On distributions of barrier crossing times as observed in single-molecule studies of biomolecules

Alexander M. Berezhkovskii[1] and Dmitrii E. Makarov[2,*]

[1]Mathematical and Statistical Computing Laboratory, Office of Intramural Research, Center for Information Technology, National Institutes of Health, Bethesda, Maryland and [2]Department of Chemistry and Biochemistry and Oden Institute for Computational Engineering and Sciences, University of Texas at Austin, Austin, Texas

ABSTRACT   Single-molecule experiments that monitor time evolution of molecular observables in real time have expanded beyond measuring transition rates toward measuring distributions of times of various molecular events. Of particular interest is the first-passage time for making a transition from one molecular configuration (a) to another (b) and conditional first-passage times such as the transition path time, which is the first-passage time from a to b conditional upon not leaving the transition region intervening between a and b. Another experimentally accessible (but not yet studied experimentally) observable is the conditional exit time, i.e., the time to leave the transition region through a specified boundary. The distributions of such times contain a wealth of mechanistic information about the transitions in question. Here, we use the first and the second (and, if desired, higher) moments of these distributions to characterize their relative width for the model in which the experimental observable undergoes Brownian motion in a potential of mean force. We show that although the distributions of transition path times are always narrower than exponential (in that the ratio of the standard deviation to the distribution's mean is always less than 1), distributions of first-passage times and of conditional exit times can be either narrow or broad, in some cases displaying long power-law tails. The conditional exit time studied here provides a generalization of the transition path time that also allows one to characterize the temporal scales of failed barrier crossing attempts.

WHY IT MATTERS   Single-molecule measurements directly visualize dynamics of proteins, DNA, and molecular motors as they cross activation barriers and carry out their biological functions, but, in contrast to the complexity of the molecular motion involving many atoms, single-molecule signals are inherently low dimensional, reporting only on a few degrees of freedom. Toward solving the inverse problem of learning about the underlying dynamics from single-molecule signals, this work discusses distributions of barrier crossing timescales expected for the important model in which the time evolution of the single-molecule signal can be described as Brownian dynamics.

INTRODUCTION

Learning about dynamics of molecules from experimental data is a difficult inverse problem. Molecular trajectories occur in a $3N$-dimensional space, where N is the number of participating atoms (including those of the solvent in the case of biomolecular processes). Experimental measurements, in contrast, typically provide a low-dimensional signal. In single-molecule studies, this signal is the time dependence $x(t)$ of a single variable x (or, in some of the state-of-the-art studies, two variables), such as the distance between two mo-

lecular groups, or it is a list of photon arrival times. For several decades, the common approach to this inverse problem has been to postulate a specific phenomenological model—that the trajectory $x(t)$ is described as one-dimensional diffusion in a potential of mean force—and to deduce the parameters of this model (such as the diffusivity D) by fitting experimental observations. But recent improvements in the spatial and temporal resolution of single-molecule experiments offer an opportunity to both refine such models (e.g., by deducing more accurate potentials or by allowing for coordinate-dependent diffusivity) and to move beyond phenomenological models toward more accurate, data-driven ones (1,2). Importantly, recent theoretical advances (to be discussed below) show that certain details about the underlying multidimensional dynamics are encoded in the low-dimensional signal

$x(t)$, thereby suggesting that progress toward solving the above-mentioned inverse problem is possible.

A phenomenon of particular significance for the field of biomolecular dynamics is that of large conformational transitions attained via the crossing of a free energy barrier (Fig. 1). More precisely, suppose that the experimentally determined potential of mean force $U(x)$ (defined through the requirement that the equilibrium distribution $p_{eq}(x)$ of the coordinate x is the Boltzmann distribution in this potential, $p_{eq}(x) \propto \exp\left(\left\{ -\frac{U(x)}{k_B T} \right\}\right)$ has a double-well shape with two wells separated by a barrier, each well corresponding to a distinct molecular species. The experimental observable then provides a natural (albeit not necessarily optimal) choice of the reaction coordinate for the inter-well transitions, or "chemical reactions." Biomolecular folding is a common example of such a process. With a more general potential $U(x)$, this picture can also be extended to such phenomena as biomolecular binding (3,4) or even to nonequilibrium processes such as the stroke of a molecular motor (5,6).

Although the theory presented below is general and does not rely on any assumptions about the shape of $U(x)$, for concreteness let us assume that we are dealing with biomolecular folding. The coordinate x, then, may correspond to the molecule's extension, as in single-molecule force spectroscopy (7) or may be related to the distance between two dyes as in fluorescence resonance energy transfer studies (8). The potential $U(x)$ then typically has two wells corresponding to the molecule's unfolded and folded states.

There are several timescales of interest in this case. If the barrier height is much greater than the thermal energy $k_B T$ then the typical time for a transition between the two wells is much longer than the equilibration time within each well, and it is meaningful to define such equilibration timescales, τ_A and τ_B, for each well.
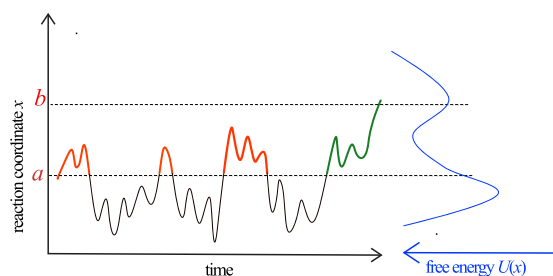

FIGURE 1 Barrier crossing dynamics: a first-passage path from a to b consists of "internal loops" entering and exiting the transition region through the same boundary (22) (*red*), "external loops" in which the system dwells inside the well region $(-\infty, a)$ (*black*), and a "transition path" in which the system traverses the transition region from one boundary to the other (*green*). Typical one-dimensional free energy landscape of folding is shown on the right.

In the field of protein folding, the latter timescales, usually referred to as the "reconfiguration" times (8−12), have attracted considerable experimental attention, particularly for the unfolded state. They can be measured, e.g., using single-molecule fluorescence correlation spectroscopy (13−18).

Reconfiguration timescale is a dynamical characteristic of thermal fluctuations around in a (meta)stable thermodynamic state. To understand the kinetics of transitions between such states, one needs information about the dynamics of barrier crossing. If we choose some point a within the folded state and a point b within the unfolded state, then the first-passage time $t_{FP}(a \to b)$ to go from a to b provides a measure of the overall time required to accomplish the transition. In general, this time of course will depend on the choice of the points a and b, but if these points are separated by a barrier that is much higher than $k_B T$ then the average value of this time is insensitive to the choice of the initial and final points. Indeed, typical trajectories originating from a or b will relax into the adjacent free energy well on a timescale that is much shorter than the escape time from the well, and the much longer escape time from the well will provide a dominant contribution to the mean first-passage time. In this limit, the inverse mean first-passage time defines the rate (coefficient) k for the transition (see, e.g., (19,20)). Moreover, for a double-well free energy landscape the distribution of this time is close to exponential, as expected for first-order chemical kinetics.

To obtain further mechanistic insight into the dynamics of interwell transitions, one can dissect a trajectory that takes the system from a to b into pieces in which the system dwells in the initial state and where it is caught midway between the initial and the final state. The latter are arguably most interesting as they determine the transition "mechanism." To make this idea precise, we call the interval (a, b) the "transition region." Typically, its boundaries are chosen to contain the free energy barrier (Fig. 1). Consider now a typical trajectory that crosses a at $t = 0$ and eventually arrives, for the first time, at b (Fig. 1). In most cases, it will not stay continuously within the interval (a, b) but rather escape it to the left well region $(-\infty, a)$. It may further reenter the transition region and loop back into the left well multiple times. Finally, it will enter the transition region for the last time and proceed to cross the point b thereby exiting the transition region toward the right well region $(b, +\infty)$. Thus, the first-passage path from a to b generally consists of "internal loops" entering and exiting the transition region through the same boundary (*red* in Fig. 1), "external loops" in which the system dwells inside the well region $(-\infty, a)$ (*black* in Fig. 1), and a "transition path" in which the system

traverses the transition region from one boundary to the other (21,22) (*green* in Fig. 1).

The dynamics in the transition region, then, can be characterized by the following:

1) The return time $t_R(a \rightarrow a)$, which is the temporal duration of failed attempts to cross the transition region (the *red trajectory pieces* in Fig. 1). This is the time that the system starting at the boundary a will take to return to this boundary conditional upon staying within the barrier region for $0 < t < t_R(a \rightarrow a)$.

