

Time-asymmetric fluctuation theorem and efficient free-energy estimation

Adrienne Zhong ^{1,2,*}, Ben Kuznets-Speck^{3,4,†} and Michael R. DeWeese ^{1,2,5}


¹Department of Physics, University of California, Berkeley, Berkeley, California 94720, USA

²Redwood Center For Theoretical Neuroscience, University of California, Berkeley, Berkeley, California 94720, USA

³Biophysics Graduate Group, University of California, Berkeley, California 94720, USA

⁴Department of Physics and Astronomy, Northwestern University, Evanston, Illinois 60208, USA

⁵Department of Neuroscience, University of California, Berkeley, Berkeley, California 94720, USA

 (Received 22 November 2023; accepted 16 August 2024; published 12 September 2024)

The free-energy difference ΔF between two high-dimensional systems is notoriously difficult to compute but very important for many applications such as drug discovery. We demonstrate that an unconventional definition of work introduced by Vaikuntanathan and Jarzynski (2008) satisfies a microscopic fluctuation theorem that relates path ensembles that are driven by protocols unequal under time reversal. It has been shown before that counterdiabatic protocols—those having additional forcing that enforces the system to remain in instantaneous equilibrium, also known as escorted dynamics or engineered swift equilibration—yield zero-variance work measurements for this definition. We show that this time-asymmetric microscopic fluctuation theorem can be exploited for efficient free-energy estimation by developing a simple (i.e., neural-network free) and efficient adaptive time-asymmetric protocol optimization algorithm that yields ΔF estimates that are orders of magnitude lower in mean squared error than the generic linear interpolation protocol with which it is initialized.

DOI: [10.1103/PhysRevE.110.034121](https://doi.org/10.1103/PhysRevE.110.034121)

I. INTRODUCTION

Free-energy differences $\Delta F = F_B - F_A$ between pairs of potential-energy functions $U_A(x)$ and $U_B(x)$ are sought after by physicists, chemists, and pharmaceutical scientists alike [1–6]. Here, $x \in \mathbb{R}^d$ is the configuration space coordinate, and the free energy for each potential is defined as $F_{A,B} = -\beta^{-1} \ln \int e^{-\beta U_{A,B}(x)} dx$, where $\beta = (k_B T)^{-1}$ is inverse temperature. For high-dimensional systems, ΔF can only be calculated numerically through sampling methods, which can be computationally costly and slow to converge [4]. Here we present an adaptive method that greatly reduces the variance of ΔF estimates based on a fluctuation theorem we derive.

One class of estimators takes work measurements as input from protocols $U(x, t)$ that “switch” $U(x, 0) = U_A(x) \rightarrow U(x, t_f) = U_B(x)$ in finite time t_f . Because the work, traditionally defined for a trajectory $X(t)|_{t \in [0, t_f]}$ [7] as

$$W_{\text{trad}}[X(t)] = \int_0^{t_f} \frac{\partial U}{\partial t}(X(t), t) dt \quad (1)$$

satisfies the Jarzynski equality

$$\langle e^{-\beta W} \rangle = e^{-\beta \Delta F}, \quad (2)$$

the Jarzynski estimator $\widehat{\Delta F}_{\text{Jar}} = -\beta^{-1} \ln(n^{-1} \sum_{i=1}^{n_s} e^{-\beta W_{\text{trad}}^i})$ may be applied to work measurements $\{W_{\text{trad}}^i, |i = 1, \dots, n_s\}$. Unfortunately this estimator can be slow to converge, because the average is often dominated by rare events.

Estimators that use bidirectional work measurements (i.e., those that also consider $U_B \rightarrow U_A$ switching processes) generally have lower variance than unidirectional work estimators [8]. In particular, Shirts *et al.* in Ref. [9] showed that, if forward $\{W_F^i | i = 1, \dots, n_s\}$ and reverse work measurements $\{W_R^i | i = 1, \dots, n_s\}$, assumed here to be equal in number for simplicity, are collected from forward and reverse protocols satisfying Crooks Fluctuation Theorem

$$\mathcal{P}_F(+W) = \mathcal{P}_R(-W)e^{\beta(W - \Delta F)}, \quad (3)$$

then the Bennett acceptance ratio estimator $\widehat{\Delta F}_{\text{BAR}}$ [10], defined implicitly as the ΔF satisfying

$$\sum_{i=1}^{n_s} \frac{1}{1 + e^{-\beta(W_F^i - \Delta F)}} - \sum_{j=1}^{n_s} \frac{1}{1 + e^{-\beta(W_R^j + \Delta F)}} = 0, \quad (4)$$

is the lowest-variance asymptotically unbiased estimator. Bidirectional measurements of W_{trad} for a pair of time-reversal-symmetric forward and reverse protocols satisfy Eq. (3) [11,12], but measurements can also be collected from *mixtures* of different measurement-protocol pairs

$$\mathcal{P}_F(\cdot) = \sum_i \alpha_i \mathcal{P}_F^i(\cdot) \text{ and } \mathcal{P}_R(\cdot) = \sum_i \alpha_i \mathcal{P}_R^i(\cdot), \quad (5)$$

with $\sum_i \alpha_i = 1$, as long as each $(\mathcal{P}_F^i, \mathcal{P}_R^i)$ pair satisfies Eq. (3).

In this paper, we consider the nonstandard definition of work introduced in Ref. [13], for which, remarkably, there exists finite-time protocols that yield zero-variance work measurements. Importantly, these zero-variance protocols include a separate counterdiabatic forcing term that effectuates a faster-than-quasistatic time evolution. We explicitly show that this nonstandard definition of work satisfies the fluctuation

*Contact author: adrizhong@berkeley.edu

†These authors contributed equally to this work.

theorem (3) for measurements that are produced from separate forward and reverse protocols that are *unequal* under time reversal. We demonstrate that the time-asymmetric fluctuation theorem for this unconventional work may be exploited for efficient free-energy estimation by proposing an algorithm that iteratively improves time-asymmetric protocols and uses measurements collected across *all* iterations. On three examples of increasing complexity, we show that 10^3 measurements made under our adaptive protocol algorithm give ΔF estimates that are a factor of $\approx 10^2$ – 10^4 lower in mean squared error than the same number of measurements made with the generic time-symmetric linear interpolation protocol with which it was initialized.

The first version of this paper was posted on arXiv in April 2023 [14]. Near-simultaneously, the preprint [15] was posted on arXiv, in which the authors independently derived the same theoretical results as we found for overdamped dynamics and demonstrated through impressive numerical results the utility of the time-asymmetric fluctuation theorem.

II. TIME-ASYMMETRIC WORK

For our setting we consider a time-varying potential energy $U_0(x, t)$ for $t \in [0, t_f]$, that begins at $U_0(x, 0) = U_A(x)$ and ends at $U_0(x, t_f) = U_B(x)$. To this we add an additional potential $U_1(x, t)$ that satisfies $U_1(x, 0) = U_1(x, t_f) = 0$. In the overdamped limit, a trajectory $X(t)$ evolves according to the Langevin equation

$$\dot{X}(t) = -\nabla(U_0 + U_1) + \sqrt{2\beta^{-1}}\eta(t) \text{ with } X(0) \sim \rho_A(\cdot). \quad (6)$$

Here, $\rho_A(x) = e^{-\beta[U_A(x) - F_A]}$ is the equilibrium distribution for $U_A(x)$, and $\eta(t)$ is an instantiation of standard d -dimensional Gaussian white noise specified by $\langle \eta_i(t) \rangle = 0$ and $\langle \eta_i(t)\eta_j(t') \rangle = \delta_{ij}\delta(t - t')$ [16]. (We consider underdamped dynamics in Appendix C.)

In Ref. [13] the authors introduced an unconventional work definition, which in our setting is the trajectory functional

$$W[X(t)] = \int_0^{t_f} \frac{\partial U_0}{\partial t} - \nabla U_0 \cdot \nabla U_1 + \beta^{-1} \nabla^2 U_1 dt \quad (7)$$

(∇^2 is the scalar Laplace operator), and demonstrated that, remarkably, $W[X(t)] = \Delta F$ for *every* trajectory $X(t)$ if $U_1(x, t)$ gives the counterdiabatic force for $U_0(x, t)$, meaning

$$\frac{\partial \rho_0}{\partial t} = \nabla \cdot (\rho_0 \nabla U_1) \text{ for } \rho_0(x, t) := e^{-\beta[U_0(x, t) - F_0(t)]}. \quad (8)$$

Here $\rho_0(x, t)$ is the instantaneous equilibrium distribution corresponding to $U_0(x, t)$, with time-dependent free energy $F_0(t) = -\beta^{-1} \ln \int e^{-\beta U_0(x, t)} dx$ satisfying $F_0(0) = F_A$ and $F_0(t_f) = F_B$. Counterdiabatic driving has been studied before in various contexts [17–21]. Under these conditions, the time-dependent probability distribution for Eq. (6) is always in instantaneous equilibrium with $U_0(x, t)$.

Indeed, expanding Eq. (8) yields

$$\frac{\partial U_0}{\partial t} - \nabla U_0 \cdot \nabla U_1 + \beta^{-1} \nabla^2 U_1 = \frac{dF_0}{dt}, \quad (9)$$

which, when plugged into Eq. (7), shows that the time-asymmetric work $W[X(t)] = \int_0^{t_f} \dot{F}_0(t) dt = F_0(t_f) - F_0(0) = \Delta F$ for *every* trajectory $X(t)$. With optimally chosen $U_0(x, t)$

and $U_1(x, t)$, the free-energy difference may be obtained from simulating a single finite-time trajectory. Unfortunately, Eq. (9) is typically infeasible to solve for multidimensional systems, and to formulate the partial differential equation (PDE), $\dot{F}_0(t)$, and therefore ΔF , must already be known.

III. TIME-ASYMMETRIC MICROSCOPIC FLUCTUATION THEOREM

In the late 1990s, Crooks [11,12] discovered that the microscopic fluctuation theorem

$$W[X(t)] = \Delta F + \beta^{-1} \ln \frac{\mathcal{P}[X(t)]}{\tilde{\mathcal{P}}[\tilde{X}(t)]} \quad (10)$$

is satisfied by the traditional work $W = W_{\text{trad}}$. Here $\mathcal{P}[X(t)]$ is the probability of observing a trajectory $X(t)$, and $\tilde{\mathcal{P}}[\tilde{X}(t)]$ is the probability of observing its time-reversed trajectory $\tilde{X}(t) = X(t_f - t)$ in a “reverse” path ensemble driven by the protocol $\tilde{U}(x, t) = U(x, t_f - t)$. In this section, we derive the microscopic fluctuation theorem satisfied by the unconventional work definition Eq. (7).

In our overdamped setting, the probability of realizing a trajectory $X(t)$ from the dynamics Eq. (6) may be formally expressed up to a normalization factor as

$$\mathcal{P}[X(t)] = \rho_A(X(0))e^{-\beta S[X(t)]}, \quad (11)$$

where

$$S[X(t)] = (\text{I}) \int_0^{t_f} \frac{|\dot{X} + \nabla(U_0 + U_1)|^2}{4} dt, \quad (12)$$

is the Onsager-Machlup action functional (see Appendix A1 for a quick review, also Ref. [22]). We use (I) to indicate that the integral is taken in an Itô sense (reviewed in Appendix A2). After Eqs. (7) and (11) are plugged into Eq. (10), straightforward manipulations under the rules of stochastic calculus (see Appendix A3) yield

$$\begin{aligned} \tilde{\mathcal{P}}[\tilde{X}(t)] &= e^{-\beta\{U_A(X(0)) - F_A + S[X(t)] + W[X(t)] - \Delta F\}} \\ &= \rho_B(\tilde{X}(0))e^{-\beta \tilde{S}[\tilde{X}(t)]}, \end{aligned} \quad (13)$$

where $\rho_B(x) = e^{-\beta[U_B(x) - F_B]}$ is the equilibrium distribution for $U_B(x)$, and

$$\tilde{S}[\tilde{X}(t)] = (\text{I}) \int_0^{t_f} \frac{|\dot{\tilde{X}} + \nabla(\tilde{U}_0 - \tilde{U}_1)|^2}{4} dt \quad (14)$$

has the form of a path action, with $\tilde{U}_{0,1}(x, t) = U_{0,1}(x, t_f - t)$ denoting the time-reversed potential energies. Equation (13) gives the probability of observing the path $\tilde{X}(t)$ under the Langevin equation

$$\dot{\tilde{X}}(t) = -\nabla(\tilde{U}_0 - \tilde{U}_1) + \sqrt{2\beta^{-1}}\eta(t) \text{ with } \tilde{X}(0) \sim \rho_B(\cdot), \quad (15)$$

which differs from Eq. (6) by a minus sign on the U_1 term. In other words, the reverse path ensemble that satisfies Eq. (13)

for the time-asymmetric work Eq. (7) is one that is driven by a protocol $\tilde{U}_0 - \tilde{U}_1$ that is *different* from the time reversal of the forward protocol $U_0 + U_1$. One can also verify that its associated definition of work,

$$\tilde{W}[\tilde{X}(t)] = \int_0^{t_f} \frac{\partial \tilde{U}_0}{\partial t} - \nabla \tilde{U}_0 \cdot \nabla(-\tilde{U}_1) + \beta^{-1} \nabla^2(-\tilde{U}_1) dt, \quad (16)$$

satisfies $\tilde{W}[\tilde{X}(t)] = -W[X(t)]$, so the same optimal $U_0(x, t)$ and $U_1(x, t)$ satisfying Eq. (8) also give $\tilde{W}[\tilde{X}(t)] = -\Delta F$ for every trajectory.

Through standard methods (see Appendix B), the fluctuation theorem Eq. (3) follows directly from the microscopic fluctuation theorem Eq. (10). Thus, the time-asymmetric work [Eq. (7)] holds a deeper significance than how it may first appear—it relates the forward and reverse path ensembles given by Eqs. (6) and (15) that are driven by time-asymmetric protocols.

