THE JOURNAL OF PHYSICAL CHEMISTRY Letters

pubs.acs.org/JPCL Letter

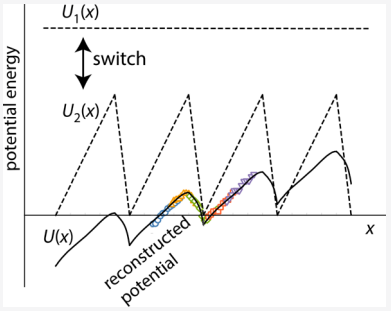
From Nonequilibrium Single-Molecule Trajectories to Underlying Dynamics

Alexander M. Berezhkovskii and Dmitrii E. Makarov*

Cite This: J. Phys. Chem. Lett. 2020, 11, 1682−1688 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Single-molecule observations of biomolecular dynamics and folding are commonly rationalized using the model of diffusive dynamics on a free-energy landscape, which is inferred via the Boltzmann inversion of the equilibrium distribution of the experimental observable. Can the same model be applied to high-resolution single-molecule trajectories of molecular machines that lack thermal equilibrium so that the Boltzmann inversion method is inapplicable? In this Letter, we discuss two approaches to reconstructing the underlying free-energy landscape in such nonequilibrium systems and explore the performance of this model in application to trajectories with complex underlying dynamics.



Observable dynamics of molecular rearrangements are usually interpreted in terms of phenomenological chemical kinetics, where discrete species interconvert according to chemical rate equations. Within this picture, for example, the kinetics of protein folding may be described using just two rate coefficients, those of folding and unfolding. Connecting this picture to a high-resolution, atomistically detailed view of the same process has been a major challenge, which is hampered by the low-dimensional nature and limited time resolution of experimental signals as well as by the inaccessibility of experimentally relevant time scales by atomistic simulations. Low-dimensional models that view molecular rearrangements as motion along collective "reaction coordinates" offer a promising middle ground between the two extremes, as they provide a coarse-grained description of simulation data and can be used to rationalize experimental chemical kinetics.

In particular, Kramers' seminal paper,[1] which considers the dynamics of barrier crossing along a selected 1D reaction coordinate x while the rest of the degrees of freedom are treated using the theory of Brownian motion, has become a paradigm for chemical dynamics in solution[2−5] and, in particular, for complex biomolecular rearrangements such as protein folding.[6,7]

When the reaction coordinate x is equated to the experimental signal, the Kramers view of chemical reaction dynamics is especially fruitful because it allows one to construct a dynamical model directly from single-molecule experimental data that report on the temporal evolution of the reaction coordinate $x(t)$. This study is motivated by such single-molecule experiments,[8−13] which have recently attained a microsecond time resolution along with a nanometer spatial resolution, allowing them to probe the barrier crossing dynamics of biomolecular folding and binding directly and thus expanding the utility of the Kramers view beyond its use as a phenomenological description. Disregarding experimental artifacts (which have been studied extensively in recent years[14−19]), the reconstruction of the underlying dynamical model from a trajectory $x(t)$ may proceed as follows. One first determines the equilibrium distribution $p_{eq}(x)$ and then uses it to recover the potential of mean force (PMF) $U(x)$ via the Boltzmann inversion

$$U(x) = -(1/\beta) \ln p_{eq}(x) \tag{1}$$

where $\beta = (k_B T)^{-1}$ is the inverse thermal energy. Following the work by Kramers and subsequent authors, the probability distribution of the reaction coordinate at time t, $p(x, t)$, is assumed to satisfy the Smoluchowski equation

$$\frac{\partial p(x, t)}{\partial t} = \frac{\partial}{\partial x} D(x) e^{-\beta U(x)} \frac{\partial}{\partial x} e^{\beta U(x)} p(x, t) \tag{2}$$

where $D(x)$ is the (possibly coordinate-dependent) diffusivity. It follows from eqs 1 and 2 that the equilibrium distribution $p_{eq}(x)$ is the stationary solution of eq 2.

The direct estimation of diffusivity $D(x)$ from experimental trajectories is a more difficult task. Although several methods have been proposed,[20−26] their experimental application so far have been limited and fraught with experimental artifacts stemming from insufficient time resolution and from coupling of the molecular dynamics to that of experimental probes.[27] As

Received: December 14, 2019
Accepted: February 4, 2020
Published: February 4, 2020

https://dx.doi.org/10.1021/acs.jpclett.9b03705
J. Phys. Chem. Lett. 2020, 11, 1682−1688

a result, the position dependence of the diffusivity is usually neglected, and the constant diffusivity D is treated as an adjustable parameter, which can be estimated by fitting, for example, the experimentally observed transition rate to that predicted by the Kramers theory.

