A gentle introduction to the non-equilibrium physics of trajectories: Theory, algorithms, and biomolecular applications

Daniel M. Zuckerman and John D. Russo

COMPUTATIONAL PHYSICS

A gentle introduction to the non-equilibrium physics of trajectories: Theory, algorithms, and biomolecular applications

Daniel M. Zuckerman[a)] and John D. Russo
Department of Biomedical Engineering, Oregon Health & Science University, Portland, Oregon 97239

Despite the importance of non-equilibrium statistical mechanics in modern physics and related fields, the topic is often omitted from undergraduate and core-graduate curricula. Key aspects of non-equilibrium physics, however, can be understood with a minimum of formalism based on a rigorous trajectory picture. The fundamental object is the ensemble of trajectories, a set of independent time-evolving systems, which easily can be visualized or simulated (e.g., for protein folding) and which can be analyzed rigorously in analogy to an ensemble of static system configurations. The trajectory picture provides a straightforward basis for understanding first-passage times, "mechanisms" in complex systems, and fundamental constraints on the apparent reversibility of complex processes. Trajectories make concrete the physics underlying the diffusion and Fokker–Planck partial differential equations. Last but not least, trajectory ensembles underpin some of the most important algorithms that have provided significant advances in biomolecular studies of protein conformational and binding processes. © *2021 Published under an exclusive license by American Association of Physics Teachers.*
https://doi.org/10.1119/10.0005603

I. INTRODUCTION

Most of the phenomena we encounter in daily life, from weather to cooking to biology, are fundamentally out of equilibrium and require physics typically not touched on in the undergraduate or even graduate physics curricula. Many physics students are alarmed at the complexity and abstraction of thermodynamics and "sadistical mechanics," and understandably would not seek out instruction in non-equilibrium statistical physics. Yet there is a surprising range of fundamental non-equilibrium material that can be made accessible in a straightforward way using *trajectories*, which are essentially movies of systems executing their natural dynamics. The trajectory picture first and foremost is fundamental;[1–3] for example, dynamics generate equilibrium, but not the other way around.[4] It can also lead, with a minimum of mathematics, to understanding key non-equilibrium phenomena (relaxation and steady states) and similarly to extremely powerful cutting-edge simulation methods (path sampling). Students deserve a taste of this material.

Why are trajectories fundamental? A trajectory is simply the sequence of phase-space points through which a system passes, recorded perhaps as a "movie" listing all atomic positions and velocities at evenly spaced time points—the "frames" of the movie. Such movies are fundamental because, as we learned from Newton, nature creates forces that lead to dynamics,[5] i.e., to trajectories. We may attempt to describe the dynamics in various average ways—e.g., using equilibrium ideas—but the trajectories are the basis of everything. Theories, such as equilibrium statistical mechanics, generally build in assumptions, if not approximations. In fact, the most fundamental definition of equilibrium itself derives from dynamics, via detailed balance,[1,4,6] whereby there must be an equal-and-opposite balance of flows between any two microstates.

Dynamical descriptions generally have more information in them than average or equilibrium theories.[4,7,8] As a simple example, perhaps you know that someone sleeps eight hours a day. However, that average hides the time at which sleep occurs as well as whether it includes an afternoon nap. In the case of diffusion, we know that particles observed in a localized region will tend to spread out over time. However, if we only observe the spatial density, we do not know which particles went where. Trajectories, which track particles over time, inherently capture this information.

A trajectory ensemble description, as described below, provides *the* key observables for transition processes: rate and mechanism. In a biomolecular context, these are essentially everything we want to know. Consider protein folding. We want to know how fast proteins fold and how folding rates change under specific mutations.[9,10] We also want to know the mechanism of folding: the conformations that are visited during the process which in turn can illuminate chemical-structure causes of rate changes due to mutation.[10,11] Other conformational processes in biomolecules arguably are of even greater interest, such as binding[12] and allostery,[10,13] due to their implications for drug design; here again, rate and mechanism are of utmost importance.[14,15]

This article will explain the theory of trajectory ensembles, starting with simple diffusion and moving to systems with complex energy landscapes. We will explore essential aspects of non-equilibrium statistical mechanics, focusing on timescale quantification via the mean first-passage time. The understanding of non-equilibrium trajectory ensembles leads directly to the "super parallel" weighted ensemble simulation methodology, widely used in computational biology,[16] which is explored in a one-dimensional pedagogical example. A number of exercises are given along with clearly demarcated more advanced material.

The statistical mechanics of trajectories has been addressed pedagogically, in different ways, in prior work. Clear, basic-level descriptions can be found in some textbooks[3,4] and path-sampling papers in the molecular-oriented literature.[17–19] Astumian and co-workers highlighted the importance of trajectories and their probabilistic description in multiple contexts[20,21] and provided important semi-microscopic, discrete-state descriptions of molecular motors,[22,23] building on the seminal work of Hill.[24,25] Ghosh and co-workers employed trajectory concepts in presenting Jaynes's maximum-caliber approach to inferring kinetics;[26] note the related work by Pressé co-workers[27] and Ghosh co-workers.[28] Swendsen's discussion of irreversibility is also of interest,[29] as is the classic treatment by Chandrasekhar.[30] The present discussion attempts to provide a more elementary discussion of trajectory physics, with a focus on computational applications not found in most prior work. Perhaps unexpectedly, the path sampling algorithms derivable from the present description are very much at the leading edge of molecular computation.[31]

II. BASICS: DYNAMICS AND TRAJECTORIES

In this section, we introduce the building blocks of our analysis, starting from one-dimensional Newtonian motion. We add fundamental stochastic elements and then develop the trajectory picture with an associated numerical recipe.

A. Stochastic dynamics

The starting point for our quantitative trajectory description is the simplest form of stochastic dynamics, often called Brownian dynamics, which we will justify starting from Newton's second law. Brownian dynamics are also known by more intimidating terminology, as overdamped Langevin dynamics, but their essence is simple to understand. As a familiar reference, we first write the one-dimensional (1D) law of classical motion,

$$m\frac{dx^2}{dt^2} = f, \tag{1}$$

where m is mass, x is position, and $f = -dU/dx$ is the force, with $U(x)$ the potential energy. Advancing one step in complexity, the 1D Langevin equation models motion in a viscous (frictional) medium by adding a damping force that always opposes the direction of motion (velocity), as well as a random force f_{rand} from collisions,[4,6] yielding

$$m\frac{d^2x}{dt^2} = f - \gamma m \frac{dx}{dt} + f_{\text{rand}}, \tag{2}$$

where $\gamma > 0$ is the friction constant, effectively a collision frequency, as can be seen by dimensional analysis. Details of

the random force will be given later. Both forces are needed; otherwise, damping would eliminate all motion.

In the *overdamped* limit, inertia is ignored. This is akin to motion in a beaker filled with thick oil: there is minimal tendency for an object to continue in any given direction in the absence of force; with a force such as gravity, terminal (constant) velocity is reached quickly; i.e., no further acceleration occurs despite the force. At microscopic scales, however, there continues to be random thermal motion due to molecular collisions. Setting the inertial term $m\,d^2x/dt^2$ to zero in Eq. (2) and re-arranging terms, the overdamped Langevin equation is[4,6]

$$\frac{dx}{dt} = \frac{1}{m\gamma}(f + f_{\text{rand}}). \tag{3}$$

This simplified equation of motion may look unusual to those unfamiliar with it, but studying its application in a numerical context will make its physical basis and relation to diffusion more clear.

B. Time-discretized overdamped dynamics and computation

We will make most use of a discrete-time picture (fixed time steps) which not only greatly simplifies the mathematics but also translates directly into simple computer implementation. If we discretize the dynamics of Eq. (3) by writing the velocity as $\Delta x/\Delta t$ and multiplying through by Δt, we arrive at a very useful equation,

$$\Delta x = \frac{\Delta t}{m\gamma}(f + f_{\text{rand}}) = \Delta x_{\text{det}} + \Delta x_{\text{rand}}, \tag{4}$$

where $\Delta x_{\text{det}} = f\Delta t/m\gamma$ is the deterministic component of the spatial step due to an external force (e.g., molecular, gravitational, or electrostatic) and Δx_{rand} is the random part due to thermal molecular collisions. At finite temperature, microscopic motion must not cease; hence, in the Langevin picture, thermal fluctuations must balance "dissipation" due to damping of the γ term.[6,32] To accomplish this, Δx_{rand} is typically assumed to follow a zero-mean Gaussian distribution which must have its variance given by[4]

$$\sigma^2 = 2k_B T \Delta t/m\gamma. \tag{5}$$

The Gaussian assumption is justified on the basis of the central limit theorem[4] because a molecule in aqueous solution can experience upwards of 10^{13} collisions per second,[6,30] and hence, a large number of collisions occur in any $\Delta t > 1$ ns. The high collision frequency also justifies the implicit assumption here that sequential Δx_{rand} values are independent, i.e., not time-correlated.

With the distribution of Δx_{rand} specified, the discrete overdamped dynamics Eq. (4) is simultaneously a prescription for computer simulation of trajectories *and* directly implies a probabilistic description of trajectories. Let us start with computer simulation, which is simpler by far. Defining $x_j = x(t = j\Delta t)$, Eq. (4) is essentially a recipe for calculating the next position $x_{j+1} \equiv x_j + \Delta x$ in a time-sequence, given x_j. For a sufficiently small time step Δt, the force f will be approximately constant over the whole time interval, so we take

$$\Delta x_{\text{det}} = f(x_j)\Delta t / m\gamma, \tag{6}$$

and Δx_{rand} is chosen from a Gaussian (normal) distribution of variance σ^2 from Eq. (5). Looping over this process yields a discrete-time *trajectory*,

$$\text{traj} = \{x_0, x_1, x_2, \ldots\}, \tag{7}$$

which is just a list of positions at intervals of Δt. We can easily recast trajectory elements in terms of spatial increments,

$$
\begin{aligned}
x_0 &= x_0 \quad \text{(arbitrary)}, \\
x_1 &= x_0 + \Delta x_1, \\
x_2 &= x_0 + \Delta x_1 + \Delta x_2 = x_1 + \Delta x_2, \\
&\cdots,
\end{aligned} \tag{8}
$$

which is useful for understanding simulation algorithms such as Eq. (4).

Trajectories of *simple diffusion* can be generated from Eq. (4) by setting $f = 0$ (hence $\Delta x_{\text{det}} = 0$). The recipe given above simplifies to choosing a Gaussian random step at each time point, i.e.,

$$\Delta x = \Delta x_{\text{rand}} \quad \text{(simple diffusion)}, \tag{9}$$

as we would expect. Schematic examples of these simplest stochastic trajectories are shown in Fig. 1. There is no directionality in simple diffusion, but only a statistical tendency to diffuse away from the starting point, as will be quantified below.