Although much more involved, the time-asymmetric fluctuation theorem may also be derived for underdamped dynamics through similar techniques. We include our derivation in Appendix C.

Significantly, the time-asymmetric fluctuation theorem may be exploited for efficient free-energy estimation. In particular, by considering optimizing two *different* protocols—one for the forward dynamics and the other for the reverse dynamics—the variance of ΔF estimates may be lowered by orders of magnitude. We now propose our algorithm demonstrating this.

IV. ALGORITHM

In this section we present an on-the-fly adaptive importance-sampling protocol optimization algorithm, inspired by Ref. [23], that uses the previously collected bidirectional samples (i.e., from both forward and reverse protocols) to iteratively discover lower-variance time-asymmetric protocols. Exploiting the mathematical structure of the Onsager-Machlup action, our algorithm requires minimal computational overhead, solely the inclusion of easily computable auxiliary variables in each trajectory's time evolution.

Concretely, we consider the objective function

$$J = J_F + J_R = \langle W \rangle_F + \langle \tilde{W} \rangle_R. \quad (17)$$

Jensen's inequality applied to Eq. (2) implies $\langle W \rangle_F \geq \Delta F$ and $\langle \tilde{W} \rangle_R \geq -\Delta F$, with equality only for zero-variance optimal protocols.

Our simulations are performed using the Euler-Mayurama method to discretize Eqs. (6) and (15). Instead of directly discretizing Eq. (7), we measure for every trajectory the expression derived from Eq. (13),

$$W[X(t)] = U_B(X(t_f)) - U_A(X(0)) + \beta^{-1} \ln \frac{\mathcal{P}[X(t)|X(0)]}{\tilde{\mathcal{P}}[\tilde{X}(t)|\tilde{X}(0)]}, \quad (18)$$

with the correct discrete-path probabilities, so as to preserve the fluctuation theorem Eq. (10). In our setting this may be written as

$$W[X(t)] = \{U_B(X(t_f)) + \tilde{S}[\tilde{X}(t)]\} - \{U_A(X(0)) + S[X(t)]\}. \quad (19)$$

From now on we use Einstein notation, where repeated upper and lower Greek indices signify summation. Let $\{U_\mu(x, t) | \mu = 1, \dots, M\}$ denote a set of time-dependent basis functions. We used a linear parametrization of the forward and reverse protocols $U_F = U_0 + U_1$ and $U_R = U_0 - U_1$:

$$U_{F,R}(x, t) = \begin{cases} U_A(x) & \text{for } t = 0 \\ \theta_{F,R}^\mu U_\mu(x, t) & \text{for } t \in (0, t_f) \\ U_B(x) & \text{for } t = t_f, \end{cases} \quad (20)$$

with parameters $\theta = (\theta_F, \theta_R) \in \mathbb{R}^{2M}$, which explicitly shows that the forward and reverse protocols are not constrained to be equal under time reversal. Having unequal $\theta_F \neq \theta_R$ yields $U_1 \neq 0$, which gives a counterdiabatic driving that is otherwise absent when forward and reverse protocols are constrained to be equal. By minimizing the work under this parametrization we find the parameters that yield an approximation to the true counterdiabatic force [i.e., a U_1 that satisfies the PDE (9)]. Under this parametrization, the Onsager-Machlup path action Eq. (12) and the time-asymmetric work Eq. (19) are quadratic in θ :

$$S[X(t); \theta] = \theta_F^\mu \theta_F^\nu \mathbf{a}_{\mu\nu} + \theta_F^\mu \mathbf{b}_\mu + \theta\text{-independent terms}, \quad (21)$$

$$W[X(t); \theta] = -(\theta_F^\mu \theta_F^\nu \mathbf{a}_{\mu\nu} + \theta_F^\mu \mathbf{b}_\mu + \mathbf{c}) + (\theta_R^\mu \theta_R^\nu \tilde{\mathbf{a}}_{\mu\nu} + \theta_R^\mu \tilde{\mathbf{b}}_\mu + \tilde{\mathbf{c}}), \quad (22)$$

where

$$\begin{aligned} \mathbf{a}_{\mu\nu}[X(t)] &= (\text{I}) \int_0^{t_f} \frac{\nabla U_\mu \cdot \nabla U_\nu}{4} dt, & \mathbf{b}_\mu[X(t)] &= (\text{I}) \int_0^{t_f} \frac{\dot{X} \cdot \nabla U_\mu}{2} dt, & \mathbf{c}[X(t)] &= U_A(X(0)), \\ \tilde{\mathbf{a}}_{\mu\nu}[X(t)] &= (\text{BI}) \int_0^{t_f} \frac{\nabla U_\mu \cdot \nabla U_\nu}{4} dt, & \tilde{\mathbf{b}}_\mu[X(t)] &= -(\text{BI}) \int_0^{t_f} \frac{\dot{X} \cdot \nabla U_\mu}{2} dt, & \tilde{\mathbf{c}}[X(t)] &= U_B(X(t_f)) \end{aligned} \quad (23)$$

are θ -independent functionals of the time-discretized trajectory $X(t)$ [24]. Here, (BI) refers to a backwards Itô integral, needed to write terms of the reverse ensemble $\tilde{S}[\tilde{X}(t)]$ as

a functional of $X(t)$. [Equations (21)–(23) apply for the reverse path ensemble $\tilde{S}[\tilde{X}(t)]$, $\tilde{W}[\tilde{X}(t)]$, through the transformation $t \rightarrow t_f - t$, $\{F, R\} \rightarrow \{R, F\}$.] These variables $\mathbf{a}, \tilde{\mathbf{a}} \in$

$\mathbb{R}^{M \times M}$, $\mathbf{b}, \tilde{\mathbf{b}} \in \mathbb{R}^M$, and $\mathbf{c}, \tilde{\mathbf{c}} \in \mathbb{R}$ are akin to the eligibility trace variables (sometimes called ‘‘Malliavin weights’’) used in reinforcement learning policy-gradient algorithms [25–29], which are dynamically evolved with each trajectory $X(t)$.

In the following two paragraphs we consider only the forward ensemble for simplicity. When a trajectory $X^i(t)$ is sampled, we calculate not only its work $W^i = W[X^i(t)]$ but also its functional values $(\mathbf{a}^i, \mathbf{b}^i, \mathbf{c}^i, \tilde{\mathbf{a}}^i, \tilde{\mathbf{b}}^i, \tilde{\mathbf{c}}^i)$ [Eq. (23)], which are saved alongside the protocol parameters $\theta^i = (\theta_F^i, \theta_R^i)$ employed to sample the trajectory. Given n_s sampled trajectories, the collected data $\{(W^i, \mathbf{a}^i, \mathbf{b}^i, \mathbf{c}^i, \tilde{\mathbf{a}}^i, \tilde{\mathbf{b}}^i, \tilde{\mathbf{c}}^i)\}_{i=1}^{n_s}$ may be used to construct a θ -dependent importance-sampling estimator for $\langle W \rangle_F$:

$$\hat{J}_F(\theta) = \frac{\sum_{i=1}^{n_s} r_F^i(\theta) \mathbf{w}_F^i(\theta)}{\sum_{i=1}^{n_s} r_F^i(\theta)}, \quad (24)$$

where the sum is over collected forward samples i , $r_F^i(\theta)$ is the likelihood ratio (i.e., the Radon–Nikodym derivative)

$$r_F^i(\theta) = \frac{\mathcal{P}[X^i(t) \text{ from } \theta]}{\mathcal{P}[X^i(t) \text{ from } \theta^i]} = e^{-\beta(S[X^i(t); \theta] - S[X^i(t); \theta^i])}. \quad (25)$$

Here, $S[X^i(t); \theta]$ and $\mathbf{w}^i(\theta) = W[X^i(t); \theta]$, defined in Eqs. (21) and (22), are the path action and time-asymmetric work for the trajectory $X^i(t)$ as if it were sampled under θ instead of the protocol θ^i it was actually sampled under. Protocols may now be optimized by using Eq. (24) as a surrogate objective function, which yields a newly optimized protocol θ^* that can be used to iteratively sample even-lower-variance trajectories.

Of course, the quality of the importance-sampling estimate Eq. (24) degrades the further away the input θ is from the set of θ^i under which samples were collected. One common heuristic of this degradation is the effective sample size [30]

$$n_F^{\text{eff}}(\theta) = \frac{[\sum_{i=1}^{n_s} r_F^i(\theta)]^2}{\sum_{i=1}^{n_s} r_F^i(\theta)^2}, \quad (26)$$

ranging from 1 (uneven r_F^i values, high degradation) to n_s (equal r_F^i values, low degradation).

We now state our algorithm (pseudocode given in Algorithm 1): at each iteration, an equal number of independent forward and reverse trajectories are simulated through Eqs. (6) and (15) using the U_F, U_R specified by current parameters θ_{curr} , with the time-asymmetric work W and auxiliary variables $(\mathbf{a}, \mathbf{b}, \mathbf{c}, \tilde{\mathbf{a}}, \tilde{\mathbf{b}}, \tilde{\mathbf{c}})$ of each trajectory dynamically calculated; then the protocol is updated through solving the nonlinear constrained optimization problem

$$\theta_{\text{next}} = \operatorname{argmin}_{\theta} \{ \hat{J}(\theta) | \{n_F^{\text{eff}}(\theta), n_R^{\text{eff}}(\theta)\} \geq f n_s \}, \quad (27)$$

for which there are efficient numerical solvers (e.g., SLSQP [31] pre-implemented in SciPy [32]). Here, $\hat{J}(\theta) = \hat{J}_F(\theta) + \hat{J}_R(\theta)$, $n_F^{\text{eff}}(\theta)$ and $n_R^{\text{eff}}(\theta)$ are constructed with the n_s forward and n_s reverse samples collected over all iterations, and $f \in [0, 1)$ is a hyperparameter specifying the constraint strength: the fraction of total samples we are constraining $n_{F,R}^{\text{eff}}$ to be greater than or equal. Finally, $\widehat{\Delta F}_{\text{BAR}}$ is calculated with the bidirectional work measurements collected across all iterations using Eq. (4), which is permitted by the satisfaction of Eq. (5).

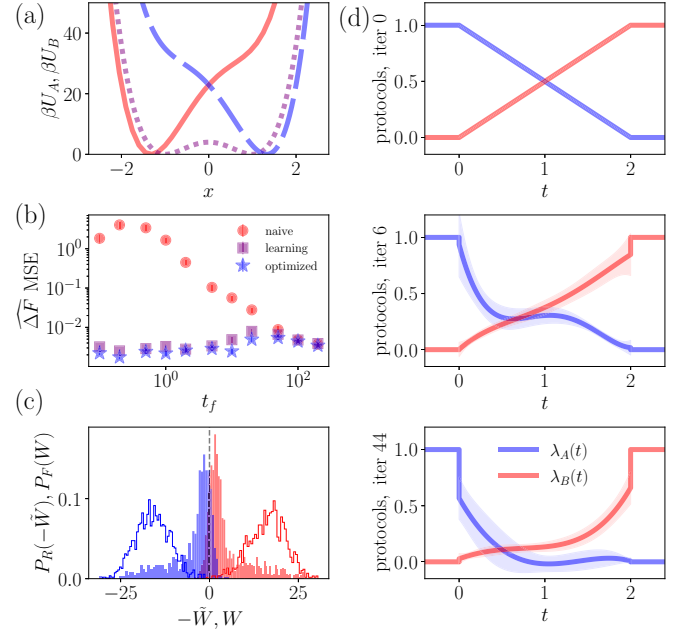


FIG. 1. (a) The potentials $U_A(x)$ (solid red) and $U_B(x)$ (dashed blue) are obtained by linearly biasing a double well (dotted purple). (b) $\widehat{\Delta F}_{\text{BAR}}$ mean squared error from 1000 bidirectional measurements drawn solely from the naive protocol (red circles), cumulatively from protocols that are adaptively optimized (‘‘learning’’) with our algorithm (purple squares), and solely from the last-iteration (‘‘optimized’’) protocols (blue stars) for various protocol times t_f . (c) Single-trial $\mathcal{P}_F(W)$ (red, right of the dashed vertical line) and $\mathcal{P}_R(-W)$ (blue, left of the dashed vertical line) work distributions for 1000 measurements from the naive protocol (unfilled), and adaptively optimized protocols (filled) for $t_f = 2$, the analytic ground truth $\Delta F = 0$ shown as a gray dashed line. Measurements made with protocol optimization have significantly more overlap, leading to lower estimator error. (d) Forward protocols $\lambda_A(t)$ [blue, going from $\lambda_A(0) = 1$ to $\lambda_A(t_f) = 0$] and $\lambda_B(t)$ [red, going from $\lambda_B(0) = 0$ to $\lambda_B(t_f) = 1$] at various iterations of protocol optimization for $t_f = 2$. Shaded region represents variability across 100 independent trials. In the optimized last-iteration protocol, $\lambda_A(t) + \lambda_B(t)$ (giving the energy scale) is greatly reduced at intermediate times, while $\lambda_B(t) - \lambda_A(t)$ (giving the linear bias) is time-asymmetrically shifted. The reverse protocols (not shown here) are similar, see Fig. 4

V. NUMERICAL EXAMPLES

In this section we report the performance of our algorithm for three test model systems.