Single-molecule studies that probe barrier crossing dynamics directly hold promise for elucidating conformational changes that eventually lead to free-energy transduction in molecular machines. For example, an experimental report on high-temporal-resolution trajectories of a molecular motor position[12] has already been published. Molecular machines, however, are nonequilibrium systems. The goal of this Letter is to explore how to extend the above approach to reconstructing the underlying dynamics from long trajectories to cyclical systems that are not in equilibrium. Can one use a high-resolution trajectory $x(t)$ of, say, a molecular motor moving along its track to build a continuous 1D model describing the underlying dynamics along x? The simplest model that captures the features of motor dynamics takes advantage of the track periodicity and assumes that the dynamics is governed by eq 2 in a tilted periodic potential

$$U(x) = U_0(x) - Fx \tag{3}$$

where $U_0(x) = U_0(x + L)$ and L is the period of the track (Figure 1). The force F mimics the thermodynamic driving

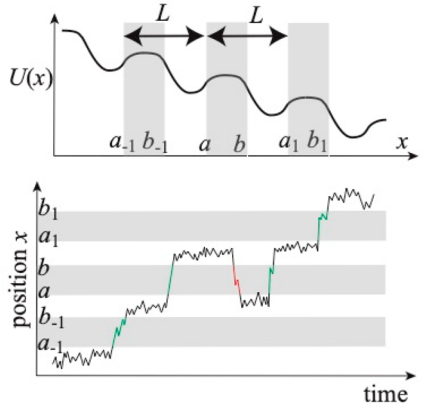


Figure 1. Tilted periodic potential model of the dynamics of a periodic walker (e.g., molecular motor). The walker undergoes an overall drift in one direction. The effective potential in this model can be reconstructed by considering dynamics within an arbitrarily chosen interval (a, b) and its periodic replicas (a_n, b_n). The green and red trajectory segments are the forward and backward transition paths traversing the intervals (a_n, b_n).

and controls the motor speed. This model has been used to describe, for example, muscle contraction,[28] and it has been the subject of numerous theoretical studies.[29] We will refer to this model as the effective potential model (EPM).

Unfortunately, such a system does not have an equilibrium distribution but rather exhibits a nonequilibrium steady state; hence eq 1 can no longer be used to recover the underlying potential $U(x)$. Several studies[30−32] explored the reconstruction of $U(x)$ from steady-state probability distributions; these methods, however, require advance knowledge of the diffusivity D. Its value can be estimated independently[32] in experimental setups where the motor dynamics is monitored via the position of a large experimental probe (for example, a bead), and the observed diffusivity is dominated by the Stokesian friction on the probe. Our interest here, however, is in experiments where the probe is small,[11,12] and so the value of D reflects the intrinsic molecular dynamics and must be inferred from the trajectory $x(t)$ itself; as previously noted, this is a difficult task. Ideally, one wishes for a method that infers the effective potential without any knowledge about the microscopic diffusivity, as can be done in the equilibrium case (eq 1). A method of this type was proposed in ref 22, but, in contrast with the present study focusing on the nonequilibrium steady-state scenario, it was concerned with relaxation toward equilibrium. The latter scenario is exemplified by the relaxation of a stretched polyprotein chain, a system for which the effective free-energy landscape has been reconstructed from experimental trajectories.[22,33] Note that polyproteins relax on time scales on the order of seconds over length scales of tens to hundreds of nanometers,[22,33] a regime distinct from the one considered here, where the relevant time scales are microseconds to milliseconds, whereas the relevant length scales are on the order of nanometers.

In what follows, we will discuss how the effective tilted periodic potential can be recovered from the observed trajectories. We will further explore how well the EPM describes the dynamics of a molecular motor that *is not* governed by eqs 2 and 3 and will discuss the applicability of the EPM to experimental data in such cases.

Estimating the Effective Potential from Local Drift Velocities and from Splitting Probabilities. We start with analyzing dynamics within some interval (a, b), as illustrated in Figure 1. The interval may or may not coincide with the periodically repeated unit cell of the system, $(0, L)$. Given that the trajectory of our periodic system is spatially unbounded, it will likely not visit this interval more than a few times, but, given the periodicity and the absence of memory in diffusive trajectories, we can consider the dynamics within all of the periodic replicas of this interval, as illustrated in Figure 1. Equivalently, we can analyze the dynamics of the modified trajectory obtained by mapping $x(t)$ onto a finite interval containing N unit cells (most commonly, $N = 1$)

$$\tilde{x}(t) = x(t) \bmod NL \tag{4}$$

where N is an integer. This modified trajectory is confined within the interval $(0, NL)$, and N, a, and b can be chosen such that we have $(a, b) \in (0, NL)$.

We now describe two ways to recover the potential $U(x)$ from an observed nonequilibrium trajectory $x(t)$:

(i) Consider discretized diffusive trajectories in the potential $U(x)$, $x_n = x(n\Delta t)$, obtained via the numerical integration of the corresponding Langevin equation or recorded experimentally. Here Δt is a time step. At sufficiently short time steps, the time evolution of the system is governed by an approximate Green's function of the form

$$G(x_{n+1}, t + \Delta t|x_n, t) \approx \frac{1}{\sqrt{4\pi D\Delta t}} \exp\left\{-\frac{[x_{n+1} - x_n - u(x_n)t]^2}{4D\Delta t}\right\} \tag{5}$$

Here $G(x_{n+1}, t + \Delta t|x_n, t)$ is the conditional probability of finding the system at point x_{n+1} at time $t + \Delta t$ given that it was located at x_n at time t, and

$$u(x_n) = -\beta U'(x_n)D \tag{6}$$

is the local drift velocity at point x_n. For simplicity, we have assumed that the diffusivity is coordinate-independent.

https://dx.doi.org/10.1021/acs.jpclett.9b03705
J. Phys. Chem. Lett. 2020, 11, 1682−1688