III. SIMPLE DIFFUSION IN THE TRAJECTORY PICTURE

The basics of diffusion, such as Fick's law and the diffusion equation, are well known, so diffusion theory is a perfect context for introducing the trajectory formulation. Students may find that following the behavior of individual particles is a more concrete exercise than visualizing probability distributions. In this section, we show that the trajectory approach yields the familiar average description of
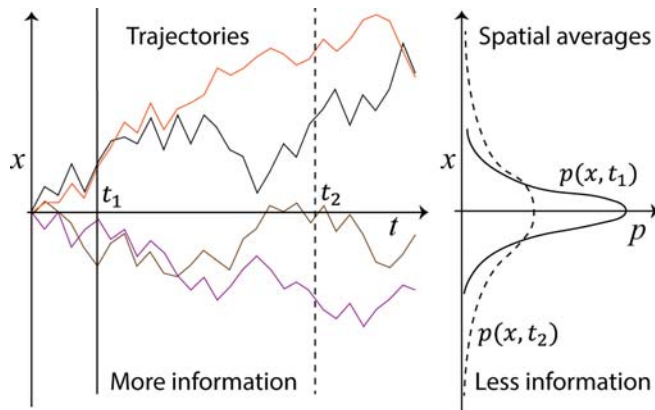

Fig. 1. Simple diffusion, two ways. At left are schematic time-discretized trajectories illustrating one-dimensional diffusion started from the initial point $x_0 = 0$. Averaging over the positions of many trajectories at specific time points t_1 and t_2 yields the distributions shown at right, with $p(x, t_i) = p(x_i | x_0)$. Averaging can aid interpretation but it also removes information, namely, the connectivity among the trajectories' sequences of points.

simple diffusion in a force-free (constant-energy) landscape. In the bigger picture, we get an explicit sense of physical details of trajectories which are averaged (integrated) out to yield the distribution picture.

A. Probabilistic picture for trajectories

We start by analyzing diffusive trajectories based on random steps where the force f has been set to zero. The procedure (Eq. (9)) of repeatedly choosing a Gaussian step with variance from Eq. (5) implicitly but precisely defines a probability distribution for an entire trajectory (Eq. (7)), which will prove of fundamental importance. First, by construction, the probability of a *single* step Δx is given by

$$p_1(\Delta x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\Delta x^2 / 2\sigma^2}. \tag{10}$$

This is the meaning of choosing a Gaussian step. Note that Eq. (10) depends only on the magnitude and not on the starting point of the specific step, which is a characteristic of simple diffusion because no forces are present.

For the full trajectory, we use the simple rule that the probability of a sequence of independent steps is simply the product of the individual step probabilities: think of a sequence of fair coin flips characterized by 1/2 to the appropriate power. Hence, for an N-step trajectory defined by Eq. (8) starting from x_0, we have

$$p(\text{traj}) = p_1(\Delta x_1) \cdot p_1(\Delta x_2), \cdots, p_1(\Delta x_N) \tag{11}$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}} \right)^N \prod_{j=1}^{N} e^{-\Delta x_j^2 / 2\sigma^2}. \tag{12}$$

A multi-dimensional distribution such as Eq. (12) may not be trivial to understand for those not used to thinking in high dimensions. First, why is it a multi-dimensional distribution? Well, it describes the distribution of a *set* of points, the trajectory $\{x_0, x_1, x_2, \ldots, x_N\}$. Note that we immediately obtain the Δx values needed for Eq. (12) from the x values using Eq. (8): $\Delta x_1 = x_1 - x_0$ and so on. So if you are given a set of (trajectory) x values, you can convert them into Δx values and plug them into Eq. (12) to get the probability of that trajectory. You can do this for *any* set of x values, even ridiculously unphysical values with gigantic jumps, but of course the probability will be tiny for unphysical trajectories. For completeness, strictly speaking, Eq. (12) is a probability *density*[4] and absolute probabilities are only obtained by integrating over a finite region.

The distribution of trajectories encodes all the information we could possibly want about diffusive behavior, although some math is needed to get it. Alternatively, as a proxy for the distribution, multiple trajectories could be simulated to quantify their average behavior. In the case of simple diffusion, however, the math of the trajectory distribution is both tractable and illuminating.

As a fascinating technical aside, note that the product of exponentials in Eq. (12) can be re-written as the exponential of a sum $(-\sum_j \Delta x_j^2 / 2\sigma^2)$, which makes the probability look somewhat like a Boltzmann factor. Indeed, consulting the definition of σ^2 in Eq. (5), we find it is proportional to $k_B T$. Of course, the argument of our exponential is not a true energy, but can be considered an effective path energy, known as the "action."[4,33]

(In the non-diffusive case, $\Delta x_{\mathrm{det}} \neq 0$ leads to an additional term in the exponent and the action; see below.) The action formulation, and the consideration of all possible paths, is the heart of the path-integral formulation of quantum mechanics.[34] The path-probability formulation is truly fundamental to physics.

B. Deriving the spatial distribution from trajectories

A key observable of interest is the distribution of x values at a fixed but arbitrary time point (Fig. 1). To build up to this, we will carefully derive the equation for the conditional probability distribution $p(x_2|x_0)$ of $x_2 = x(2\Delta t)$ values, i.e., the distribution for a fixed starting point x_0. The critical idea is that we can obtain the probability of any given x_2 value by summing (i.e., integrating) over all possible two-step trajectories that reach the particular value starting from x_0. Because both forward and backward motion are possible for the intermediate step, we must consider all possible x_1 values. Mathematically, this amounts to

$$p(x_2|x_0) = \int_{-\infty}^{\infty} dx_1 \; p_1(\Delta x_1) \; p_1(\Delta x_2),$$
$$= \int_{-\infty}^{\infty} d\Delta x_1 \; p_1(\Delta x_1) \cdot p_1(x_2 - (x_0 + \Delta x_1)), \quad (13)$$

where we have used Eq. (11) to start and then Eq. (8) to substitute for Δx_2.

We can evaluate the integral in Eq. (13) exactly. Plugging in the expression for p_1 from Eq. (10) and setting $y = \Delta x_1$, we have

$$p(x_2|x_0) = \frac{1}{2\pi\sigma^2} \int_{-\infty}^{\infty} dy \, e^{-y^2/2\sigma^2} e^{-(x_2-x_0-y)^2/2\sigma^2},$$
$$= \frac{1}{2\pi\sigma^2} e^{-(x_2-x_0)^2/4\sigma^2} \int_{-\infty}^{\infty} dy \, e^{-[y-(x_2-x_0)/2]^2/\sigma^2},$$
$$= \frac{1}{\sqrt{2\pi}(\sqrt{2}\sigma)} e^{-(x_2-x_0)^2/2\cdot 2\sigma^2}, \quad (14)$$

where the second line is derived by completing the square in the exponent and the third line is derived by performing the Gaussian integral shown.

The result Eq. (14) for the distribution of positions after $2\Delta t$ is very informative, especially by comparison to the single-step distribution Eq. (10). The distribution of possible outcomes is still a Gaussian of mean x_0, but the variance is doubled; equivalently, the standard deviation has increased by a factor of $\sqrt{2}$. See Fig. 1. It is important that *we derived this distribution by averaging (integrating) over the ensemble of two-step trajectories*. As promised, the information was indeed encoded in the original trajectory distribution Eq. (12).

From here, it is not hard to generalize to an arbitrary number of steps by repeating the integration process. The result is that the distribution of $x_n = n\Delta t$ values is also a Gaussian with mean x_0, but with variance $n\sigma^2$,

$$p(x_n|x_0) = \frac{1}{\sqrt{2n\pi}\sigma} e^{-(x_n-x_0)^2/2n\sigma^2}. \quad (15)$$

Equation (15) embodies the usual description of diffusion, as we will see in two ways, but it also contains *less* information than our initial trajectory description.

C. Confirming the probabilistic description of diffusion

Have we really recapitulated the usual description of diffusion? As a first check, we immediately recover the expected linear time dependence of the mean-squared displacement[4] based on Eq. (15). This is because the variance *is* the mean-squared displacement or deviation (MSD), and the number of steps $n = t/\Delta t$ is simply proportional to time. By the definition of a Gaussian distribution, the variance implicit in Eq. (15) is $n\sigma^2$, and we therefore have

$$\mathrm{MSD} \equiv (x_n - x_0)^2 = \int_{-\infty}^{\infty} dx_n (x_n - x_0)^2 p(x_n|x_0),$$
$$= n\sigma^2 = (t/\Delta t)\sigma^2. \quad (16)$$

If we define the diffusion constant via $\mathrm{MSD} = 2Dt$ (in one dimension), then from Eqs. (5) and (16), we derive $D = k_B T/m\gamma$, which is a well-known result.[4]

Second, by renaming the variable $x_n \to x = x(t)$ in Eq. (15) and noting that time $t = n\Delta t$, we can see that Eq. (15) describes the time-evolving probability distribution of positions $p(x, t)$, which is the well-known solution to the 1D diffusion equation,

$$\frac{\partial p}{\partial t} = D \frac{\partial^2 p}{\partial x^2}. \quad (17)$$

This can be verified by direct differentiation, but see Sec. VIII for a hint. The agreement with the *continuous*-time diffusion equation implies that time discretization is irrelevant, but be warned that this is not always the case, as discussed in Sec. VIII.

D. What is missing from the standard description of diffusion?

Because the distribution of positions (Eq. (15)) is known for any time and provides the exact solution to the diffusion equation, it may seem there is nothing more to know. However, the key observables—the timescale (or rate) and mechanism of any particular process—either are not available at all from the positional distribution or not easily available.[35,36]

These shortcomings stem from the information missing from the spatial distribution. Even if we know the spatial distribution at two times, *we still do not know how any given diffusing particle went from one place to another*. That is, although we know the fraction of particles that will be located between any x and $x + dx$, we do not know which came from left or right and exactly from where. This information is encoded in the dynamics and recorded in the distribution of trajectories Eq. (11), which is essentially a distribution of paths taken through position space. It is fair to say, therefore, that the trajectory distribution *is* the mechanism, assuming that all trajectories considered conform to criteria of interest (e.g., starting at $x = 0$ and perhaps reaching a value $x > a$ after n steps.)