We chose our basis set in order to represent protocols of the form

$$U(x, t) = \lambda_A(t)U_A(x) + \lambda_B(t)U_B(x) + \lambda_C(t)U_C(x), \quad (28)$$

where $U_C(x)$ is an additional quasicounterdiabatic potential

$$U_C(x) = -c \cdot x \text{ with } c = \frac{\langle x \rangle_B - \langle x \rangle_A}{t_f} \quad (29)$$

that provides a spatially uniform forcing proportional to the difference in equilibrium mean positions $\langle x \rangle_{A,B} = \int x \rho_{A,B}(x) dx$ [33], which is needed to, e.g., spatially translate an equilibrium distribution (see Appendix A in Ref. [34]). The

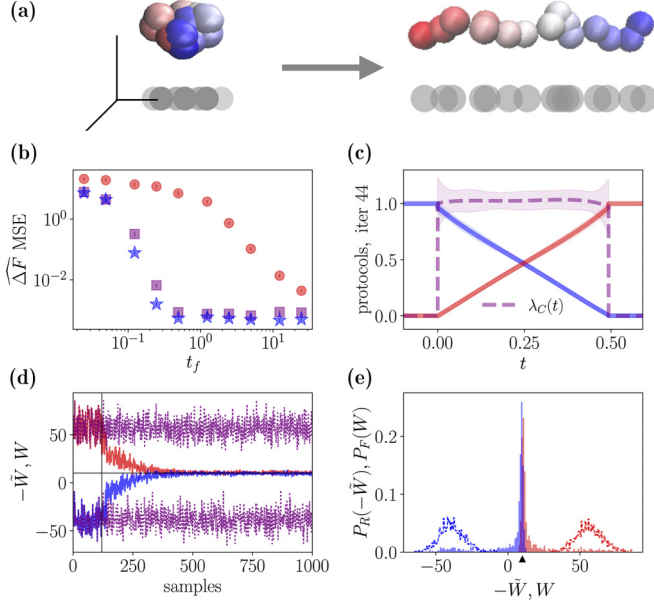


FIG. 2. (a) A Rouse polymer is stretched from a collapsed state to an extended one. (b) $\widehat{\Delta F}_{\text{BAR}}$ mean squared error versus protocol time, points colored as in Fig. 1. (c) For moderate protocol times ($t_f = 0.5\tau_R$ displayed here) the optimized protocol ($\lambda_A(t)$, $\lambda_B(t)$, $\lambda_C(t)$) learned in 44 iterations is the counterdiabatic protocol Eq. (31). Here, $\lambda_C(t)$ is drawn with the dashed purple lines. (d) Single-trial bidirectional work samples from the naive protocol (W and $-\tilde{W}$ dotted purple) and adaptively optimized protocols (W red starting above zero, $-\tilde{W}$ blue starting below zero) for $t_f = 0.5\tau_R$. Vertical line demarcates start of protocol optimization. The analytic ground truth value ΔF is shown as a horizontal line. (e) Work distributions corresponding to the samples in panel (d). The ground truth is indicated by the triangular arrow. Cumulative measurements made under protocol optimization (filled) have dramatically greater overlap than measurements made under the naive protocols (unfilled), leading to lower estimator error.

basis set is given by

$$\left\{ U_\ell(x) p_m \left(\frac{2t}{t_f} - 1 \right) \middle| \ell \in \{A, B, C\}, 0 \leq m \leq m_{\max} \right\}, \quad (30)$$

where $p_m(\cdot)$ denotes the m th Legendre polynomial.

For all numerical examples, $m_{\max} = 4$ was used; we found that using a larger m_{\max} did not substantially improve performance, while drastically increasing computational runtime (see discussion section below, where we discuss how computational runtime scales with number of parameters). The algorithm was initialized with 120 bidirectional samples drawn from a generic naive linear interpolation protocol $\lambda_A(t) = 1 - t/t_f$, $\lambda_B(t) = t/t_f$, $\lambda_C(t) = 0$ for both U_F and U_R . At each iteration Eq. (27) was solved for $n_{\text{mb}} = 20$ independently subsampled minibatches of size $n_s^{\text{mb}} = 80$ with $f = 0.3$; the protocol was then updated to the minibatch-averaged θ_{next} ; finally, 20 additional bidirectional samples were drawn with the new protocol. In total, 44 iterations were performed, giving 1000 bidirectional samples.

A. Linearly biased double well

The first system we consider is a one-dimensional quartic double well with a linear bias [Fig. 1(a)]. The potentials are $U_A(x) = E_0[(x^2 - 1)^2/4 - x]$, $U_B(x) = E_0[(x^2 - 1)^2/4 + x]$ (cf. Ref. [35] for optimal protocols minimizing $\langle W_{\text{trad}} \rangle_F$). We set $U_C(x) = 0$ because $U_B(x) - U_A(x)$ is already linear in x . We use $\beta = 1$, $E_0 = 16$, and a time step $\Delta t = 1 \times 10^{-3}\tau$, where $\tau = 1$ is the natural timescale (here the length scale, inverse temperature, and friction coefficient are all unity $\ell = \beta = \gamma = 1$).

Figure 1(b) displays the $\widehat{\Delta F}_{\text{BAR}}$ estimator mean squared error for 1000 bidirectional work measurements collected solely from the naive protocol (red), the 1000 measurements collected cumulatively over on-the-fly protocol optimization (purple), and 1000 measurements collected solely from the last iteration (blue). Each dot represents the empirical average over 100 independent trials. Note that the mean squared error is up to 1600 times lower under protocol optimization compared with under the naive protocol (obtained at $t_f = 0.2$). For $t_f \gtrsim 10$ the algorithm does not converge within the 1000 measurements (see Fig. 4), leading to less improvement. Figure 1(c) shows that bidirectional work measurements collected under the protocol optimization algorithm have significantly more overlap than measurements collected from the naive protocol, leading to reduced estimator error. Figure 1(d) gives snapshots on how the optimal protocol is adaptively learned. Note that the optimized time-asymmetric protocols feature discontinuous jumps occurring at $t = 0$ and $t = t_f$, which has been observed to be ubiquitous for time-symmetric protocols [35–44] (also see Ref. [45] for a recently proposed explanation for this phenomenon).

B. Rouse polymer

Next we consider a $(N + 1)$ -bead Rouse polymer [Fig. 2(a)] with stiffness k and intrinsic energy given by $U_{\text{Rouse}}(x_0, x_1, \dots, x_N) = \sum_{n=0}^{N-1} (k/2)(x_{n+1} - x_n)^2$ from harmonic bonds between adjacent beads [46,47]. We estimate ΔF between a collapsed state (fixing $x_0 = x_N = 0$) and an extended state (fixing $x_0 = 0$, $x_N = \lambda_f$), so our configuration space is $x \in \mathbb{R}^{N-1}$ with potential energies $U_A(x_1, \dots, x_{N-1}) = U_{\text{Rouse}}(0, x_1, \dots, x_{N-1}, 0)$ and $U_B(x_1, \dots, x_{N-1}) = U_{\text{Rouse}}(0, x_1, \dots, x_{N-1}, \lambda_f)$. Equilibrium averages $\langle x_n \rangle_A = 0$, $\langle x_n \rangle_B = n\lambda_f/N$ give $U_C(x) = -(\lambda_f/Nt_f) \sum_{n=1}^N nx_n$, and the analytic ground truth free energy (i.e., the true free-energy difference against which we can compare algorithm estimates) is given by $\Delta F = F_B - F_A = k\lambda_f^2/(2N)$. It may be verified that, for this problem, the time-varying potential energies

$$U_0(x, t) = \left(1 - \frac{t}{t_f}\right) U_A(x) + \left(\frac{t}{t_f}\right) U_B(x),$$

$$U_1(x, t) = U_C(x) \quad (31)$$

solve Eq. (9) and are thus counterdiabatic.

We use $\beta = k = 1$, $N = 20$, and time step $\Delta t = 2.5 \times 10^{-5}\tau_R$, where $\tau_R = \beta N^2/\pi^2$ is the Rouse relaxation time [46]. Initial conditions for $\rho_A(x)$ and $\rho_B(x)$ were drawn from a

normal-modes basis as described in Appendix D. Figure 2(b) shows an improvement of up to 8300 (for $t_f = 0.5\tau_R$) in $\widehat{\Delta F}_{\text{BAR}}$ mean squared error between naive and optimized protocols. The counterdiabatic solution Eq. (31) corresponds to $\lambda_A(t) = (1 - t/t_f)$, $\lambda_B(t) = t/t_f$, and $\lambda_C(t) = 1$, which what the algorithm learns for $t_f = \tau_R$ as depicted in Fig. 2(c). (This was generally the case for $t_f \geq 0.5\tau_R$. For $t_f < 0.5\tau_R$ the algorithm learns a suboptimal local solution that still provides some improvement, see Fig. 5.) Figure 2(d) shows single-trial traces of bidirectional work measurements for the naive protocol (purple) and adaptively optimized protocols (red for W , blue for $-\tilde{W}$), for $t_f = 0.5\tau_R$. The protocol converges in ≈ 20 iterations (requiring ≈ 500 measurements), and then consistently gives work measurements closely centered at the ground truth free energy (gray horizontal line). Histograms of these traces (filled) are shown in Fig. 2(e), exhibiting a remarkable increase in the overlap compared with their naive counterparts (unfilled).

C. Worm-like chain with attractive linker

We now consider a $(N + 1)$ -bead worm-like chain model (WLC) in two dimensions with an added Lennard-Jones interaction between the first and last beads (similar to the third example of Ref. [48]). Fixing $(x_0, y_0) = (0, 0)$, the configuration space is $\vec{\phi} \in \mathbb{R}^N$, where ϕ_n is the angle of the n th bond with respect to the x axis, with $(x_n(\vec{\phi}), y_n(\vec{\phi})) = (\sum_{m=1}^n \cos \phi_m, \sum_{m=1}^n \sin \phi_m)$. The angular potential $U_\phi = k \sum_{n=1}^{N-1} [1 - \cos(\phi_{n+1} - \phi_n)]$ penalizes the bending of adjacent bonds, and $U_{\text{LJ}} = 4\epsilon_{\text{LJ}}[(\sigma_{\text{LJ}}/r_N)^{12} - (\sigma_{\text{LJ}}/r_N)^6]$ specifies the interaction between first and last beads, where $r_N = (x_N^2 + y_N^2)^{1/2}$ is the end-to-end distance. We take $k = 6$, $\beta = 1$, $\epsilon_{\text{LJ}} = 8$, $\sigma_{\text{LJ}} = 4$, and $N = 15$.

Figure 3(a) displays the conditioned free energy $F(R/N) := -\beta^{-1} \ln \rho_{\text{eq}}(r_N = R)$, where ρ_{eq} is the equilibrium probability of observing the end-to-end distance under $U = U_\phi + U_{\text{LJ}}$ [49] (constructed from 10^7 equilibrium samples of U_ϕ , obtained with the Metropolis-adjusted Langevin Algorithm [50], that were reweighted by U_{LJ}). $F(R/N)$ exhibits a deep well for $R_A \approx 2^{1/6}\sigma_{\text{LJ}}$ (trapped or bent state) and a shallow well at large $R_B \approx 0.9N$ (free or relaxed state), separated by a barrier; their difference in value $\Delta F \approx 4.18$ may be calculated by estimating the ΔF between $U_A(\vec{\phi}) = U_\phi + U_{\text{LJ}} + (k_{\text{ext}}/2)(r_N - \lambda_i)^2$ and $U_B(\vec{\phi}) = U_\phi + U_{\text{LJ}} + (k_{\text{ext}}/2)(r_N - \lambda_f)^2$ for $\lambda_i = 2^{1/6}\sigma_{\text{LJ}}$, $\lambda_f = 0.9N$, and $k_{\text{ext}} \gg 1$.

We calculate the ΔF between U_A and U_B for $k_{\text{ext}} = 200$. We use time step $\Delta t = 1.41 \times 10^{-4}\tau_{\text{LJ}}$, where $\tau_{\text{LJ}} = (\epsilon_{\text{LJ}}/\sigma^2)^{1/2}$ is the Lennard-Jones timescale. We use $U_C(\vec{\phi}) = -\sum_n c_n r_n$, radially pulling on each individual bead, constructed with $c_n = (\langle r_n \rangle_B - \langle r_n \rangle_A)/t_f$ from equilibrium samples of ρ_A and ρ_B . Figure 3(b) displays single-trial work histograms for $t_f = 0.71\tau_{\text{LJ}}$, showing that work measurements made under our protocol optimization algorithm are much closer to the ground truth $\Delta F \approx 4.18$ (numerically obtained via the Metropolis-adjusted Langevin Algorithm, see previous paragraph), as opposed to using the naive protocol. Figure 3(c) shows the updating $\widehat{\Delta F}_{\text{BAR}}$ estimator over 100 independent trials converges substantially faster to the ground

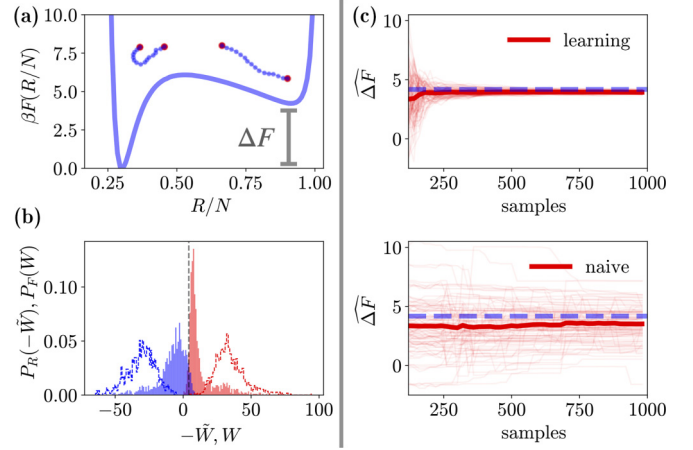


FIG. 3. Worm-like chain with an attractive linker. (a) Numerically obtained ground truth free-energy surface relative to its value at $R_A = 2^{1/6}\sigma_{\text{LJ}}$; the left well corresponds to the ends of the chain bound to one another and the right well corresponds to a nearly straight configuration, with a free-energy difference $\Delta F \approx 4.18$. (b) Work distributions before (unfilled) and during (filled) optimization for $t_f = 0.71$, the ground truth shown as a gray dashed line. (Same coloring as Fig. 1.) (c) The $\widehat{\Delta F}_{\text{BAR}}$ estimator updated over the 1000 samples converges to the ground truth value (dashed blue line) much more quickly under protocol optimization than under the naive protocol. At 1000 samples the protocol optimization free-energy estimate was $\widehat{\Delta F} = 3.94 \pm 0.11$, while for the naive protocol was $\widehat{\Delta F} = 3.52 \pm 1.48$ (cf. ground truth value of $\Delta F = 4.18$). It took 200 total samples for the mean squared error to drop below $1.00(k_B T)$ under protocol optimization

truth. With 1000 bidirectional samples under the naive protocol the mean squared error is $1.62(k_B T)$; under protocol optimization, only 160 samples (i.e., just after two iterations of protocol optimization) are required to have a smaller mean squared error. Over various t_f , the mean squared error is up to 120 times lower under protocol optimization compared with under the naive protocol, see Fig. 7.