2) The transition path time $t_{TP}(a \rightarrow b)$, which is the temporal duration of successful attempts to cross the barrier (the *green trajectory piece* in Fig. 1). It is the time that the system starting at the boundary a will take to reach the other boundary b conditional upon staying within the barrier region for $0 < t < t_{TP}(a \rightarrow b)$.

The transition path time has received considerable theoretical (2,20–37) and experimental (3,38–44) attention in the last decade, whereas the return time, to our knowledge, has not yet been studied experimentally. As will be seen below, both of these times, as well as the first-passage time, are limiting cases of the more general conditional exit time (45,46), which is the time to exit the transition region through a given boundary (a or b) having started from a given point x_0 within the transition region.

The purpose of this study is twofold. First, we propose that the more general yet experimentally accessible conditional exit time offers additional information about barrier crossing dynamics, particularly about failed barrier crossing attempts not captured by the transition path time. Second, we would like to explore the shape of the distribution of the conditional exit time (and its limiting cases such as the return and transition path time). In particular, the distribution width can be characterized by the ratio of the distribution variance to its mean (known as the coefficient of variation), which, for the case of diffusive dynamics in a potential of mean force, can be calculated analytically. Experimentally, the coefficient of variation, which only requires the first and the second moments of the distribution, is a more robust statistical measure than the distribution itself. Note that in the context of chemical kinetics and molecular biophysics, the first-passage-time distribution, particularly for the model of one-dimensional diffusive motion, is well studied (20,45,47,48), and the expression for the first moment of the transition-path-time distribution has also been derived more recently (25,49). In contrast, higher moments of the distribution of the transition path time, which are required to quantify the distribution's width, have only received attention over the past year (35,37,50).

The importance of the distribution shape and width is exemplified by several recent developments: first, it was shown that the short-time behavior of the first-passage-time distribution contains information about the number of kinetic intermediates of a process (51,52). Second, the shape of the distribution of the transition path time measured in a fluorescence resonance energy transfer study of protein binding suggested the existence of an intermediate encounter complex (i.e., potential well in the transition region) that could not be observed directly (3). Third, it was shown (35,37,50) that the relative width of the distribution of the transition path time, as quantified by its coefficient of variation, cannot be too large if the system's dynamics along the coordinate x is diffusive; thus, observation of broad distributions of transition path times can only be explained by a higher-dimensional free energy landscape allowing for parallel pathways. If transition-path-time distributions for diffusive dynamics are always narrow, what can we say about the distributions of other characteristic times of interest? In all of these cases, distributions of the observed times inform one about global features of the underlying free energy landscapes. Our hope, then, is that studying the distributions of various characteristic barrier crossing times is a viable step toward solving the inverse problem of learning about underlying multidimensional landscapes from low-dimensional signals.

We will further show here that these distributions display interesting (and sometimes pathological) behavior in some limiting cases. In particular, the mean return time $t_R(a \rightarrow a)$ is identically zero in the case of diffusive dynamics. This result has an important experimental implication; because of the finite spatial and temporal resolution of any experiment, the precise crossing of a boundary cannot be detected. As a result, an attempt to measure $t_R(a \rightarrow a)$ will lead to a finite time that depends on, e.g., experimental resolution, data sampling rate, and/or the smoothing procedure used to reduce the noise in the experimental signal. We propose that such artifacts can be avoided if, instead, one considers the conditional exit time $t_E(x_0 \rightarrow a)$, where the system starts from a point $x_0 \in (a, b)$ rather than from a. This resulting time has a well-defined distribution for any $x_0 > a$, and its behavior as x_0 approaches the left boundary a (i.e., when it becomes the return time) is quite interesting, as the distribution becomes "arbitrarily broad" for a sufficiently small distance $(x_0 - a)$ from the boundary in the sense that its mean becomes infinitely smaller than the standard deviation from the mean (note, however, that when inertial effects are neglected as in the Smoluchowski equation model, the absolute values of the standard deviation and the mean both approach 0 as $x_0 \rightarrow a$). In this regard, the distribution of the return

time is drastically different from that of the transition path time. The physical origins of this behavior will be explained.

Likewise, we will show that distributions of first-passage times can be broad, with a coefficient of variation that can be arbitrarily large. In contrast, distributions of the transition path time are always narrow, with a coefficient of variation below 1.

MATERIALS AND METHODS

General formalism

Distributions of conditional exit times and first-passage times

To determine the statistical properties of conditional exit times, we envisage a large ensemble of particles, each starting, at $t = 0$, from a point x_0 that belongs to the transition region (a, b). The stochastic trajectories of each particle may be described by the usual overdamped Langevin equation ([20]), in which the force acting on the particle includes, in addition to the deterministic component $-U'(x)$, a friction force as well as a random noise component. We monitor each particle until it reaches either boundary a or boundary b, record the time it took to reach this boundary (and which boundary was reached), and then discard it from the ensemble. We seek the conditional probability density $\rho_{a(b)}(t|x_0)$, which describes the distribution of the time to exit the transition region for particles exiting through the boundary $a(b)$. Note that $\rho_{a(b)}(t|x_0)$ can be thought of as the probability density of the first-passage time from x_0 to $a(b)$ conditional upon not crossing the point $b(a)$. Further note that

$$p_{a \to b}^{TP}(t) = p_{b \to a}^{TP}(t) = \lim_{x_0 \to a} \rho_b(t|x_0) = \lim_{x_0 \to b} \rho_a(t|x_0) \quad (1)$$

is the distribution of the transition path time for the transition paths starting in a and ending in b, which is the same as the distribution of the transition path time for paths going from b to a because of the time-reversal symmetry of Brownian dynamics ([19,26,27]). Moreover, the limit

$$p_{a \to a}^{R}(t) = \lim_{x_0 \to a} \rho_a(t|x_0) \quad (2)$$

gives the distribution of the return time for recrossing the boundary a. We will, however, show that this limit is pathological for the model of diffusive dynamics.

The ensemble of trajectories originating from x_0 and terminating at the transition region boundaries can be described by the Smoluchowski equation. Specifically, the (conditional) probability density $G(x, t|x_0)$ of finding the particle at point $x \in (a, b)$ at time t (Green's function) obeys the equation

$$\frac{\partial G}{\partial t} = \frac{\partial}{\partial x} D(x) e^{-\beta U(x)} \frac{\partial}{\partial x} e^{\beta U(x)} G, \quad (3)$$

with the initial condition

$$G(x, 0|x_0) = \delta(x - x_0) \quad (4)$$

and the absorbing boundary conditions at the boundaries

$$G(a, t|x_0) = G(b, t|x_0) = 0. \quad (5)$$

The probability density $\rho_{a(b)}(t|x_0)$ is proportional to the flux of particles $j(a, t|x_0)$ or $j(b, t|x_0)$ exiting through the boundary a or b:

$$\rho_a(t|x_0) = -\frac{j(a, t|x_0)}{\phi(x_0 \to a)}$$
$$= \frac{1}{\phi(x_0 \to a)} D(a) \frac{\partial G(x, t|x_0)}{\partial x}\bigg|_{x=a} \quad (6)$$

and

$$\rho_b(t|x_0) = \frac{j(b, t|x_0)}{\phi(x_0 \to b)} = -\frac{1}{\phi(x_0 \to b)} D(b) \frac{\partial G(x, t|x_0)}{\partial x}\bigg|_{x=b}. \quad (7)$$

Here, $\phi(x_0 \to a) = -\int_0^\infty dt j(a, t|x_0)$ and $\phi(x_0 \to b) = 1 - \phi(x_0 \to a) = \int_0^\infty dt j(b, t|x_0)$ are the fractions of particles in the ensemble exiting through a and b; those are known as the splitting probabilities (i.e., the probabilities to exit through the boundaries a and b) ([48]).

The probability distribution of the first-passage time from x_0 to b is calculated analogously, with the only difference being that no absorbing boundary condition is imposed at $x = a$, as the trajectory $x(t)$ can cross the point a any number of times before being terminated at $x = b$ (see Fig. 1). In this case, we have

$$p_{x_0 \to b}^{FP}(t) = -D(b) \frac{\partial G(x, t|x_0)}{\partial x}\bigg|_{x=b} \quad (8)$$

for any starting point $x_0 < b$. We note that for a potential that has the property $U(x) \to \infty$ for $x \to -\infty$, this first-passage-time distribution can be obtained as the $a \to -\infty$ limit of the conditional exit time distribution $\rho_b(t|x_0)$. In other words, because the absorbing boundary at $a \to -\infty$ is never reached, the first-passage times from x_0 to b are identical to the exit times. Thus, first-passage times can be viewed as a limiting case of conditional exit times.