E. Beyond simple diffusion in one dimension

Before we move beyond a single dimension, a useful reference for developing intuition is the generalization of the single-step distribution Eq. (10) when a force is present. Reframing the procedure Eq. (4) probabilistically, the

distribution for overdamped dynamics of a 1D particle in the presence of a spatially varying potential $U(x)$ is a different Gaussian,

$$p_1(\Delta x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(\Delta x - \Delta x_{\text{det}})^2 / 2\sigma^2}. \tag{18}$$

In contrast to the simple-diffusion case Eq. (10), the distribution of possibilities is centered on the deterministic (force-driven or "drift") step Δx_{det} defined by Eq. (6). That is, the particle tends to move in the direction of the force, albeit stochastically.

Equation (18) should guide your intuition for single-step motion of a stochastic system: there is a distribution of possibilities centered on the deterministic step. The deterministic component generally could depend on inertia and/or force although in the overdamped case there is no inertia. Note that Δx_{det} in Eq. (18) implicitly depends on the starting position for the step: see Eq. (6). Below, we will make the position dependence more explicit.

IV. THE NON-EQUILIBRIUM STEADY STATE (NESS) AND THE HILL RELATION FOR RATES

Probably the most important observable in a dynamical process, at least in biomolecular studies, is the rate for a process. As we will see, the rate is closely related to a specific non-equilibrium steady state, which is essential to understand but also quite accessible.

Physicists often quantify a rate via the *mean first-passage time* (MFPT).[32,35–37] The first-passage time is simply the time required for a process from start to finish, e.g., the time required for a protein to fold, starting from when it is initialized in an unfolded state. In Fig. 2, this is the time from initiation in "source" state A to absorption in "sink" state B. (We are thus employing source-sink boundary conditions.) Chemists and biochemists quantify kinetics via the "rate constant" for a conformational process like protein folding, which has units of s^{-1} and can be defined as the reciprocal MFPT, although chemists prefer a definition based on directly measurable "relaxation times."[4,38,39] Our discussion will focus solely on the MFPT for simplicity.

The MFPT can be directly obtained from a steady-state trajectory ensemble, so we will start by defining a source-sink non-equilibrium steady state (NESS) as sketched in
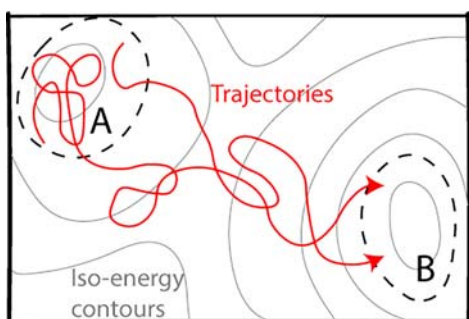

Fig. 2. Source-sink non-equilibrium steady state. Trajectories (red curves, color online) are initiated in state A and terminated upon reaching state B, with states bounded by dashed contours. Importantly, trajectories that reach B are then re-initiated from A. Such a system will reach a non-equilibrium steady state after a transient "relaxation" period. Gray solid lines show iso-energy contours of a schematic landscape.

Fig. 2. Independent trajectories are initiated in the source macrostate A (e.g., the set of unfolded protein configurations) according to a specified distribution p_0 (e.g., a single configuration or the equilibrium distribution over A). A second, non-overlapping sink macrostate B is an absorbing state in that trajectories reaching B are terminated, although in our source-sink setup they are immediately restarted in A selected according to the p_0 distribution. If this process is allowed to run for long enough so that each trajectory has reached B and been recycled back to A many times, the system will reach a non-equilibrium steady state. Without a sink state or recycling, the system will relax to equilibrium, which is also a steady state. (See Sec. VIII to explore the difference between equilibrium and other steady states.)

The MPFT is derivable from a NESS trajectory ensemble in a direct way, which will seem obvious once we are aware of it. The derivation is simple, but requires some thought. Imagine we have a large number $M \gg 1$ of independent systems that together make up the source-sink NESS (Fig. 2). By construction, the NESS is characterized by a *constant flow* of trajectories into B. We can simply count the number of trajectories arriving during some time interval τ and call this count m. Thus, a fraction m/M of the total probability arrives in time τ.

To continue our derivation, we can estimate this same fraction of trajectories arriving based solely on the meaning of the MFPT. By definition, the average amount of time a trajectory requires to traverse from A to B *is* the MFPT, so the (average) probability for any given trajectory to arrive during an interval τ is precisely τ/MFPT, which in turn is the same as the fraction expected to arrive in τ. In other words, $m/M = \tau$/MFPT, and we have derived the *Hill relation*,[4,25]

$$\frac{1}{\text{MFPT}} = \frac{m/M}{\tau} = \text{Flux}(A \rightarrow B \,|\text{NESS}), \tag{19}$$

where the flux is the probability arriving to B per unit time in the NESS. Equation (19) is an exact relation with no hidden assumptions, although not surprisingly the MFPT is particular to the initiating distribution p_0 of the particular NESS in which the flux is measured. That is, the MFPT depends on where in state A trajectories are initiated.

The Hill relation hints at a remarkable possibility: estimation of a long timescale (the MFPT) based on an arbitrary short period of observation (τ). If this could be done routinely, it would represent a major accomplishment in computational physics.[31] In Sec. VI, we describe a simple algorithm that can leverage the Hill relation for practical computations in many systems. We also explain the challenges involved.

V. MORE ADVANCED DISCUSSION OF ENSEMBLES AND THERMODYNAMIC STATES

This section describes additional fundamental concepts in non-equilibrium physics, but the discussion necessarily becomes more technical. Readers can skip this section without compromising their ability to understand subsequent material.

A. Notation and nomenclature for multi-dimensional systems

We will frame our discussion a bit more generally in the context of multi-dimensional systems. Fortunately, this

extension adds only incremental conceptual and mathematical complexity. To keep notation as simple as possible, we will use $\vec{x}$ to represent *all* microscopic coordinates—the phase-space vector consisting of all positions and velocities of all atoms in our classical representation. In some cases, such as overdamped dynamics (Eq. (3)), velocities may be excluded from the description, but the $\vec{x}$ notation remains valid. A *macrostate* is defined to be a *set* of $\vec{x}$ points. These macrostates are not to be confused with thermodynamic states such as equilibrium at some constant temperature or a non-equilibrium steady state.

As with our discussion of simple diffusion above, we will strictly use discrete time: $t = 0, \Delta t, 2\Delta t, \dots$. Discrete time greatly simplifies our description of trajectory probabilities without sacrificing any physical insights. In a trivial extension of Eq. (7), we therefore write a trajectory as

$$\text{traj} = \{\vec{x}_0, \vec{x}_1, \vec{x}_2, \dots\}, \tag{20}$$

where $\vec{x}_j$ is the phase point at time $t = j\Delta t$.

B. The initialized trajectory ensemble in multiple dimensions

The probabilistic description of a multi-dimensional trajectory follows logic almost identical to the 1D diffusion formulation of Eq. (11), except for two details. First, we now include the possibility that the initial system phase point $\vec{x}_0$ itself is chosen from some distribution p_0. Second, in contrast to simple diffusion, where the distribution Eq. (10) of outcomes p_1 for any single step depends only on the magnitude of Δx, more generally the outcome depends on the starting point of the step because the force may vary in space. We therefore adopt a notation which makes this explicit: $p_1(\vec{x}_{j-1} \rightarrow \vec{x}_j) = p_1(\vec{x}_j | \vec{x}_{j-1})$ is the (conditional) probability distribution for $\vec{x}_j$ values, given the prior position $\vec{x}_{j-1}$. The probability of a full trajectory is then the product of the initial distribution and the sequence of stepwise distributions,

$$\begin{aligned} p(\text{traj}) &= p_0(\vec{x}_0) \cdot p_1(\vec{x}_0 \rightarrow \vec{x}_1) \\ &\quad \cdot p_1(\vec{x}_1 \rightarrow \vec{x}_2) \cdots p_1(\vec{x}_{N-1} \rightarrow \vec{x}_N). \end{aligned} \tag{21}$$

The mathematical form of p_1 must now account for multi-dimensional aspects of the system, as well as any forces or inertia if present: see Eq. (18) and the discussion following it. Although specifying p_1 in generality is beyond the scope of our discussion, we should note that the form of Eq. (21) indicates we have assumed Markovian behavior: the distribution of outcomes p_1 at any time depends only on the immediately preceding time point.

We must be careful to specify our system without ambiguity. A given physical system, such as a particular protein molecule in a specified solvent at known temperature and pressure, can be considered in a variety of thermodynamic states, such as equilibrium or a non-equilibrium state. The system and the thermodynamic conditions both must be specified. Conveniently, the two aspects are described by different parts of the trajectory distribution equation Eq. (21): the intrinsic physical properties such as forces and dynamics are encoded in the single-step p_1 factors, while the thermodynamic state or ensemble is determined by the initial distribution p_0 along with boundary conditions. Some boundary conditions will be discussed below.

The distribution Eq. (21) describes the *initialized trajectory ensemble*, the set of trajectories originating from a specified phase-point distribution p_0 at time $t = 0$. For instance, p_0 could represent a single unfolded protein configuration (making p_0 a Dirac delta function), a set of unfolded configurations, a solid in a metastable state, or the set of initial positions of multiple dye molecules in a solvent. Figure 1 illustrates the one-dimensional trajectory ensemble initialized from $p_0(x) = \delta(x)$.

As with simple diffusion, we can revert to the simpler, averaged description of a spatial distribution that evolves in time due to the dynamics. That is, in principle, we can calculate the distribution of phase points at time $t = N\Delta t$ starting from p_0, denoted $p(\vec{x}_N | p_0)$. When forces are present, the diffusion (partial differential) equation Eq. (17) must be generalized to account for the tendency of a particle to move a certain direction, leading to the Fokker–Planck/Smoluchowski picture.[32,35,36] Appendix A describes the corresponding Smoluchowski equation that governs overdamped motion with forces. However, as with simple diffusion, the spatial distribution represents an average over the information-richer trajectories.

C. Connection to relaxation, state populations, and thermodynamics

It is important to note that, in general, an initialized system will "relax" away from its initial distribution p_0. For systems of interest here, the system's phase-point distribution $p(\vec{x}_N | p_0)$ will tend to relax toward a *steady state* dependent only on the boundary conditions. In a constant-temperature system with no particle exchange, for example, the distribution will approach equilibrium as embodied in the Boltzmann factor: $\lim_{N \rightarrow \infty} p(\vec{x}_N | p_0) \propto \exp(-H(\vec{x})/k_B T)$, where $H(\vec{x})$ is the total energy of point $\vec{x}$, k_B is Boltzmann's constant, and T is the absolute temperature. In general, whether equilibrium or not, the steady state that is reached typically will be independent of p_0 after sufficient time for a "well-behaved" system. In Sec. IV, we explored *non-equilibrium* steady states critical to understanding conformational transitions.

