}$. Suppose that $F(\mathbf{q})$ has a saddle point at $\mathbf{q}^{\ddagger}$. The free energy can be expanded around the saddle point as $\beta F(\mathbf{q}^{\ddagger} + \delta\mathbf{q}) = \beta F_{\ddagger} + \delta\mathbf{q}^{\mathrm{T}}\mathbf{A}\delta\mathbf{q}$ where $\mathbf{A}$ is a matrix of second derivatives of the free energy with respect to $\mathbf{q}$ at location $\mathbf{q}^{\ddagger}$. Multidimensional extensions of Kramers theory[59] by Langer,[60] and Berekzhkovskii and Szabo,[6] show that the matrix products **AD** and also **DA** will have exactly one negative eigenvalue. The negative eigenvalues of **AD** and **DA** correspond to eigenvectors that point along the maximum flux[60] and the reaction coordinate[6] directions, respectively. Berezhkovskii and Szabo showed that, when the free energy is quadratic and when the diffusion tensor is coordinate independent, a variational maximization of the mean first passage time identifies the optimal one-dimensional reaction coordinate direction in the space of $\mathbf{q}$.[6] Their result also confirms that the optimal reaction coordinate is the committor.[6] However, application of their elegant result requires knowledge of the multidimensional free energy $F(\mathbf{q})$ and the multidimensional diffusion tensor $\mathbf{D}(\mathbf{q})$. Both objects are difficult to obtain for more than pairs of two coordinates. Thus it has not been possible in practice to use the variational KLBS theory to identify reaction coordinates from hundreds of candidate coordinates.

A different strategy for obtaining low dimensional Smoluchowski models combines the unbiased dynamical trajectories from transition path sampling with likelihood maximization to identify an accurate one-dimensional reaction coordinate.[5,12,14–16,23] The strategy based on transition path sampling, such as the KLBS theory, identifies a reaction coordinate that accurately predicts the committor. However, as noted in the Introduction, transition path sampling is not feasible for all systems. We propose that the optimal reaction coordinate can alternatively be identified by screening candidate coordinates for short-time dynamical self-consistency. Before discussing the algorithm, we must show when and how dynamical self-consistency is equivalent to the more familiar concept that reaction coordinates should accurately parameterize the committor.[5,7,8]

The short time Greens functions for the Smoluchowski equations (1) and (2) are well known. Examples can be found in work on coarse molecular dynamics,[64,69] and diffusion maps.[70,71] Most treatments construct the short time Green's functions from Gaussians with a mean drift velocity $d\langle q\rangle/dt = -D\partial\beta F/\partial q$, and a time dependent variance, determined by the rate at which a swarm of trajectories spreads around the drifting mean, $\langle[q\text{-}\langle q\rangle]^2\rangle = 2Dt$.[69] However, complications can arise when using the Gaussian Greens functions explicitly. First, extra terms appear when diffusivity is coordinate dependent. Thorough outlines of coordinate dependent diffusion in one and higher dimensions are given in recent works by Hummer and co-workers.[68,72] Second, even with constant diffusivity D, short time estimates for the mean and variance in q become inaccurate at times long enough that trajectories see changes in the gradient of the free energy.[38] These complications are circumvented in our treatment by working with *abstract* numerical Greens functions, instead of Gaussian approximations.

To explain and justify our treatment we need to define several distributions. First, we define the conditional equilibrium probability distribution $\rho_{\mathrm{EQ}}(\mathbf{x_0}|q_0)$ for configurations $\mathbf{x_0}$ on the q_0 isosurface of a collective variable $q(\mathbf{x})$ as

$$\rho_{EQ}(\mathbf{x}_0|q_0) = \frac{\rho_{EQ}(\mathbf{x_0})\delta[q(\mathbf{x_0}) - q_0]}{\int d\mathbf{x}\rho_{EQ}(\mathbf{x})\delta[q(\mathbf{x}) - q_0]}, \quad (5)$$

where $\delta[q]$ denotes the Dirac delta function and $\rho_{\mathrm{EQ}}(\mathbf{x}) = \exp(-\beta\mathrm{E}(\mathbf{x}))/\int d\mathbf{x}\exp(-\beta\mathrm{E}(\mathbf{x}))$ is the equilibrium Boltzmann distribution, with $\mathrm{E}(\mathbf{x})$ the potential energy, $\beta = 1/k_BT$ the reciprocal temperature, and k_B Boltzmann's constant. We also define the transition probability $\rho(\mathbf{x},t|\mathbf{x_0})$ as the distribution observed after a time t for a swarm of trajectories initiated

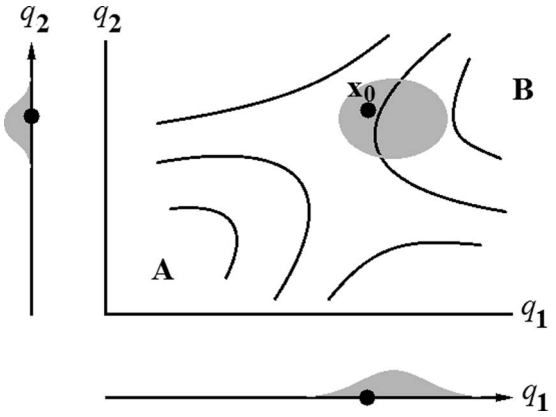

FIG. 1. Schematic depicting a projection of short trajectory swarm data (gray) from an initial point $\mathbf{x}_0$ (black) to specific coordinates $\mathbf{q}_1$ and $\mathbf{q}_2$ at later times.

from a configuration $\mathbf{x_0}$. $\rho(\mathbf{x},t|\mathbf{x_0})$ will not appear explicitly in the final algorithm, but it plays an important role in constructing the projected transition probabilities. For any collective variable q the time-evolved swarm from configuration $\mathbf{x}_0$ to configurations $\mathbf{x}$ at time t can be projected onto q as

$$P(q,t|\mathbf{x_0}) = \int d\mathbf{x}\delta[q - q(\mathbf{x})]\rho(\mathbf{x},t|\mathbf{x_0}), \tag{6}$$

which is the transition probability from $\mathbf{x_0}$ to q($\mathbf{x}$) after a time t. Figure 1 illustrates this projection.

In special cases where a one-dimensional Smoluchowski model is valid for the reaction coordinate q, then $P(q, t|\mathbf{x}_0)$ would simply be the Green's function solution of the Smoluchowski equation. Thus for these special cases the transition probabilities along $q(\mathbf{x})$ have a dynamical self-consistency property

$$P(q,t|q_0) = P(q,t|\mathbf{x}_0) \quad \forall \mathbf{x}_0 \; on \; q(\mathbf{x}_0) = q_0. \tag{7}$$

Conversely, the hypothetical self-consistency in Eq. (7) can help identify reaction coordinates in special cases where the reaction coordinate follows a one-dimensional Smoluchowski equation. We note that Eq. (7) can be viewed as a type of dynamical closure property. However, closure usually implies an approximation to a hierarchy of equations. Here we do not invoke a closure approximation, but rather determine the extent to which the dynamics of different variables naturally display a "closure."

Short-time dynamical self-consistency (when it exists) is related to the earlier committor definitions of the reaction coordinate. The committor analysis procedure of Du *et al.*[25] and Geissler *et al.*[58] examines whether isosurfaces of a trial reaction coordinate $q(\mathbf{x})$ coincide with isocommittor surfaces. A committor distribution for configurations on an isosurface of q that is not sharply peaked indicates that q alone is not sufficient to predict the *long* time dynamics along the q-axis.[4,73] In the committor analysis context, a long time is sufficient for relaxation into one of the stable basins A or B, but still much shorter than the reaction time on which trajectories might re-emerge from A or B (i.e., the mean first passage time).[1] Instead of using Eq. (3) or the long time dynamics to compute the committor along an accurate reaction coordinate, one might use the short time dynamics in combination with the Chapman-Kolmogorov equation.[29] If there exists a $q(\mathbf{x})$ for which the dynamics between A and B can be entirely reduced to a one-dimensional Smoluchowski equation, then for that $q(\mathbf{x})$ the long time transition probabilities are given formally by

$$P(q, n\Delta t|q_0) = \int \cdots \int P(q, \Delta t|q^{(n-1)}) \cdots P(q^{(2)}, \Delta t|q^{(1)}) \times P(q^{(1)}, \Delta t|q_0) dq^{(1)} dq^{(2)} \cdots dq^{(n-1)}. \tag{8}$$

After a time nΔt that is longer than the transition path time but much shorter than the mean first passage time, the committor can then be evaluated as

$$p_B(q_0) = \int P(q, n\Delta t|q_0) h_B(q) dq, \tag{9}$$

where $h_B(q) = 1$ for values of q in state B and otherwise $h_B(q) = 0$. Thus instead of seeking coordinates that give a sharply peaked committor distribution, our method seeks coordinates q for which swarms launched from any point on a $q(\mathbf{x}) = q_0$ isosurface give similar projections onto the q-axis. This dynamical self-consistency criterion is illustrated in Figures 2 and 3.