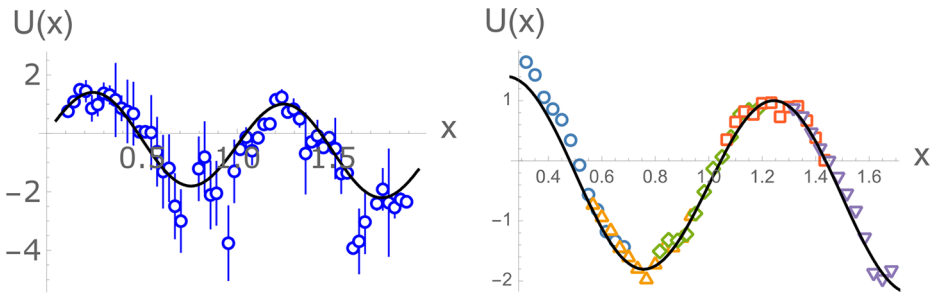


Figure 2. Tilted periodic potential reconstructed using eq 11 from a Langevin trajectory using (a) one large window and (b) five smaller overlapping windows. Open symbols show the numerically estimated right-hand side of eq 11, with different colors corresponding to different windows. Estimated statistical errors in this case are comparable to the symbol size. The solid black line is the true potential.

Equation 5 is simply the solution of eq 2 in the case of a linear potential (that is, the potential $U(x)$ approximated by its Taylor expansion, to first order, in the vicinity of x_n); for a short enough time step Δt, the typical displacement $x_{n+1} - x_n$ is small enough that the variation of the force acting on the system can be neglected, making this approximation increasingly accurate with a decreasing time step. From eq 5, one readily finds the first and the second moments of the particle's displacement during a single time step

$$\langle x_{n+1} - x_n \rangle / \Delta t \approx u(x_n) \tag{7}$$

$$\langle (x_{n+1} - x_n)^2 \rangle \approx 2D\Delta t + [u(x_n)\Delta t]^2 \tag{8}$$

Now, for all points of the modified trajectory that are found in the vicinity of some given point x, $x_n \approx x$ (in practice one can choose these points according to the criterion $|x_n - x| < \varepsilon$, where ε is a small parameter), one can compute the average of the displacement moments and use eqs 6−8 to obtain the values of the diffusivity D and of the deterministic force $F(x) = -U'(x) \approx -U'(x_n) = u(x_n)/(\beta D)$. The diffusivity D obtained in this manner is local near the point x, and so this approach can be extended to the case where the diffusivity is coordinate-dependent. Finally, the potential $U(x)$ is obtained via the numerical integration of the force

$$U(x) = U(0) - \int_0^x \mathrm{d}x' F(x') = U(0) - \frac{1}{\beta D}\int_0^x \mathrm{d}x' u(x') \tag{9}$$

We note that a similar approach has been developed, in the context of equilibrium systems, in refs 34 and 35. Other approaches to finding $U(x)$ and their relationship to the methods described here are discussed in the Supporting Information. To provide an accurate estimate of the potential and the diffusivity, the time step Δt must be small enough that the short-time approximation, eq 5, is valid. In practice, this severely limits the applicability of this approach to experimental observations, whose time resolution is limited and is usually in the microsecond range, and even to trajectories obtained via atomistic simulations, where the data sampling rate is usually constrained by data storage limitations.

(ii) The second approach considers the splitting probability $\Phi(x \to b|a)$ defined as the probability that the system that starts in x, $x \in (a, b)$, will reach the interval boundary b before it reaches a. This probability is described by the well-known expression[36]

$$\Phi(x \to b|a) = 1 - \Phi(x \to a|b) = \frac{\int_a^x D^{-1}(y)\mathrm{e}^{\beta U(y)}\,\mathrm{d}y}{\int_a^b D^{-1}(y)\mathrm{e}^{\beta U(y)}\,\mathrm{d}y} \tag{10}$$

If we assume that the diffusion coefficient is coordinate-independent, then the potential $U(x)$ is readily recovered for any point x, $x \in (a, b)$, from eq 10

$$U(x) = \beta^{-1} \ln[\mathrm{d}\Phi(x \to b|a)/\mathrm{d}x] + c_{ab} \tag{11}$$

where c_{ab} is independent of x. Equation 11 has already been used to reconstruct PMFs in equilibrium systems,[37,38] but it also holds for nonequilibrium systems such as the one described by the tilted periodic potential shown in Figure 1. Moreover, given the assumed Markovian character of the trajectory $x(t)$, the splitting probability is completely determined by the dynamics within the interval (a, b) and is independent of the dynamics (or the shape of the potential) outside this interval. Notably, eq 11 does not depend on the diffusivity (except via an unimportant additive constant). This is an important and useful feature of the method that it shares with the Boltzmann inversion (eq 1) because the accurate estimation of the diffusivity directly from trajectories is challenging.[27]

A special case of eq 10 allows one to reconstruct the tilting force F in a different way: Consider the splitting probabilities $\Phi_+(x) = \Phi(x \to x + L|x - L)$ and $\Phi_-(x) = \Phi(x \to x - L|x + L)$, which are the transition probabilities from x to $x \pm L$ in the next periodic cell to the right and to the left.[39] Using eqs 3 and 10, one finds[39]

$$\Phi_+(x)/\Phi_-(x) = e^{\beta F L} \tag{12}$$

and thus the tilting force can be estimated by simply counting the numbers of forward (to the right) and backward (to the left) steps of length L. We note that this formula is correct as long as L is the period (or multiple of periods) of the system regardless of how complex the shape of the potential $U(x)$ might be within the period.