Whether the system is in the relaxation or steady regime, the phase-point distribution $p(\vec{x})$, obtainable from the trajectory picture, directly connects to observable and thermodynamic properties. Most simply, the time-dependent macrostate population can be obtained as the integral of $p(\vec{x})$ over a region of phase space: this is the fraction of probability in the state which evolves in time with p. At a system-wide level, both the entropy and average energy can be obtained from well-known integrals over p.[4,7] These also evolve with time, directly leading to the entropy production picture. Further detail on these topics is beyond the scope of the present discussion, and interested readers should consult suitable Refs. 7 and 27.

D. The ensemble of trajectories and the meaning of equilibrium

When we speak of an "ensemble" of trajectories, the word has the same meaning as in ordinary statistical mechanics,[4,6] namely, a set of *fully independent* trajectories generated under the conditions of interest (see below). That is, each member of the ensemble is a replica of the same physical system but is initiated from a phase point that typically will differ from others in the ensemble.

An ensemble in principle can be generated according to any process and under any conditions we care to specify. The dynamics of these trajectories could be governed by simple constant-temperature diffusion or there could be a temperature gradient, forces, or both. Trajectories could additionally be subject to certain boundary conditions: for example, they might be assumed to reflect off some boundary in phase space or be absorbed on reaching a certain "target" region as we considered in Sec. IV. The full set of rules governing a set of trajectories defines the ensemble by determining the weights of each trajectory as in Eq. (21), and we are often interested in ensemble or average behavior because this is what is usually observed experimentally although single-molecule studies are by now a well-established and important field of study.[21,40,41]

It is critical to appreciate that an individual trajectory generally cannot be considered to be of equilibrium or non-equilibrium character in an intrinsic sense. (A possible exception is an extremely long trajectory which itself fully embodies all defining criteria of the ensemble.[4]) Generally, it is the *distribution* of trajectories that determines whether a system is in equilibrium and, if not, what ensemble it represents. Two finite-length trajectories that have the same weight in the equilibrium ensemble might have different weights in a non-equilibrium ensemble. The trajectory distribution will be determined by the initial phase-point distribution p_0 in conjunction with the imposed boundary and thermodynamic conditions such as temperature.

Let us consider equilibrium in the trajectory ensemble picture. For simplicity, we will assume that our initial phase point distribution is already Boltzmann-distributed: $p_0(\vec{x}_0) \propto \exp\left(-H(\vec{x}_0)/k_B T\right)$. As trajectories evolve in time from their initial points, the system will remain in equilibrium if the thermodynamic and boundary conditions remain the same. Thus, dynamics underlie equilibrium. We can say dynamics *define* equilibrium through detailed balance: if we count transitions occurring between small volumes around phase points $\vec{x}_i$ and $\vec{x}_j$ over any interval of time, the counts $i \to j$ and $j \to i$ will be identical within noise; the same is true for any size volumes in equilibrium.[1,4] This detailed balance property not only keeps the distribution stationary in time, but it means there are no net flows anywhere in phase space. Detailed balance further implies there is no net flow along any trajectory-like path—i.e., the forward and exactly time-reversed trajectories will occur an equal number of times.[23,42]

Note that our discussion here applies to thermal (constant-temperature) equilibrium for systems whose full configurations or phase points may include real-space coordinates and/or chemical degrees of freedom. That is, the trajectory picture of equilibrium applies for conformational processes in molecules, such as isomerization or folding; for simple diffusion or diffusion with possibly space-varying "drift" forces; for molecular binding, which may include both translational and conformational processes; for chemical processes involving electronic degrees of freedom such as bond formation and breakage; and for any combination of these whether modeled in full detail or approximately, so long as there is no implicit addition or removal of energy or particles. The trajectory picture does not apply for mechanical equilibrium, the balance of forces.

VI. POWERFUL SIMULATION METHODOLOGY BASED ON TRAJECTORY ENSEMBLES

A. Goals and challenges of computation

To consider computational strategies, we should first understand the goals of computation. As we do so, keep in mind a concrete process like protein folding or another spontaneous transition from a metastable state to a more stable one, such as a conformational change in a protein, a change in crystal lattice form, or a re-arrangement of a molecular cluster. For any of these transitions, we might be interested in the following:

(i) The "kinetics"—the MFPT or some other measure of rate for the transition.
(ii) The "mechanism" or pathways of the process—the sequence(s) of states exhibited during the transition.
(iii) The "relaxation" process—the timescales and mechanisms of describing the transient way the system "settles in" to a steady state.

We will first consider a simple, though typically impractical, way to calculate any or all of the above. As sketched in Fig. 3, the naive "brute force" implementation would simply be to initiate a large number of trajectories using an initial distribution of interest p_0 and wait until all trajectories have made the transition of interest. From this set of trajectories, we could (i) average their durations to obtain the MFPT or (ii) analyze the states occurring during transitions to quantify the mechanism.[43] For (iii) relaxation, we could wait still longer until the spatial/configurational distribution becomes stationary (using "recycling" if studying a constant-temperature NESS) and quantify the relaxation time as well as the mechanism, perhaps via probability shifts that occur. However, the strategy of waiting for multiple spontaneous transitions will only work for the simplest systems, such as low-dimensional toy models. (See Sec. VIII.)

In general, the brute force approach will *not* be practical for complex systems, and if a system is complicated and directly pertinent to real-world problems, it is likely to be too expensive to permit thorough brute-force simulation. We can quantify the challenges with a back-of-the-envelope calculation. For the system of interest, say you can afford a total of M simulations of duration $t_{\max}$. This means, roughly, that
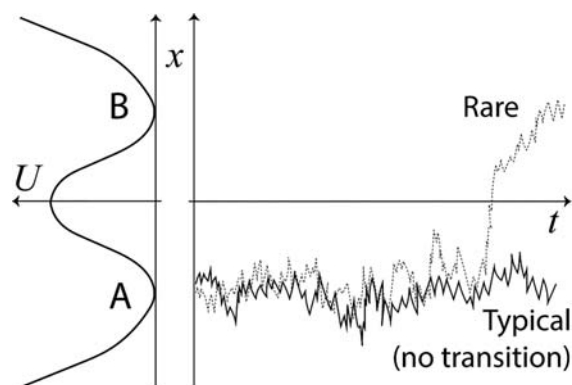

Fig. 3. The challenge of rare-event sampling in computation. Trajectories are initiated in state A, but in challenging systems most will remain in state A (solid trajectory). Transitions (dotted line) may be extremely unlikely or effectively unobservable in realistic, high-complexity systems such as protein conformational changes. Hence, typical "brute force" simulations can be both wasteful and expensive.

you can determine the distribution of phase points at any time $t < t_{max}$, denoted $p(\vec{x}, t)$, to a precision of $1/M$; typically, you will not have knowledge of behavior beyond t_{max}. As a point of reference in biomolecules, current hardware limits t_{max} to 1–10 μs in most systems (and to ms for small systems with extraordinary resources[9]), whereas most biological phenomena occur on a timescale of at least 100 μs and more typically on ms–s scales.

B. Efficient simulation via the weighted ensemble approach

Fortunately, there are now methods[31,44–47] that can sidestep the $1/M$ limitation just described, and we will focus on the most straightforward of these, known as the *weighted ensemble* (WE) strategy.[16,48,49] WE is a multi-trajectory "splitting method" based on a proposal credited to von Neumann[50] that can provide information on relaxation and steady-state behavior. WE can provide this information using less *overall* computing than naive simulation, i.e., the product Mt_{max} is smaller. It achieves this by re-allocating computing effort (trajectories) away from easy-to-sample regions of phase space toward rarer regions. WE is also an unbiased method: on average, it exactly recapitulates trajectory ensemble behavior and hence the time-evolution of the spatial distribution $p(\vec{x}, t)$;[49] the latter property reflects consistency with the Fokker–Planck equation,[35,36] which is briefly described in Appendix A.

WE simulation follows a fairly simple procedure, schematized in Fig. 4, which promotes the presence of trajectories in relatively rare regions of an energy landscape. In a basic implementation,[48] phase space is divided into non-overlapping bins of the user's construction, and a target number of trajectories per bin is set—say, 2, for concreteness.

The bins should finely subdivide difficult-to-sample regions such as energy barriers to enable "statistical ratcheting" up hills if trajectories are examined frequently enough. That is, because short trajectories always have some probability to move uphill in energy, brief unbiased fluctuations can be "captured" for ratcheting and effectively concatenated to study otherwise rare events, sidestepping the $1/M$ limitation. Trajectories are started at the user's discretion; let us assume two trajectories are started in a bin of state A, with the goal of sampling transitions to B.

Trajectories in WE are run in parallel for brief intervals of time τ (with MFPT $\gg \tau \gg \Delta t$, where Δt is the simulation time step), then stopped and restarted according to simple probabilistic rules. In our example, each of the two trajectories is initially given a weight 1/2 at $t = 0$ and the essential idea is to ensure probability moves in an unbiased way, thus preserving the trajectory ensemble behavior and $p(\vec{x}, t)$. If a trajectory is found to occupy an otherwise empty bin after one of the τ intervals, *two* "child" trajectories are initiated from the final phase point of the "parent" trajectory, and each child inherits 1/2 of the parent's weight—a process called *splitting*. The two child trajectories in the previously unvisited bin create the ratcheting effect: there is twice the likelihood to explore that region, and to continue to still rarer regions, than if we did not replicate trajectories. Stochastic dynamics must be used; otherwise, child trajectories will evolve identically.