We emphasize that dynamical self-consistency alone is not sufficient to ensure reaction coordinate accuracy. Examples can be constructed where unimportant coordinates can also predict their own dynamical evolution. Consider for example overdamped dynamics with constant and isotropic diffusion on a two-dimensional bistable potential energy surface $\beta V(x, y) = K(x^2-1)^2 + y^2$ where $K \gg 1$. Clearly x is the reaction coordinate for transitions between $x = -1$ and $x = 1$. Indeed x does predict, regardless of y, its own short time evolution, i.e., $P(x, t|x_0, y_0) = P(x, t|x_0)$. However, the variable y also predicts its own evolution. Specifically, y relaxes at the same rate towards a distribution centered at $y = 0$ when initiated from any isosurface $y = y_0$, regardless of x. Yet, y does not even separate the stable basins at $x = -1$ and

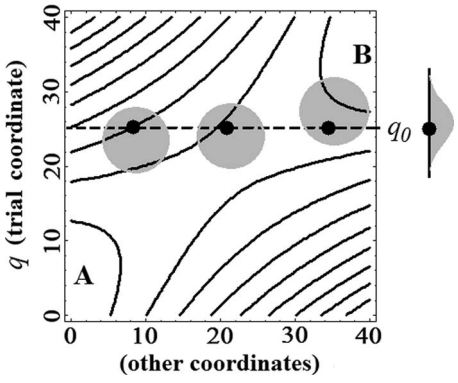

FIG. 2. For a reaction coordinate whose dynamics follow a one-dimensional Smoluchowski equation, swarms of trajectories from different individual configurations on each isosurface will drift and diffuse similarly. Therefore, for each isosurface, the projection of the individual swarms should each resemble the combined projection of all swarms (depicted to the right of the free energy surface). In the case depicted above, the individual swarms behave differently from each other, and therefore differently from their combined projection.

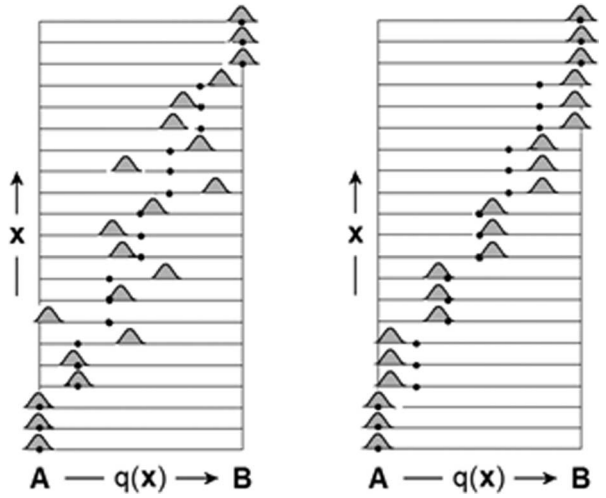

FIG. 3. If the dynamics of q follow a one-dimensional Smoluchowski equation, dynamical self-consistency should apply at all isosurfaces of q. (Left) For a poor reaction coordinate $q(\mathbf{x})$ the value of q alone is not a good predictor of drift or diffusion from the configuration $\mathbf{x}$. (Right) If q shows dynamical self-consistency for all isosurfaces of q from reactant A to product B, then q is an accurate reaction coordinate.

$x = 1$. The coordinate y in this example is a dynamically closed but spurious coordinate. The simple requirements of bistability and lack of bistability on isosurfaces of λ can eliminate many spurious candidates like the variable y in this example. For a good reaction coordinate, the drift direction along the reaction coordinate should "split," drifting away from some isosurface $q = q_0$ between states A and B. Spurious coordinates like y may be more difficult to recognize in realistic applications. Thus, as with all other coordinate identification methods, a final committor analysis is recommended to validate the optimal reaction coordinates.[11]

METHOD

Similar to the likelihood maximization methods,[12–14] the new method makes no *a priori* assumptions about the reaction coordinate q or the subspace of component variables $\mathbf{q}$. We do, however, require a collection of initial trajectory swarms. An initial order parameter that may be a poor reaction coordinate helps to focus the swarm data in the relevant part of configuration space. Let this initial order parameter be denoted $\lambda(\mathbf{x})$. The free energy as a function of $\lambda(\mathbf{x})$ is then

$$\beta F_\lambda(\lambda) = -\ln \int d\mathbf{x}\, e^{-\beta E(\mathbf{x})} \delta[\lambda - \lambda(\mathbf{x})]. \qquad (10)$$

Some caveats should be considered in choosing $\lambda(\mathbf{x})$. As noted above, within the stable reactant and product basins, many coordinates will self-consistently evolve toward the stable minima. The initial coordinate will help eliminate unimportant dynamics within the stable basin from consideration. The initial variable $\lambda(\mathbf{x})$ should therefore give a bistable free energy landscape with a maximum at some location $\lambda_{\max}$ that clearly separates basins A and B. Additionally, the ensemble of configurations on the surface $\lambda(\mathbf{x}) = \lambda_{\max}$ should not be bimodal. In some cases a suitable coordinate $\lambda(\mathbf{x})$ is easily found, and as shown below $\lambda(\mathbf{x})$ can even be orthogonal to the true reaction coordinate in some cases. In other cases, these requirements on $\lambda(\mathbf{x})$ are, unfortunately, only met when $\lambda(\mathbf{x})$ is already a reasonable reaction coordinate.

When coordinate dependent diffusion is important, then it is the modified free energy $\beta F_\lambda(\lambda) - \ln D(\lambda)$ that should be bistable with a clear separation between A and B. Methods to obtain the coordinate dependent diffusivity $D(\lambda)$ are described elsewhere.[38,49–51,53] Once $\lambda(\mathbf{x})$ and the modified free energy $\beta F_\lambda(\lambda) - \ln D(\lambda)$ have been obtained, we define a dimensionless and scale invariant bias potential $\Lambda(\mathbf{x})$:

$$\Lambda(\mathbf{x}) \equiv \beta E(\mathbf{x}) - 2\beta F_\lambda(\lambda(\mathbf{x})) + \ln D(\lambda(\mathbf{x})). \qquad (11)$$

The bias potential $\Lambda(\mathbf{x})$ weights configurations in the same way that they would enter a mean first passage time calculation with λ as the reaction coordinate. It is easily verified that $\int e^{-\Lambda(\mathbf{x})} \delta[\lambda - \lambda(\mathbf{x})] d\mathbf{x} = e^{+\beta F(\lambda)}/D(\lambda)$, and that the stable basins have negligible support in the density $\exp[-\Lambda(\mathbf{x})]$. An important feature of the bias potential $\Lambda(\mathbf{x})$ in Eq. (11) is its invariance[74] for transformations between monotonically related coordinates λ_1 and λ_2 (see the Appendix for a derivation). This feature diminishes the dependence of the resulting sampling of initial points on the choice of the initial coordinate $\lambda(\mathbf{x})$. In contrast, using a simpler biasing potential such as windows between hard walls will introduce such dependence. The potential $\Lambda(\mathbf{x})$ is illustrated in Figure 4 for a model potential energy surface.

The distribution $\exp[-\Lambda(\mathbf{x})]$ should be sampled to harvest initial points for the trajectory swarm calculations. The swarms at points sampled from $\exp[-\Lambda(\mathbf{x})]$ can be used to test many alternative trial reaction coordinates, and thus the data collection step will not have to be repeated. The remainder of the manuscript gives all formulas with quantities weighted by the factor $\exp[-\Lambda(\mathbf{x})]$. Note that in practical applications these weighting factors represent data points sampled from the distribution $\exp[-\Lambda(\mathbf{x})]$.

Suppose we wish to test a trial reaction coordinate q for dynamical self-consistency. Using the data harvested during

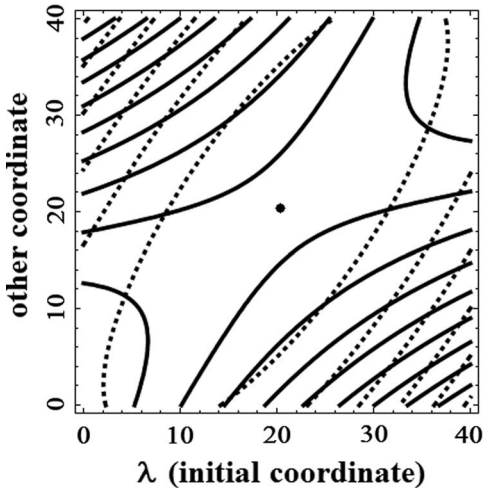

FIG. 4. The solid curves are contours of the actual free energy landscape $\beta F(q_1, q_2)$ with a saddle point at the round dot. The coordinate q_1 has been used as the initial coordinate, i.e., $\lambda(\mathbf{x}) = q_1(\mathbf{x})$. The dotted contours are curves of constant Λ. Note that Λ is peaked where the original landscape had a saddle point.

the sampling of $\exp[-\Lambda(\mathbf{x})]$ we can construct (a sample of) the conditional distribution

$$\rho_\lambda(\mathbf{x}_0|q_0) = \frac{e^{-\Lambda(\mathbf{x_0})}\delta[q(\mathbf{x_0}) - q_0]}{\int d\mathbf{x}\, e^{-\Lambda(\mathbf{x})}\delta[q(\mathbf{x}) - q_0]}. \quad (12)$$