Recursive equations for the distribution moments

Similarly to an approach known in the literature (see, e.g., ([48])), we now outline a general procedure for the calculation of the moments of the distributions of the conditional exit times introduced above. For concreteness, let us focus on the distribution $\rho_a(t|x_0)$ of the exit time to the boundary a. We are interested in its n-th moment,

$$\langle t^n(x_0 \to a) \rangle = \int_0^\infty dt t^n \rho_a(t|x_0)$$
$$= \frac{1}{\phi(x_0 \to a)} \int_0^\infty dt t^n j(a, t|x_0) \quad (9)$$
$$\equiv \frac{1}{\phi(x_0 \to a)} \langle \tau^n(x_0 \to a) \rangle,$$

where we have introduced the moments of an unnormalized distribution

$$\langle \tau^n(x_0 \to a) \rangle = \int_0^\infty dt t^n j(a, t|x_0)$$
$$= \int_0^\infty dt t^n D(a) \frac{\partial G(x, t|x_0)}{\partial x}\bigg|_{x=a}. \quad (10)$$

To find such moments, we start with the adjoint Smoluchowski equation (see, e.g., (20,53)), which considers $G(x, t|x_0)$ as a function of the starting point x_0:

$$\frac{\partial G}{\partial t} = e^{\beta U(x_0)} \frac{\partial}{\partial x_0} D(x_0) e^{-\beta U(x_0)} \frac{\partial G}{\partial x_0}, \tag{11}$$

with the initial condition of Eq. 4 and the boundary conditions

$$G(x, t|a) = G(x, t|b) = 0. \tag{12}$$

Multiplying both sides of this equation by t^n and integrating over time, we obtain

$$e^{\beta U(x_0)} \frac{\partial}{\partial x_0} D(x_0) e^{-\beta U(x_0)} \frac{\partial F_n}{\partial x_0} = -n F_{n-1}, n > 0 \tag{13}$$

and

$$e^{\beta U(x_0)} \frac{\partial}{\partial x_0} D(x_0) e^{-\beta U(x_0)} \frac{\partial F_0}{\partial x_0} = -\delta(x - x_0), \tag{14}$$

where we have introduced the auxiliary functions

$$F_n(x|x_0) = \int_0^\infty dt\, t^n G(x, t|x_0), \tag{15}$$

which satisfy the boundary condition

$$F_n(x|a) = F_n(x|b) = 0. \tag{16}$$

From Eq. 10, the moments $\langle \tau^n(x_0 \rightarrow a) \rangle$ now can be written as

$$\langle \tau^n(x_0 \rightarrow a) \rangle = D(a) \frac{\partial F_n(x|x_0)}{\partial x} \bigg|_{x=a}. \tag{17}$$

Using Eqs. 13 and 17, we find that these moments satisfy the following equation:

$$e^{\beta U(x_0)} \frac{\partial}{\partial x_0} D(x_0) e^{-\beta U(x_0)} \frac{\partial \langle \tau^n(x_0 \rightarrow a) \rangle}{\partial x_0} = \\ - n \langle \tau^{n-1}(x_0 \rightarrow a) \rangle, n > 0, \tag{18}$$

which should be supplemented with the boundary conditions

$$\langle \tau^n(x_0 \rightarrow a) \rangle \big|_{x=a} = \langle \tau^n(x_0 \rightarrow a) \rangle \big|_{x=b} = 0. \tag{19}$$

Thus, if the $(n - 1)$-th moment, $\langle \tau^{n-1}(x_0 \rightarrow a) \rangle$, is known, then the next moment $\langle \tau^n(x_0 \rightarrow a) \rangle$ can be obtained by integrating Eq. 18 twice.

To obtain the n-th moment of interest, $\langle \tau^n(x_0 \rightarrow a) \rangle$, the moment $\langle \tau^n(x_0 \rightarrow a) \rangle$ needs to be divided by the splitting probability (Eq. 9). As this probability is the zeroth-order moment of the unnormalized distribution, it can be obtained by solving Eq. 14 with the boundary conditions of Eq. 16, resulting in the known result (48)

$$\phi(x_0 \rightarrow a) = D(a) \frac{\partial}{\partial x} \left[\int_0^\infty dt\, G(x, t|x_0) \right]_{x=a}$$

$$= D(a) \frac{\partial F_0(x|x_0)}{\partial x} \bigg|_{x=a} = \frac{\int_{x_0}^b dx\, e^{\beta U(x)} \big/ D(x)}{\int_a^b dx\, e^{\beta U(x)} \big/ D(x)}. \tag{20}$$

Note that although the equation hierarchy, Eq. 18, is similar to that derived in the literature for first-passage times (37,48), there is an important difference: because $\rho_{a(b)}(t|x_0)$ is a conditional distribution, Eq. 18 is satisfied by the moments $\langle \tau^n(x_0 \rightarrow a) \rangle$ and not by $\langle t^n(x_0 \rightarrow a) \rangle$.

RESULTS

Exit times

We now give the general analytical solutions for the first and second moments of the exit times. Starting with the unnormalized distribution's "zeroth moment" (which is the splitting probability), integrating Eq. 18 twice, and dividing by the splitting probability (Eq. 9), we find the first moment

$$\langle t(x_0 \rightarrow a) \rangle_E = \frac{1}{\int_{x_0}^b dx\, e^{\beta U(x)} \big/ D(x)} \int_a^{x_0} \frac{dx\, e^{\beta U(x)}}{D(x)} \\ \times \int_a^b \frac{dy\, e^{\beta U(y)}}{D(y)} \int_x^y dz\, \phi(z \rightarrow a) e^{-\beta U(z)}, \tag{21}$$

where the splitting probability $\phi(z \rightarrow a)$ is defined in Eq. 20. (Note that here we use the notation $\langle \ldots \rangle_E$ to indicate averaging over the distribution of conditional exit times.) Similarly, the second moment now can be expressed in terms of the first moment by integrating Eq. 18 twice. This gives

$$\langle t^2(x_0 \rightarrow a) \rangle_E = \frac{2}{\int_{x_0}^b dx\, e^{\beta U(x)} \big/ D(x)} \int_a^{x_0} \frac{dx\, e^{\beta U(x)}}{D(x)} \int_a^b \frac{dy\, e^{\beta U(y)}}{D(y)} \\ \times \int_x^y dz\, \phi(z \rightarrow a) \langle t(z \rightarrow a) \rangle_E e^{-\beta U(z)}. \tag{22}$$

The expression for the n-th moment, $\langle t^n(x_0 \rightarrow a) \rangle_E$, $n = 3$, 4..., is obtained by replacing $2\langle t(z \rightarrow a) \rangle_E$ in Eq. 22 by $n\langle t^{n-1}(z \rightarrow a) \rangle_E$.

We are particularly interested in the distribution's relative width, which is conventionally quantified by

its coefficient of variation equal to the ratio of the standard deviation to the distribution's mean:

$$C = \frac{[\langle t^2 \rangle - \langle t \rangle^2]^{\frac{1}{2}}}{\langle t \rangle}. \tag{23}$$

Narrower distributions have smaller values of the coefficient of variation, and the coefficient of variation for an exponential distribution is 1. The value of the coefficient of variation encodes important information about the underlying dynamics and energy landscape. For example, in the case of barrier crossing by transition paths, lower values of this coefficient correspond to higher barriers (35), whereas a value of 1 (corresponding to a single-exponential distribution) suggests (inasmuch as the model of diffusive dynamics is applicable) that transition paths cross a potential well (intermediate) that traps the system (3). Even more interestingly, regardless of the underlying potential of mean force, this coefficient cannot possibly exceed 1 for the model of one-dimensional diffusive dynamics (35,37); thus, experimental observation of C exceeding one automatically invalidates such a model. But what about coefficients of variation for the more general exit time?

For the double-well potential shown in Fig. 2, Fig. 3 shows how the coefficient of variation of the exit time through the left boundary depends on the starting point x_0. For x_0 approaching the right boundary, $x_0 \rightarrow b$, the exit time approaches the transition path time $t_{TP}(b \rightarrow a)$; it has been shown previously (35,37) that the coefficient of variation for this time is always less than 1 (assuming the validity of the diffusive dynamics model).