If more than two WE trajectories are found in a bin, pruning (or *merging*) is performed in a pairwise fashion: a random number is generated to select one of an arbitrary pair for continuation with probabilities proportional to their weights, and the selected trajectory absorbs the weight of the other trajectory, which is discontinued. In this fashion, energy minima do not collect large numbers of trajectories
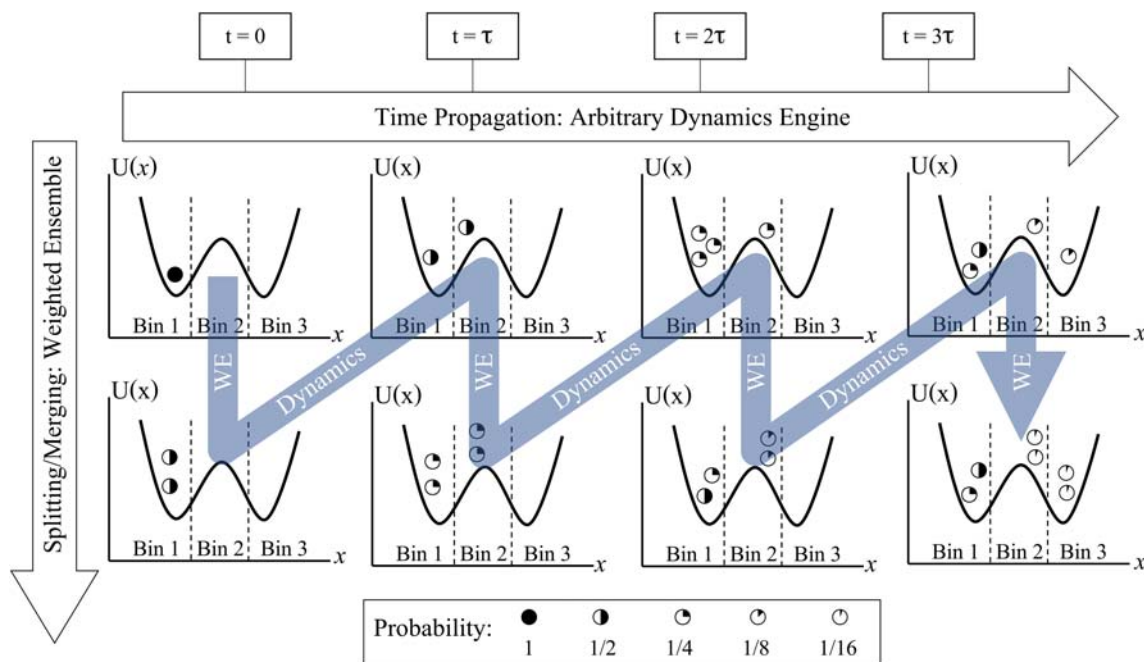

Fig. 4. Efficient simulation via the weighted ensemble (WE) method (Ref. 51). Phase space is divided into bins, and trajectories are started according to an initial distribution of interest (far left). Dynamics are run briefly, allowing trajectories to visit other bins, after which the WE steps of "splitting" (replication) and "merging" (pruning) are performed. Weights of parent trajectories are shared among children from splitting events, permitting the estimation of very low-probability events. In this example, a target of two trajectories per bin has been set. Reproduced with permission from Donovan *et al*, PLoS Comput. Biol. **12**(2), e1004611 (2016). Copyright 2016 Authors, licensed under a Creative Commons Attribution (CC BY) license.

which would add cost to the simulation but provide minimal statistical value. The processes described, in fact, constitute unbiased statistical *resampling*.[49] (See Sec. VIII.) In WE, the total trajectory cost is limited to the number of bins multiplied by the number of trajectories per bin and the trajectory length. This amounts to $M t_{max}$ in our case, given $M/2$ bins.

Although the total simulation cost is bounded by $M t_{max}$ (plus overhead for splitting/merging), events *much* rarer than $1/M$ can be seen because of the splitting procedure. Indeed, exponentially rare processes are elicited as WE produces an unbiased estimate of the trajectory ensemble and $p(\vec{x}, t)$. A dramatic example is shown in Fig. 5 for diffusion and binding in a 3D box, where the distribution of possible binding outcomes extends *tens of orders of magnitude below what standard simulation provides*. For monitoring the transient time evolution of a system, WE is almost like a "magic bullet."

Obtaining the MFPT from WE simulation is more challenging than characterizing $p(\vec{x}, t)$ in many cases. To use the Hill relation Eq. (19), the system must relax to steady state and this relaxation is *not* accelerated by WE for the very reason it is so successful in characterizing $p(\vec{x}, t)$; i.e., because it is unbiased. To see this more concretely, let t_{SS} be the average time required for a given system to relax to steady state. Then, because WE runs M copies of the system, the total cost for observing a WE simulation relax to steady state is $\sim M t_{SS}$, which will be prohibitive in some though not all systems.[52] Even when $M t_{SS}$ is a prohibitive cost, the MFPT can be obtained from transient data ($t < t_{SS}$) available in WE
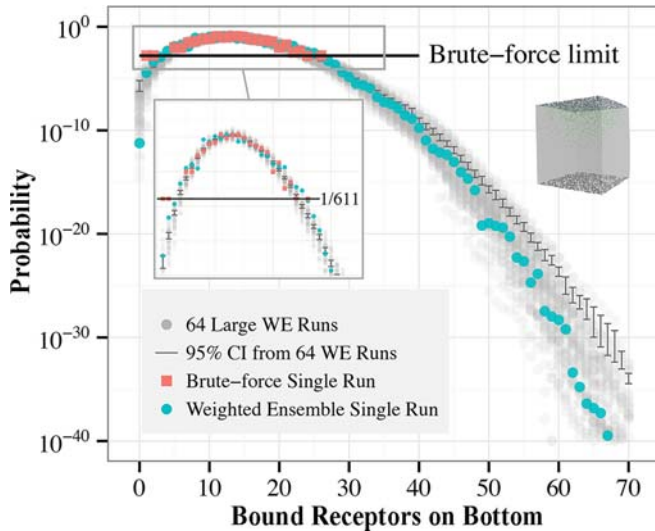

Fig. 5. Weighted ensemble simulation of extremely rare diffusion and binding events (Ref. 51). Particles are initiated at the top of a three-dimensional box (upper right inset) and allowed to diffuse without bias. Any particles that reach the bottom surface of the box can bind to receptors located there. The graph shows the probability distribution of bound receptors after a short time interval—i.e., the likelihood of different outcomes that would result from a single brute-force diffusion simulation. WE enables sampling deep into the tails of the distribution because more trajectories are allotted to rarer outcomes, whereas an equivalent amount of "brute force" sampling cannot detect events rarer than the reciprocal of the number of trajectories, as shown by solid horizontal lines. WE simulations used simulation time equivalent to 611 brute force trajectories, as indicated in the left inset. The grey dots represent independent WE runs (of which green (online) is a representative) and solid vertical bars give the confidence interval based on the grey data—which appears to be skewed upward because of the logarithmic scale. Reproduced with permission from Donovan *et al.*, PLoS Comput. Biol. **12**(2), e1004611 (2016). Copyright 2016 Authors, licensed under a Creative Commons Attribution (CC BY) license.

simulation: although the details are beyond the scope of this discussion, the idea is to use much finer-grained and faster-relaxing bins (than were used to run the WE simulation) in a quasi-Markov approximation.[53] Below, we apply WE directly for MFPT calculation in a simple system.

Like any advanced computational method, WE has its subtleties and limitations. Most important are correlations. Although WE trajectories are independent (non-interacting), exactly as assumed in the trajectory-ensemble definitions, correlations arise in the overall WE protocol due to the splitting and merging steps. After all, when a trajectory is "split," by construction the child trajectories are identical until the split point. Therefore, assessing statistical uncertainty in WE estimates requires great care, even though the method is unbiased.[54]

C. Applying the weighted ensemble to a simple model

To illustrate the power and validity of the weighted ensemble method, we employ it to estimate the transition rate over a high energy barrier in a simple system. We use the WESTPA implementation[55] of WE and apply it to a simple 1D double-well potential under overdamped Langevin dynamics (Eq. (3)) with parameters chosen to approximate the behavior of a small molecule in water. We assume a mass of $100 \, u$, temperature $T = 300 \, K$, a barrier height of $10 k_B T$, and a friction coefficient $\gamma = 24.94 \, ps^{-1}$ which is reasonable for water and corresponds to a diffusion constant of $10^{-6} \, cm^2/s$. The simulation is run with a timestep of 3 ps, and all simulation code is available on Github.[56]

The WE simulation is set up with walkers beginning in the rightmost basin, and with the two basin macrostates defined as $x > 20 \, nm$ and $x < -20 \, nm$, as shown in Fig. 6. Twenty uniform bins of width 2 nm uniformly span from $x = -20.0$ to 20.0 nm, with two additional bins on either end reaching to $\pm \infty$. The WE simulation is run with a resampling time of $\tau = 60 \, ps$ and a target count of ten trajectories per bin, so roughly 200 trajectories will run during each τ. Walkers that reach the left basin are "recycled" and restarted from $x = 20 \, nm$ to generate a non-equilibrium steady state and exploit the Hill relation Eq. (19).

To quantify the effectiveness of WE simulation for this case, we can compare the cost for computing the rate constant (i.e., flux or reciprocal MFPT) from WE simulation to brute-force simulation of overdamped Langevin dynamics. Note from Fig. 6 that WE simulations reach steady values after ~ 3000 iterations, which corresponds to $\sim 12 \times 10^6$ steps of total simulation (for a single WE run, accounting for all ~ 200 trajectories) or a total of 36 μs of simulated time. Also from Fig. 6 and Eq. (19), the MFPT is $\sim 1000 \, \mu s$. Thus, we see that WE simulation has generated the *average* first-passage time using an overall amount of computing that is only a small fraction (~ 0.04) of the time needed to yield a *single* transition event via direct simulation, let alone to generate a reliable MFPT estimate from multiple events.

VII. CONCLUDING DISCUSSION

The trajectory arguably is the most fundamental object in classical statistical mechanics, particularly for non-equilibrium phenomena, and this article has attempted to connect trajectory physics with more familiar topics in the traditional physics curriculum. By focusing in depth on the simplest possible example—diffusion—we have been able to formalize and visualize the probabilistic/ensemble picture
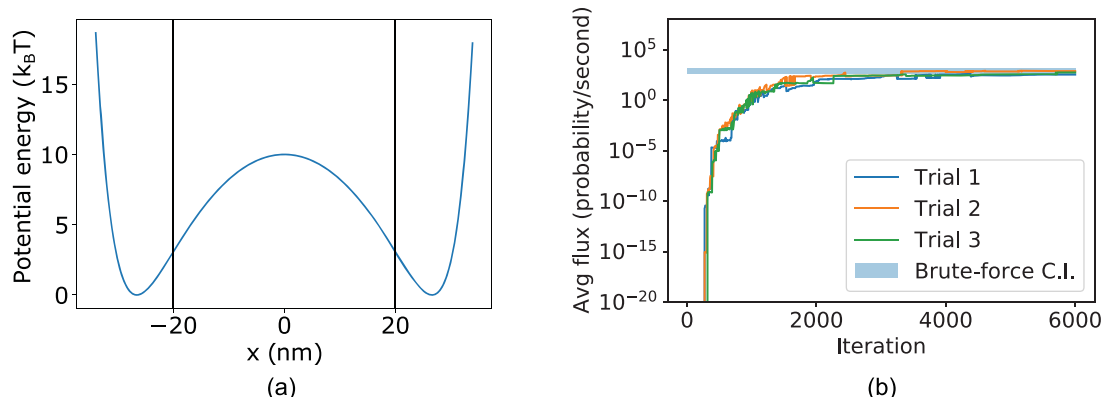
Fig. 6. Weighted ensemble estimation of the rate of a rare event: high-barrier crossing. (a) potential energy function used for double-well simulation with 10 k_BT barrier and state boundaries indicated by the vertical black lines. (b) average flux into the left basin state for simulations started from the right basin, as computed from three independent weighted ensemble simulations (colored lines). The average flux estimates the inverse MFPT by Eq. (19), yielding ∼1 ms. For reference, an independent estimate of the flux is computed using a very long "brute force" simulation (horizontal line). The brute force confidence interval (C.I.) is shown as a blue shaded region, which is ± twice the standard error of the mean based on 11 transitions.

and connect it with simpler spatial distributions. We have further been able to connect these ensembles with observable populations, kinetics, and thermodynamic states, as well as understand a modern, practical path-sampling approach.