The combined projection of all trajectory swarms initiated from the $q(\mathbf{x}) = q_0$ isosurface (weighted by their probability in the $\exp[-\Lambda(\mathbf{x})]$ distribution) is given by

$$P^{(\lambda)}(q, t|q_0) = \int d\mathbf{x}_t \int d\mathbf{x}_0\, \delta[q(\mathbf{x}_t) - q]p(\mathbf{x}_t, t|\mathbf{x}_0)\rho_\lambda(\mathbf{x}_0|q_0). \quad (13)$$

If q is a reaction coordinate for which a valid one-dimensional Smoluchowski equation exists, then Eq. (7) holds for all points on isosurface $q(\mathbf{x}) = q_0$, i.e., the projections of individual swarms onto q should be identical to the distribution $P^{(\lambda)}(q,t|q_0)$. The similarity or difference between two probability distributions can be quantified by the Kullback-Leibler divergence Δ_{KL},[75]

$$\Delta_{KL}[P_1||P_2] = \int d\mathbf{x} P_1(\mathbf{x}) \ln(P_1(\mathbf{x})/P_2(\mathbf{x})). \quad (14)$$

Note that $\Delta_{\mathrm{KL}}[P_1\|P_2] \neq \Delta_{\mathrm{KL}}[P_2\|P_1]$ unless $P_1 = P_2$. The reference (P_2) and comparison (P_1) distributions should be chosen (when possible) so that $P_2(\mathbf{x})$ is non-zero at all $\mathbf{x}$ where $P_1(\mathbf{x})$ is non-zero. The analysis below uses this convention to ensure a non-singular formulation. Mathematically, the Kullback-Leibler divergence is the likelihood of a probability distribution given some "binned" data.[76] In the context of equilibrium statistical mechanics, $\Delta_{\mathrm{KL}}[P_1\|P_2] \geq 0$ is the Gibbs inequality.[77] Δ_{KL} has been used to optimize protocols for non-equilibrium work calculations,[78] and to optimize coarse grained models from all-atom simulations.[79–81] To detect dynamical self-consistency at q_0, test whether each individual projected swarm from $\mathbf{x}$ on the isosurface $q(\mathbf{x}) = q_0$ has zero Kullback-Leibler divergence relative to the combined projection of swarms from different points on the isosurface $q(\mathbf{x}) = q_0$. The $\exp[-\Lambda(\mathbf{x})]$ weighted average Kullback-Leibler divergence between the individual swarms $P(\mathrm{q}, \mathrm{t}|\mathbf{x}_0)$ and their combined projection $P^{(\lambda)}(\mathrm{q}, \mathrm{t}|\mathrm{q}_0)$ is

$$\Delta_{KL}^{(\lambda)}[q_0] = \int d\mathbf{x}_0 \rho_\lambda(\mathbf{x}_0|q_0)\Delta_{KL}[P(q, t|\mathbf{x}_0)||P^{(\lambda)}(q, t|q_0)]. \quad (15)$$

When $\Delta_{\mathrm{KL}}{}^{(\lambda)}[q_0] = 0$, the projected dynamics from the specific isosurface $q(\mathbf{x}) = q_0$ have perfect self-consistency. Some positive deviation from zero will always occur because of sampling noise, but additional deviations from zero indicate a loss of information about the dynamics upon projection to the variable q. To obtain a global measure of dynamical self-consistency, the contributions from each isosurface of $q(\mathbf{x})$ must be integrated,

$$\langle\Delta_{KL}[q(\mathbf{x})]\rangle_\Lambda = \int dq_0 \rho^{(\lambda)}(q_0)\Delta_{KL}^{(\lambda)}[q_0]. \quad (16)$$

The notation $\langle\Delta_{\mathrm{KL}}[q(\mathbf{x})]\rangle_\Lambda$ indicates that the quantity in Eq. (16) is a functional of the map $q(\mathbf{x})$ from any configuration $\mathbf{x}$ to the reaction coordinate q. Note, however, that all $\mathbf{x}$-dependence was integrated away in Eq. (15). The distribution $\rho^{(\lambda)}(q)$ in Eq. (16) is a projection of the distribution $\exp[-\Lambda(\mathbf{x})]$ onto trial coordinate q

$$\rho^{(\lambda)}(q_0) = Z_\lambda^{-1} \int d\mathbf{x}\, e^{-\Lambda(\mathbf{x})}\delta[q(\mathbf{x}) - q_0]. \quad (17)$$

with $Z_\lambda = \int d\mathbf{x} \exp[-\Lambda(\mathbf{x})]$. To see that Eq. (16) is an $\exp[-\Lambda]$-weighted average of Δ_{KL} for projections of the swarms onto the trial coordinate q, assemble all of the pieces together to obtain

$$\langle\Delta_{KL}[q(\mathbf{x})]\rangle_\Lambda = Z_\lambda^{-1} \int dq_0 \left\{ \int d\mathbf{x}_0 e^{-\Lambda(\mathbf{x}_0)}\delta[q(\mathbf{x}_0) - q_0] \times \Delta_{KL}[P(q, t|\mathbf{x}_0)||P^{(\lambda)}(q, t|q_0)] \right\}. \quad (18)$$

A reaction coordinate with the dynamical self-consistency property should be a functional minimizer of $\langle\Delta_{\mathrm{KL}}[q(\mathbf{x})]\rangle_\Lambda$. An algorithm for computing $\langle\Delta_{\mathrm{KL}}[q(\mathbf{x})]\rangle_\Lambda$ is given below.

ALGORITHM

1) Umbrella sample with a bias potential $\mathrm{V_{bias}(x)}$ along an initial coordinate $\lambda(\mathbf{x})$ to obtain $\beta\mathrm{F}_\lambda(\lambda)$ and if applicable $\mathrm{lnD}(\lambda)$. $\beta\mathrm{F}_\lambda(\lambda)$ should be bistable with a maximum between A and B at some position λ_{max}. Additionally, the distribution of configurations on surface $\lambda = \lambda_{\mathrm{max}}$ should not be bimodal. Data from regions beyond the reactant and product minima should not be included. Store the configurations from each sampled window.
2) In each λ-window, the data was originally sampled according to the weight $\exp(-\beta\mathrm{E}(\mathbf{x}) - \beta\mathrm{V_{bias}}(\lambda))$ where $\mathrm{V_{bias}}(\lambda)$ was the bias used for umbrella sampling. Reweight the data by assigning each configuration the weight $\exp(\beta\mathrm{V_{bias}}(\lambda) + 2\beta\mathrm{F}_\lambda(\lambda) - \mathrm{lnD}(\lambda))/\mathrm{N}$ where N is the number of configurations in the λ-window.
3) Sample $\exp(-\Lambda(\mathbf{x}))$ by resampling the aggregate umbrella sampling data according to the weights assigned in step (2). Because of the factors of N in the reweighting, the data from all of the λ-windows can be combined for sampling.
4) For each sampled configuration in step (3), generate a swarm of short trajectories. We report calculations with 1000 trajectories per swarm in this study, but 100 trajectories per swarm gave similar results. The swarm trajectories should be of much longer duration than the velocity-velocity correlation time so that a description based on diffusive dynamics is appropriate. To minimize computational effort, the swarm trajectories should also be much shorter than the time required to relax to reactant and product states.
5) Select a trial coordinate q and divide the range of q into uniformly and narrowly spaced windows. Bin the data sampled in step (3) into the q-windows. (It has already been sampled according to the appropriate distribution). Within each q-window, combine all of the swarm data into a projected swarm using Eq. (13). For each configuration in the q-window, compute the Kullback-Leibler divergence between its individual swarm and the projected swarm, Eq. (15). Sum the total Kullback-Leibler

divergence over all sampled configurations from all windows. Divide by the total number of individual swarms. This is the expression in Eq. (18).

For applications involving molecular simulation, only steps (1) and (4) of the algorithm will involve simulations for data generation. The data generation steps involve a free energy calculation and the generation of swarms. The cost of these steps should be comparable to the cost of other methods that reconstruct free energy landscapes and optimize reaction paths based on swarms.[64,67]

We recommend some pre-screening tests for each trial reaction coordinate. In real applications the importance or unimportance of some coordinates may not be obvious *a priori*, so a quantitative pre-screening of trial coordinates is important. $\lambda(\mathbf{x})$, when properly defined, is a good order parameter, which means it should be able to distinguish between A and B. Data obtained in calculating $F_\lambda(\lambda)$ can therefore be used to pre-screen trial reaction coordinates before using the dynamical self-consistency test in step (5) of the algorithm. For example, a correlation coefficient between the trial coordinate and the initial coordinate λ can be computed using only data from states A and B. The absolute value of this correlation coefficient should be near unity for viable trial coordinates. Here, data from basins A and B should contribute equally regardless of their relative free energies, otherwise the correlation coefficient will reflect only correlations within the more stable basin. One could alternatively require that projections of states A and B onto each trial coordinate give non-overlapping distributions. Using these pre-screening tests, many coordinates related to irrelevant saddles and shallow minima can be immediately eliminated from consideration.