VI. DISCUSSION

In this paper, we derived the time-asymmetric microscopic fluctuation theorem for the unconventional work introduced by Ref. [13]. We then demonstrated its practical utility for free-energy estimation by presenting an adaptive time-asymmetric protocol optimization algorithm, whose effectiveness we illustrated with three toy models of varying complexity. Time-asymmetric protocols have been considered before [34,51,52], but we use $\widehat{\Delta F}_{\text{BAR}}$ on bidirectional work measurements from adaptive time-asymmetric protocols. A clear next step is to test our algorithm on more physically realistic systems. This work should be straightforward to implement with JAX-MD [53]. In principle our algorithm should work with underdamped dynamics [51], and it should also be possible to adaptively optimize the protocol time t_f and sampling ratio n_F/n_R . Another future direction to pursue is to differentially weight early versus later samples in the estimator to account for differences in the variance of work

measurements, as the algorithm more closely approximates a counterdiabatic protocol.

The fast convergence in our method comes from exploiting of the quadratic structure of the Onsager-Machlup path action to construct $\hat{J}(\theta)$, which allows all samples to be used in each optimization step. Typically the most computationally expensive step in a molecular-dynamics simulation is calculating potential-energy gradients ∇U to evolve $X(t)$, which does not need to be repeated to evolve $(\mathbf{a}_{\mu\nu}, \mathbf{b}_\mu, \dots)$. A valid critique of our algorithm is that the number of auxiliary variables included with each trajectory scales quadratically with the number of basis functions, becoming prohibitively large when considering, for example, a separate control force on each particle of a many particle system. However, we have shown that a small number of basis functions to represent Eq. (28) already produces a substantial improvement in efficiency for our three examples. That said, it is straightforward to add additional basis functions [cf. Eq. (2) of Ref. [54]], which may be useful for more complex and realistic systems. It would be interesting to apply recent methods [55] to learn the optimal set of additional basis functions, that apply force along specific coordinates: bonds, angles, native contacts and other collective variables to further improve performance for larger-scale systems.

As mentioned at the end of the introduction, our results were independently derived in Ref. [15] within a machine-learning context. It is noteworthy that stochastic thermodynamics has shown to be a useful theoretical framework not only for nonequilibrium statistical physics, but also for machine learning in flow-based diffusion models [56–59]. In particular, we recognize significant ties between our work, and that of “stochastic normalizing flows” [60], wherein authors also consider constructing counterdiabatic protocols under the name “deterministic invertible functions.” It can be shown that counterdiabatic protocols are perfect stochastic normalizing flows, and they report (after sufficient neural-network training) excellent numerical results for sampling and free-energy estimation. The primary difference is that, in their work, they fix $U_0 = (1 - t/t_f)U_A + (t/t_f)U_B$ and use a neural-network ansatz, whereas here we use an adaptive importance sampling algorithm with a linear spatiotemporal basis ansatz for both U_0 and U_1 . Likewise, we note that Ref. [61] explores time-asymmetric Markovian processes for sampling, building off of entropy-regularized optimal transport wherein solutions of the continuous-time Schrödinger bridge problem involve asymmetrically controlled diffusion processes [62] (see also Refs. [63,64]). This is intriguing because solving for optimal time-*symmetric* protocols has been shown to be equivalent to solving the continuous-time formulation [65] of standard optimal transport [35,66–69]. In light of all this, we suspect deep theoretical connections between stochastic thermodynamics and machine learning may be further uncovered through the time-asymmetric fluctuation theorem.

Documented code for this project may be found online [70].

ACKNOWLEDGMENTS

The authors would like to thank Nilesh Tripuraneni, Hunter Akins, Chris Jarzynski, Gavin Crooks, Steve Strong, and the participants of the Les Houches “Optimal Transport Theory and Applications to Physics” and Flatiron “Measure Transport, Diffusion Processes and Sampling” workshops for useful discussions; Jorge L. Rosa-Raíces for helpful comments on an earlier paper version; and Evie Pai for lending personal computing resources. This research used the Savio computational cluster resource provided by the Berkeley Research Computing program at the University of California, Berkeley (supported by the University of California Berkeley Chancellor, Vice Chancellor for Research, and Chief Information Officer). A.Z. is supported by the Department of Defense (DoD) through the National Defense Science & Engineering Graduate (NDSEG) Fellowship Program. B.K.-S. is supported by the Kavli Energy Nanoscience institute through the Philomathia Foundation Fellowship. M.R.D. thanks Steve Strong and Fenrir LLC for supporting this project. This work was supported in part by the U.S. Army Research Laboratory and the U.S. Army Research Office under Contract W911NF-20-1-0151.

APPENDIX A: MICROSCOPIC FLUCTUATION THEOREM

1. The Onsager-Machlup action

For overdamped Langevin dynamics for $X(t) \in \mathbb{R}^d$,

$$\dot{X}(t) = -\nabla U(X(t), t) + \sqrt{2\beta^{-1}}\eta(t) \text{ with } X(0) \sim \rho(\cdot), \quad (\text{A1})$$

where $\eta(t)$ is an instantiation of standard Gaussian white noise with statistics $\langle \eta_i(t) \rangle = 0$ and $\langle \eta_i(t)\eta_j(t') \rangle = \delta_{ij}\delta(t-t')$, and $\rho(\cdot)$ is its initial distribution, the formal expression for the probability of a path’s realization is (up to a multiplicative factor)

$$\mathcal{P}[X(t)] = \rho(X(0))e^{-\beta S[X(t)]}. \quad (\text{A2})$$

Here $S[X(t)]$ is the Onsager-Machlup path action functional

$$S[X(t)] = (\text{I}) \int_0^{t_f} \frac{|\dot{X}(t) + \nabla U(X(t), t)|^2}{4} dt, \quad (\text{A3})$$

which comes from the path discretization into N time steps with time step $\Delta t = t_f/N$: $X(t) \rightarrow [X_0, X_1, \dots, X_N]$, with $X_n \approx X(t_n)$, $N = t_f/\Delta t$, and $t_n = n\Delta t$, generated from Euler-Maruyama dynamics,

$$X_0 \sim \rho(\cdot), \quad (\text{A4})$$

$$X_{n+1} = X_n - \nabla U(X_n, t_n)\Delta t + \sqrt{2\beta^{-1}}\Delta B_n, \quad (\text{A5})$$

where $\Delta B_n \sim \mathcal{N}(0, \Delta t I_d)$ is a d -dimensional Gaussian random variable (i.e., Brownian increment) [22].

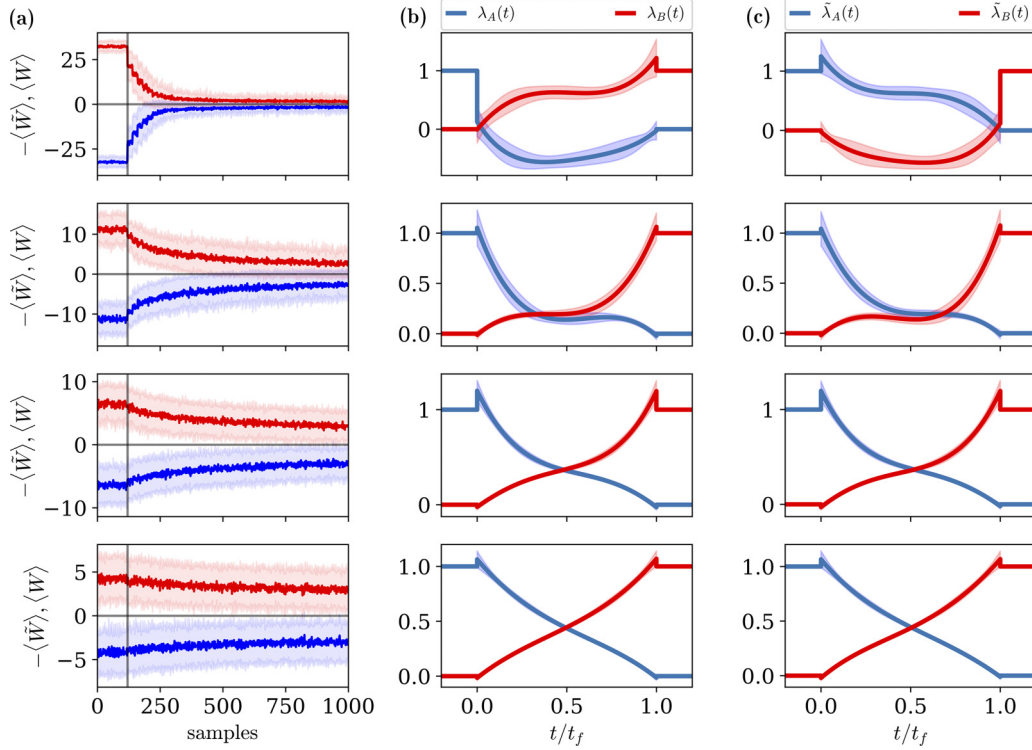


FIG. 4. (a) For the linearly biased double well, ensemble-averaged work across 100 trials, as a function of sample number. Same coloring as Fig. 4(d). Parameter optimization begins at 120 samples, and happens every 20 samples. The rows correspond to protocol times $t_f = 0.2, 5.0, 20.0$, and 50.0 , respectively. Convergence is slower for larger t_f ; for $t_f = 50.0$ the protocol may not have converged within 1000 samples. (b) The forward protocols $U_F(\cdot, t) = \lambda_A(t)U_A(\cdot) + \lambda_B(t)U_B(\cdot)$ after 1000 samples. Same coloring as Fig. 1(d). Rows correspond to the same t_f . (c) The reverse protocols $U_R(\cdot, t) = \tilde{\lambda}_A(t)U_A(\cdot) + \tilde{\lambda}_B(t)U_B(\cdot)$ after 1000 samples. The reverse protocols appear to satisfy $\tilde{\lambda}_A(t) = \lambda_B(t_f - t)$ and $\tilde{\lambda}_B(t) = \lambda_A(t)$, which is due to the symmetry of $U_A(\cdot)$ and $U_B(\cdot)$ in the problem.

The probability of the realization of a particular path is then

$$\begin{aligned}
 \mathcal{P}(X_0, X_1, \dots, X_N) &= \mathcal{P}(X_0)\mathcal{P}(X_1|X_0)\mathcal{P}(X_2|X_1) \cdots \mathcal{P}(X_N|X_{N-1}) \\
 &= \rho(X_0) \prod_{n=0}^{N-1} (4\pi\beta^{-1}\Delta t)^{-d/2} \exp\left(-\frac{|X_{n+1} - X_n + \nabla U(X_n, t_n)\Delta t|^2}{4\beta^{-1}\Delta t}\right) \\
 &\propto \rho(X_0) \exp\left(-\beta \sum_{n=0}^{N-1} \frac{|\frac{X_{n+1}-X_n}{\Delta t} + \nabla U(X_n, t_n)|^2}{4} \Delta t\right), \tag{A6}
 \end{aligned}$$

where the normalization factor $(4\pi\beta^{-1}\Delta t)^{-Nd/2}$ is hidden in the last line.

Taking $N \rightarrow \infty$ with $\Delta t = t_f/N \rightarrow 0$, the sum within the exponential becomes

$$\sum_{n=0}^{N-1} \frac{|\frac{X_{n+1}-X_n}{\Delta t} + \nabla U(X_n, t_n)|^2}{4} \Delta t \longrightarrow \text{(I)} \int_0^{t_f} \frac{|\dot{X}(t) + \nabla U(X(t), t)|^2}{4} dt = S[X(t)], \tag{A7}$$

which yields the formal expression Eq. (A2).