Numerical Aspects of Estimating the Potentials from Splitting Probabilities. In view of its experimental relevance, the method of potential reconstruction from splitting probabilities deserves further discussion. By choosing the interval length to coincide with one period, $b - a = L$, one can use eq 11 to reconstruct

https://dx.doi.org/10.1021/acs.jpclett.9b03705
J. Phys. Chem. Lett. 2020, 11, 1682−1688

the potential within an entire periodic cell, and thus the full tilted periodic potential can be recovered. In practice, however, accurate estimation of the splitting probabilities in the vicinity of the interval boundaries is difficult and limited by the time resolution of the trajectory $x(t)$. Indeed, the accurate estimation of, say, $\Phi(b - \delta \to b|a)$, where δ is small, requires that the details of the trajectory traveling between $b-\delta$ and b are resolved. To avoid errors in using eq 11 near the boundaries, it is more practical to choose the interval (a, b) to be longer than the period; for example, one can choose the boundaries such that $a < 0 < L < b$ and only use the splitting probabilities within its subinterval $(0, L)$ to reconstruct the potential.

The longer the interval (a, b), however, the more demanding the requirements for the length of the trajectory used. Indeed, if the interval length is such that $\beta F(b - a) \gg 1$, then very few trajectories starting from x and reaching the left boundary (a) will be observed (unless x is close to a), resulting in statistical errors in estimating the splitting probabilities that will be amplified when taking their numerical derivatives as in eq 11. This issue is thus particularly important to consider in cases of strongly biased dynamics, as, for example, for a molecular motor that almost never makes a step back.

A better strategy, which works even for strongly biased dynamics, is to use smaller overlapping intervals to reconstruct the local shape of the potential within each interval and to recover the global shape by matching the reconstructed potentials in overlapping intervals; in other words, the constants c_{ab} in eq 11 are chosen such that the reconstructed potential is a continuous function.

Figure 2 shows a numerical example of reconstructing the potential $\beta U(x) = 1.5 \sin(2\pi x/L) - 0.4x/L$ from an overdamped Langevin trajectory using eq 11. In what follows, we report distances in units of the period L; times in units of L^2/D, where D is the diffusivity; and energies in units of the thermal energy $\beta^{-1} = k_B T$. The Langevin equation was integrated using the Euler method, with a time step of $\Delta t = 0.000025$ over a total time of $\tau = 500$. In Figure 2a, we show the result obtained using a single large window, $(a, b) = (0.05, 1.95)$, that is greater than the period. The splitting probability was evaluated on a grid containing 61 points, and the derivatives of the splitting probability were evaluated using the finite difference method. In Figure 2b, we show the result obtained from the same trajectory using five overlapping windows, (0.25, 0.75), (0.5, 1.0), (0.75, 1.25), (1.0, 1.5), and (1.25, 1.75), with 15 grid points per window. Clearly the method employing shorter overlapping windows is in this example superior for the same trajectory length. We note that the potential $U(x)$ is also readily recovered from the same trajectory using eqs 6–9 (data not shown).

What Happens When the Two Approaches Are Applied to "Wrong Dynamics"? Equations 6–9 or eq 11 can be applied to any trajectory $x(t)$ generated by periodic dynamics, resulting in two estimates for an effective potential $U(x)$, but it is not clear that the two methods would predict the same potential. Moreover, what is the relation between the dynamics in this potential and the true dynamics (and how would we know if there is a difference)?

Here we explore the ability of the effective potential model to describe the dynamics that was generated by a different model, specifically, by the two-landscape model of a molecular motor. In the latter model, which has been widely used in the literature,[40] binding of, say, an ATP molecule to the motor is modeled as a discrete switch to a different potential landscape. As a result, the system undergoes diffusion in a potential that switches, stochastically, between two periodic functions, $U_1(x)$ and $U_2(x)$.

Our choice for these functions is the well-studied case[41,42] in which $U_1(x) = 0$ and $U_2(x)$ is a periodic piecewise linear function of period L

$$U_2(x) = \begin{cases} U_0 x/l, & 0 \le x < l \\ U_0(L - x)/(L - l), & l \le x < L \end{cases} \tag{13}$$

where the parameter l is the location of the barrier within the interval $(0, L)$. Diffusive dynamics in each individual potential, $U_1(x)$ and $U_2(x)$, does not exhibit unidirectional motion. However, when the system stochastically switches between these two potentials according to the first-order kinetic scheme

$$U_1 \underset{k}{\overset{k}{\rightleftarrows}} U_2 \tag{14}$$

where k is a switching rate, unidirectional transport ensues, provided that the potential $U_2(x)$ is asymmetric ($l \neq L/2$).

We have applied the above two methods to a trajectory $x(t)$ generated by the two-landscape model ($l = 0.8$, $k = 100$, $U_0 = 10 k_B T$), assuming the same diffusivity, D, in both potentials, to reconstruct an effective potential $U(x)$. Figure 3 shows the

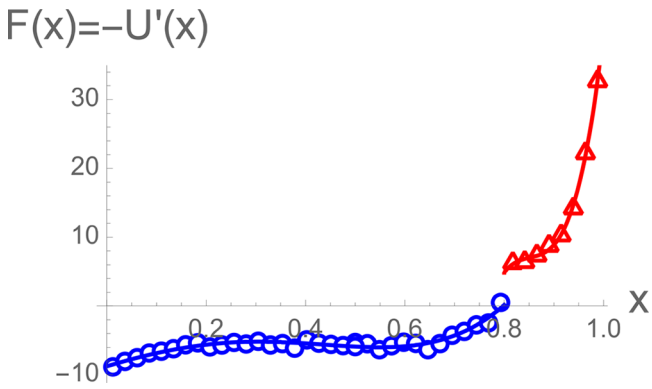


Figure 3. Numerical estimation of the effective deterministic force $F(x) = -U'(x)$ by applying eqs 6 and 7 to a trajectory generated by the two-landscape model. Estimated statistical errors are smaller than the symbols. Solid lines show polynomial fits to the numerical data used to integrate the force and obtain the effective potential through eq 9.

effective force $F(x) = -U'(x)$ reconstructed numerically from a trajectory using eqs 6–8. Similarly to the force generated by the potential $U_2(x)$, this effective force exhibits discontinuities at $x = 0$, L and at $x = l$. The potential obtained by integrating the force (eq 9) is shown as a solid black line in Figure 4, and the potential obtained from splitting probabilities (eq 11) using five windows, as in the example previously discussed, is shown with open markers, with different colors corresponding to different windows.