In the opposite limit, $x_0 \rightarrow a$, the exit time becomes the return time $t_R(a \rightarrow a)$, and we observe that the coefficient of variation diverges. Mathematically, this divergence arises because both the first and the second moments grow linearly with $x_0 - a$ in the limit $x_0 - a \rightarrow 0$.

FIGURE 3 Coefficient of variation for the exit time distribution $\rho_a(t|x_0)$ as a function of the starting point x_0. The transition region boundaries are shown as dashed lines. Inset shows the mean exit time as a function of x_0. This time is measured in dimensionless units, with L^2/D (where $2L$ is the distance between the two potential minima and D the diffusivity) being the unit of time. The unit of distance is L.

Broadening of the exit time distribution as the starting point x_0 approaches the absorbing boundary a is also observed directly in the distribution's shape shown, for three values of x_0, in Fig. 4. These distributions were obtained using Eq. 6, from a numerical solution of the Smoluchowski equation that uses the spectral expansion method (see 27 for further details). Note that because both the first and the second moments of this distribution increase with increasing x_0 (see the *inset* of Fig. 3, in which the first moment of the distribution is plotted as a function of x_0), the time in each distribution is rescaled by its mean value. In other words, the first moments of each distribution as plotted in Fig. 4 are the same and equal to one.

At first glance, the distribution $\rho_a(t|x_0)$ obtained for a value x_0 that is closest to the left boundary seem more
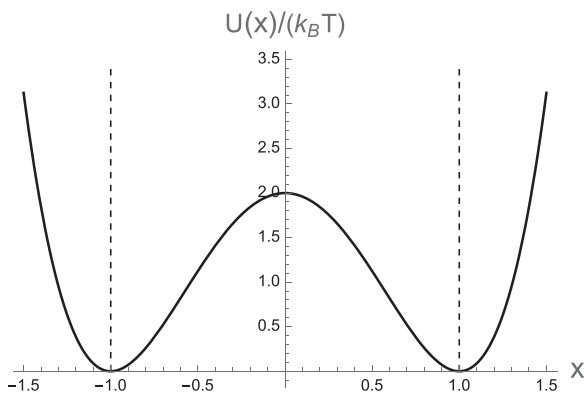
FIGURE 2 Double-well potential $U(x) = 2k_B T(x^2 - 1)^2$ used to illustrate the results here. The transition region boundaries, $a = -1$, $b = 1$, are shown as dashed lines.
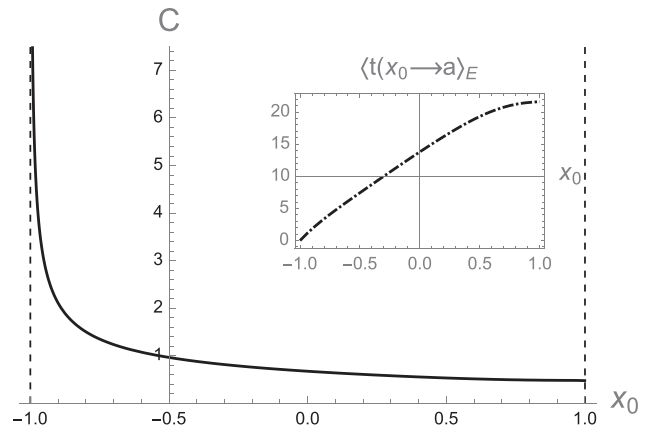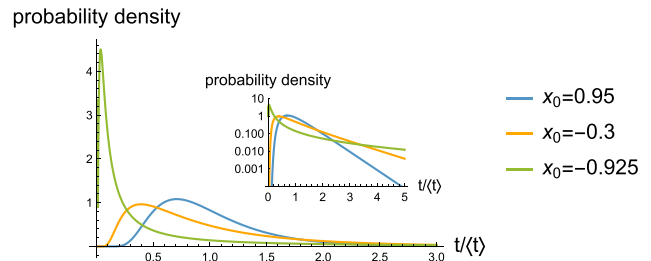
FIGURE 4 Probability distribution $\rho_a(t|x_0)$ of the exit time to the left boundary (Eq. 6) for different values of the initial position x_0, as indicated in the legend. The time in each case is normalized by the distribution's mean. The potential $U(x)$ and the transition region boundaries are the same as in Figures. 2 and 3. Of the three distributions shown here, the one corresponding to the point x_0 closest to the left boundary is the broadest (i.e., has the largest coefficient of variation) because of its long tail. This tail is easy to see when the same data are plotted on a logarithmic scale (*inset*).

"narrow" than the other two, showing a sharper rise at short times. But this is not so (inasmuch as the coefficient of variation is a good measure of width) because of the long tail exhibited by this distribution. This tail is more readily observed when the same distribution is plotted on a logarithmic scale (Fig. 4, inset).

To further understand the distribution broadening as x_0 approaches the boundary a, it is instructive to consider the case of zero potential, $U(x) = 0$, with a coordinate-independent diffusion coefficient $D(x) = D$. Setting, without loss of generality, $a = 0$, the integrals of Eqs. 21 and 22 can be evaluated analytically to give

$$\langle t(x_0 \to 0) \rangle_E = \frac{(2b - x_0)x_0}{6D} \qquad (24)$$

and

$$\langle t^2(x_0 \to 0) \rangle_E = \frac{x_0\left(8b^3 + 8b^2 x_0 - 12b x_0^2 + 3x_0^3\right)}{180D^2}. \qquad (25)$$

Both functions are proportional to x_0 in the limit $x_0 \to 0$, thus leading to $x_0^{-1/2}$ divergence of the coefficient of variation (Eq. 23).

Moreover, the distribution of the exit time for the case of zero potential can be calculated analytically. Although this problem of free diffusion on an interval with absorbing boundaries was studied in the literature—see, e.g., (45)—we provide a derivation for $\rho_0(t|x_0)$ in Appendix A for completeness and as a simple illustration of the approach here. Specifically, the Laplace transform of $\rho_0(t|x_0)$ is given by

$$\tilde{\rho}_0(s|x_0) = \frac{\sinh\left[(b - x_0)\sqrt{\frac{s}{D}}\right]}{(b - x_0)\sinh\left[b\sqrt{\frac{s}{D}}\right]}. \qquad (26)$$

Consider now the case $x_0 \ll b$, $b\sqrt{\frac{s}{D}} \gg 1$. In the time domain, the second inequality implies that we are considering timescales much shorter than the characteristic diffusion time on the segment $(0, b)$, i.e., $t \ll \frac{b^2}{D}$. In this limit, we obtain $\tilde{\rho}_0(s|x_0) \approx \frac{1}{b - x_0}e^{-x_0\sqrt{\frac{s}{D}}}$, and taking the inverse Laplace transform, we find

$$\rho_0(t|x_0) \approx \frac{x_0}{2\sqrt{\pi D t^3}}e^{-\frac{x_0^2}{4Dt}}. \qquad (27)$$

In fact, Eq. 27 provides an accurate description of the exit time distribution even in the presence of a nonzero potential $U(x)$ as long as the time t is not too long; see Fig. 5. This is because, at short enough timescales, diffusion dominates over the drift, and the presence of the potential is immaterial. At longer times, however,

probability density
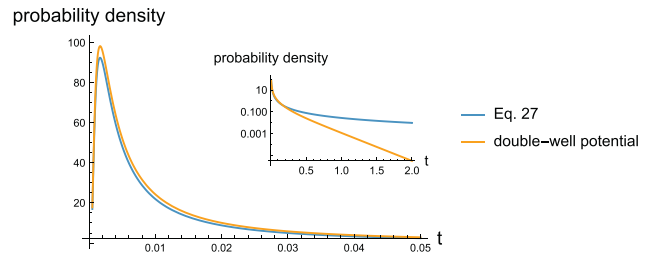
probability density

— Eq. 27
— double-well potential

FIGURE 5 Eq. 27 versus the probability distribution for the exit time to the left boundary, $\rho_a(t|x_0)$, in the double-well potential of Fig. 2. In both cases, the distance $x_0 - a$ is much shorter than the length $b - a$ of the transition region and is equal to 0.1. The time is measured in dimensionless units, with L^2/D (where $2L$ is the distance between the two potential minima and D the diffusivity) being the unit of time. The distance L provides a unit of distance.

the exit-time distribution exhibits an exponential tail, in contrast to a power-law tail predicted by Eq. 27. A true power-law tail will be observed in the absence of a potential when the right boundary is removed to infinity (see First-passage times).