RESULTS

We illustrate the dynamical self-consistency test for identifying accurate reaction coordinates with three simple examples employing two-dimensional model potentials with anisotropic diffusion.

Example I: Nucleation with anisotropic structure formation kinetics

The model

The first example is taken from research on the effects of diffusion anisotropy in nucleation kinetics. Sear[61] employed a simple model free energy landscape of the form,

$$\beta F(n_F, n_S) = \beta \Delta F_\ddagger \{1 - (n_F + n_S - n_\ddagger)^2 / n_\ddagger^2\} - \chi n_F n_S, \tag{19}$$

where n_F and n_S denote the number of fast and slow spins in the nucleus (see Figure 5). Thus, n_F represents a fast coordinate and n_S represents a slow coordinate. $n_\ddagger$ is the location of the barrier top along either n_F when n_S is frozen or along n_S when n_F is frozen. Note that n_F, $n_S > 0$ because cluster sizes are non-negative quantities. $\Delta F_\ddagger$ represents the free energy barrier if either of the two coordinates n_F or n_S are frozen at zero. The $\chi n_F n_S$ term lowers the free energy barrier when the coordinates evolve together along the $n_F = n_S$ line. Sear used parameters that yield a small 8 kT free energy barrier with extremely large diffusion anisotropy.[61] As outlined by Berezhkovskii and Zitserman,[31] the escape pathways in this limit become dependent on the initial conditions. In particular, trajectories started at a given value of n_S, whether large or small, will escape before n_S can change. In some cases, the diffusion anisotropies used by Sear were factors of a million different for motion in different directions.[61] Our parameters were chosen to avoid the Berezhkovskii-Zitserman regime[31] of extreme diffusion anisotropy, where the reaction coordinate is thought not to exist. We use $\beta \Delta F_\ddagger = 32$, $n_\ddagger = 64$, and $\chi = 3/160$, which results in a saddle point at location (n_F, n_S) = (20, 20). The model free energy landscape is shown in Figure 5.

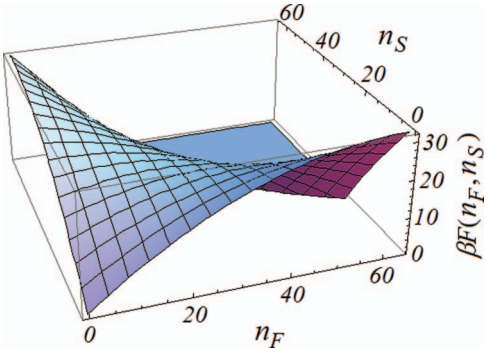

FIG. 5. The free energy landscape for a model of nucleation where the nucleus size can change along either fast (n_F) and slow (n_S) mobility directions.

The dynamics on the free energy landscape is modeled using kinetic Monte Carlo[82,83] with separate attachment and detachment rates for the slow (n_S) and fast (n_F) degrees of freedom. Following Sear,[61] the kinetics for n_S evolution are given relative to the kinetics for n_F evolution. The rate of attachment for the n_F coordinate, i.e., the rate constant for $n_F \rightarrow n_F + 1$, is τ^{-1}. The rate of detachment, i.e., the process $n_F \rightarrow n_F - 1$, is given by $\tau^{-1}\exp[\beta F(n_F, n_S) - \beta F(n_F - 1, n_S)]$. The corresponding attachment rate for the n_S degree of freedom is slower by a factor of s, and is thus equal to $s\tau^{-1}$. Analogous to the n_F variable, the detachment rate for the n_S variable is equal to $s\tau^{-1}\exp[\beta F(n_F, n_S) - \beta F(n_F, n_S - 1)]$. The factor s therefore controls diffusion anisotropy on the free energy landscape.

From the diffusion anisotropy and the quadratic free energy, the KLBS theory can provide accurate reaction coordinates for purposes of comparison. Ingredients in this theory are the diffusion tensor $\mathbf{D}$, which, apart from a multiplicative scalar that has no effect on the reaction coordinate, is

$$\mathbf{D} = \begin{bmatrix} 1 & 0 \\ 0 & s \end{bmatrix}, \tag{20}$$

and the second derivative matrix $\mathbf{A}$ of the free energy

$$\mathbf{A} = \begin{bmatrix} \partial^2 \beta F / \partial n_F^2 & \partial^2 \beta F / \partial n_F \partial n_S \\ \partial^2 \beta F / \partial n_S \partial n_F & \partial^2 \beta F / \partial n_S^2 \end{bmatrix}_\ddagger. \tag{21}$$

$\mathbf{A}$ and $\mathbf{D}$ were used only to provide exact reaction coordinates from KLBS theory. The dynamical self-consistency test does not require either quantity.

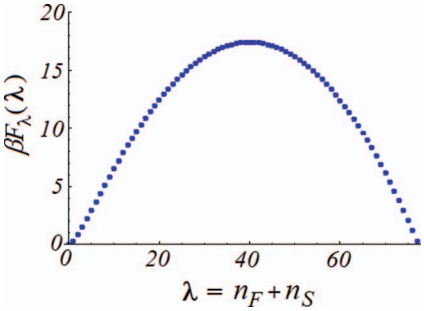

FIG. 6. The free energy as a function of the initial coordinate $\lambda = n_F + n_S$.

Simulations

The collective variable used to compute the weighting for initial conditions of the swarms was chosen along the diagonal in the free energy landscape, i.e., $\lambda = n_F + n_S$. Figure 6 shows the projected free energy $\beta F_\lambda(\lambda)$ as a function of λ. The activation barrier is approximately 17 kT. Because the diffusion tensor is coordinate independent, the $\ln D(q)$ term is a constant in the construction of the potential $\Lambda(n_F, n_S) = \beta F(n_F, n_S) - 2\beta F_\lambda(n_F + n_S)$. We performed simulations with different settings of the anisotropy $s = 1.0$, 0.3, 0.1, and 0.03. For each point in the discrete space (n_F, n_S) we initialized a swarm. Individual swarms contain 1000 short trajectories, each of duration 4.0τ, but the method gave similar results with 100 trajectories per swarm. For a few initial points, Figure 7 shows histograms of the endpoints of 1000 trajectories (a "swarm") after they have evolved for a duration 2.0τ. (A shorter time 2.0τ was used for the graphic to prevent the spreading packets from overlapping in the figure.)

Optimization of the reaction coordinate

Next, we use the dynamical self-consistency test to optimize the reaction coordinate. We represent trial reaction coordinates q by the angle θ with the n_F-axis as shown in Figure 8. The dynamical self-consistency test results are shown in Figure 9. With isotropic diffusion, i.e., $s = 1.0$, the initial coordinate $\lambda = n_F + n_S$, is the correct reaction coordinate, so the optimum predicted by our method should be at corresponding $\theta = \pi/4$. As s becomes increasingly small the optimal reaction coordinate rotates toward the n_S direction, i.e., the optimal angle θ grows toward $\pi/2$.

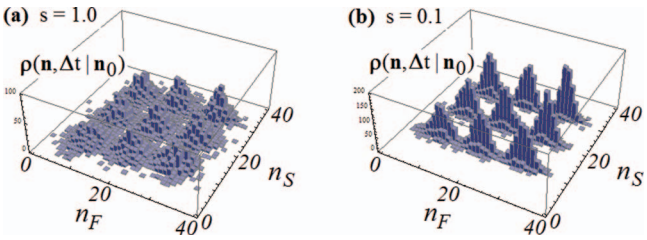

FIG. 7. Histograms of endpoints of 1000 trajectories 2.0τ after initiation. Each figure shows nine swarms initiated at nine different points on the free energy landscape. In (a) the anisotropy is $s = 1.0$, and in (b) the anisotropy is $s = 0.1$. Differences in the way the swarms drift are difficult to visually discern, but the dynamical self-consistency test can detect differences.

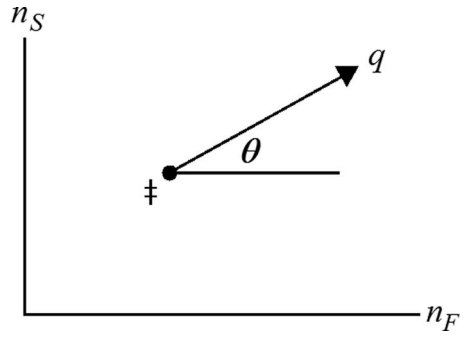

FIG. 8. The trial reaction coordinate q can be represented by the angle θ between the direction of progress along q and the n_F-axis.

As expected from KLBS theory, the reaction coordinate rotates toward n_S for stronger diffusion anisotropy. However, the nature of the free energy surface does not allow the reaction coordinate to rotate as far as the n_S or n_F axes ($\theta = \pi/2$ or $\theta = 0$). If q is progress along the vector $(\cos\theta, \sin\theta)$, then the dividing surface must be parallel to the vector $\mathbf{v}_\perp = (\sin\theta, -\cos\theta)$. The quadratic form $\mathbf{v}_\perp{}^{\mathrm{T}}\mathbf{A}\mathbf{v}_\perp$ must be positive to construct a dividing surface partition function. The saddle point in the free energy of Eq. (19) is broad in the direction normal to the minimum free energy path, i.e., the pathway is not a narrow tube. Because of the broad saddle, the optimal θ must always be in the interval $0.425\pi < \theta < 0.751\pi$. Figure 9 shows that for $s = 0.03$, the optimal θ is approaching the upper bound of 0.425π.