2. Stochastic integrals and Itô's formula

Here we briefly review the rules of stochastic calculus. For a stochastic path (i.e., a ‘‘Brownian motion’’) $X(t)|_{t \in [0, t_f]}$ from Eq. (A1) and some vector-valued function $b(x, t)$, the three following choices for the time-discretization of the integral

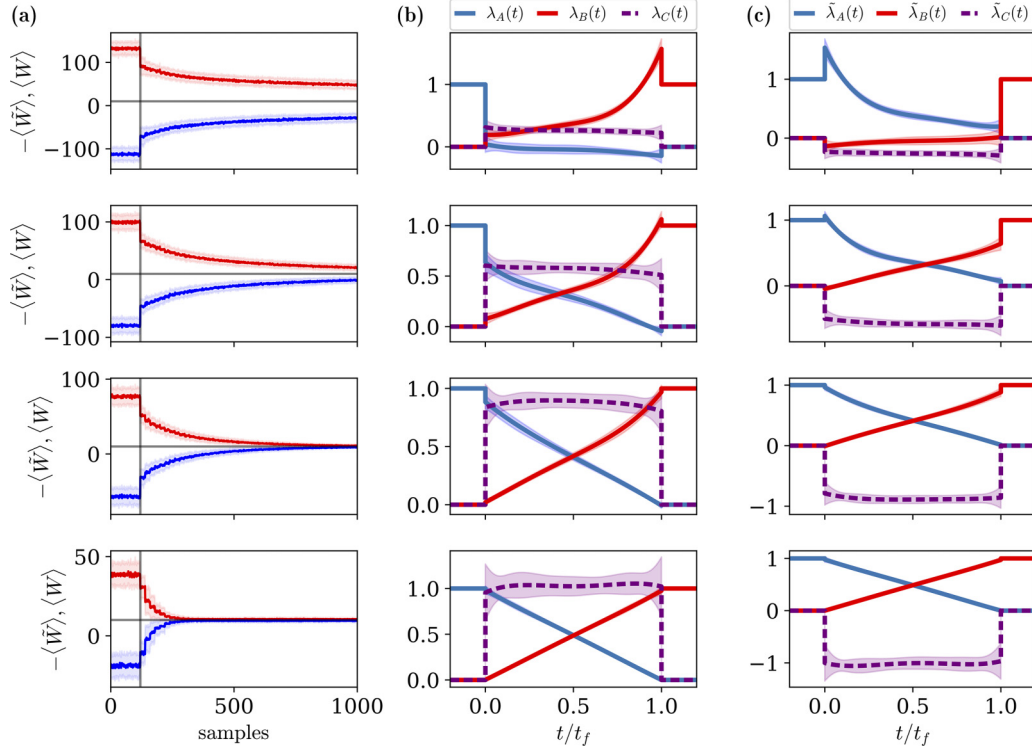


FIG. 5. (a) For the Rouse polymer, ensemble-averaged work across 100 trials, as a function of sample number. Parameter optimization begins at 120 samples, and happens every 20 samples. Same coloring as Fig. 4(d). The rows correspond to protocol times $t_f = 0.05\tau_R, 0.12\tau_R, 0.25\tau_R,$ and $1.23\tau_R,$ respectively. Unlike for the double well, convergence here is slower for smaller t_f . (b) The forward protocols $U_F(\cdot, t) = \lambda_A(t)U_A(\cdot) + \lambda_B(t)U_B(\cdot) + \lambda_C(t)U_C(\cdot)$ after 1000 samples. Same coloring as Fig. 4(c). Rows correspond to the same t_f . For $t_f = 0.05\tau_R$ and $t_f = 0.12\tau_R,$ the protocol has not yet converged to the counterdiabatic solution $\lambda_A(t) = (1 - t/t_f), \lambda_B(t) = t/t_f, \lambda_C = 1.$ (c) The reverse protocols $U_R(\cdot, t) = \tilde{\lambda}_A(t)U_A(\cdot) + \tilde{\lambda}_B(t)U_B(\cdot) + \tilde{\lambda}_C(t)U_C(\cdot)$ after 1000 samples. The reverse protocols appear to satisfy $\tilde{\lambda}_A(t) = \lambda_B(t_f - t), \tilde{\lambda}_B(t) = \lambda_A(t),$ and $\tilde{\lambda}_C(t) = -\lambda_C(t),$ which is due to the symmetry of $U_A(\cdot), U_B(\cdot),$ and $U_C(\cdot)$ in the problem.

$$\int_0^{t_f} b(X(t), t) \cdot \dot{X}(t) dt:$$

$$\sum_{n=0}^{N-1} b(X_n, t_n) \cdot \Delta X_n,$$

$$\sum_{n=0}^{N-1} b(X_{n+\frac{1}{2}}, t_{n+\frac{1}{2}}) \cdot \Delta X_n,$$

and

$$\sum_{n=0}^{N-1} b(X_{n+1}, t_{n+1}) \cdot \Delta X_n$$

[here $\Delta X_n = (X_{n+1} - X_n), X_{n+\frac{1}{2}} = (X_n + X_{n+1})/2,$ and $t_{n+\frac{1}{2}} = (t_n + t_{n+1})/2$] do not necessarily converge to the same value under the $N \rightarrow \infty, \Delta t = t_f/N \rightarrow 0$ limit. This is in contrast with the case where $X(t)|_{t \in [0, t_f]}$ is continuously differentiable, e.g., the solution of a well-behaved *deterministic* differential equation, in which case the above three time discretizations do converge to the same integral value under the limit [71].

Therefore, for trajectories $X(t)|_{t \in [0, t_f]}$ obtained through the stochastic differential equation Eq. (A1), we must define each

of these as distinct integrals;

$$(I) \int_0^{t_f} b(X(t), t) \cdot \dot{X} dt := \lim_{N \rightarrow \infty} \sum_{n=0}^{N-1} b(X_n, t_n) \cdot \Delta X_n,$$

$$(S) \int_0^{t_f} b(X(t), t) \cdot \dot{X} dt := \lim_{N \rightarrow \infty} \sum_{n=0}^{N-1} b(X_{n+\frac{1}{2}}, t_{n+\frac{1}{2}}) \cdot \Delta X_n,$$

and

$$(BI) \int_0^{t_f} b(X(t), t) \cdot \dot{X} dt := \lim_{N \rightarrow \infty} \sum_{n=0}^{N-1} b(X_{n+1}, t_{n+1}) \cdot \Delta X_n,$$

which are the Itô, Stratonovich, and backward Itô integrals, respectively. They are related to one another by Itô's lemma [72]

$$(I) \int_0^{t_f} b(X(t), t) \cdot \dot{X} + \beta \nabla \cdot b(X(t), t) dt$$

$$= (S) \int_0^{t_f} b(X(t), t) \cdot \dot{X} dt$$

$$= (BI) \int_0^{t_f} b(X(t), t) \cdot \dot{X} - \beta \nabla \cdot b(X(t), t) dt. \quad (A8)$$

The Stratonovich integration convention (i.e., with the time-symmetric midpoint-rule discretization) is particularly

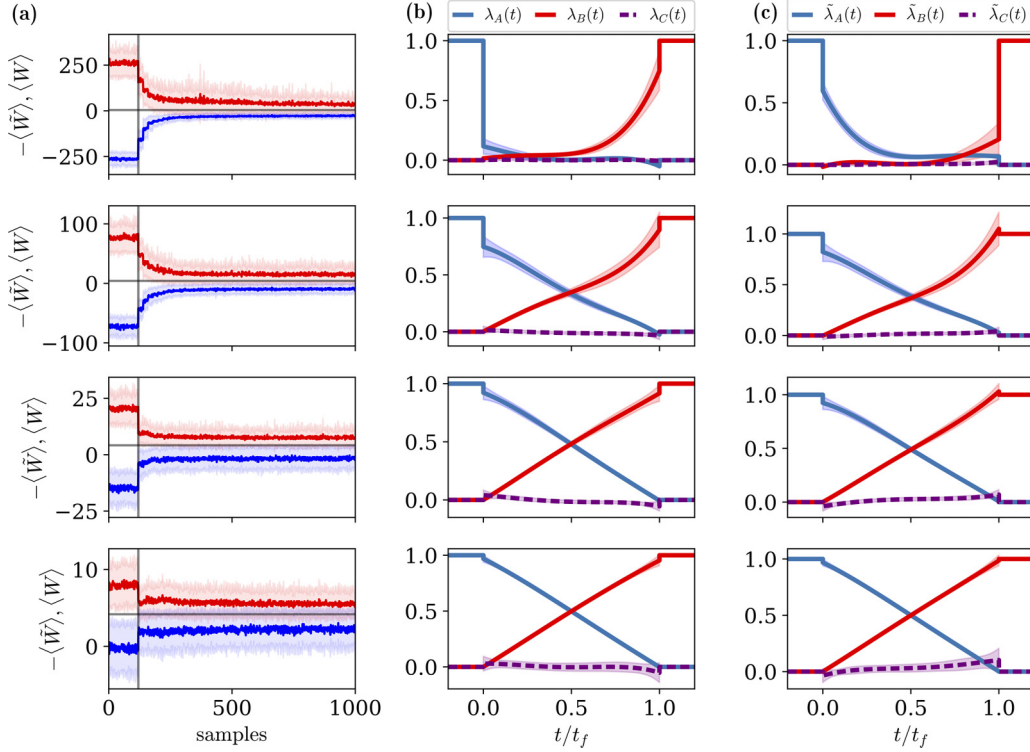


FIG. 6. (a) For the worm-like chain, ensemble-averaged work across 100 trials, as a function of sample number. Parameter optimization begins at 120 samples, and happens every 20 samples. The rows correspond to protocol times $t_f = 0.07\tau_{LJ}$, $0.28\tau_{LJ}$, $1.41\tau_{LJ}$, and $7.07\tau_{LJ}$, respectively. It appears convergence is reached rapidly, within the 1000 samples for all cases. Same coloring as Fig. 4(d). (b) The forward protocols $U_F(\cdot, t) = \lambda_A(t)U_A(\cdot) + \lambda_B(t)U_B(\cdot) + \lambda_C(t)U_C(\cdot)$ after 1000 samples. Same coloring as Fig. 4(c). Rows correspond to the same t_f . For small t_f , the protocol has reduced magnitude. This corresponds to lowering the potential, or raising the temperature [i.e., smaller $\beta U(\cdot, t)$]. (c) The reverse protocols $U_R(\cdot, t) = \tilde{\lambda}_A(t)U_A(\cdot) + \tilde{\lambda}_B(t)U_B(\cdot) + \tilde{\lambda}_C(t)U_C(\cdot)$ after 1000 samples. Due to the intrinsic asymmetry of the problem between pulled and collapsed states, the resulting reverse protocols do not obey the symmetries observed in the double-well and Rouse polymer problems.

convenient because ordinary calculus rules (e.g., the chain rule, product rule, etc.) apply.

Note that the three separate time-discretizations of integrals of the form $\int_0^{t_f} f(X(t), t) dt$:

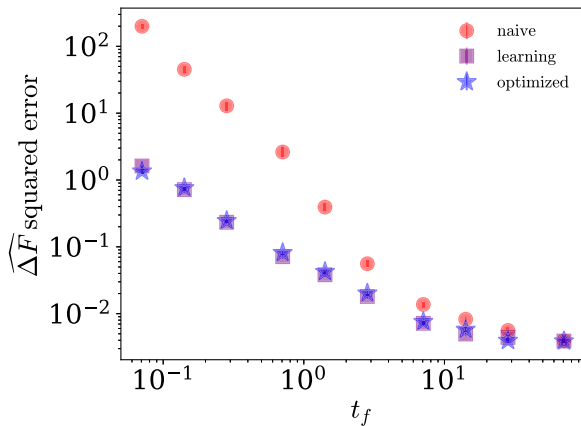


FIG. 7. Performance plot for the worm-like chain, same coloring as Fig. 1(b). That the mean squared error for protocol learning is near equal to the optimized protocol implies convergence occurs quickly within protocol optimization, cf. Fig. 6(a). At $t_f = 0.07\tau_{LJ}$, the MSE is 123.3 times lower under protocol optimization than under the naive protocol.

$$\sum_{n=0}^{N-1} f(X_n, t_n) \Delta t, \quad \sum_{n=0}^{N-1} f(X_{n+\frac{1}{2}}, t_{n+\frac{1}{2}}) \Delta t,$$

and

$$\sum_{n=0}^{N-1} f(X_{n+1}, t_{n+1}) \Delta t,$$

do converge to the same value under the $N \rightarrow \infty$ with $\Delta t = t_f/N \rightarrow 0$ limit, thus (I) $\int_0^{t_f} f(X(t), t) dt =$ (S) $\int_0^{t_f} f(X(t), t) dt =$ (BI) $\int_0^{t_f} f(X(t), t) dt$.

3. Microscopic fluctuation theorem derivation

Here, we use the stochastic calculus reviewed above to derive Eq. (10), i.e., the equivalence of its first and second lines. We start by manipulating the expression within

the exponent

$$\begin{aligned}
U_A(X(0)) + S[X(t)] + W[X(t)] &= U_B(X(t_f)) + (\text{I}) \int_0^{t_f} \left\{ \frac{|\dot{X} + \nabla(U_0 + U_1)|^2}{4} - (\dot{X} + \nabla U_1) \cdot \nabla U_0 + \beta^{-1} \nabla^2(U_1 - U_0) \right\} dt \\
&= U_B(X(t_f)) + \frac{1}{4}(\text{I}) \int_0^{t_f} |\dot{X}|^2 dt + \frac{1}{4}(\text{I}) \int_0^{t_f} |\nabla(U_1 - U_0)|^2 dt + \frac{1}{2}(\text{I}) \int_0^{t_f} \dot{X} \cdot \nabla(U_1 - U_0) \\
&\quad + 2\beta \nabla^2(U_1 - U_0) dt \\
&= U_B(X(t_f)) + \frac{1}{4}(\text{BI}) \int_0^{t_f} |\dot{X}|^2 dt + \frac{1}{4}(\text{BI}) \int_0^{t_f} |\nabla(U_1 - U_0)|^2 dt + \frac{1}{2}(\text{BI}) \int_0^{t_f} \dot{X} \cdot \nabla(U_1 - U_0) dt \\
&= U_B(X(t_f)) + (\text{BI}) \int_0^{t_f} \frac{|\dot{X} + \nabla(U_1 - U_0)|^2}{4} dt \\
&= U_B(\tilde{X}(0)) + (\text{I}) \int_0^{t_f} \frac{|\dot{X} + \nabla(\tilde{U}_1 - \tilde{U}_0)|^2}{4} dt \\
&= U_B(\tilde{X}(0)) + \tilde{S}[\tilde{X}(t)], \text{ where } \tilde{S}[\tilde{X}(t)] = (\text{I}) \int_0^{t_f} \frac{|\dot{X} + \nabla(\tilde{U}_0 - \tilde{U}_1)|^2}{4} dt. \tag{A9}
\end{aligned}$$

The first equality comes from using Itô's lemma $U_0(X(t_f), t_f) - U_0(X(0), 0) = (\text{I}) \int_0^{t_f} \partial_t U_0 + \dot{X} \cdot \nabla U_0 + \beta^{-1} \nabla^2 U_0 dt$. The second equality follows from standard algebraic manipulation. The third equality comes from converting the forward Itô integrals (I) to backward Itô integrals (BI) using Itô's formula (A8). The fourth equality results from standard algebraic manipulation. The fifth equality comes from the time-reversal transformation $t \rightarrow t_f - t$, with the backward Itô integral becoming a forward Itô integral under time reversal.