Returning to the case of zero potential, when $x_0^2 \ll Dt$, the exponential in Eq. 27 can be replaced by 1. On the other hand, when $Dt \ll b^2$, the particle originating from x_0 does not have time to reach the right boundary b. When both of these conditions are satisfied, $\frac{x_0^2}{D} \ll t \ll \frac{b^2}{D}$, Eq. 27 predicts a power law $\rho_0(t|x_0) \propto t^{-\frac{3}{2}}$, a behavior characteristic of a "broad" distribution displaying multiple timescales. Note, however, that the absolute width of the distribution, as opposed to the relative width quantified by the coefficient of variation, decreases as $x_0 \to 0$. Indeed, at $x_0 = 0$, the Laplace transform of this distribution is constant, and the distribution is the δ function, $\rho_0(t|x_0) = \delta(t)$, indicating that a particle cannot leave the absorbing boundary $x_0 = 0$. In other words, the return time is identically zero in this case.

This δ-function distribution of the return time should be expected of diffusive dynamics (or equivalently, of the overdamped Langevin equation in which the inertial term containing acceleration is omitted). It is pathological: if inertial effects are taken into account, the system crossing the left boundary a with a positive velocity will travel to the right over some finite time. Thus, it will take a finite time to return to the boundary. The typical return time, then, would be comparable to the velocity correlation time, which is typically very short and thus unresolvable by current single-molecule techniques (54). This conclusion has important experimental implications: an attempt to measure the return time will likely be confounded by experimental artifacts such as the limited spatial and temporal resolution or the smoothing of the trajectory (which introduces an artificial "velocity" that would be absent in true diffusive

dynamics). These difficulties can be avoided by measuring the exit-time distribution $\rho_a(t|x_0)$ instead, with the initial point x_0 being within the limits allowed by the experiment's spatial resolution.

First-passage times

We now consider the distribution of the first-passage time $p_{x_0 \to b}^{FP}(t)$ to arrive at the (target) boundary b starting from $x = x_0 < b$. The first moment of this distribution is given by the known expression (20,45,47)

$$\langle t(x_0 \to b) \rangle_{FP} = \int_{x_0}^{b} \frac{dx e^{\beta U(x)}}{D(x)} \int_{-\infty}^{x} dy e^{-\beta U(y)}, \quad (28)$$

and the second moment can be obtained similarly to the approach of Exit times by integrating Eq. 18:

$$\langle t^2(x_0 \to b) \rangle_{FP} = 2 \int_{x_0}^{b} \frac{dx e^{\beta U(x)}}{D(x)} \int_{-\infty}^{x} dy e^{-\beta U(y)} \langle t(y \to b) \rangle_{FP}. \quad (29)$$

The dependence of the coefficient of variation on the initial position x_0 for the double-well potential shown in Fig. 2 and for a coordinate-independent $D(x)$ is illustrated in Fig. 6. For x_0 located not too far from the left potential minimum (located at $x = -1$), we observe that its value is close to 1. This is easy to understand: thermally activated escape from a potential well separated from the target boundary b by a sufficiently high barrier is governed by first-order kinetics, resulting in an exponential distribution of escape times whose coefficient of variation is 1. Even though the barrier in this case is only $2k_BT$, the coefficient of variation is only slightly below 1.
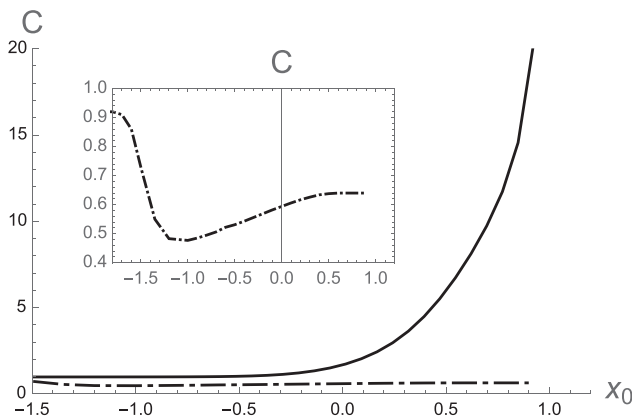

FIGURE 6 Coefficient of variation C for the first-passage time from x_0 to b (*solid line*) and for the transition path time from x_0 to b (*dashed-dotted line*) as a function of the starting point x_0. Inset: same data for the transition path time.

When the initial position x_0 moves toward b, however, the coefficient of variation increases, and it diverges as the starting point approaches the target, $x_0 \to b$. This behavior is similar to that of the exit time discussed in Exit times and illustrated in Fig. 3. The physical origins of this behavior will be further discussed below. Mathematically, the divergent behavior of the coefficient of variation follows from Eqs. 28 and 29 (and from the definition of the coefficient of variation, Eq. 23). Indeed, based on these equations, both the first and the second moments of the distribution, to lowest order in the distance to the target $(b - x_0)$, are proportional to $(b - x_0)$, and thus, the coefficient of variation scales as $C \propto (b - x_0)^{-\frac{1}{2}}$.

To gain further insight into the shapes of the first-passage-time distributions, consider the analytically tractable case of a linear potential,

$$U(x) = -Fx, \quad (30)$$

describing a particle subjected to a constant force $F \geq 0$. Using Eqs. 28 and 29, we find, for $x_0 < b$,

$$\langle t(x_0 \to b) \rangle_{FP} = \frac{b - x_0}{\beta F D} \quad (31)$$

and

$$\langle t^2(x_0 \to b) \rangle_{FP} = \frac{2}{(\beta F)^3 D^2} \left[(b - x_0) + \frac{\beta F(b - x_0)^2}{2} \right], \quad (32)$$

from which, using Eq. 23, we find

$$C = \left[\frac{2}{\beta F(b - x_0)} \right]^{1/2} \quad (33)$$

showing divergent behavior as $x_0 \to b$.

A particularly interesting case is that of free diffusion, which is obtained by setting the force F to zero. Formally, Eq. 33 predicts that the coefficient of variation should diverge for any initial position x_0, but of course, both the first and the second moments of the distribution diverge in this case, so the coefficient of variation is not well defined at exactly zero force (see Eq. 23). Nevertheless, the distribution of the first-passage time itself is well defined for free diffusion. It is given by (see, e.g., (45); Appendix B provides a derivation for completeness)

$$p_{x_0 \to b}^{FP}(t) = \frac{b - x_0}{2\sqrt{\pi D t^3}} e^{-\frac{(b - x_0)^2}{4Dt}}. \quad (34)$$

Equation 34 is identical to Eq. 27 (provided that x_0 is replaced by $b - x_0$; both of these quantities are the distances to the absorbing boundary in each case).

However, unlike Eq. 27, which only holds at intermediate timescales, Eq. 34 has a power-law tail decaying as $t^{-3/2}$ at arbitrarily long times. This difference is due to the fact that diffusion on an infinitely long segment, unlike diffusion within a transition region of finite length, does not possess a longest characteristic timescale. More specifically, as noted above, the power law in Eq. 27 holds only at times short enough that the particle does not have enough time to reach the right boundary b. In contrast, no such boundary exists in this case, diffusion takes place on a semi-infinite segment, and arbitrarily long excursions from the point x_0 are possible. Although the distribution of Eq. 34 is itself normalized, all of its moments diverge.

These findings show that first-passage time distributions can be broad, with a coefficient of variation exceeding 1, or even with its moments being infinite. This is different from the distribution of the transition path time, which, for one-dimensional diffusive dynamics, is always below 1, as further discussed below.

Transition path times

The first and second moments of the distribution of the transition path time can be found from Eqs. 21 and 22 as the limiting case in which the starting point x_0 coincides with the absorbing boundary b. Note that this gives moments of the transition path time from b to a, but given the time-reversal symmetry of transition paths, they are the same as those for transition paths from a to b. The result for the first moment has been derived previously (25,49,55):

$$\langle t(a \to b) \rangle_{TP} = \langle t(b \to a) \rangle_{TP} = \left(\int_a^b \frac{dx e^{\beta U(x)}}{D(x)} \right) \times \int_a^b dz \phi(z \to x_0) \phi(z \to b) e^{-\beta U(z)}. \quad (35)$$

For the second moment, we arrive at

$$\langle t^2(a \to b) \rangle_{TP} = \langle t^2(b \to a) \rangle_{TP} = 2 \left(\int_a^b \frac{dx e^{\beta U(x)}}{D(x)} \right) \times \int_a^b dz \phi(z \to x_0) \phi(z \to b) \langle t(z \to b) \rangle_E e^{-\beta U(z)}, \quad (36)$$

where $\langle t(x_0 \to b) \rangle_E$ is the mean exit time through boundary b conditional upon not crossing the boundary a. This result was reported earlier in (35) for coordinate-independent diffusivity and in (37).