Reactive trajectories for $s = 0.03$ resemble two-step nucleation trajectories. Fluctuations to large n_S occur with lower frequency than fluctuations to equivalently large n_F. However, when fluctuations to large n_S do occur, they persist for a long time, giving the n_F degree of freedom ample time to cross the effectively lowered barrier. Once the saddle region has been crossed the trajectory can proceed downward on the free energy landscape moving primarily in the n_F direction. The system gets farther and farther from the minimum free energy pathway as it proceeds. Peters has suggested this effect of dynamical anisotropy as a mechanism for metastable polymorph selection.[40] The red trajectory in the schematic Figure 10 illustrates this dynamical two-step behavior.

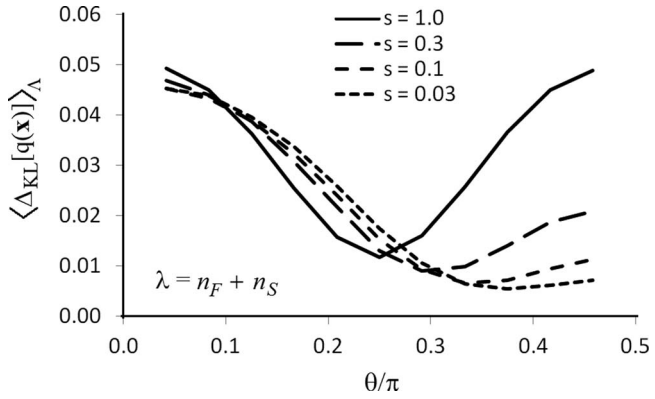

FIG. 9. $\exp[-\Lambda]$ weighted Kullback-Leibler divergences for different trial reaction coordinates (represented by θ) and for different values of the diffusion anisotropy, $s = 1.0$, 0.3, 0.1, 0.03.

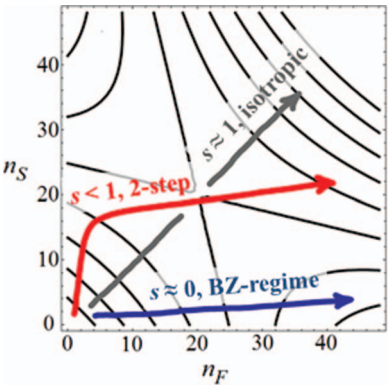

FIG. 10. Illustrating three mechanistic regimes that prevail for different degrees of diffusion anisotropy. When $s \approx 1$, the diffusion tensor is isotropic and most pathways follow the minimum free energy path. When $s < 1$, but not extremely small, a two step nucleation mechanism prevails with initial motion along the slow coordinate before escape in the n_F-direction. Finally, when s is extremely small the Berezhkovskii-Zitserman (BZ) regime prevails. In the BZ-regime, trajectories can escape in the n_F-direction with the n_S degree of freedom frozen.

In the Berezhkovskii-Zitserman (BZ) regime, i.e., when $s \approx 0$ becomes extremely small, escaping trajectories will entirely avoid the saddle point as seen in the work by Sear.[61] The blue trajectory in Figure 10 schematically illustrates an escape path in the BZ-regime. Finally, when $s = 1$, we have isotropic diffusion, and typical trajectories follow the minimum free energy path (MFEP). The gray line in Figure 10 approximates the MFEP for the free energy landscape of Eq. (19).

Example II: A narrow tube landscape

The model

This second example tests the method for a narrow tube free energy landscape and with different choices for the initial collective variable λ. Again the model has slow and fast degrees of freedom n_S and n_F. The free energy landscape now has the coordinate n_S bilinearly coupled to the n_F degree of freedom,

$$\beta F(n_F, n_S) = \beta \Delta F_{\ddagger}\{1 - (n_F - n_{\ddagger})^2/n_{\ddagger}^2\} + \chi(n_S - n_F)^2. \tag{22}$$

The parameters are $\beta \Delta F_{\ddagger} = 18$, $n_{\ddagger} = 20$, and $\chi = 1/16$. In this model, $\Delta F_{\ddagger}$ is the free energy barrier when n_S relaxes adiabatically in response to changes in n_F. But because n_S is the slower degree of freedom adiabatic relaxation cannot actually occur. For these parameters, the free energy landscape (shown in Figure 11) confines the transition paths to a narrow tube. The differences between the free energy landscapes of Figures 5 and 11 have important consequences for the dynamics. When n_S is dynamically frozen in the broad landscape of Figure 5, escape can still occur in the n_F direction. In contrast, when n_S is dynamically frozen in narrow tube landscape, n_F is dynamically caged within a small range of n_F-values which are thermally accessible at constant n_S. Because of the harmonic valley shape of the narrow tube, motion toward the barrier top by either coordinate induces the other coordinate to follow. Counter-intuitively, this property means that narrow tubes admit a wider range of admissible trial reaction coordinates.

The effect of the choice of λ

Using the above-mentioned narrow tube model we tested the sensitivity of our method to the choice of the initial collective variable λ. Figures 12(a) and 12(b) show projections of the free energy onto the choices $\lambda = n_S$ and $\lambda = n_F$ coordinates, respectively. Dynamical self-consistency optimizations were separately performed, first using $\lambda = n_S$, and then using $\lambda = n_F$.

Swarms were generated at each point of the (n_F, n_S) landscape. The two initial collective variables yield different $\Lambda(n_F, n_S)$ potentials, so the data is weighted differently in the two cases. The dynamical self-consistency test given in Eq. (17) was applied to different reaction coordinates, again represented using the angle θ, for different degrees of the diffusion anisotropy s (see Figure 13).

The two choices of λ give optimal coordinates that are very similar as shown in Table I. Note that the dynamical self-consistency test for the narrow tube landscape can identify coordinates that are effectively orthogonal to the initial collective variable λ. For example, in the case where $s = 0.03$, the optimal reaction coordinate corresponds to $\theta \approx \pi/2$, which is nearly orthogonal to the initial choice $\lambda = n_F$.

An analogy can be made between Grote-Hynes theory[63] and the results of this example. The Grote-Hynes theory employs a generalized Langevin equation[29] to model non-adiabatic inertial dynamics in barrier crossings.[32,63,84] Pollak[85] and others[86] have noted that the one-dimensional

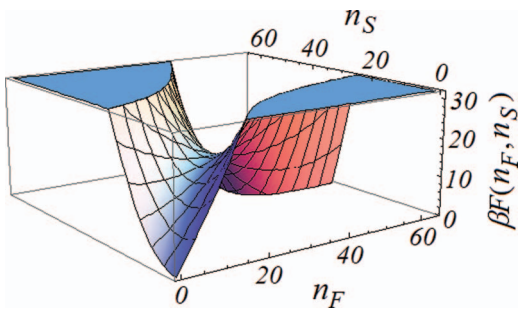

FIG. 11. Narrow tube type free energy landscape as a function of fast (n_F) and slow (n_S) coordinates. The saddle point is at $(n_F, n_S) = (20, 20)$.

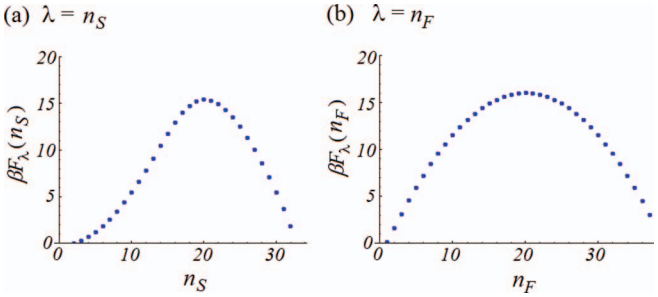

FIG. 12. Projections of the free energy onto different initial coordinates. (a) The initial coordinate was $\lambda = n_S$. (b) The initial coordinate was $\lambda = n_F$.

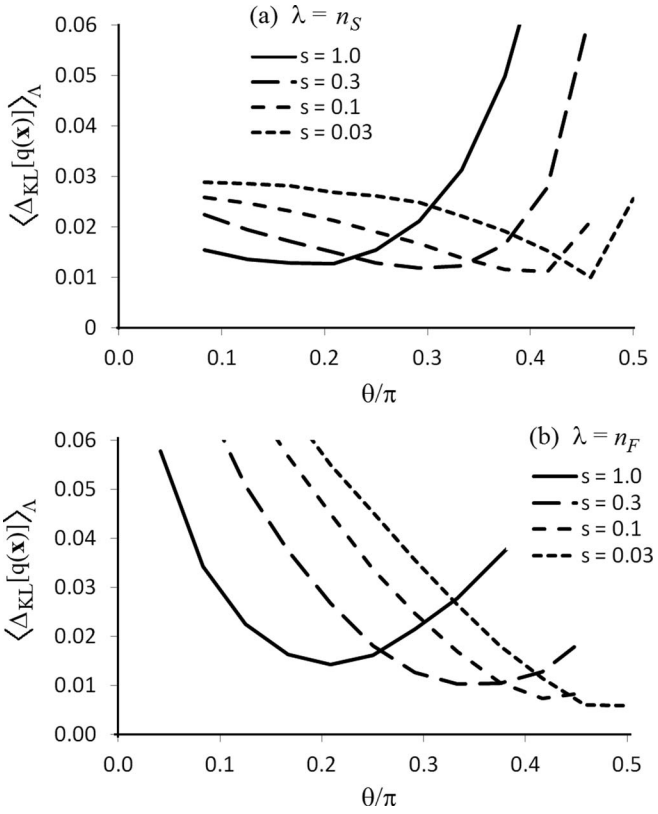

FIG. 13. $\langle\Delta_{KL}[q]\rangle_\Lambda$ for different coordinates (represented by the angle θ as shown in Figure 8). (a) The initial coordinate was $\lambda = n_S$. (b) The initial coordinate was $\lambda = n_F$. In both cases the optimal coordinate (the minimum) rotates toward n_S as the diffusion tensor becomes more anisotropic.