Finally, we plug in the above to the first line of Eq. (13) to obtain

$$\begin{aligned}
\tilde{\mathcal{P}}[\tilde{X}(t)] &= e^{-\beta\{U_A(X(0)) - F_A + S[X(t)] + W[X(t)] - \Delta F\}} \\
&= e^{-\beta\{U_B(\tilde{X}(0)) - F_B + \tilde{S}[\tilde{X}(t)]\}} \\
&= \rho_B(\tilde{X}(0)) e^{-\beta \tilde{S}[\tilde{X}(t)]}, \tag{A10}
\end{aligned}$$

thus completing our derivation.

APPENDIX B: DERIVING THE FLUCTUATION THEOREM FROM THE MICROSCOPIC FLUCTUATION THEOREM

In this section we derive the Crooks Fluctuation Theorem

$$\frac{\mathcal{P}_F(+W)}{\mathcal{P}_R(-W)} = e^{+\beta(W - \Delta F)} \tag{B1}$$

from the microscopic fluctuation theorem

$$\mathcal{P}[X(t)] e^{-\beta W[X(t)]} = \tilde{\mathcal{P}}[\tilde{X}(t)] e^{-\beta \Delta F}. \tag{B2}$$

We begin by recalling that $\mathcal{P}_F(\cdot)$, giving the probability of observing a particular work value in the forward ensemble, is defined as

$$\begin{aligned}
\mathcal{P}_F(w) &= \langle \delta(W - w) \rangle_F \\
&= \int DX(t) \mathcal{P}[X(t)] \delta(W[X(t)] - w), \tag{B3}
\end{aligned}$$

where we write w to distinguish the argument of $\mathcal{P}_F(\cdot)$ from the path-functional work $W = W[X(t)]$. Here $DX(t)$ denotes an integral over all paths $X(t)|_{t \in [0, t_f]}$, $\mathcal{P}[X(t)]$ is the probability of its realization, and $\delta(\cdot)$ is the Dirac δ function. Plugging Eq. (B2) into the above expression, we get

$$\begin{aligned}
\mathcal{P}_F(w) &= \int DX(t) \tilde{\mathcal{P}}[\tilde{X}(t)] e^{+\beta\{W[X(t)] - \Delta F\}} \delta(W[X(t)] - w) \\
&= e^{+\beta(w - \Delta F)} \int DX(t) \tilde{\mathcal{P}}[\tilde{X}(t)] \delta(W[X(t)] - w) \\
&= e^{+\beta(w - \Delta F)} \int D\tilde{X}(t) \tilde{\mathcal{P}}[\tilde{X}(t)] \delta(-\tilde{W}[\tilde{X}(t)] - w) \\
&= e^{+\beta(w - \Delta F)} \mathcal{P}_R(-w), \tag{B4}
\end{aligned}$$

where in the second line we pull out the exponential using the Dirac δ function, in the third line we consider the coordinate change $X(t) \rightarrow \tilde{X}(t)$ [using also $\tilde{W}[\tilde{X}(t)] = -W[X(t)]$, see Eq. (16) and the text that follows it in the main text], and in the fourth line we have recognized that the path integral expression is equivalent to the probability of observing the work value $-w$ in the reverse path ensemble.

APPENDIX C: TIME-ASYMMETRIC FLUCTUATION THEOREM FOR UNDERDAMPED DYNAMICS

In this section, we generalize the time-asymmetric fluctuation theorem to underdamped dynamics. To begin, we review three types of dynamics: overdamped, underdamped, and deterministic.

At inverse temperature β , overdamped Langevin dynamics for a stochastic trajectory $X(t) \in \mathbb{R}^d$ are given by the overdamped Langevin equation

$$\dot{X} = -\mu \nabla_x U_0(X(t), t) + \sqrt{2\mu\beta^{-1}} \eta(t), \tag{C1}$$

where μ is the mobility, and $\eta(t)$ is an instantiation of standard d -dimensional Gaussian white noise, i.e., with statistics $\langle \eta_i(t) \rangle = 0$, $\langle \eta_i(t) \eta_j(t') \rangle = \delta_{ij} \delta(t - t')$. Note that, for

ALGORITHM 1. Time-asymmetric protocol optimization via adaptive importance sampling.

```

1: inputs  $\beta, U_A(x), U_B(x)$ ; step size  $dt$ , number of time steps  $N$ , basis functions  $\{U_\mu(x, n)|_{n=1, \dots, N}\}$ , initial guess  $\theta_{\text{mit}}$ 
2: parameters Samples per iteration  $N_s$ , minibatches per iteration  $N_{\text{mb}}$ , minibatch size  $n_s^{\text{mb}}$ , constraint strength  $f$ 
3: given Methods  $\text{DrawSampleA}(), \text{DrawSampleB}()$  that return equilibrium samples from  $\rho_A, \rho_B$ 
4: output Iteratively updated  $\widehat{\Delta F}$  estimate
5:
6: function RunTrajFparameters  $\theta = (\theta_F, \theta_R)$  ▷ Euler-Maruyama method
7:   Obtain  $x_0 \leftarrow \text{DrawSampleA}()$ 
8:   Initialize  $x, \mathbf{a}_{\mu v}, \mathbf{b}_\mu, \tilde{\mathbf{a}}_{\mu v}, \tilde{\mathbf{b}}_\mu \leftarrow x_0, 0, 0, 0, 0$ 
9:   for  $n = 1, \dots, N$  do
10:     Evaluate  $\nabla U_\mu \leftarrow \nabla U_\mu(x, n)$  for each  $\mu$ 
11:     Calculate  $dx \leftarrow -\theta_F^\mu \nabla U_\mu dt + \sqrt{2\beta^{-1}} dB$ , where  $dB \sim \mathcal{N}(0, dt \times I_d)$  is a  $d$ -dimensional normal random variable
12:     Evaluate  $\nabla \tilde{U}_\mu \leftarrow \nabla U_\mu(x + dx, n)$  for each  $\mu$ 
13:     Evolve  $x \leftarrow x + dx$ 
14:     Evolve  $\mathbf{a}_{\mu v}, \mathbf{b}_\mu \leftarrow \mathbf{a}_{\mu v} + \nabla U_\mu \cdot \nabla U_v dt/4, \mathbf{b}_\mu + \nabla U_\mu \cdot dx/2$ 
15:     Evolve  $\tilde{\mathbf{a}}_{\mu v}, \tilde{\mathbf{b}}_\mu \leftarrow \tilde{\mathbf{a}}_{\mu v} + \nabla \tilde{U}_\mu \cdot \nabla \tilde{U}_v dt/4, \tilde{\mathbf{b}}_\mu - \nabla \tilde{U}_\mu \cdot dx/2$  ▷ This holds because  $d\tilde{x} = -dx$ .
16:   end for
17:   Evaluate  $\mathbf{c}, \tilde{\mathbf{c}} \leftarrow U_A(x_0), U_B(x)$ 
18:   Calculate  $W \leftarrow -(\theta_F^\mu \theta_F^v \mathbf{a}_{\mu v} + \theta_F^\mu \mathbf{b}_\mu + \mathbf{c}) + (\theta_R^\mu \theta_R^v \tilde{\mathbf{a}}_{\mu v} + \theta_R^\mu \tilde{\mathbf{b}}_\mu + \tilde{\mathbf{c}})$ 
19:   return  $W, \mathbf{a}_{\mu v}, \mathbf{b}_\mu, \mathbf{c}, \tilde{\mathbf{a}}_{\mu v}, \tilde{\mathbf{b}}_\mu, \tilde{\mathbf{c}}, \theta$ 
20: end function
21:
22: function RunTrajRparameters  $\theta = (\theta_F, \theta_R)$ 
23:   Obtain  $\tilde{x}_0 \leftarrow \text{DrawSampleB}()$ 
24:   Initialize  $\tilde{x}, \tilde{\mathbf{a}}_{\mu v}, \tilde{\mathbf{b}}_\mu, \mathbf{a}_{\mu v}, \mathbf{b}_\mu \leftarrow \tilde{x}_0, 0, 0, 0, 0$ 
25:   for  $n = 1, \dots, N$  do
26:     Evaluate  $\nabla \tilde{U}_\mu \leftarrow \nabla U_\mu(\tilde{x}, N + 1 - n)$  for each  $\mu$  ▷ because  $\nabla \tilde{U}_\mu(\cdot, n) = \nabla U_\mu(\cdot, N + 1 - n)$ 
27:     Calculate  $d\tilde{x} \leftarrow -\theta_R^\mu \nabla \tilde{U}_\mu dt + \sqrt{2\beta^{-1}} dB$ , where  $dB \sim \mathcal{N}(0, dt \times I_d)$  is a  $d$ -dimensional normal random variable
28:     Evaluate  $\nabla U_\mu \leftarrow \nabla U_\mu(\tilde{x} + d\tilde{x}, N + 1 - n)$  for each  $\mu$ 
29:     Evolve  $\tilde{x} \leftarrow \tilde{x} + d\tilde{x}$ 
30:     Evolve  $\tilde{\mathbf{a}}_{\mu v}, \tilde{\mathbf{b}}_\mu \leftarrow \tilde{\mathbf{a}}_{\mu v} + \nabla \tilde{U}_\mu \cdot \nabla \tilde{U}_v dt/4, \tilde{\mathbf{b}}_\mu + \nabla \tilde{U}_\mu \cdot d\tilde{x}/2$ 
31:     Evolve  $\mathbf{a}_{\mu v}, \mathbf{b}_\mu \leftarrow \mathbf{a}_{\mu v} + \nabla U_\mu \cdot \nabla U_v dt/4, \mathbf{b}_\mu - \nabla U_\mu \cdot d\tilde{x}/2$ 
32:   end for
33:   Evaluate  $\tilde{\mathbf{c}}, \mathbf{c} \leftarrow U_B(\tilde{x}_0), U_A(\tilde{x})$ 
34:   Calculate  $\tilde{W} \leftarrow -(\theta_R^\mu \theta_R^v \tilde{\mathbf{a}}_{\mu v} + \theta_R^\mu \tilde{\mathbf{b}}_\mu + \tilde{\mathbf{c}}) + (\theta_F^\mu \theta_F^v \mathbf{a}_{\mu v} + \theta_F^\mu \mathbf{b}_\mu + \mathbf{c})$ 
35:   return  $\tilde{W}, \tilde{\mathbf{a}}_{\mu v}, \tilde{\mathbf{b}}_\mu, \tilde{\mathbf{c}}, \mathbf{a}_{\mu v}, \mathbf{b}_\mu, \mathbf{c}, \theta$ 
36: end function
37:
38: function UpdateThetaforward samples  $\mathcal{S}_F$ , reverse samples  $\mathcal{S}_R$ 
39:   Initialize  $\mathcal{S}_{\theta, \text{mb}} \leftarrow \{\}$ 
40:   repeat  $N_{\text{mb}}$  times ▷ Use larger  $N_{\text{mb}}$  for larger  $|\mathcal{S}_F|$ .
41:     Randomly select  $\mathcal{S}_F^{\text{mb}} \subset \mathcal{S}_F$  of size  $n_s^{\text{mb}}$  without replacement
42:     Randomly select  $\mathcal{S}_R^{\text{mb}} \subset \mathcal{S}_R$  of size  $n_s^{\text{mb}}$  without replacement
43:      $\theta^* \leftarrow \text{argmin}_\theta \{ \hat{J}_F(\theta; \mathcal{S}_F^{\text{mb}}) + \hat{J}_R(\theta; \mathcal{S}_R^{\text{mb}}) \mid n_F^{\text{eff}}(\theta; \mathcal{S}_F^{\text{mb}}) \geq f n_s^{\text{mb}}, n_R^{\text{eff}}(\theta; \mathcal{S}_R^{\text{mb}}) \geq f n_s^{\text{mb}} \}$ 
44:      $\mathcal{S}_{\theta, \text{mb}}.insert(\theta^*)$ 
45:   end
46:   return  $\text{mean}(\mathcal{S}_{\theta, \text{mb}})$ 
47: end function
48:
49: procedure Main
50:   Initialize parameters  $\theta \leftarrow \theta_{\text{mit}}$  and sample arrays  $\mathcal{S}_F, \mathcal{S}_R \leftarrow \{\}, \{\}$ 
51:   repeat
52:     repeat  $N_s$  ▷ Use  $N_{\text{mb}} + N_s$  on first iteration.
53:        $\mathcal{S}_F.insert(\text{RunTrajF}(\theta))$ 
54:        $\mathcal{S}_R.insert(\text{RunTrajR}(\theta))$ 
55:     end
56:     Update estimate  $\widehat{\Delta F} \leftarrow \widehat{\Delta F}_{\text{BAR}}(\mathcal{S}_F, \mathcal{S}_R)$ 
57:      $\theta \leftarrow \text{UpdateTheta}(\mathcal{S}_F, \mathcal{S}_R)$ 
58:   until out of computer time
59: endprocedure

```

generality, the mobility is *not* set to one, in contrast with the main text.