As proven in (35,37), the coefficient of variation of the transition-path-time distribution (Eq. 23) always remains below 1 (Fig. 6). For the potential shown in Fig. 2 and for initial points $x_0 < -1$ located to the left of the potential minimum at $x = -1$, the coefficients of variation for both transition-path-time and first-passage-time distributions are close to each other and to the value $C = 1$ expected for an exponential distribution (Fig. 6). This is not surprising; if the starting point $x_0 < -1$ is located on a steep left wall of a potential well, a trajectory originating from x_0 will likely proceed toward the right of the starting point, and thus, a first-passage and a transition path time should be nearly the same. Moreover, their distributions should be close to exponential, with $C \approx 1$, as these times should be close to the time of thermally activated escape from the left potential well.

As the initial point x_0 approaches the target b, the coefficient of variation for the transition-path-time distribution stays below 1, whereas the coefficient of variation for the first-passage-time distribution increases and diverges for $x_0 \to b$, a behavior explained in First-passage times.

DISCUSSION

Our findings show that the distribution of the transition path time is special in that, in the case of purely diffusive dynamics, its coefficient of variation cannot exceed 1, in contrast to the distributions of the first-passage and conditional exit times. In other words, transition-path-time distributions are narrow (narrower than exponential), whereas first-passage-time and exit-time distributions may be arbitrarily broad, with their coefficients of variations being arbitrarily large.

What is the physical origin of this difference? It is instructive to consider the free diffusion case discussed in First-passage times. Consider a Brownian particle starting to the left of the target at $x_0 < b$. It may proceed directly toward the target b, in which case its trajectory will be a transition path from x_0 to b, or it may exercise an arbitrarily long detour to the left of the starting point before finally finding the target at $x = b$ (note that in one dimension, the particle will eventually find the target with certainty). In a sense, these two scenarios may be considered as two parallel pathways contributing to the first-passage-time distribution, but only with the first scenario contributing to the transition-path-time distribution. As shown in (35,50), it is the existence of such parallel pathways with disparate characteristic timescales that leads to broad distributions, and indeed, Eq. 34 exhibits a power-law tail describing long-lasting events.

Exit times (Eqs. 21 and 22) provide a more general description of barrier crossing dynamics than the

better-studied transition path time. Exit times have been discussed previously in the context of finding practical reaction coordinates (46). It is widely believed that the splitting probability (committor) has attractive mathematical properties that makes it in a certain sense an optimal reaction coordinate (25,53,56–62). The splitting probability answers the following question: starting from a certain point in phase or configuration space, what is the probability to get to the reaction product before getting to the reactant? In our one-dimensional model, the function $\phi(x_0 \to b)$ (x_0 being the starting point) gives the answer to this question, provided that the segment $(-\infty, a)$ is regarded as the reactant and $(b, +\infty)$ as the product. The conditional exit time is a complementary quantity that answers a question that is concerned with the transition time-scale: starting from a certain point in phase or configuration space, what is the (mean) time to get to the product conditional upon reaching the product before getting to the reactant? In our case, this question is answered by the mean exit time $\langle t(x_0 \to b)\rangle_E$. The exit time generalizes the notion of the transition path time: in the limit in which the initial point x_0 coincides with the boundary a, this time becomes the transition path time $t_{TP}(a \to b)$. Another limiting case of the conditional exit time is the (mean) return time $\langle t(x_0 \to a)\rangle_E|_{x_0 \to a} = \langle t(a \to a)\rangle_R$, which is the mean time it takes to return to the boundary a starting from this boundary and conditional upon not exiting through the boundary b. For the case of diffusive dynamics considered here, however, this time is identically equal to 0.

This work has focused on the model of diffusive dynamics along a one-dimensional reaction coordinate. Given the many successes of this model in application to biomolecular folding (see, e.g., (63–65)), we view elaboration on further consequences of this model for the dynamics in the transition region a worthy pursuit, but it also gives us an opportunity to delineate potential limitations of this model in application to experimental data. Indeed, the following three properties are strictly satisfied by this model, yet they may be violated when the dynamics along the reaction coordinate are influenced by memory effects and/or when multidimensionality is essential (66,67):

1) Locality of exit times: Distributions of exit times (and transition path times in particular) are independent of the potential $U(x)$ outside the transition region (a, b). This is obvious from the derivation outline described in Distributions of conditional exit times and first-passage times, as the properties of the potential outside the transition interval do not enter into the picture (of course, Markovianity of diffusion, allowing one to disregard any knowledge

of the system's past before its arrival at the point x_0 is key here). Note that this local property is not true for first-passage times from a to b, which depend on the properties of the potential $U(x)$ for $x < a$.

2) The distribution of the transition path time has a coefficient of variation that cannot exceed 1, regardless of the potential and the transition region boundaries. This property is a direct consequence of Eqs. 23, 35, and 36 (see 35).

3) Transition path times exhibit forward and backward symmetry, $p_{a \to b}^{TP}(t) = p_{b \to a}^{TP}(t)$, as already stated in Eq. 1. Note that this property is true even for systems that are not in equilibrium (68) because of property 1.

Although strictly true for diffusive dynamics, some (or all) of these properties may not be true if the dynamics along x is not diffusive. For example, property 1 cannot be generally true for non-Markovian dynamics, property 2 has already been found to be violated for the dynamics of reaction coordinates in protein folding (35), and property 3, although true for any equilibrium system, was found to be violated for systems that are simultaneously nonequilibrium and non-Markovian (68). Violation of any of the above exact predictions would invalidate, with certainty, the diffusive picture of the dynamics along a reaction coordinate and call for a more accurate model.

We conclude this work with comments on how experimental time resolution may affect information that can be deduced from the distributions of the various times discussed here. Generally, limited time resolution may result in missing short-time events, thereby skewing the apparent distribution toward longer times while simultaneously making it narrower. More formally, we may write the true distribution $p(t)$ of the time of interest as

$$p(t) = f p_{un}(t) + (1-f) p_{obs}(t), \qquad (37)$$

where f is the fraction of unobserved short-time events with a normalized distribution $p_{un}(t)$, and $1 - f$ the fraction of observed events with a normalized distribution $p_{obs}(t)$. The latter distribution is the one observed experimentally. Thus, we have

$$C^2 + 1 = \frac{\langle t^2 \rangle}{\langle t \rangle^2} = \frac{f \langle t^2 \rangle_{un} + (1-f) \langle t^2 \rangle_{obs}}{[f \langle t \rangle_{un} + (1-f) \langle t \rangle_{obs}]^2}, \qquad (38)$$

where $\langle t \rangle_{un}$, $\langle t \rangle_{obs}$ and $\langle t^2 \rangle_{un}$, $\langle t^2 \rangle_{obs}$ are the first and second moment of the unobserved and observed distributions, as indicated by the subscripts. Given that unobserved events are shorter than the observed ones, we have $\langle t \rangle_{un} < \langle t \rangle_{obs}$, and it is obvious that the

true distribution mean $\langle t \rangle = f\langle t \rangle_{un} + (1-f)\langle t \rangle_{obs}$ is always shorter than the observed value $\langle t \rangle_{obs}$.

Let us now focus on the case of measuring transition path times. Given that higher barriers (for the same transition region boundaries) correspond to shorter mean transition path times (19,25), this longer observed value of the transition path time could lead the experimentalist to deduce a lower transition barrier. On the other hand, as follows from Eq. 38, the true coefficient of variation C is greater than the observed one, $C_{obs} = \frac{[\langle t^2 \rangle_{obs} - \langle t \rangle_{obs}^2]^{\frac{1}{2}}}{\langle t \rangle_{obs}}$, and a narrower observed distribution of the transition path time, with a greater coefficient of variation, corresponds to a higher barrier. Thus, analysis of experimental distribution widths may reveal experimental artifacts.