Grote-Hynes model is equivalent to the inertial dynamics near a saddle point in a multidimensional harmonic energy landscape. In the multidimensional model of Pollak,[85] dynamical caging emerges as the effect of slow solvent degrees of freedom that non-adiabatically trap the faster coordinate of the generalized Langevin equation. By redefining the reaction coordinate to include slower solvent components, Pollak[85] obtained a reaction coordinate whose (inertial) dynamics were independent of all other collective variables.[85,86]

The Grote-Hynes and Pollak models consider inertial dynamics, while we considered purely diffusive dynamics. However, the narrow tube and large anisotropy of the model system in Figure 11 correspond to strong coupling between slow and fast modes in Pollak's analysis.[85] Like Pollak we find an optimal reaction coordinate that is redirected toward the slower mode, n_S. Like Pollak we also find the optimal coordinate as that with (diffusive) dynamics that are independent of the other coordinates.

TABLE I. Comparison between reaction coordinates from the dynamical self-consistency test and from KLBS theory. Coordinates identified by dynamical self-consistency are summarized for two different initial coordinates. For narrow tube potential energy landscapes the dynamical self-consistency test can correctly identify accurate reaction coordinates even for an inaccurate initial coordinate λ.

	Optimal θ		Optimal θ
s	$\Delta_{KL}(\lambda = n_S)$	$\Delta_{KL}(\lambda = n_F)$	(KLBS)
1.00	0.21π	0.21π	0.195π
0.30	0.29π	0.33π	0.338π
0.10	0.42π	0.42π	0.425π
0.03	0.46π	0.50π	0.472π

Example III: When isosurfaces of λ give a bimodal distribution

In this example, we consider a two-dimensional empirical valence bond (EVB) potential. The potential is constructed from a Morse function: $m(x) = 100(\exp[-2(x-1)] - 2\exp[-(x-1)])$ and an exp-4 function: $g(y) = \exp[-y^4/256]$. These are used to obtain a potential energy surface as an adiabatic EVB ground state V(x, y) by solving the secular equation

$$\begin{vmatrix} g(\phi y - 1.5)m(\phi x + 2) - V & c \\ c & g(\phi y + 1.5)m(2 - \phi x) - V \end{vmatrix} = 0$$

with c = 20.0 and ϕ = 0.025. Simple Metropolis Monte Carlo trajectories at $k_BT = 1.0$ with trial moves that are unit steps in ±x or ±y directions at each step were used for the dynamics. The discrete random walk, similar to an overdamped Langevin dynamics after several steps, drifts in the direction $-\partial V/\partial(x, y)$. The potential and the minimum energy path from the saddle to the two minima are shown in Figure 14. This example is complicated because the barrier is extremely narrow in the reaction coordinate direction and wide in the transverse direction. The saddle region is therefore a broad sharp ridge rather than a narrow tube. A second source of difficulty is that the optimal reaction coordinate is not a linear combination of x and y, but rather a nonlinear vertical energy gap between the diabatic states. A third difficulty is that the stable basins are much longer in the y-direction than in the x-direction. Therefore *within the basins*, y decorrelates more slowly than x. If y were rescaled to obtain nearly circular basins, then the diffusion would become anisotropic with the variable y being slow. But even though y decorrelates more slowly than x *within the basins*, the dynamical bottleneck (and hence the *slowest* decorrelation

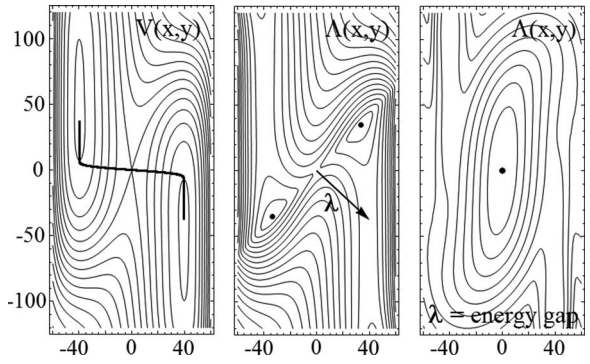

FIG. 14. Contour plots of (a) the EVB potential V(x, y) with the minimum energy path between the reactant and product minima, (b) the potential Λ(x, y) constructed with λ chosen as the end-to-end direction along the minimum energy path, and (c) the potential Λ(x, y) constructed with the ideal energy gap coordinate for λ. Contour spacings are $5k_BT$ in all plots.

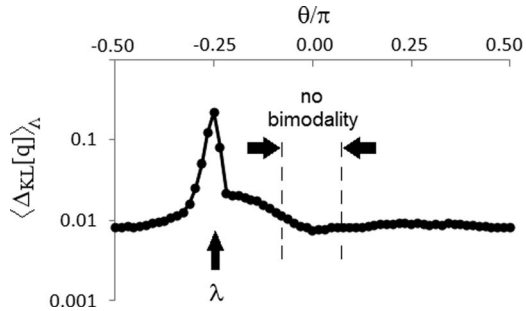

FIG. 15. $\langle \Delta_{KL}[q] \rangle_\Lambda$ for different linear trial coordinates q(x, y) represented by the angle θ. The initial coordinate λ(x, y) gives a hysteretic free energy $F_\lambda(\lambda)$ if sampled imperfectly. Coordinates similar to λ clearly do not have dynamical self-consistency, but other coordinates are not clearly distinguished in $\langle \Delta_{KL}[q] \rangle_\Lambda$.

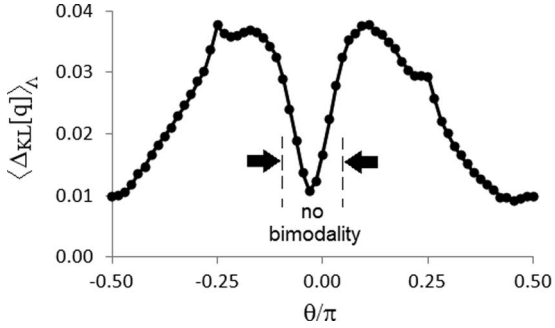

FIG. 16. $\langle \Delta_{KL}[q] \rangle_\Lambda$ for different linear trial coordinates q(x, y) represented by the angle θ. The initial coordinate λ(x, y) is the vertical energy gap between diabatic states of the EVB model.

time) arises from the high barrier *between the basins* in the x direction.

We first illustrate what happens when the initial coordinate λ does not meet the aforementioned requirements. The initial coordinate λ = 0.732x–0.681y changes along the end-to-end vector along the minimum energy path. Figure 14(b) shows an arrow in direction $u_\lambda = \nabla\lambda/||\nabla\lambda||$, and the potential Λ(x, y) which results from the choice λ = 0.732x–0.681y. Λ(x, y) has two minima because the distribution on the dividing surface λ = 0, i.e. $\langle \delta[\lambda(x, y)] \rangle^{-1} \exp[-\beta V(x, y)] \delta[\lambda(x, y)]$, is bimodal. The bimodality indicates that λ = 0.732x–0.681y does not fully remove the unstable direction from the λ isosurfaces. Neither of the minima in the potential Λ(x, y) coincide with the saddle point region of the original potential. Figure 15 shows the value of $\langle \Delta_{KL}[q] \rangle_\Lambda$ for each trial reaction coordinate q from $\theta = -\pi/2$ to $\theta = \pi/2$. Coordinates which are similar to the initial coordinate λ (indicated by the vertical arrow) have isosurfaces with support near both minima of Λ(x, y). These coordinates are clearly identified as being incorrect because swarms in these two basins evolve differently. In contrast, isosurfaces of coordinates that are very different from λ can have support in one of the two basins, but never both. These coordinates are essentially indistinguishable. Thus the dynamical self-consistency test fails to identify a reaction coordinate for the same reasons that lead to a bimodal distribution on λ = 0. (The correct coordinate, by a tiny margin, does minimize $\langle \Delta_{KL}[q] \rangle_\Lambda$, but there seems to be no physical reason to trust that result.)

As already indicated by the narrow range of q with unimodal dividing surface distributions in Figure 15, there is only a narrow range of admissible choices for λ. Because the saddle is a broad sharp ridge, λ must already be an accurate reaction coordinate to satisfy the no-bimodality requirement. The present method therefore cannot help identify coordinates for passage over such broad sharp ridges except perhaps to make fine adjustments to a coordinate that is already approximately correct.