On the other hand, underdamped dynamics for position and momentum variables $X(t) \in \mathbb{R}^d$, $P(t) \in \mathbb{R}^d$ are given by the underdamped Langevin equation

$$\dot{X} = \frac{P(t)}{m},$$

$$\dot{P} = -\nabla_x U_0(X(t), t) - \gamma P(t) + \sqrt{2\gamma\beta^{-1}}\zeta(t), \quad (\text{C2})$$

where m is the mass, γ is the friction coefficient, and $\zeta(t)$ is also an instantiation of standard d -dimensional Gaussian white noise with $\langle \zeta_i(t) \rangle = 0$, $\langle \zeta_i(t)\zeta_j(t') \rangle = \delta_{ij}\delta(t-t')$ [73]. Importantly, this underdamped dynamics has the Hamiltonian

$$H(x, p, t) = \frac{|p|^2}{2m} + U_0(x, t). \quad (\text{C3})$$

Note that taking the limit $\gamma \rightarrow 0$ reproduces standard Hamiltonian mechanics.

$$\dot{X} = b_1^x(X(t), P(t), t) + \frac{P(t)}{m} + \{-\mu\nabla_x U_0(X(t), t) + \sqrt{2\mu\beta^{-1}}\eta(t)\},$$

$$\dot{P} = b_1^p(X(t), P(t), t) - \nabla_x U_0(X(t), t) - \gamma P(t) + \sqrt{2m\gamma\beta^{-1}}\zeta(t),$$

$$\text{with } X(0), P(0) \sim \rho_A,$$

(C5)

where ρ_A is the equilibrium distribution corresponding to the Hamiltonian $H_A(x, p) = H(x, p, 0)$ at time $t = 0$, to be specified below. By considering m , μ , and γ as independent parameters, overdamped dynamics are reproduced by taking the limit $\gamma \rightarrow 0$, $m \rightarrow \infty$ with the assumption $b_1^x(x, p, t)$ has no p dependence, while underdamped dynamics are reproduced under the limit $\mu \rightarrow 0$, with further taking $\gamma \rightarrow 0$ yielding deterministic Hamiltonian dynamics.

2. Time-asymmetric work and path action

In this general setting, we consider protocols that “switch” the Hamiltonian between

$$H(x, p, 0) = H_A(x, p) = \frac{|p|^2}{2m} + U_A(x)$$

$$\rightarrow H(x, p, t_f) = H_B(x, p) = \frac{|p|^2}{2m} + U_B(x),$$

by switching the potential energy $U(x, 0) = U_A(x) \rightarrow U(x, t_f) = U_B(x)$. The equilibrium distributions for H_A and H_B are given by $\rho_A(x, p) = e^{-\beta[H_A(x, p) - F_A]}$ and $\rho_B(x, p) = e^{-\beta[H_B(x, p) - F_B]}$, with $F_A = -\beta^{-1} \ln \int e^{-\beta H_A} dx dp$ and $F_B = -\beta^{-1} \ln \int e^{-\beta H_B} dx dp$ denoting the equilibrium free energy. For ease of notation, from here on out we denote a phase-space trajectory with variable $Z(t) := (X(t), P(t))$ [not to be

Finally, we explicitly write out the deterministic dynamics under a flow field $b_1(x, p, t) \in \mathbb{R}^{2d}$ that applies to both position and momentum variables

$$\begin{aligned} \dot{X} &= b_1^x(X(t), P(t), t), \\ \dot{P} &= b_1^p(X(t), P(t), t). \end{aligned} \quad (\text{C4})$$

This vector flow field b_1 is a generalization of the time-asymmetric (gradient) force provided by $-\nabla U_1$ in the main text.

1. Generalized Langevin dynamics

For our derivation, we define a hybridized Langevin equation that combines all of the above three dynamics:

confused with the partition function $\exp(-\beta F)$ often denoted by the same variable].

The unconventional work for a stochastic trajectory $Z(t)|_{t \in [0, t_f]}$ is given by

$$W[Z(t)] = \int_0^{t_f} \frac{\partial H}{\partial t} + b_1 \cdot \nabla H - \beta^{-1} \nabla \cdot b_1 dt \quad (\text{C6})$$

(we use notation $b_1 \cdot \nabla H = b_1^x \cdot \nabla_x H + b_1^p \cdot \nabla_p H$ and $\nabla \cdot b_1 = \nabla_x \cdot b_1^x + \nabla_p \cdot b_1^p$) [13], while the probability of observing the trajectory under Eq. (C5) is

$$\mathcal{P}[Z(t)] = \rho_A(Z(0))e^{-\beta S[Z(t)]}, \quad (\text{C7})$$

with the path action given by the Itô integral

$$S[Z(t)] = (\text{I}) \int_0^{t_f} \frac{|\dot{X} - b_1^x - P/m + \mu\nabla_x U_0|^2}{4\mu} + \frac{|\dot{P} - b_1^p + \nabla_x U_0 + \gamma P|^2}{4m\gamma} dt.$$

This may be seen by considering the probability of obtaining the particular realization of the noise terms $\eta(t)$ and $\zeta(t)$ that produce the trajectory $Z(t) = (X(t), P(t))$.

3. Derivation of microscopic fluctuation theorem

As with the detailed derivation for the overdamped case in Appendix A, we manipulate the sum

$$H_A(Z(0)) + S[Z(t)] + W[Z(t)] = H_B(Z(t_f)) + (\text{I}) \int_0^{t_f} \frac{|\dot{X} - b_1^x - P/m + \mu\nabla_x U_0|^2}{4\mu} + \frac{|\dot{P} - b_1^p + \nabla_x U_0 + \gamma P|^2}{4m\gamma} - (\dot{Z} - b_1) \cdot \nabla H - \beta^{-1} (\nabla \cdot b_1 + \mu\nabla_x^2 H + m\gamma\nabla_p^2 H) dt$$

$$\begin{aligned}
&= H_B(Z(t_f)) + (\text{I}) \int_0^{t_f} \frac{|\dot{X}|^2}{4\mu} + \frac{|\dot{P}|^2}{4m\gamma} dt + (\text{I}) \int_0^{t_f} \frac{|-b_1^x - P/m + \mu \nabla_x U_0|^2}{4\mu} \\
&\quad + \frac{|-b_1^p + \nabla_x U_0 + \gamma P|^2}{4m\gamma} dt + \frac{1}{2} (\text{I}) \int_0^{t_f} \dot{X} \cdot \left(\frac{-b_1^x - P/m + \mu \nabla_x U_0}{\mu} \right) \\
&\quad + \dot{P} \cdot \left(\frac{-b_1^p + \nabla_x U_0 + \gamma P}{m\gamma} \right) dt + (\text{I}) \int_0^{t_f} (b_1^x - \dot{X}) \cdot \nabla_x U_0 + (b_1^p - \dot{P}) \cdot (P/m) \\
&\quad - \beta^{-1} \{ \nabla_x \cdot (b_1^x + \mu \nabla_x U_0) + \nabla_p \cdot [b_1^p + m\gamma(P/m)] \} dt. \tag{C8}
\end{aligned}$$

The first equality comes from plugging in the work and path action definitions and applying the total time derivative with Itô's lemma $H_B(Z(t_f)) - H_A(Z(0)) = (\text{I}) \int_0^{t_f} \partial_t H + Z \cdot \nabla H + \beta^{-1} (\mu \nabla_x^2 H + m\gamma \nabla_p^2 H) dt$ [here the divergence terms come from the Gaussian white noise on $X(t)$ and $P(t)$, see Appendix A], while the second equality is obtained by expanding out each of the squared terms.

Continuing with our derivation, we insert $\nabla_x \cdot p = \nabla_p \cdot \nabla_x U_0 = 0$ into the integral, as well as apply the following two algebraic manipulations:

$$\begin{aligned}
&(\text{I}) \int_0^{t_f} \frac{|-b_1^x - P/m + \mu \nabla_x U_0|^2}{4\mu} + \frac{|-b_1^p + \nabla_x U_0 + \gamma P|^2}{4m\gamma} dt + (\text{I}) \int_0^{t_f} b_1^x \cdot \nabla_x U_0 + b_1^p \cdot (P/m) dt \\
&= (\text{I}) \int_0^{t_f} \frac{|(b_1^x + P/m) + \mu \nabla_x U_0|^2}{4\mu} + [b_1^x + (P/m)] \cdot \nabla_x U_0 - (\nabla_x U_0 - b_1^p) \cdot (P/m) + \frac{|(-b_1^p + \nabla_x U_0) + \gamma P|^2}{4m\gamma} dt \\
&= (\text{I}) \int_0^{t_f} \frac{|b_1^x + P/m + \mu \nabla_x U_0|^2}{4\mu} + \frac{|-b_1^p + \nabla_x U_0 - \gamma P|^2}{4m\gamma} dt, \tag{C9}
\end{aligned}$$

and

$$\begin{aligned}
&\frac{1}{2} (\text{I}) \int_0^{t_f} \dot{X} \cdot \left(\frac{-b_1^x - P/m + \mu \nabla_x U_0}{\mu} \right) + \dot{P} \cdot \left(\frac{-b_1^p + \nabla_x U_0 + \gamma P}{m\gamma} \right) dt - (\text{I}) \int_0^{t_f} \dot{X} \cdot \nabla_x U_0 + \dot{P} \cdot (P/m) dt \\
&= \frac{1}{2} (\text{I}) \int_0^{t_f} (-\dot{X}) \cdot \left(\frac{b_1^x + P/m + \mu \nabla_x U_0}{\mu} \right) + \dot{P} \cdot \left(\frac{-b_1^p + \nabla_x U_0 - \gamma P}{m\gamma} \right) dt, \tag{C10}
\end{aligned}$$

ultimately yielding

$$\begin{aligned}
&H_A(Z(0)) + S[Z(t)] + W[Z(t)] \\
&= H_B(Z(t_f)) + (\text{I}) \int_0^{t_f} \frac{|\dot{X}|^2}{4\mu} + \frac{|\dot{P}|^2}{4m\gamma} dt + (\text{I}) \int_0^{t_f} \frac{|b_1^x + P/m + \mu \nabla_x U_0|^2}{4\mu} + \frac{|-b_1^p + \nabla_x U_0 - \gamma P|^2}{4m\gamma} dt \\
&\quad + \frac{1}{2} (\text{I}) \int_0^{t_f} (-\dot{X}) \cdot \left(\frac{b_1^x + P/m + \mu \nabla_x U_0}{\mu} \right) + \dot{P} \cdot \left(\frac{-b_1^p + \nabla_x U_0 - \gamma P}{m\gamma} \right) \\
&\quad - 2\beta^{-1} \left\{ \mu \nabla_x \cdot \left(\frac{b_1^x + \mu \nabla_x U_0 + P/m}{\mu} \right) - m\gamma \nabla_p \cdot \left(\frac{-b_1^p - \gamma P + \nabla_x U_0}{m\gamma} \right) \right\} dt \tag{C11}
\end{aligned}$$

$$\begin{aligned}
&= H_B(Z(t_f)) + (\text{BI}) \int_0^{t_f} \frac{|\dot{X}|^2}{4\mu} + \frac{|\dot{P}|^2}{4m\gamma} dt + (\text{BI}) \int_0^{t_f} \frac{|b_1^x + P/m + \mu \nabla_x U_0|^2}{4\mu} + \frac{|-b_1^p + \nabla_x U_0 - \gamma P|^2}{4m\gamma} dt \\
&\quad + \frac{1}{2} (\text{BI}) \int_0^{t_f} (-\dot{X}) \cdot \left(\frac{b_1^x + P/m + \mu \nabla_x U_0}{\mu} \right) + \dot{P} \cdot \left(\frac{-b_1^p + \nabla_x U_0 - \gamma P}{m\gamma} \right) dt \\
&= H_B(Z(t_f)) + (\text{BI}) \int_0^{t_f} \frac{|-\dot{X} + b_1^x + P/m + \mu \nabla_x U_0|^2}{4\mu} + \frac{|\dot{P} - b_1^p + \nabla_x U_0 - \gamma P|^2}{4m\gamma} dt \tag{C12}
\end{aligned}$$

$$= H_B(\tilde{Z}(0)) + (\text{I}) \int_0^{t_f} \frac{|\dot{\tilde{X}} + \tilde{b}_1^x - \tilde{P}/m + \mu \nabla_x \tilde{U}_0|^2}{4\mu} + \frac{|\dot{\tilde{P}} + \tilde{b}_1^p + \nabla_x \tilde{U}_0 + \gamma \tilde{P}|^2}{4m\gamma} dt =: H_B(\tilde{Z}(0)) + \tilde{S}[\tilde{Z}(t)]. \tag{C13}$$

Here, in the second equality we use Itô's lemma to transform from forward to backward Itô integrals; in the third equality we express the terms back into the squared expression; and in the fourth equality we perform the time-reversal change of variables $\tilde{X}(t) = X(t_f - t)$, $\tilde{P}(t) = -P(t_f - t)$ with $\tilde{b}_1^x(x, p, t) = b_1^x(x, -p, t_f - t)$, $\tilde{U}_0(x, t) = U_0(x, t_f - t)$, and $\tilde{b}_1^p(x, p, t) = -b_1^p(x, -p, t_f - t)$, which transforms the backward Itô integral back into a forward Itô integral.

Here, $\tilde{S}[\tilde{Z}(t)]$ is the path action for the generalized Langevin dynamics,

$$\begin{aligned}\dot{\tilde{X}} &= -\tilde{b}_1^x(\tilde{X}(t), \tilde{P}(t), t) + \frac{\tilde{P}(t)}{m} + \{-\mu \nabla_x \tilde{U}_0(\tilde{X}(t), t) + \sqrt{2\mu\beta^{-1}}\eta(t)\}, \\ \dot{\tilde{P}} &= -\tilde{b}_1^p(\tilde{X}(t), \tilde{P}(t), t) - \nabla_x \tilde{U}_0(\tilde{X}(t), t) - \gamma \tilde{P}(t) + \sqrt{2m\gamma\beta^{-1}}\zeta(t),\end{aligned}\quad (\text{C14})$$

which from inspection differs from Eq. (C5) by having minus signs in front of the b_1^x and b_1^p terms. After specifying the initial conditions $\tilde{X}(0), \tilde{P}(0) \sim \rho_B$, we have a path ensemble for which the probability of observing a particular trajectory $\tilde{Z}(t)|_{t=0}^{t_f}$ satisfies

$$\begin{aligned}\hat{P}[\tilde{Z}(t)] &= \rho_B(\tilde{Z}(0))e^{-\beta\tilde{S}[\tilde{Z}(t)]} \\ &= e^{-\beta\{H_B(\tilde{Z}(0)) - F_B + \tilde{S}[\tilde{Z}(t)]\}} \\ &= e^{-\beta\{H_A(\tilde{Z}(0)) + S[\tilde{Z}(t)] + W[\tilde{Z}(t)] - F_B\}} \\ &= \rho_A(\tilde{Z}(0))e^{-\beta\{S[\tilde{Z}(t)] + F_A - F_B + W[\tilde{Z}(t)]\}} \\ &= P[\tilde{Z}(t)]e^{-\beta\{W[\tilde{Z}(t)] - \Delta F\}},\end{aligned}\quad (\text{C15})$$

namely, the time-asymmetric microscopic fluctuation theorem for the generalized dynamics Eq. (C5).