APPENDIX A: EXIT TIME DISTRIBUTION FOR FREE DIFFUSION

Following the approach outlined in Distributions of conditional exit times and first-passage times, we write the diffusion equation describing the time evolution of Green's function $G(x, t|x_0)$ of a free particle (i.e., Eq. 3 with constant diffusivity D and with $U(x) = 0$)

$$\frac{\partial G}{\partial t} = D \frac{\partial^2 G}{\partial x^2}, \qquad (A1)$$

with the initial condition

$$G(x, 0|x_0) = \delta(x - x_0) \qquad (A2)$$

and with absorbing boundary conditions (cf. Eq. 5)

$$G(a = 0, t|x_0) = G(b, t|x_0) = 0. \qquad (A3)$$

To solve Eq. A1, we rewrite it in Laplace space:

$$s\widehat{G} - \delta(x - x_0) = D \frac{\partial^2 \widehat{G}}{\partial x^2}, \qquad (A4)$$

where

$$\widehat{G}(x, s|x_0) = \int_0^\infty dt e^{-st} G(x, t|x_0). \qquad (A5)$$

The solution of Eq. A4 satisfying the absorbing boundary conditions can be written in the form

$$\widehat{G}(x, s|x_0) = \begin{cases} A \sinh\left[x\sqrt{\dfrac{s}{D}}\right], & 0 \leq x < x_0 \\ B \sinh\left[(b-x)\sqrt{\dfrac{s}{D}}\right], & x_0 < x \leq b \end{cases}. \qquad (A6)$$

The coefficients A and B can be determined from the continuity of the Green's function at $x = x_0$,

$$\widehat{G}(x_0 + 0, s|x_0) = \widehat{G}(x_0 - 0, s|x_0), \qquad (A7)$$

and from the condition

$$D \frac{\partial \widehat{G}}{\partial x}\bigg|_{x_0 - 0}^{x_0 + 0} = -1, \qquad (A8)$$

which is obtained by integrating Eq. A4 over an infinitesimal interval from $x_0 - 0$ to $x_0 + 0$.

The exit time distribution can now be found from Eq. 6 rewritten in Laplace space:

$$\widehat{\rho}_0(t|x_0) = \frac{1}{\phi(x_0 \to 0)} D \frac{\partial \widehat{G}(x, t|x_0)}{\partial x}\bigg|_{x=0}, \qquad (A9)$$

where the splitting probability $\phi(x_0 \to 0)$ is the integral of the flux exiting through the left boundary,

$$\begin{aligned} \phi(x_0 \to 0) &= D \int_0^\infty dt \frac{\partial G(x, t|x_0)}{\partial x}\bigg|_{x=0} \\ &= D \frac{\partial \widehat{G}(x, t|x_0)}{\partial x}\bigg|_{x=0} \end{aligned} \qquad (A10)$$

The expression in Eq. 26 is then obtained using Eqs. A6, A7, A8, A9, and A10.

APPENDIX B: FIRST-PASSAGE TIME DISTRIBUTION FOR FREE DIFFUSION

Following the approach outlined in Distributions of conditional exit times and first-passage times, we seek the solution $G(x, t|x_0)$ of the diffusion equation

$$\frac{\partial G}{\partial t} = D \frac{\partial^2 G}{\partial x^2}, \qquad (B1)$$

with the initial condition

$$G(x, 0|x_0) = \delta(x - x_0) \qquad (B2)$$

and with absorbing boundary condition at $x = b$:

$$G(b, t|x_0) = 0. \qquad (B3)$$

We assume that $x_0 < b$. The solution for $x < b$ is easily constructed from Green's function of the freely diffusing particle using the method of images:

$$G(x, t|x_0) = \frac{1}{\sqrt{4\pi Dt}}\left[e^{-\frac{(x-x_0)^2}{4Dt}} - e^{-\frac{(x-2b+x_0)^2}{4Dt}}\right]. \qquad (B4)$$

The probability distribution of the first-passage time from x_0 to the boundary b is equal to the flux at the target boundary,

$$p_{x_0 \to b}^{FP}(t) = -D\frac{\partial G(x,t|x_0)}{\partial x}\bigg|_{x=b} = \frac{b - x_0}{2\sqrt{\pi D}}t^{-\frac{3}{2}}e^{-\frac{(b-x_0)^2}{4Dt}},$$

$$(B5)$$

which is the expression in Eq. 34. Note that this distribution is automatically normalized because the integral of the flux over time, equal to the probability of crossing $x = b$ at any time, is 1.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Kilic, Z., I. Sgouralis, and S. Pressé. 2021. Generalizing HMMs to continuous time for fast kinetics: hidden Markov jump processes. *Biophys. J.* 120:409–423.

2. Makarov, D. E. 2021. Barrier crossing dynamics from single-molecule measurements. *J. Phys. Chem. B.* 125:2467–2476.

3. Sturzenegger, F., F. Zosel, …, B. Schuler. 2018. Transition path times of coupled folding and binding reveal the formation of an encounter complex. *Nat. Commun.* 9:4708.

4. Schreiber, G., G. Haran, and H. X. Zhou. 2009. Fundamental aspects of protein-protein association kinetics. *Chem. Rev.* 109:839–860.

5. Kolomeisky, A. B. 2015. Motor Proteins and Molecular Motors, First Edition. CRC Press, Boca Raton, FL.

6. Berezhkovskii, A. M., and D. E. Makarov. 2020. From nonequilibrium single-molecule trajectories to underlying dynamics. *J. Phys. Chem. Lett.* 11:1682–1688.

7. Hoffer, N. Q., and M. T. Woodside. 2019. Probing microscopic conformational dynamics in folding reactions by measuring transition paths. *Curr. Opin. Chem. Biol.* 53:68–74.

8. Schuler, B., and W. A. Eaton. 2008. Protein folding studied by single-molecule FRET. *Curr. Opin. Struct. Biol.* 18:16–26.

9. Schuler, B., E. A. Lipman, and W. A. Eaton. 2002. Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. *Nature.* 419:743–747.

10. Hagen, S. J., J. Hofrichter, …, W. A. Eaton. 1996. Diffusion-limited contact formation in unfolded cytochrome c: estimating the maximum rate of protein folding. *Proc. Natl. Acad. Sci. USA.* 93:11615–11617.

11. Bieri, O., and T. Kiefhaber. 1999. Elementary steps in protein folding. *Biol. Chem.* 380:923–929.

12. Bieri, O., J. Wirz, …, T. Kiefhaber. 1999. The speed limit for protein folding measured by triplet-triplet energy transfer. *Proc. Natl. Acad. Sci. USA.* 96:9597–9601.

13. Wang, Z., and D. E. Makarov. 2003. Nanosecond dynamics of single polypeptide molecules revealed by photoemission statistics of fluorescence resonance energy transfer: a theoretical study. *J. Phys. Chem. B.* 107:5617–5622.

14. Gopich, I. V., D. Nettels, …, A. Szabo. 2009. Protein dynamics from single-molecule fluorescence intensity correlation functions. *J. Chem. Phys.* 131:095102.

15. Hoffmann, A., A. Kane, …, B. Schuler. 2007. Mapping protein collapse with single-molecule fluorescence and kinetic synchrotron radiation circular dichroism spectroscopy. *Proc. Natl. Acad. Sci. USA.* 104:105–110.

16. Nettels, D., I. V. Gopich, …, B. Schuler. 2007. Ultrafast dynamics of protein collapse from single-molecule photon statistics. *Proc. Natl. Acad. Sci. USA.* 104:2655–2660.

17. Soranno, A., B. Buchli, …, B. Schuler. 2012. Quantifying internal friction in unfolded and intrinsically disordered proteins with single-molecule spectroscopy. *Proc. Natl. Acad. Sci. USA.* 109:17800–17806.

18. Haas, E., and I. Z. Steinberg. 1984. Intramolecular dynamics of chain molecules monitored by fluctuations in efficiency of excitation energy transfer. A theoretical study. *Biophys. J.* 46:429–437.

19. Makarov, D. E. 2015. Single Molecule Science: Physical Principles and Models. CRC Press, Boca Raton, FL.

20. Elber, R., D. E. Makarov, and H. Orland. 2020. Molecular Kinetics in Condensed Phases: Theory, Simulation, and Analysis, First Edition. Wiley and Sons, Hoboken, NJ.

21. Berezhkovskii, A. M., L. Dagdug, and S. M. Bezrukov. 2017. First passage, looping, and direct transition in expanding and narrowing tubes: effects of the entropy potential. *J. Chem. Phys.* 147:134104.

22. Berezhkovskii, A. M., L. Dagdug, and S. M. Bezrukov. 2017. Mean direct-transit and looping times as functions of the potential shape. *J. Phys. Chem. B.* 121:5455–5460.