Notably, the estimates of the potential $U(x)$ obtained by the two different methods are identical within numerical error. We have, then, simulated overdamped Langevin dynamics using the polynomial fit for the force obtained from the first method (Figure 3) and the same value of the diffusivity D. Remarkably, the dynamics in the reconstructed potential $U(x)$ closely resembles the original dynamics in the two-landscape model. Specifically, a particularly useful quantity that characterizes the

https://dx.doi.org/10.1021/acs.jpclett.9b03705
J. Phys. Chem. Lett. 2020, 11, 1682–1688

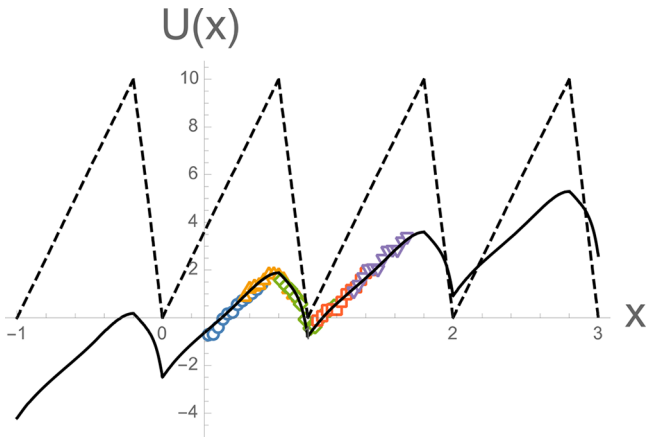


Figure 4. In the two-landscape model, the potential of the system switches between the periodic potential of eq 13 ($U_0 = 10k_BT$) shown as a dashed line and zero, with a switching rate of $k = 100$, resulting in an overall drift to the left. The dynamics of the two-landscape model was mapped onto diffusion in a single tilted periodic potential using the two methods described above. The solid line shows a numerical fit of the potential reconstructed using eqs 6−9. (Its derivative is shown in Figure 3.) Colored symbols show the numerical reconstruction of the effective potential from splitting probabilities using eq 11 with five overlapping windows. Estimated statistical errors are smaller than the symbols.

dynamics of a motor is the mean number of steps, per unit time, that it makes in the forward and backward directions. A "step" is defined as follows: We map the trajectory $x(t)$ onto a nearest-neighbor continuous time random walk[39] by considering a succession of equally spaced milestones placed at $x_n = nL$. The state of the trajectory is described by the discrete index $n = n(t)$, which changes only when a new milestone is crossed that is different from the one previously crossed. Note that this, in a sense, is equivalent to the milestoning procedure.[43] A forward (backward) step then corresponds to a transition $n \rightarrow n \pm 1$. For both the original trajectory generated by simulating the two-landscape model and for the reconstructed single-landscape model, we have measured the step frequencies $\nu_\pm$ (i.e., the number of forward and backward steps per unit time). For the original trajectory, we find $\nu_+ = 0.23$, $\nu_- = 1.14$, close to the values $\nu_+ = 0.18$, $\nu_- = 1.0$ estimated for the dynamics in the reconstructed potential.

Correspondingly, the drift velocity is $u = \nu_+ - \nu_- = -0.91$ for the original trajectory and $u = -0.82$ for the reconstructed dynamics. Note that in the EPM, transport is accomplished via thermally activated hopping, and so the result is exponentially sensitive to the barrier height; the ∼10% difference between the reconstructed frequencies and the original ones is comparable to the numerical reconstruction errors.

The finding that the dynamics in the reconstructed potential is close to that in the original model is surprising given that the underlying physical mechanisms causing unidirectional transport in the two models are quite different: In the reconstructed potential it is caused by thermally activated transitions between adjacent potential minima, whereas in the original two-landscape model, the large barrier height of the potential $U_2(x)$ ($U_0 = 10k_BT$) makes thermally activated transitions highly improbable, and the transitions between adjacent cells predominantly happen thanks to free diffusion in the barrierless state with the potential $U_1(x) = 0$. The observation that an effective 1D model reconstructed from a trajectory can predict the properties of the original system accurately despite the differences between physical mechanisms causing unidirectional motion in each case raises two questions. First, why do two physically distinct models predict similar dynamics along x? Second, can a more careful analysis of an experimental trajectory $x(t)$ differentiate between different models or, at least, warn one of the potential inadequacy of the effective potential model?