Even when the dynamical self-consistency test cannot help find reaction coordinates, it can still determine whether an accurate reaction coordinate obtained by another method can accurately reduce the dynamics to a one-dimensional Smoluchowski equation. To illustrate this point, we construct Λ(x, y) by using the vertical energy gap for the coordinate λ(x). The resulting potential Λ(x, y) is shown in Figure 14(c). Figure 16 again shows $\langle \Delta_{KL}[q] \rangle_\Lambda$ for each trial reaction coordinate q from $\theta = -\pi/2$ to $\theta = \pi/2$. There are two minima in $\langle \Delta_{KL}[q] \rangle_\Lambda$. One is the spurious coordinate y which does not separate the two basins and for which all surfaces $y = y_0$ have bimodal distributions. y would therefore be discarded according to the recommendations for screening trial coordinates. The other minimum, within the "no bimodality" brackets, is the optimal linear reaction coordinate: q = 0.9952x–0.098y. Interestingly, the value of $\langle \Delta_{KL}[q] \rangle_\Lambda$ for the optimal linear coordinate is similar to values obtained for the exactly reducible quadratic free energy landscapes in examples 1 and 2 above. This suggests that dynamical self-consistency and reduction to a one-dimensional Smoluchowski equation may be achievable even in systems that initially appear to require more variables.

For future directions in tests for dynamical self-consistency, we note that there is a close link between the bias function Λ(x, y) in this work and the factor $\exp[-\beta V(x)]^* |\nabla p_B|^2$ which appears in the functional for optimizing the committor function in transition path theory (TPT).[26,87,88] Consider the case where p_B depends only on the collective variable λ with constant diffusion D, then by Eq. (3) we have $dp_B/d\lambda = \exp[\beta F(\lambda)]$ within a constant. Therefore $\exp[-\beta V(x)] |dp_B/d\lambda|^2 = \exp[-\lambda\ (x)]$. However, an important difference is that TPT can begin with several important initial coordinates[26] $\lambda_1, \lambda_2, \ldots, \lambda_n$ while the method in this work requires a single initial coordinate. Instead of the $\exp[-\Lambda(\mathbf{x})]$ weights used in this work, a global metric for dynamical self-consistency based on TPT might weight origins of the swarms by $\exp[-\beta V(x)]^* |\nabla p_B|^2$ for a set of several approximate initial coordinates. In cases like example III, where a single suitable initial coordinate is difficult to identify, the more general multivariable framework of TPT might help find accurate one-dimensional coordinates with dynamical self-consistency. However, we reiterate that reduction to a single coordinate with dynamical self-consistency is only possible in special cases.

We also note another future direction which is suggested by example III. When the single initial coordinate is chosen poorly, Λ(x) develops two minima as seen in Figure 14(b). Maximum information entropy measures could help identify

the coordinate which separates these two minima (as one of the many directions normal to λ). The additional coordinate, in combination with λ, spans a two-dimensional space which does contain the correct coordinate, as suggested by Figure 14(b). Certainly, a systematic reduction from many dimensions to two would also be useful.

CONCLUSIONS

This work introduces a novel way to identify optimal reaction coordinate by a dynamical self-consistency test. Specifically, the dynamics of short trajectories launched from an ensemble of configurations on isosurfaces of a trial coordinate are projected back onto the trial coordinate. If the dynamics of individual swarms at each point evolve like swarms initiated from all other points on the same trial coordinate isosurface, then the trial coordinate has a dynamically self-consistent projection property. For certain systems where the reaction coordinate evolution can be described by a one-dimensional Smoluchowski model, coordinates which are dynamically self-consistent in the transition pathway can be identified as accurate reaction coordinates. We use the Kullback-Leibler divergence[75] to screen trial reaction coordinates for dynamical self-consistency. The swarms of short dynamical trajectories used to screen trial coordinates are generated only once by sampling along an initial coordinate. The initial coordinate should be a good order parameter for which the free energy calculation gives no bimodality. Once initial swarm data are generated, they can be used to test other trial reaction coordinates, much like the strategy of the earlier likelihood maximization approach.[12–14]

For free energy landscapes of the "narrow tube" type as shown in Figure 11, the method is robust and insensitive to the choice of the initial coordinate. The results including dynamics with anisotropic diffusion tensors were consistent with the exact reaction coordinates from KLBS theory.[6] The method also performed well for models with moderately broad saddles that have been considered in the context of nucleation.[60,61,89] However, the proposed method performs poorly for broad sharp ridges, i.e., when the barrier is narrow along the reaction coordinate and very wide in transverse directions. For this type of landscape nearly all initial coordinates give bimodal dividing surface distributions, so the only viable choice for the initial coordinate becomes the unknown reaction coordinate. The method is therefore less general than methods based on full-length transition paths. Nevertheless, the method may be useful in certain applications where transition path sampling simulations are not feasible. Furthermore, the framework provides insight on the relationship between reaction coordinates, committors, and short-time dynamics in those special cases where dynamics along the reaction coordinate can be described by a one-dimensional Smoluchowski equation. Beyond finding reaction coordinates, the dynamical self-consistency test can also assess whether reaction coordinates obtained by other methods can yield an accurate one-dimensional Smoluchowski model for the multidimensional dynamics.

ACKNOWLEDGMENTS

This work was supported by NSF CAREER (Grant No. 0955502) and by a NSF Computational Discovery and Innovation (CDI) (Grant No. 1125235). We thank Professor Scott Shell whose work stimulated our thinking about the Kullback-Leibler divergence. We thank the reviewers for helpful suggestions. Finally, we thank David Wu, Amadeu Sum, Gregg Beckham, Valeria Molinero, and Frank Noe for stimulating discussions.

APPENDIX: AN INVARIANCE OF Λ(x)

Equation (11) proposes the construction of a bias potential from an initial collective variable $\lambda_1(x)$. Suppose instead an alternative collective variable λ_2 was used with λ_2 being some monotonic function $\lambda_2 = \lambda_2(\lambda_1)$. As λ_1 and λ_2 depend monotonically on each other, there is a one-to-one mapping of the isosurfaces of the bias potential $\mathbf{\Lambda(x)} = \text{constant}$. An ideal biasing function should therefore weight configurations identically under either of these two initial coordinate choices. Equation (11) has such an invariance property[74] for monotonically related coordinates. Consider the bias potentials

$$\Lambda_1(\mathbf{x}) \equiv \beta E(\mathbf{x}) - 2\beta F_1(\lambda_1(\mathbf{x})) + \ln D_1(\lambda_1(\mathbf{x})), \tag{A1}$$

and

$$\Lambda_2(\mathbf{x}) \equiv \beta E(\mathbf{x}) - 2\beta F_2(\lambda_2(\mathbf{x})) + \ln D_2(\lambda_2(\mathbf{x})). \tag{A2}$$

Now suppose that a monotonic function relates λ_2 and λ_1, with Jacobian corrections for the change of coordinates the two free energies are related,

$$\beta F_2(\lambda_2) = \beta F_1(\lambda_1) - \ln \left| \frac{d\lambda_1}{d\lambda_2} \right|. \tag{A3}$$

Similarly, coordinate transformation of the local diffusion constant gives

$$D_2(\lambda_2) = D_1(\lambda_1) \left| \frac{d\lambda_2}{d\lambda_1} \right|^2. \tag{A4}$$

Inserting Eqs. (A3) and (A4) into Eq. (A2) gives Eq. (A1). Therefore the biasing function as defined in Eq. (11) is invariant among physically equivalent initial coordinates.

[1] P. G. Bolhuis, D. Chandler, C. Dellago, and P. G. Geissler, Annu. Rev. Phys. Chem. **53**, 291 (2002).
[2] E. Rosta, H. L. Woodcock, B. R. Brooks, and G. Hummer, J. Comput. Chem. **30**, 1634 (2009).
[3] T. S. van Erp, J. Chem. Phys. **125**, 174106 (2006).
[4] B. Peters, Molec. Sim. **36**, 1265 (2010).
[5] A. Ma and A. R. Dinner, J. Phys. Chem. B **109**, 6769–6779 (2005).
[6] A. Berezhkovskii and A. Szabo, J. Chem. Phys. **122**, 014503 (2005).
[7] Y. M. Rhee and V. S. Pande, J. Phys. Chem. B **109**, 6780 (2005).
[8] W. E, W. Q. Ren, and E. Vanden-Eijnden, Chem. Phys. Lett. **413**, 242–247 (2005).
[9] L. Maragliano, A. Fischer, E. Vanden-Eijnden, and G. Ciccotti, J. Chem. Phys. **125**, 024106 (2006).
[10] F. Noe, C. Schutte, E. Vanden-Eijnden, L. Reich, and T. R. Weikl, Proc. Natl. Acad. Sci. U.S.A. **106**, 19011–19016 (2009).
[11] B. Peters, J. Chem. Phys. **125**, 241101 (2006).
[12] B. Peters and B. L. Trout, J. Chem. Phys. **125**, 054108 (2006).
[13] B. Peters, G. T. Beckham, and B. L. Trout, J. Chem. Phys. **127**, 034109 (2007).
[14] W. Lechner, J. Rogal, J. Juraszek, and P. G. Bolhuis, J. Chem. Phys. **133**, 174110 (2010).