APPENDIX D: INITIAL SAMPLES FOR ROUSE POLYMER FROM NORMAL-MODE DECOMPOSITION

For the Rouse polymer, random samples may be drawn by exploiting a normal-mode decomposition. We can write

$$\begin{aligned}U_{A,B}(x_1, x_2, \dots, x_{N-1}) \\ = \tilde{U}\left(x_1 - \frac{\lambda_{i,f}}{N}, x_2 - \frac{2\lambda_{i,f}}{N}, \dots, x_{N-1} - \frac{(N-1)\lambda_{i,f}}{N}\right) \\ + \frac{k\lambda_{i,f}^2}{2N},\end{aligned}\quad (\text{D1})$$

where $\tilde{U}(\tilde{y}) = \tilde{y}^T K \tilde{y}/2$ with

$$K = \begin{pmatrix} 2k & -k & & & & \\ -k & 2k & -k & & & \\ & -k & & \ddots & & \\ & & & \ddots & -k & \\ 0 & & & & -k & 2k & -k \\ & & & & & -k & 2k \end{pmatrix}.\quad (\text{D2})$$

Then, we can do an eigenmode decomposition of K , writing

$$\tilde{U}(\tilde{y}) = \hat{U}(\tilde{z}) = \sum_{n=1}^{N-1} \frac{\kappa_n z_n^2}{2},$$

where

$$\kappa_n = 2\left[1 - \cos\left(\frac{\pi n}{N}\right)\right],\quad (\text{D3})$$

$$z_n = \sqrt{\frac{2}{N}} \sum_{m=1}^{N-1} \sin\left(\frac{2\pi nm}{N}\right) y_m.\quad (\text{D4})$$

Finally, for an individual initial condition, we draw the normal random variable $z_n \sim \mathcal{N}(\mu = 0, \sigma^2 = (\beta\kappa_n)^{-1})$ for each n because $\hat{\rho}(\tilde{z}) \propto \prod \exp(-\beta\kappa_n z_n^2/2)$; then we convert from \tilde{z} to \tilde{y} coordinates via

$$y_n = \sqrt{\frac{2}{N}} \sum_{m=1}^{N-1} \sin\left(\frac{2\pi nm}{N}\right) z_m,\quad (\text{D5})$$

before finally adding $x_n = y_n + n\lambda_{i,f}/N$ to get our initial condition.

Incidentally, the $k\lambda_{i,f}^2/2N$ in the expression comparing $U_{A,B}(\tilde{x})$ to $U(\tilde{y})$ is, up to an additive constant, the free energy F_A or F_B .

-
- [1] Z. Cournia, B. Allen, and W. Sherman, *J. Chem. Inf. Model.* **57**, 2911 (2017).
[2] G. Tawa, I. Topol, S. Burt, R. Caldwell, and A. Rashin, *J. Chem. Phys.* **109**, 4852 (1998).
[3] C. P. Kelly, C. J. Cramer, and D. G. Truhlar, *J. Phys. Chem. B* **110**, 16066 (2006).
[4] C. Chipot and A. Pohorille, *Free Energy Calculations*, Vol. 86 (Springer, Berlin, Heidelberg, 2007).
[5] M. R. Shirts, D. L. Mobley, and S. P. Brown, *Drug Design* **1**, 61 (2010).
[6] Z. Cournia, C. Chipot, B. Roux, D. M. York, and W. Sherman, in *Free Energy Methods in Drug Discovery: Current State and Future Directions* (ACS Publications, 2021), pp. 1–38.
[7] In this paper, to avoid confusion we use lower case x to denote configurations, and upper case $X(t)$ to denote (stochastic) trajectories.
[8] M. R. Shirts and V. S. Pande, *J. Chem. Phys.* **122**, 144107 (2005).
[9] M. R. Shirts, E. Bair, G. Hooker, and V. S. Pande, *Phys. Rev. Lett.* **91**, 140601 (2003).
[10] C. H. Bennett, *J. Comput. Phys.* **22**, 245 (1976).
[11] G. E. Crooks, *J. Stat. Phys.* **90**, 1481 (1998).
[12] G. E. Crooks, *Phys. Rev. E* **60**, 2721 (1999).
[13] S. Vaikuntanathan and C. Jarzynski, *Phys. Rev. Lett.* **100**, 190601 (2008).
[14] A. Zhong, B. Kuznets-Speck, and M. R. DeWeese, [arXiv:2304.12287](https://arxiv.org/abs/2304.12287).
[15] F. Vargas and N. Nüsken, [arXiv:2307.01050](https://arxiv.org/abs/2307.01050).
[16] Typically there is a factor of the friction coefficient γ multiplying the left side of (6). For notational simplicity we have set it to one, which given that it is equal for all dimensions may always be done through a time rescaling.

- [17] A. del Campo, *Phys. Rev. Lett.* **111**, 100502 (2013).
- [18] D. Guéry-Odelin, A. Ruschhaupt, A. Kiely, E. Torrontegui, S. Martínez-Garaot, and J. G. Muga, *Rev. Mod. Phys.* **91**, 045001 (2019).
- [19] E. Ilker, Ö. Güngör, B. Kuznets-Speck, J. Chiel, S. Deffner, and M. Hinczewski, *Phys. Rev. X* **12**, 021048 (2022).
- [20] S. Iram, E. Dolson, J. Chiel, J. Pelesko, N. Krishnan, Ö. Güngör, B. Kuznets-Speck, S. Deffner, E. Ilker, J. G. Scott *et al.*, *Nat. Phys.* **17**, 135 (2021).
- [21] A. G. Frim, A. Zhong, S.-F. Chen, D. Mandal, and M. R. DeWeese, *Phys. Rev. E* **103**, L030102 (2021).
- [22] A. B. Adib, *J. Phys. Chem. B* **112**, 5910 (2008).
- [23] J. Tang and P. Abbeel, On a connection between importance sampling and the likelihood ratio policy gradient, in *Proceedings of the 23rd International Conference on Neural Information Processing Systems*, Vol. 1, pp. 1000–1008 (2010).
- [24] The astute reader might point out that, under the rules of stochastic calculus, how they are written out $\mathbf{a} = \tilde{\mathbf{a}}$ for all $X(t)$. However, they are different under a finite time discretization, and thus must both be kept track.
- [25] R. J. Williams, *Mach. Learn.* **8**, 5 (1992).
- [26] J. Peters and S. Schaal, *Neural Netw.* **21**, 682 (2008).
- [27] P. B. Warren and R. J. Allen, *Phys. Rev. Lett.* **109**, 250601 (2012).
- [28] A. Das and D. T. Limmer, *J. Chem. Phys.* **151**, 244123 (2019).
- [29] A. Das, B. Kuznets-Speck, and D. T. Limmer, *Phys. Rev. Lett.* **128**, 028005 (2022).
- [30] A. B. Owen, *Monte Carlo Theory, Methods and Examples* (2013), <https://artowen.su.domains/mc/>.
- [31] D. Kraft, *Forschungsbericht-Deutsche Forschungs- und Versuchsanstalt für Luft- und Raumfahrt* (1988).
- [32] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright *et al.*, *Nat. Methods* **17**, 261 (2020).
- [33] The first two terms are a generalization of the protocol form $U_\lambda(x) = (1 - \lambda)U_A(x) + \lambda U_B(x)$ commonly considered in stochastic thermodynamics. Typically, linear combinations of just $U_A(x)$ and $U_B(x)$ are not effective in transporting $\rho_A(x) \rightarrow \rho_B(x)$, motivating our inclusion of the linear potential $U_C(x)$. We found that empirically, by including this easy-to-compute $U_C(x)$, we achieved much higher performance than using only $U_A(x)$ and $U_B(x)$.
- [34] G. Li, H. T. Quan, and Z. C. Tu, *Phys. Rev. E* **96**, 012144 (2017).
- [35] A. Zhong and M. R. DeWeese, *Phys. Rev. E* **106**, 044135 (2022).
- [36] T. Schmiedl and U. Seifert, *Phys. Rev. Lett.* **98**, 108301 (2007).
- [37] H. Then and A. Engel, *Phys. Rev. E* **77**, 041105 (2008).
- [38] M. V. S. Bonança and S. Deffner, *Phys. Rev. E* **98**, 042103 (2018).
- [39] P. Nazé, S. Deffner, and M. V. Bonança, *J. Phys. Commun.* **6**, 083001 (2022).
- [40] S. Blaber, M. D. Louwerse, and D. A. Sivak, *Phys. Rev. E* **104**, L022101 (2021).
- [41] S. Whitelam, *Phys. Rev. E* **108**, 044138 (2023).
- [42] A. Rolandi, M. Perarnau-Llobet, and H. J. Miller, *New J. Phys.* **25**, 073005 (2023).
- [43] M. C. Engel, J. A. Smith, and M. P. Brenner, *Phys. Rev. X* **13**, 041032 (2023).
- [44] M. Esposito, R. Kawai, K. Lindenberg, and C. Van den Broeck, *Europhys. Lett.* **89**, 20003 (2010).
- [45] A. Zhong and M. R. DeWeese, *Phys. Rev. Lett.* **133**, 057102 (2024).
- [46] M. Doi, S. F. Edwards, and S. F. Edwards, *The Theory of Polymer Dynamics* (Oxford University Press, 1988), Vol. 73.
- [47] Because of the harmonic nature of the interbead potential, the dynamics of the Rouse polymer separates in its three spatial dimensions; without loss of generality, we can consider just a single spatial dimension.
- [48] B. Kuznets-Speck and D. T. Limmer, *Biophys. J.* **122**, 1659 (2023).
- [49] From the change of coordinates $R = (x_N^2 + y_N^2)^{1/2}$, $\vartheta = \tan^{-1}(y_N/x_N)$ and the radial symmetry of U_ϕ and U_{LJ} , one has $\rho_{\text{eq}}(R) \propto (2\pi R)^{-1} \rho_{\text{eq}}(x_N = R, y_N = 0)$.
- [50] G. O. Roberts and R. L. Tweedie, *Bernoulli* **2**, 341 (1996).
- [51] G. Li and Z. C. Tu, *Phys. Rev. E* **100**, 012127 (2019).
- [52] S. Blaber and D. A. Sivak, *J. Chem. Phys.* **153**, 244119 (2020).
- [53] S. Schoenholz and E. D. Cubuk, *Adv. Neural Inf. Process. Syst.* **33**, 11428 (2020).
- [54] L. N. Naden, T. T. Pham, and M. R. Shirts, *J. Chem. Theory Comput.* **10**, 1128 (2014).
- [55] A. N. Singh and D. T. Limmer, *J. Chem. Phys.* **159**, 024124 (2023).
- [56] J. Sohl-Dickstein, E. Weiss, N. Maheswaranathan, and S. Ganguli, in *International Conference on Machine Learning* (PMLR, 2015), pp. 2256–2265.
- [57] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole, Score-based generative modeling through stochastic differential equations, in *International Conference on Learning Representations* (2021).
- [58] A. Doucet, W. Grathwohl, A. G. Matthews, and H. Strathmann, *Adv. Neural Inf. Process. Syst.* **35**, 21482 (2022).
- [59] M. S. Albergo, N. M. Boffi, and E. Vanden-Eijnden, [arXiv:2303.08797](https://arxiv.org/abs/2303.08797).
- [60] H. Wu, J. Köhler, and F. Noé, *Adv. Neural Inf. Process. Syst.* **33**, 5933 (2020).
- [61] E. Bernton, J. Heng, A. Doucet, and P. E. Jacob, [arXiv:1912.13170](https://arxiv.org/abs/1912.13170).
- [62] Y. Chen, T. T. Georgiou, and M. Pavon, *SIAM Rev.* **63**, 249 (2021).
- [63] V. De Bortoli, J. Thornton, J. Heng, and A. Doucet, *Adv. Neural Inf. Process. Syst.* **34**, 17695 (2021).
- [64] J. Berner, L. Richter, and K. Ullrich, *Trans. Mach. Learn. Res.* (2024).
- [65] J.-D. Benamou and Y. Brenier, *Numer. Math.* **84**, 375 (2000).
- [66] E. Aurell, C. Mejía-Monasterio, and P. Muratore-Ginanneschi, *Phys. Rev. Lett.* **106**, 250601 (2011).
- [67] Y. Chen, T. T. Georgiou, and A. Tannenbaum, *IEEE Trans. Autom. Control* **65**, 2979 (2019).
- [68] M. Nakazato and S. Ito, *Phys. Rev. Res.* **3**, 043093 (2021).
- [69] S. Chennakesavalu and G. M. Rotskoff, *Phys. Rev. Lett.* **130**, 107101 (2023).
- [70] <https://github.com/adrizhong/dF-protocol-optimization>.
- [71] C. C. Pugh, *Real Mathematical Analysis* (Springer, New York/Heidelberg/Berlin, 2002), Vol. 2011.
- [72] B. Oksendal, *Stochastic Differential Equations: An Introduction with Applications* (Springer Science & Business Media, 2013).
- [73] D. A. Sivak, J. D. Chodera, and G. E. Crooks, *Phys. Rev. X* **3**, 011007 (2013).