Although we do not fully understand the answer to the first question, the following considerations suggest the plausibility of the EPM (when applied to dynamics generated by a two-landscape model) if the drift velocity of the motion is the central quantity of interest. For the purpose of computing the drift velocity, it is sufficient to consider the continuous-time random walk $n = n(t)$, as described above. The properties of the random walk in the case where $x(t)$ is Markovian are completely determined by the distribution $\rho(\tau)$ of the lifetime τ of a discrete state (i.e., the time between two consecutive steps) and by the splitting probabilities $\Phi_+(0)$ and $\Phi_-(0) = 1 - \Phi_+(0)$ (cf. eq 12) of making steps to the right and to the left (i.e., from a state n to states $n \pm 1$).[39] Moreover, the drift velocity is given by a simple formula[39] that depends on the distribution $\rho(\tau)$ only through the average lifetime of a state, $\langle\tau\rangle = \int_0^\infty d\tau\, \tau\rho(\tau) = (\nu_+ + \nu_-)^{-1}$

$$u = [\Phi_+(0) - \Phi_-(0)]L/\langle\tau\rangle \tag{15}$$

Because the effective potential obtained from eq 11 is defined so as to reproduce the splitting probabilities exactly, the effective potential model will yield the exact drift velocity if its prediction for the mean lifetime $\langle\tau\rangle$ is the same as that for the original dynamics.

We are unaware of any rigorous argument suggesting that the mean lifetime in the two models must be the same if their splitting probabilities are the same. We note, however, that, in the case of diffusive dynamics with a constant diffusivity, D, the position dependence of the splitting probability $\Phi_+(x) = \Phi(x \rightarrow x + L|x - L) = 1 - \Phi_-(x)$ and the value of the diffusivity completely determine the mean lifetime $\langle\tau\rangle$, which is given by

$$\langle\tau\rangle = \frac{1}{D}\left[\Phi_-(0)\int_{-L}^{0}\Phi_+(x)\frac{dx}{d\Phi_+(x)/dx} + \Phi_+(0)\int_0^L \Phi_-(x)\frac{dx}{d\Phi_-(x)/dx}\right] \tag{16}$$

See the Supporting Information for a derivation of eq 16. It appears plausible, then, that the same coordinate dependence of the splitting probability results in a mean lifetime that is close to that in the original system.

We now turn to the second question: How can one tell whether diffusive dynamics in the reconstructed potential is a good description of the original trajectory? One possibility is to consider the system's response to an external force:[31] Even if the EPM predicts the dynamics of the unperturbed system accurately, it may fail to reproduce, for example, the correct dependence of the drift velocity $u(f)$ on the applied force f.[31] If, however, only unperturbed dynamics is accessible experimentally, then it may still exhibit signatures of a more complex underlying process. One such signature is concerned with the time reversibility of transition paths, that is, barrier crossing events that have already been studied experimentally under equilibrium conditions.[8,10] A transition path is defined as a trajectory segment that traverses an interval (a_n, b_n), entering

https://dx.doi.org/10.1021/acs.jpclett.9b03705
J. Phys. Chem. Lett. 2020, 11, 1682−1688

it through one interval boundary and exiting it through its other boundary, without exiting the interval in between these two events. For example, an a-to-b transition path (Figure 1, green trajectory segment) enters an interval (a_n, b_n) by crossing $x(t) = a_n$ and exits it by crossing $x(t + t_{TP}) = b_n$, where t_{TP} is the transition path time.

The transition path times for the 1D diffusion in an effective potential satisfy a fundamental symmetry: The distributions of the a-to-b and b-to-a transition path times are identical.[44] In contrast, this forward/backward symmetry is, in general, broken when the dynamics is non-Markovian.[44] In particular, such symmetry breaking has been experimentally demonstrated[45] in driven systems whose potential landscape stochastically switches between two functions (as a result of, e.g., the application of a time-dependent external force). Forward−backward symmetry breaking may thus be an indication that the effective potential model is inadequate.

Applicability of the Effective Potential Model to Experimental Trajectories. Mechanistic details of how unidirectional motion is generated by molecular machines are often unclear or even controversial,[46] and single-molecule measurements offer a key window on mechanisms of transport in such periodic nonequilibrium systems (see, e.g., refs 40 and 47). In particular, single-molecule studies with high spatial and temporal resolution, as they become increasingly available, will likely provide significant insights into the nonequilibrium dynamics of machines beyond discrete kinetic models. EPM offers a simple picture of nonequilibrium dynamics as manifested by experimental reaction coordinates.

Of the above two approaches to the reconstruction of the effective potential from observed trajectories $x(t)$, we expect the second one, which relies on the measurement of splitting probabilities, to be more promising in application to experiments. Indeed, unlike the first approach, it does not require high temporal resolution of the data. Moreover, it has already been successfully applied to equilibrium biomolecular folding data.[37,38]

In addition to the effective potential, the complete dynamical description in terms of the EPM requires an estimate of the diffusivity D. Given sufficient time resolution, this can be accomplished by examining the short-time behavior of the mean-squared displacement along the trajectory and using eq 8. Note that for the two-landscape system considered above, the dynamics is diffusive in both potentials (U_1 and U_2), with the same diffusivity D; as a result, eq 8 holds for this system at short enough time lags, Δt, such that the second term can be neglected, and we have $\langle (x_{n+1} - x_n)^2 \rangle \approx 2D\Delta t$. Therefore, choosing the same value of the diffusivity for the dynamics in the effective potential was justified in this case.

In experimental applications, however, the time resolution/data sampling rate may be insufficient for using eq 8, and so the "first-principle" estimation of the diffusion coefficient may be a challenge similar to that encountered in single-molecule studies of equilibrium systems.[27]

Finally, we note that although eq 11 can be applied to obtain an effective potential $U(x)$ for any dynamics (Markovian or non-Markovian), the resulting potential may depend on the choice of the interval (a, b), and such a dependence would be another indication that the 1D diffusion model is an inadequate description of the true dynamics.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpclett.9b03705.