[15]R. B. Best and G. Hummer, Proc. Natl. Acad. Sci. U.S.A. **102**, 6732 (2005).
[16]D. Antoniou and S. D. Schwartz, J. Phys. Chem. B **115**, 2465 (2011).
[17]B. Peters, J. Phys. Chem. B **115**, 12671 (2011).
[18]J. D. Chodera and V. S. Pande, Phys. Rev. Lett. **107**, 098102 (2011).
[19]S. Krivov, S. Muff, and A. Caflisch, J. Phys. Chem. B **112**, 8701 (2008).
[20]S. Muff and A. Caflisch, J. Chem. Phys. **130**, 125104 (2009).
[21]S. Krivov, J. Phys. Chem. B **115**, 11382 (2011).
[22]M. A. Rohrdanz, W. Zheng, M. Maggioni, and C. Clementi, J. Chem. Phys. **134**, 124116 (2011).
[23]E. E. Borrero and F. A. Escobedo, J. Chem. Phys. **127**, 164101 (2007).
[24]A. Berezhkovskii, G. Hummer, and A. Szabo, J. Chem. Phys. **130**, 205102 (2009).
[25]R. Du, V. S. Pande, A. Y. Grosberg, T. Tanaka, and E. S. Shakhnovich, J. Chem. Phys. **108**, 334 (1998).
[26]W. E and E. Vanden-Eijnden, Annu. Rev. Phys. Chem. **61**, 391 (2010).
[27]L. Onsager, Phys. Rev. **54**, 554 (1938).
[28]C. Dellago, P. G. Bolhuis, and D. Chandler, J. Chem. Phys. **108**, 1964 (1998).
[29]A. Nitzan, *Chemical Dynamics in Condensed Phases* (Oxford University Press, Oxford, 2006).
[30]D. Chandler, J. Chem. Phys. **68**, 2959 (1978).
[31]A. Berezhkovskii and V. Y. Zitserman, Chem. Phys. Lett. **158**, 369 (1989).
[32]J. P. Bergsma, B. J. Gertner, K. R. Wilson, and J. T. Hynes, J. Chem. Phys. **86**, 1356 (1987).
[33]O. A. Karim and J. A. McCammon, Chem. Phys. Lett. **132**, 219 (1986).
[34]G. Ciccotti, M. Ferrario, J. T. Hynes, and R. Kapral, J. Chem. Phys. **93**, 7137 (1990).
[35]G. S. Buchner, R. D. Murphy, N.-V. Buchete, and J. Kubelka, Biochem. Biophys. Acta - Proteins and Proteomics **1814**, 1001 (2011).
[36]K. A. Dill, B. S. Ozkan, M. S. Shell, and T. R. Weikl, Annu. Rev. Biophys. **37**, 289 (2008).
[37]J. Juraszek, J. Vreede, and P. G. Bolhuis, Chem. Phys. **396**, 30 (2012).
[38]B. Knott, N. C. Duff, M. F. Doherty, and B. Peters, J. Chem. Phys. **131**, 224112 (2009).
[39]J. Wedekind and D. Reguera, J. Phys. Chem. B **112**, 11060 (2008).
[40]B. Peters, J. Chem. Phys. **131**, 244103 (2009).
[41]A. C. Pan and D. Chandler, J. Phys. Chem. B **108**, 19681–19686 (2004).
[42]G. T. Beckham and B. Peters, J. Phys. Chem. Lett. **2**, 1133 (2011).
[43]D. Moroni, P. R. ten Wolde, and P. G. Bolhuis, Phys. Rev. Lett. **94**, 235703 (2005).
[44]Z.-J. Wang, C. Valeriani, and D. Frenkel, J. Phys. Chem. B **113**, 3776 (2009).
[45]G. M. Torrie and J. P. Valleau, J. Comput. Phys. **23**, 187 (1977).
[46]M. Sprik and G. Ciccotti, J. Chem. Phys. **109**, 7737 (1998).
[47]A. Laio, A. Rodriguez-Fortea, F. L. Gervasio, M. Ceccarelli, and M. Parrinello, J. Phys. Chem. B **109**, 6714 (2005).
[48]Y. Sugita and Y. Okamoto, Chem. Phys. Lett. **314**, 141 (1999).
[49]T. B. Woolf and B. Roux, J. Am. Chem. Soc. **116**, 5916 (1994).
[50]W. Im and B. Roux, J. Mol. Biol. **319**, 1177 (2002).
[51]G. Hummer, New J. Phys. **7**, 34 (2005).
[52]B. J. Berne, M. Borkovec, and J. E. Straub, J. Phys. Chem. **92**, 3711 (1988).
[53]A. Ma, A. Nag, and A. R. Dinner, J. Chem. Phys. **124**, 144911 (2006).
[54]B. Pan, M. S. Ricci, and B. L. Trout, J. Phys. Chem. B **114**, 4389 (2010).
[55]S. Manas, E. E. Santiso, and B. L. Trout, J. Phys. Chem. B **115**, 10400 (2011).
[56]N. C. Duff and B. Peters, J. Chem. Phys. **131**, 184101 (2009).
[57]M. R. Walsh, C. A. Koh, E. D. Sloan, D. T. Wu, and A. K. Sum, Science **326**, 1095 (2009).
[58]P. G. Geissler, C. Dellago, and D. Chandler, J. Phys. Chem. B **103**, 3706 (1999).
[59]H. A. Kramers, Physica **7**, 284 (1940).
[60]J. S. Langer, Ann. Phys. (N.Y.) **54**, 258 (1969).
[61]R. P. Sear, J. Chem. Phys. **128**, 214513 (2008).
[62]E. Sanz, C. Valeriani, D. Frenkel, and M. Dijkstra, Phys. Rev. Lett. **99**, 055501 (2007).
[63]R. F. Grote and J. T. Hynes, J. Chem. Phys. **73**, 2715 (1980).
[64]G. Hummer and I. G. Kevrekidis, J. Chem. Phys. **118**, 10762 (2003).
[65]G. Hummer, J. Chem. Phys. **120**, 516 (2004).
[66]P. Hanggi, P. Talkner, and M. Borkovec, Rev. Mod. Phys. **62**, 251 (1990).
[67]A. C. Pan, D. Sezer, and B. Roux, J. Phys. Chem. B **112**, 3432–3440 (2008).
[68]R. Best and G. Hummer, Proc. Natl. Acad. Sci. U.S.A. **107**, 1088 (2010).
[69]S. Yang, J. N. Onuchic, A. E. Garcia, and H. Levine, J. Mol. Biol. **372**, 756 (2007).
[70]A. Singer, R. Erban, I. G. Kevrekidis, and R. R. Coifman, Proc. Natl. Acad. Sci. U.S.A. **106**, 16090 (2009).
[71]W. W. Zheng, M. A. Rohrdanz, M. Maggioni, and C. Clementi, J. Chem. Phys. **134**, 144109 (2011).
[72]M. E. Johnson and G. Hummer, J. Phys. Chem. B **116**, 8573 (2012).
[73]B. Peters, Chem. Phys. Lett. **494**, 100 (2010).
[74]S. Krivov and M. Karplus, Proc. Natl. Acad. Sci. U.S.A. **105**, 13841 (2008).
[75]S. Kullback and R. Leibler, Ann. Math. Stat. **22**, 79 (1951).
[76]J. Shlens, G. D. Field, J. L. Gauthier, M. I. Grivich, D. Petrusca, A. Sher, A. M. Litke, and E. J. Chichilnisky, J. Neurosci. **26**, 8254 (2006).
[77]J. W. Gibbs, *Elementary Principles in Statistical Mechanics* (Dover, New York, 1960).
[78]D. D. L. Minh, J. Chem. Phys. **130**, 204102 (2009).
[79]D. Ming and M. E. Wall, Phys. Rev. Lett. **95**, 198103 (2005).
[80]M. S. Shell, J. Chem. Phys. **129**, 144108 (2008).
[81]J. F. Rudzinski and W. G. Noid, J. Chem. Phys. **135**, 214101 (2011).
[82]A. B. Bortz, M. H. Kalos, and J. L. Lebowitz, J. Comput. Phys. **17**, 10 (1975).
[83]D. T. Gillespie, J. Phys. Chem. **81**, 2340 (1977).
[84]G. van der Zwan and J. T. Hynes, J. Chem. Phys. **78**, 4174 (1983).
[85]E. Pollak, J. Chem. Phys. **85**, 865 (1986).
[86]D. Kohen and D. J. Tannor, J. Chem. Phys. **103**, 6013 (1995).
[87]W. E and E. Vanden-Eijnden, J. Stat. Phys. **123**, 503 (2006).
[88]E. Vanden-Eijnden, in *Computer Simulations in Condensed Matter: From Materials to Chemical Biology*, edited by M. Ferrario, G. Ciccotti, and K. Binder (Springer-Verlag, Berlin, 2006), Vol. 2, pp. 439–478.
[89]M. Iwamatsu, J. Chem. Phys. **136**, 204702 (2012).