A key lesson is that theoretical physics can view a given process at different levels of "magnification," from most microscopic to most averaged (Fig. 7). Trajectories are the most detailed and encompass all system coordinates at all times—which is usually too much to grasp. Trajectories can be averaged spatially at fixed times to yield more familiar probability distributions. Trajectory flows across surfaces of interest can also be averaged to yield probability fluxes: in equilibrium, all such fluxes are zero, whereas in transient regimes or non-equilibrium steady states (NESS's), such flows provide key information. Notably, the Hill relation (Eq. (19)) yields the mean first-passage time (MFPT) from the flux in an appropriate NESS, and furthermore, conditions on reversibility can be derived from flux arguments (Appendix B). Finally, averaging—i.e., integrating—over spatial distributions can yield observable thermodynamic information on state populations;[6,16] see also entropy production and fluctuation relations.[7,42,57,58]

This report has only given a taste of the value of the trajectory picture, which goes much further. Trajectory ideas, for example, are used to develop the Jarzynski relation.[58–60] They provide a direct connection with the path-integral formulation of quantum mechanics.[34] Trajectories offer a unique window into the often misunderstood issue of "reversibility."[61] (See Appendix B.) Not surprisingly, trajectories and their applications are still an area of active research.[7,52,62–65]

VIII. EXERCISES

(1) Confirm by differentiation that Eq. (15) is the exact solution to the diffusion equation Eq. (17), after setting $x_n = x$ and $n = t/\Delta t$. Note that t occurs both in the prefactor *and* the exponent, so differentiation requires the product rule.

(2) Time-discretization generally introduces an error into dynamics computed via Eqs. (4) and (6). Explain why there is an error and how it might be mitigated in computer simulation. For what special case is there no error even if $f \neq 0$?

(3) Implement overdamped dynamics simulation Eq. (4) of the double-well system specified in Sec. VI C. Calculate the MFPT of the system for a range of barrier
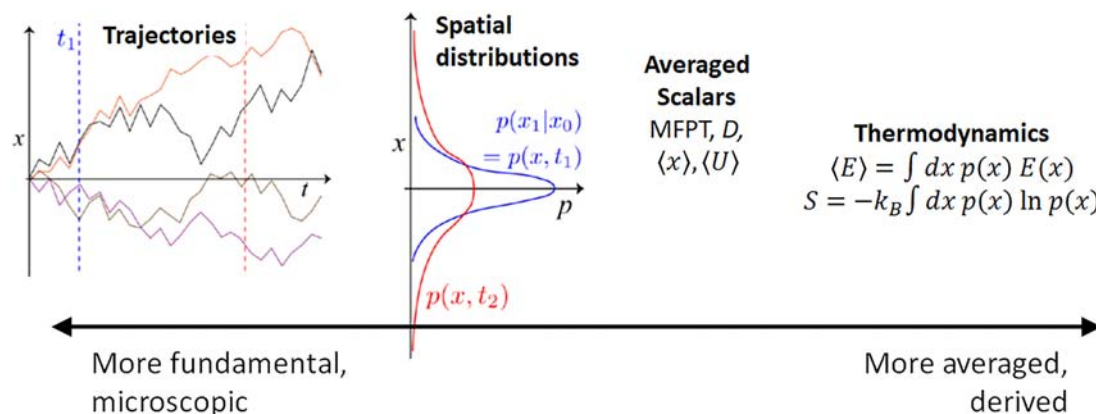

Fig. 7. From the fundamental ensemble of trajectories to more averaged observables. Because trajectories embody the dynamics that fully specifies a system, they are the most fundamental. Averaging or analysis can be performed at fixed time points, including the $t \rightarrow \infty$ stationary point. Quantities that can be calculated include the phase-space distribution $p(x, t)$, the mean first-passage time (MFPT), diffusion constant (D), average coordinates or properties (e.g., $\langle x \rangle$, $\langle U \rangle$), or system-wide thermodynamic properties, in or out of equilibrium. Although simple diffusive trajectories are pictured, the same principles apply in the case of non-zero forces.

heights, starting with a low barrier, by simple averaging of $\sim$10 observed first-passage times. Compare these values to the expected Arrhenius behavior.[4]

(4) Using the ODLD module of the WESTPA implementation of weighted ensemble, implement a *triple* well system. Consider the left-most basin to be the initial state (A) and the right-most basin the target (B). Examine the relaxation of the probability into the target state as a function of time. For cleanest data, average over multiple WE runs. Vary the depth of the middle well and explain the observed behavior.

(5) Write down the trajectory probability, the analog of Eq. (12), for a system with constant force, sometimes called simple drift. Explain in words the meaning of the distribution. If you can, integrate out intermediate time points to show that the behavior remains Gaussian with constant drift.

(6) For a simple diffusive system described by Eq. (12), obtain the distribution for x_3 by a suitable integration of Eq. (14).

(7) Write down the equations that define (i) a steady state and (ii) equilibrium for a discrete-state system in terms of steady probabilities p_i and state-to-state transition probabilities $T_{i \to j}$ for some fixed time interval. Note that equilibrium is defined by detailed balance. Show that detailed balance implies steady state but not the reverse. A counter-example suffices to disprove a hypothesis.

(8) By studying the theory underlying weighted ensemble,[49] explain in statistical terms why the "resampling" procedure for "merging" trajectories does not bias the time-evolving probability distribution $p(\vec{x}, t)$.

(9) Write pseudocode for a weighted ensemble simulation of an arbitrary system with pre-defined bins. If you are ambitious, implement your pseudocode for 1D overdamped dynamics in the double-well system in Exercise 3.

(10) Understand the continuity equation Eq. (A1) by integrating it over an interval in x from a to b. Integrating the probability density over this region gives the total probability in it. How does this probability change in time, based on the current, and why does the result make sense? Remember the one-dimensional current is defined to be positive in the right-ward direction.

(11) Show that stationary distribution of the Smoluchowski equation Eq. (A3), i.e., when $\partial p / \partial t = 0$, is the expected equilibrium distribution based on the Boltzmann factor.

APPENDIX A: THE FOKKER–PLANCK PICTURE AND SMOLUCHOWSKI EQUATION IN ONE DIMENSION

The Fokker–Planck and related equations[35,36] are essential for understanding non-equilibrium statistical mechanics.

These equations generalize the diffusion equation (17), but they perform essentially the same role: they quantify the way a spatial and/or configurational distribution changes over time based on a given energy landscape. The key point is that this is a very general concept that applies not only to center-of-mass diffusive motion but also to configurational motions internal to a molecule or system. For example, if a protein is started in a certain configuration, where is it likely to be later? The distribution $p(\vec{x}, t)$ quantifies the distribution of configurations $\vec{x}$ at any time t.

Here, we focus on the Smoluchowski equation, which is the Fokker–Planck equation specific for the overdamped, non-inertial dynamics Eq. (3) studied above. The Smoluchowski equation is easiest to grasp starting from the continuity equation, given by

$$\frac{\partial p}{\partial t} = -\frac{\partial J}{\partial x}, \tag{A1}$$

in one dimension, where $p = p(x, t)$ is the probability density at time t and $J = J(x, t)$ is the probability current, i.e., the (average) probability per unit time moving in the $+x$ direction. Note that this is the average over trajectories moving in both directions, so it is the *net* current. The continuity equation simply ensures that the change of probability in any region is the difference between incoming and outgoing probability. For students who are new to the continuity equation, Exercise 10 will clarify its meaning.

To complete the Smoluchowski equation, we need the current corresponding to overdamped dynamics (Eq. (3)). As noted above, overdamped dynamics includes both (simple) diffusion and "drift" (motion due to force). From Eq. (17), we can already infer that the diffusive current is $-D \, \partial p / \partial x$, which is Fick's law indicating that particles/probability will diffuse down their gradients in a linear fashion on average. When a force is present, Eq. (3) indicates that there is also motion linearly proportional to the force, leading to a total current

$$J(x, t) = -D\frac{\partial p}{\partial x} + \frac{D}{k_B T}f(x) \; p(x, t), \tag{A2}$$

where we have assumed $D = k_B T / m\gamma$ is constant in space.

We obtain the full Smoluchowski equation in one-dimension for fixed D by substituting the current Eq. (A2) into the continuity equation Eq. (A1), yielding

$$\frac{\partial p}{\partial t} = D\frac{\partial^2 p}{\partial x^2} - \frac{D}{k_B T}\frac{\partial}{\partial x}f \; p. \tag{A3}$$

The diffusion equation has been augmented by a term dependent on the force. Equation (A3) can be solved to find the steady-state behavior of p both out of or in equilibrium (see Sec. VIII) or to follow the time-dependent behavior as the distribution p relaxes toward its limiting steady profile.