Proof of eq 16 and comparison with other potential reconstruction methods (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Dmitrii E. Makarov − *Department of Chemistry and Oden Institute for Computational Engineering and Sciences, University of Texas at Austin, Austin, Texas 78712, United States;* orcid.org/0000-0002-8421-1846; Email: makarov@cm.utexas.edu

Author

Alexander M. Berezhkovskii − *Mathematical and Statistical Computing Laboratory, Office of Intramural Research, Center for Information Technology, National Institutes of Health, Bethesda, Maryland 20892, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpclett.9b03705

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are grateful to Keir Neuman and Jing Xu for helpful discussions. A.M.B. was supported by the Intramural Research Program of the NIH, Center for Information Technology. D.E.M. was supported by the Robert A. Welch Foundation (grant no. F-1514) and the National Science Foundation (grant no. CHE 1566001).

■ REFERENCES

(1) Kramers, H. A. Brownian Motion in a Field of Force and the Diffusion Model of Chemical Reactions. *Physica* **1940**, *7*, 284−304.

(2) Pollak, E.; Talkner, P. Reaction rate theory: what it was, where is it today, and where is it going? *Chaos* **2005**, *15*, 026116.

(3) Szabo, A.; Schulten, K.; Schulten, Z. First passage time approach to diffusion controlled reactions. *J. Chem. Phys.* **1980**, *72*, 4350.

(4) Zwanzig, R. *Nonequilibrium Statistical Mechanics*; Oxford University Press: 2001.

(5) Hanggi, P.; Talkner, P.; Borkovec, M. 50 years after Kramers. *Rev. Mod. Phys.* **1990**, *62*, 251.

(6) Socci, N. D.; Onuchic, J. N.; Wolynes, P. G. Diffusive dynamics of the reaction coordinate for protein folding funnels. *J. Chem. Phys.* **1996**, *104*, 5860−5868.

(7) Klimov, D.; Thirumalai, D. Viscosity Dependence of the Folding Rates of Proteins. *Phys. Rev. Lett.* **1997**, *79*, 317.

(8) Hoffer, N. Q.; Woodside, M. T. Probing microscopic conformational dynamics in folding reactions by measuring transition paths. *Curr. Opin. Chem. Biol.* **2019**, *53*, 68−74.

(9) Edwards, D. T.; Faulk, J. K.; Sanders, A. W.; Bull, M. S.; Walder, R.; LeBlanc, M. A.; Sousa, M. C.; Perkins, T. T. Optimizing 1-mus-Resolution Single-Molecule Force Spectroscopy on a Commercial Atomic Force Microscope. *Nano Lett.* **2015**, *15*, 7091−8.

(10) Chung, H. S.; Eaton, W. A. Protein folding transition path times from single molecule FRET. *Curr. Opin. Struct. Biol.* **2018**, *48*, 30−39.

(11) Ye, W.; Gotz, M.; Celiksoy, S.; Tuting, L.; Ratzke, C.; Prasad, J.; Ricken, J.; Wegner, S. V.; Ahijado-Guzman, R.; Hugel, T.; Sonnichsen, C. Conformational Dynamics of a Single Protein Monitored for 24 h at Video Rate. *Nano Lett.* **2018**, *18*, 6633−6637.