23. Zhang, B. W., D. Jasnow, and D. M. Zuckerman. 2007. Transition-event durations in one-dimensional activated processes. *J. Chem. Phys.* 126:074504.

24. Zuckerman, D. M., and T. B. Woolf. 2002. Transition events in butane simulations: similarities across models. *J. Chem. Phys.* 116:2586–2591.

25. Hummer, G. 2004. From transition paths to transition states and rate coefficients. *J. Chem. Phys.* 120:516–523.

26. Berezhkovskii, A. M., G. Hummer, and S. M. Bezrukov. 2006. Identity of distributions of direct uphill and downhill translocation times for particles traversing membrane channels. *Phys. Rev. Lett.* 97:020601.

27. Chaudhury, S., and D. E. Makarov. 2010. A harmonic transition state approximation for the duration of reactive events in complex molecular rearrangements. *J. Chem. Phys.* 133:034118.

28. Pollak, E. 2016. Transition path time distribution and the transition path free energy barrier. *Phys. Chem. Chem. Phys.* 18:28872–28882.

29. Pollak, E. 2017. Transition path time distribution, tunneling times, friction, and uncertainty. *Phys. Rev. Lett.* 118:070401.

30. Carlon, E., H. Orland, …, C. Vanderzande. 2018. Effect of memory and active forces on transition path time distributions. *J. Phys. Chem. B.* 122:11186–11194.

31. Laleman, M., E. Carlon, and H. Orland. 2017. Transition path time distributions. *J. Chem. Phys.* 147:214103.

32. Cossio, P., G. Hummer, and A. Szabo. 2018. Transition paths in single-molecule force spectroscopy. *J. Chem. Phys.* 148:123309.

33. Malinin, S. V., and V. Y. Chernyak. 2010. Transition times in the low-noise limit of stochastic dynamics. *J. Chem. Phys.* 132:014504.

34. Medina, E., R. Satija, and D. E. Makarov. 2018. Transition path times in non-markovian activated rate processes. *J. Phys. Chem. B.* 122:11400–11413.

35. Satija, R., A. M. Berezhkovskii, and D. E. Makarov. 2020. Broad distributions of transition-path times are fingerprints of multidimensionality of the underlying free energy landscapes. *Proc. Natl. Acad. Sci. USA.* 117:27116–27123.

36. Satija, R., A. Das, and D. E. Makarov. 2017. Transition path times reveal memory effects and anomalous diffusion in the dynamics of protein folding. *J. Chem. Phys.* 147:152707.

37. Hartich, D., and A. Godec. 2021. Emergent memory and kinetic hysteresis in strongly driven networks. *Phys. Rev. X* https://journals.aps.org/prx/accepted/ac07cK82Qa412f0474b779c61ec66ed07c08303b2.

38. Chung, H. S., and W. A. Eaton. 2018. Protein folding transition path times from single molecule FRET. *Curr. Opin. Struct. Biol.* 48:30–39.

39. Chung, H. S., J. M. Louis, and W. A. Eaton. 2009. Experimental determination of upper bound for transition path times in protein folding from single-molecule photon-by-photon trajectories. *Proc. Natl. Acad. Sci. USA.* 106:11837–11844.

40. Chung, H. S., K. McHale, …, W. A. Eaton. 2012. Single-molecule fluorescence experiments determine protein folding transition path times. *Science.* 335:981–984.

41. Neupane, K., D. A. Foster, …, M. T. Woodside. 2016. Direct observation of transition paths during the folding of proteins and nucleic acids. *Science.* 352:239–242.

42. Neupane, K., D. B. Ritchie, …, M. T. Woodside. 2012. Transition path times for nucleic Acid folding determined from energy-landscape analysis of single-molecule trajectories. *Phys. Rev. Lett.* 109:068102.

43. Yu, H., A. N. Gupta, …, M. T. Woodside. 2012. Energy landscape analysis of native folding of the prion protein yields the diffusion constant, transition path time, and rates. *Proc. Natl. Acad. Sci. USA.* 109:14452–14457.

44. Kim, J. Y., and H. S. Chung. 2020. Disordered proteins follow diverse transition paths as they fold and bind to a partner. *Science.* 368:1253–1257.

45. Redner, S. 2001. A Guide to First Passage Processes, First Edition. Cambridge University Press, Cambridge, UK.

46. Ma, P., R. Elber, and D. E. Makarov. 2020. Value of temporal information when analyzing reaction coordinates. *J. Chem. Theory Comput.* 16:6077–6090.

47. Szabo, A., K. Schulten, and Z. Schulten. 1980. First passage time approach to diffusion controlled reactions. *J. Chem. Phys.* 72:4350–4357.

48. Gardiner, C. W. 1983. Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences. Springer-Verlag, Berlin, Germany.

49. Chung, H. S., and I. V. Gopich. 2014. Fast single-molecule FRET spectroscopy: theory and experiment. *Phys. Chem. Chem. Phys.* 16:18644–18657.

50. Berezhkovskii, A. M., S. M. Bezrukov, and D. E. Makarov. 2021. Localized potential well vs binding site: mapping solute dynamics in a membrane channel onto one-dimensional description. *J. Chem. Phys.* 154:111101.

51. Li, X., and A. B. Kolomeisky. 2013. Mechanisms and topology determination of complex chemical and biological network systems from first-passage theoretical approach. *J. Chem. Phys.* 139:144106.

52. Thorneywork, A. L., J. Gladrow, …, U. F. Keyser. 2020. Direct detection of molecular intermediates from first-passage times. *Sci. Adv.* 6:eaaz4642.

53. Peters, B. 2017. Reaction Theory and Rare Events, First Edition. Elsevier, Amsterdam, the Netherlands.

54. Berezhkovskii, A. M., and D. E. Makarov. 2018. Communication: transition-path velocity as an experimental measure of barrier crossing dynamics. *J. Chem. Phys.* 148:201102.

55. Berezhkovskii, A. M., M. A. Pustovoit, and S. M. Bezrukov. 2002. Channel-facilitated membrane transport: transit probability and interaction with the channel. *J. Chem. Phys.* 116:9952–9956.

56. E, W. and E. Vanden-Eijnden. 2006. Towards a theory of transition paths. *J. Stat. Phys.* 123:503–523.

57. Lu, J., and E. Vanden-Eijnden. 2014. Exact dynamical coarse-graining without time-scale separation. *J. Chem. Phys.* 141:044109.

58. Vanden-Eijnden, E. 2006. Transition path theory. *In* Computer Simulations in Condensed Matter: From Materials to Chemical Biology. M. M. Ferrario, G. Ciccotti, and K. Binder, eds. Springer, pp. 453–493.

59. Peters, B. 2016. Reaction coordinates and mechanistic hypothesis tests. *Annu. Rev. Phys. Chem.* 67:669–690.

60. Berezhkovskii, A. M., and A. Szabo. 2013. Diffusion along the splitting/commitment probability reaction coordinate. *J. Phys. Chem. B.* 117:13115–13119.

61. Banushkina, P. V., and S. V. Krivov. 2015. Nonparametric variational optimization of reaction coordinates. *J. Chem. Phys.* 143:184108.

62. Banushkina, P. V., and S. V. Krivov. 2015. High-resolution free energy landscape analysis of protein folding. *Biochem. Soc. Trans.* 43:157–161.

63. Socci, N. D., J. N. Onuchic, and P. G. Wolynes. 1996. Diffusive dynamics of the reaction coordinate for protein folding funnels. *J. Chem. Phys.* 104:5860–5868.

64. Klimov, D., and D. Thirumalai. 1997. Viscosity Dependence of the Folding Rates of Proteins. *Phys. Rev. Lett.* 79:317–320.

65. Neupane, K., A. P. Manuel, and M. Woodside. 2016. Protein folding trajectories can be described quantitatively by one-dimensional diffusion over measured energy landscapes. *Nat. Phys.* 12:700–703.

66. Zhuravlev, P. I., M. Hinczewski, …, D. Thirumalai. 2016. Force-dependent switch in protein unfolding pathways and transition-state movements. *Proc. Natl. Acad. Sci. USA.* 113:E715–E724.

67. Avdoshenko, S. M., and D. E. Makarov. 2016. Reaction coordinates and pathways of mechanochemical transformations. *J. Phys. Chem. B.* 120:1537–1545.

68. Berezhkovskii, A. M., and D. E. Makarov. 2019. On the forward/backward symmetry of transition path time distributions in nonequilibrium systems. *J. Chem. Phys.* 151:065102.