APPENDIX B: ADVANCED TOPIC: MACROSCOPIC REVERSIBILITY BY DECOMPOSING THE EQUILIBRIUM TRAJECTORY ENSEMBLE

Many of us are aware of the intrinsic time reversibility of Newtonian mechanics, whereby any constant-energy trajectory $\vec{x}(t)$ can be "played backwards" to yield another
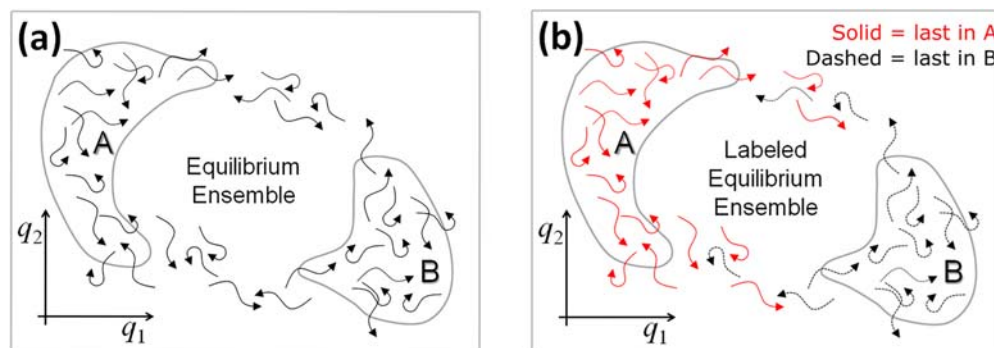
Fig. 8. An exact decomposition of the equilibrium trajectory ensemble. (a) The equilibrium ensemble, consisting of a large number of *independent* trajectories projected onto the schematic coordinates q_1 and q_2. Transitions between macrostates A and B (gray outlines) occur via two pathways, upper and lower. (b) Decomposition of the equilibrium ensemble based on which of two macrostates, A and B, has been visited most recently. These two directional ensembles, "last in A" or "A-to-B" (red, solid lines) and "last in B" or "B-to-A" (black, dashed lines), are non-equilibrium steady states. Arrow tips represent the same time point for all trajectories and arrow tails represent the most recent history, but all history is assumed to be known. Reprinted with permission from Bhatt and Zuckerman, J. Chem. Theory Comput. **7**(8), 2520–2527 (2011). Copyright 2011 American Chemical Society (Ref. [61]).

physically valid trajectory. There is an analogous condition on a stochastic trajectory, which can be derived from detailed balance.[60] However, the conditions for reversibility under more realistic circumstances involving a *distribution* of initial and final configurations require the trajectory ensemble picture.[61]

We start by considering an equilibrium ensemble of trajectories: see Fig. 8(a). The equilibrium trajectory ensemble is defined by a set of completely independent systems/trajectories for times $t > t_0$, given that at t_0, the set of phase-space points $\vec{x}(t_0)$ is equilibrium-distributed, i.e., according to the Boltzmann factor. (We don't need to worry about how equilibrium was produced.) If the phase points are equilibrium-distributed at time t_0, they will remain equilibrium-distributed thereafter. This is because the Markovian stochastic dynamics that generates equilibrium also maintains it, which is why we call it equilibrium in the first place.[4]

As sketched in Fig. 8, the equilibrium ensemble at any time t can be *exactly* decomposed into two parts based on a history-labeling process.[46,61] Specifically, based on two arbitrary non-overlapping macrostates A and B, each trajectory can be assigned to the A-to-B set—a.k.a "last-in-A" set—if it currently occupies state A or was more recently in A than B, with the remaining trajectories in the B-to-A direction. This construction requires "omniscience," in the sense of knowing the full history of each trajectory, so it is something of a thought experiment. Note that each of these directional trajectory subsets is automatically maintained as a *non-equilibrium* steady state: when an A-to-B trajectory enters B, its label switches to B-to-A, but the overall equilibrium condition ensures that equal numbers of trajectories will switch labels per unit time.[61]

We are now in a position to understand reversibility, building on the defining process of equilibrium: detailed balance.[4] As a reminder, detailed balance implies there is zero net flow between *any* pair of "microstates," i.e., small phase-space volumes. In the context of the two uni-directional steady states (A-to-B and B-to-A), detailed balance gives us a tool to consider two non-overlapping mechanistic "pathways"—arbitrary tubes of phase points connecting A and B—e.g., upper and lower pathways in Fig. 8. If we place a (hyper-)surface transecting each tube, then there is a certain probability flowing per second through each surface in, say, the A-to-B steady state; call these σ_1 and σ_2. By detailed

balance, there is no net flow through either surface in equilibrium, and so the flows in the B-to-A state must be equal and opposite. Mechanistically, the ratio σ_1/σ_2 is the same in both directions: the fraction of events taking each pathway must be the same in both directions. This is mechanistic reversibility. Fuller details and illustrations can be found in earlier work.[61]

A key point is that the preceding discussion is strictly based on the detailed-balance property of equilibrium. Thus, systems out of equilibrium should *not* be expected to exhibit mechanistic reversibility. This is true experimentally and theoretically. Examples of systems not obeying reversibility would be if A and B states were prepared under different conditions (e.g., temperature, pH,…) or, even under the same conditions, if the initial distribution in A or B did not mimic the process for constructing the directional steady states derived from equilibrium. Specifically, in the A-to-B direction, trajectories should be initiated on the surface of A according to the distribution with which they would arrive from B in equilibrium, which is known as the "EqSurf" construction.[61] To put this informally, state A needs to be "tricked" into behaving as it would in equilibrium, so trajectories are started at the boundary of A as if they had arrived from B (i.e., were last in B) in equilibrium.

a)Electronic mail: zuckermd@ohsu.edu, ORCID: 0000-0001-7662-2031.
[1]Lars Onsager, "Reciprocal relations in irreversible processes. I," Phys. Rev. **37**(4), 1505–1512 (1931).
[2]Eric Vanden-Eijnden *et al.*, "Transition-path theory and path-finding algorithms for the study of rare events," Annu. Rev. Phys. Chem. **61**, 391–420 (2010).
[3]Ron Elber, Dmitrii E. Makarov, and Henri Orland, *Molecular Kinetics in Condensed Phases: Theory, Simulation, and Analysis* (John Wiley & Sons, New York, 2020).
[4]Daniel M. Zuckerman, *Statistical Physics of Biomolecules: An Introduction* (CRC, Boca Raton, FL, 2010).
[5]Douglas C. Giancoli, *Physics for Scientists and Engineers with Modern Physics* (Pearson Education, London, 2008).
[6]Frederick Reif, *Fundamentals of Statistical and Thermal Physics* (McGraw-Hill, New York, 1965).
[7]Udo Seifert, "Stochastic thermodynamics, fluctuation theorems and molecular machines," Rep. Prog. Phys. **75**(12), 126001 (2012).
[8]Daniel M. Zuckerman, "Key biology you should have learned in physics class: Using ideal-gas mixtures to understand biomolecular machines," Am. J. Phys. **88**(3), 182–193 (2020).
[9]Kresten Lindorff-Larsen, Stefano Piana, Ron O. Dror, and David E. Shaw, "How fast-folding proteins fold," Science **334**(6055), 517–520 (2011).

[10] Alan Fersht *et al.*, *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding* (Macmillan, New York, 1999).

[11] J. Juraszek and P. G. Bolhuis, "Sampling the multiple folding mechanisms of Trp-cage in explicit solvent," Proc. Natl. Acad. Sci. **103**(43), 15859–15864 (2006).

[12] David L. Mobley and Michael K. Gilson, "Predicting binding free energies: Frontiers and benchmarks," Annu. Rev. Biophys. **46**, 531–558 (2017).

[13] K. Gunasekaran, Buyong Ma, and Ruth Nussinov, "Is allostery an intrinsic property of all dynamic proteins?," Proteins **57**(3), 433–443 (2004).

[14] Matthew C. Zwier, Adam J. Pratt, Joshua L. Adelman, Joseph W. Kaus, Daniel M. Zuckerman, and Lillian T. Chong, "Efficient atomistic simulation of pathways and calculation of rate constants for a protein–peptide binding process: Application to the MDM2 protein and an intrinsically disordered p53 peptide," J. Phys. Chem. Lett. **7**(17), 3440–3445 (2016).

[15] Robert A. Copeland, David L. Pompliano, and Thomas D. Meek, "Drug–target residence time and its implications for lead optimization," Nat. Rev. Drug Discovery **5**(9), 730–739 (2006).

[16] Daniel M. Zuckerman and Lillian T. Chong, "Weighted ensemble simulation: Review of methodology, applications, and software," Annu. Rev. Biophys. **46**, 43–57 (2017).

[17] Lawrence R. Pratt, "A statistical method for identifying transition states in high dimensional problems," J. Chem. Phys. **85**(9), 5045–5048 (1986).

[18] Christoph Dellago, Peter G. Bolhuis, Félix S. Csajka, and David Chandler, "Transition path sampling and the calculation of rate constants," J. Chem. Phys. **108**(5), 1964–1977 (1998).

[19] Daniel M. Zuckerman and Thomas B. Woolf, "Dynamic reaction paths and rates through importance-sampled stochastic dynamics," J. Chem. Phys. **111**(21), 9475–9484 (1999).

[20] Martin Bier, Imre Derényi, Marcin Kostur, and R. Dean Astumian, "Intrawell relaxation of overdamped Brownian particles," Phys. Rev. E **59**(6), 6422–6432 (1999).

[21] R. Dean Astumian, "The unreasonable effectiveness of equilibrium theory for interpreting nonequilibrium experiments," Am. J. Phys. **74**(8), 683–688 (2006).

[22] R. Dean Astumian, "Thermodynamics and kinetics of molecular motors," Biophys. J. **98**(11), 2401–2409 (2010).

[23] R. Dean Astumian, Cristian Pezzato, Yuanning Feng, Yunyan Qiu, Paul R. McGonigal, Chuyang Cheng, and J. Fraser Stoddart, "Non-equilibrium kinetics and trajectory thermodynamics of synthetic molecular pumps," Mater. Chem. Front. **4**(5), 1304–1314 (2020).

[24] Terrell L. Hill, "The linear Onsager coefficients for biochemical kinetic diagrams as equilibrium one-way cycle fluxes," Nature **299**(5878), 84–86 (1982).

[25] Terrell L. Hill, *Free Energy Transduction and Biochemical Cycle Kinetics* (Dover, Mineola, NY, 2004).

[26] Kingshuk Ghosh, Ken A. Dill, Mandar M. Inamdar, Effrosyni Seitaridou, and Rob Phillips, "Teaching the principles of statistical dynamics," Am. J. Phys. **74**(2), 123–133 (2006).

[27] Steve Pressé, Kingshuk Ghosh, Julian Lee, and Ken A. Dill, "Principles of maximum entropy and maximum caliber in statistical physics," Rev. Mod. Phys. **85**(3), 1115–1141 (2013).

[28] Kingshuk Ghosh, Purushottam D. Dixit, Luca Agozzino, and Ken A. Dill, "The maximum caliber variational principle for nonequilibria," Annu. Rev. Phys. Chem. **71**, 213–238 (2020).