(12) Andrecka, J.; Takagi, Y.; Mickolajczyk, K. J.; Lippert, L. G.; Sellers, J. R.; Hancock, W. O.; Goldman, Y. E.; Kukura, P. Interferometric Scattering Microscopy for the Study of Molecular Motors. *Methods Enzymol.* **2016**, *581*, 517−539.
(13) Yu, H.; Siewny, M. G.; Edwards, D. T.; Sanders, A. W.; Perkins, T. T. Hidden dynamics in the unfolding of individual bacteriorhodopsin proteins. *Science* **2017**, *355*, 945−950.
(14) Neupane, K.; Woodside, M. T. Quantifying Instrumental Artifacts in Folding Kinetics Measured by Single-Molecule Force Spectroscopy. *Biophys. J.* **2016**, *111*, 283−6.
(15) Woodside, M. T.; Lambert, J.; Beach, K. S. Determining intrachain diffusion coefficients for biopolymer dynamics from single-molecule force spectroscopy measurements. *Biophys. J.* **2014**, *107*, 1647−53.
(16) Makarov, D. E. Communication: Does force spectroscopy of biomolecules probe their intrinsic dynamic properties? *J. Chem. Phys.* **2014**, *141*, 241103.
(17) Nam, G. M.; Makarov, D. E. Extracting intrinsic dynamic parameters of biomolecular folding from single-molecule force spectroscopy experiments. *Protein Sci.* **2016**, *25*, 123−134.
(18) Cossio, P.; Hummer, G.; Szabo, A. On artifacts in single-molecule force spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 14248−53.
(19) Cossio, P.; Hummer, G.; Szabo, A. Transition paths in single-molecule force spectroscopy. *J. Chem. Phys.* **2018**, *148*, 123309.
(20) Mugnai, M. L.; Elber, R. Extracting the diffusion tensor from molecular dynamics simulation with Milestoning. *J. Chem. Phys.* **2015**, *142*, 014105.
(21) Lannon, H.; Haghpanah, J. S.; Montclare, J. K.; Vanden-Eijnden, E.; Brujic, J. Force-clamp experiments reveal the free-energy profile and diffusion coefficient of the collapse of protein molecules. *Phys. Rev. Lett.* **2013**, *110*, 128301.
(22) Zhang, Q.; Brujic, J.; Vanden-Eijnden, E. Reconstructing Free Energy Profiles from Nonequilibrium Relaxation Trajectories. *J. Stat. Phys.* **2011**, *144*, 344−366.
(23) Hinczewski, M.; von Hansen, Y.; Dzubiella, J.; Netz, R. R. How the diffusivity profile reduces the arbitrariness of protein folding free energies. *J. Chem. Phys.* **2010**, *132*, 245103.
(24) Berezhkovskii, A. M.; Makarov, D. E. Communication: Coordinate-dependent diffusivity from single molecule trajectories. *J. Chem. Phys.* **2017**, *147*, 201102.
(25) Best, R. B.; Hummer, G. Coordinate-dependent diffusion in protein folding. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 1088−93.
(26) Best, R. B.; Hummer, G. Diffusion models of protein folding. *Phys. Chem. Chem. Phys.* **2011**, *13*, 16902−11.
(27) Foster, D. A. N.; Petrosyan, R.; Pyo, A. G. T.; Hoffmann, A.; Wang, F.; Woodside, M. T. Probing Position-Dependent Diffusion in Folding Reactions Using Single-Molecule Force Spectroscopy. *Biophys. J.* **2018**, *114*, 1657−1666.
(28) Hill, T. L. *Free Energy Transduction and Biochemical Cycle Kinetics*; Springer-Verlag: New York, 1989.
(29) Reimann, P. Brownian motors: noisy transport far from equilibrium. *Phys. Rep.* **2002**, *361*, 57−265.
(30) Wang, H. Several Issues in Modeling Molecular Motors. *J. Comput. Theor. Nanosci.* **2008**, *5*, 2311−2345.
(31) Lopez-Alamilla, N. J.; Jack, M. W.; Challis, K. J. Reconstructing free-energy landscapes for cyclic molecular motors using full multidimensional or partial one-dimensional dynamic information. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.* **2019**, *100*, 012404.
(32) Lopez-Alamilla, N. J.; Jack, M. W.; Challis, K. J. Analysing single-molecule trajectories to reconstruct free-energy landscapes of cyclic motor proteins. *J. Theor. Biol.* **2019**, *462*, 321−328.
(33) Berkovich, R.; Fernandez, V. I.; Stirnemann, G.; Valle-Orero, J.; Fernandez, J. M. Segmentation and the Entropic Elasticity of Modular Proteins. *J. Phys. Chem. Lett.* **2018**, *9*, 4707−4713.
(34) de Oliveira, R. J. Stochastic diffusion framework determines the free-energy landscape and rate from single-molecule trajectory. *J. Chem. Phys.* **2018**, *149*, 234107.
(35) Freitas, F. C.; Lima, A. N.; Contessoto, V. G.; Whitford, P. C.; Oliveira, R. J. Drift-diffusion (DrDiff) framework determines kinetics and thermodynamics of two-state folding trajectory and tunes diffusion models. *J. Chem. Phys.* **2019**, *151*, 114106.
(36) Berezhkovskii, A. M.; Pustovoit, M. A.; Bezrukov, S. M. Channel-facilitated membrane transport: Transit probability and interaction with the channel. *J. Chem. Phys.* **2002**, *116*, 9952−9956.
(37) Manuel, A. P.; Lambert, J.; Woodside, M. T. Reconstructing folding energy landscapes from splitting probability analysis of single-molecule trajectories. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 7183−8.
(38) Covino, R.; Woodside, M. T.; Hummer, G.; Szabo, A.; Cossio, P. Molecular free energy profiles from force spectroscopy experiments by inversion of observed committors. *J. Chem. Phys.* **2019**, *151*, 154115.
(39) Dagdug, L.; Berezhkovskii, A. M. Drift and diffusion in periodic potentials: upstream and downstream step times are distributed identically. *J. Chem. Phys.* **2009**, *131*, 056101.
(40) Kolomeisky, A. B. *Motor Proteins and Molecular Motors*; CRC Press: 2015.
(41) Astumian, R. D.; Bier, M. Fluctuation Driven Ratchets: Molecular Motors. *Phys. Rev. Lett.* **1994**, *72*, 1766−1769.
(42) Julicher, F.; Ajdari, A.; Prost, J. Modeling molecular motors. *Rev. Mod. Phys.* **1997**, *69*, 1269−1282.
(43) Elber, R. Perspective: Computer simulations of long time dynamics. *J. Chem. Phys.* **2016**, *144*, 060901.
(44) Berezhkovskii, A. M.; Makarov, D. E. On the forward/backward symmetry of transition path time distributions in nonequilibrium systems. *J. Chem. Phys.* **2019**, *151*, 065102.
(45) Gladrow, J.; Ribezzi-Crivellari, M.; Ritort, F.; Keyser, U. F. Experimental evidence of symmetry breaking of transition-path times. *Nat. Commun.* **2019**, *10*, 55.
(46) Astumian, R. D. Microscopic reversibility as the organizing principle of molecular machines. *Nat. Nanotechnol.* **2012**, *7*, 684−8.
(47) Mugnai, M. L.; Hyeon, C.; Hinczewski, M.; Thirumalai, D. Theoretical Perspectives on Biological Machines. *Rev. Mod. Phys.***2020**, accepted for publication.

https://dx.doi.org/10.1021/acs.jpclett.9b03705
J. Phys. Chem. Lett. 2020, 11, 1682−1688