[29] Robert H. Swendsen, "Explaining irreversibility," Am. J. Phys. **76**(7), 643–648 (2008).

[30] Subrahmanyan Chandrasekhar, "Stochastic problems in physics and astronomy," Rev. Mod. Phys. **15**(1), 1–89 (1943).

[31] Lillian T. Chong, Ali S. Saglam, and Daniel M. Zuckerman, "Path-sampling strategies for simulating rare events in biomolecular systems," Curr. Opin. Struct. Biol. **43**, 88–94 (2017).

[32] Nicolaas Godfried Van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, New York, 1992), Vol. 1.

[33] Lars Onsager and Stefan Machlup, "Fluctuations and irreversible processes," Phys. Rev. **91**(6), 1505–1512 (1953).

[34] J. J. Sakurai, *Modern Quantum Mechanics* (Addison Wesley, Reading, MA, 1985).

[35] Crispin Gardiner, *Stochastic Methods* (Springer-Verlag, Berlin, 2009), Vol. 4.

[36] Hannes Risken and Till Frank, *The Fokker-Planck Equation: Methods of Solution and Applications* (Springer Science & Business Media, New York, 1996), Vol. 18.

[37] Sidney Redner, *A Guide to First-Passage Processes* (Cambridge U. P., Cambridge, 2001).

[38] David Chandler, *Introduction to Modern Statistical Mechanics* (Oxford U. P., Oxford, UK, 1987).

[39] John D. Chodera, Phillip J. Elms, William C. Swope, Jan-Hendrik Prinz, Susan Marqusee, Carlos Bustamante, Frank Noé, and Vijay S. Pande, "A robust approach to estimating rates from time-correlation functions," arXiv:1108.2304 (2011).

[40] Ashok A. Deniz, Samrat Mukhopadhyay, and Edward A. Lemke, "Single-molecule biophysics: At the interface of biology, physics and chemistry," J. R. Soc. Interface **5**(18), 15–45 (2008).

[41] Helen Miller, Zhaokun Zhou, Jack Shepherd, Adam J. M. Wollman, and Mark C. Leake, "Single-molecule techniques in biophysics: A review of the progress in methods and applications," Rep. Prog. Phys. **81**(2), 024601 (2017).

[42] Gavin E. Crooks, "Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences," Phys. Rev. E **60**(3), 2721–2726 (1999).

[43] Ernesto Suárez and Daniel M. Zuckerman, "Pathway histogram analysis of trajectories: A general strategy for quantification of molecular mechanisms," arXiv:1810.10514 (2018).

[44] Rosalind J. Allen, Patrick B. Warren, and Pieter Rein Ten Wolde, "Sampling rare switching events in biochemical networks," Phys. Rev. Lett. **94**(1), 018104 (2005).

[45] Anton K. Faradjian and Ron Elber, "Computing time scales from reaction coordinates by milestoning," J. Chem. Phys. **120**(23), 10880–10889 (2004).

[46] Titus S. van Erp, Daniele Moroni, and Peter G. Bolhuis, "A novel path sampling method for the calculation of rate constants," J. Chem. Phys. **118**(17), 7762–7774 (2003).

[47] Aryeh Warmflash, Prabhakar Bhimalapuram, and Aaron R. Dinner, "Umbrella sampling for nonequilibrium processes," J. Chem. Phys. **127**(15), 114109 (2007).

[48] Gary A. Huber and Sangtae Kim, "Weighted-ensemble Brownian dynamics simulations for protein association reactions," Biophys. J. **70**(1), 97–110 (1996).

[49] Bin W. Zhang, David Jasnow, and Daniel M. Zuckerman, "The "weighted ensemble" path sampling method is statistically exact for a broad class of stochastic processes and binning procedures," J. Chem. Phys. **132**(5), 054107 (2010).

[50] Herman Kahn and Theodore E. Harris, "Estimation of particle transmission by random sampling," Natl. Bur. Stand. Appl. Math. Ser. **12**, 27–30 (1951).

[51] Rory M. Donovan, Jose-Juan Tapia, Devin P. Sullivan, James R. Faeder, Robert F. Murphy, Markus Dittrich, and Daniel M. Zuckerman, "Unbiased rare event sampling in spatial stochastic systems biology models using a weighted ensemble of trajectories," PLoS Comput. Biol. **12**(2), e1004611 (2016).

[52] Upendra Adhikari, Barmak Mostofian, Jeremy Copperman, Sundar Raman Subramanian, Andrew A. Petersen, and Daniel M. Zuckerman, "Computational estimation of microsecond to second atomistic folding times," J. Am. Chem. Soc. **141**(16), 6519–6526 (2019).

[53] Jeremy Copperman and Daniel M. Zuckerman, "Accelerated estimation of long-timescale kinetics from weighted ensemble simulation via non-Markovian 'microbin' analysis," J. Chem. Theory Comput. **16**(11), 6763–6775 (2020).

[54] Barmak Mostofian and Daniel M. Zuckerman, "Statistical uncertainty analysis for small-sample, high log-variance data: Cautions for bootstrapping and Bayesian bootstrapping," J. Chem. Theory Comput. **15**(6), 3499–3509 (2019).

[55] Matthew C. Zwier, Joshua L. Adelman, Joseph W. Kaus, Adam J. Pratt, Kim F. Wong, Nicholas B. Rego, Ernesto Suárez, Steven Lettieri, David W. Wang, Michael Grabe, Daniel M. Zuckerman, and Lillian T. Chong, "WESTPA: An interoperable, highly scalable software package for weighted ensemble simulation and analysis," J. Chem. Theory Comput. **11**(2), 800–809 (2015).

[56] John Russo (2021). "jdrusso/doublewell: More physical parameters (version 1.1)," Zenodo. https://doi.org/10.5281/zenodo.4706088

[57] G. N. Bochkov and Yu. E. Kuzovlev, "Nonlinear fluctuation-dissipation relations and stochastic models in nonequilibrium thermodynamics: I. Generalized fluctuation-dissipation theorem," Physica A **106**(3), 443–479 (1981).

[58] Christopher Jarzynski, "Nonequilibrium equality for free energy differences," Phys. Rev. Lett. **78**(14), 2690–2693 (1997).

[59] C. Jarzynski, "Equilibrium free energies from nonequilibrium processes," Acta Phys. Pol., Ser. B **29**(6), 1609–1622 (1998).

[60] Gavin E. Crooks, "Nonequilibrium measurements of free energy differences for microscopically reversible Markovian systems," J. Stat. Phys. **90**(5–6), 1481–1487 (1998).

[61] Divesh Bhatt and Daniel M. Zuckerman, "Beyond microscopic reversibility: Are observable nonequilibrium processes precisely reversible?," J. Chem. Theory Comput. **7**(8), 2520–2527 (2011).

[62] Erik H. Thiede, Dimitrios Giannakis, Aaron R. Dinner, and Jonathan Weare, "Galerkin approximation of dynamical quantities using trajectory data," J. Chem. Phys. **150**(24), 244111 (2019).

[63] David Aristoff and Daniel M. Zuckerman, "Optimizing weighted ensemble sampling of steady states," Multiscale Model. Simul. **18**(2), 646–673 (2020).

[64] Grant M. Rotskoff and Eric Vanden-Eijnden, "Dynamical computation of the density of states and Bayes factors using nonequilibrium importance sampling," Phys. Rev. Lett. **122**(15), 150602 (2019).

[65] John D. Russo, Jeremy Copperman, and Daniel M. Zuckerman, "Iterative trajectory reweighting for estimation of equilibrium and non-equilibrium observables," arXiv:2006.09451 (2020).
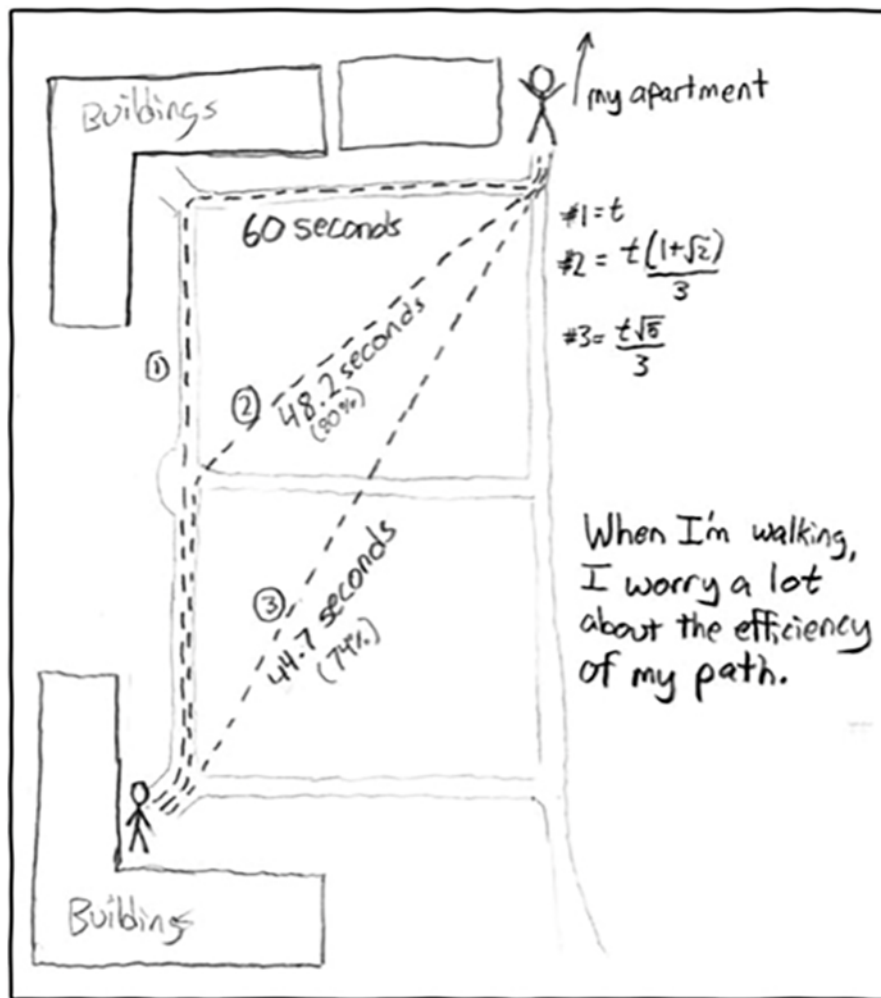
Paths

It's true, I think about this all the time. (Source: https://xkcd.com/85/)