Accelerated Jarzynski estimator with deterministic virtual trajectories

Nobumasa Ishida ●[*] and Yoshihiko Hasegawa ●[†]

Department of Information and Communication Engineering, Graduate School of Information Science and Technology,
The University of Tokyo, Tokyo 113-8656, Japan

The Jarzynski estimator is a powerful tool that uses nonequilibrium statistical physics to numerically obtain partition functions of probability distributions. The estimator reconstructs partition functions with trajectories of the simulated Langevin dynamics through the Jarzynski equality. However, the original estimator suffers from slow convergence because it depends on rare trajectories of stochastic dynamics. In this paper, we present a method to significantly accelerate the convergence by introducing deterministic virtual trajectories generated in augmented state space under the Hamiltonian dynamics. We theoretically show that our approach achieves second-order acceleration compared to a naive estimator with the Langevin dynamics and zero variance estimation on harmonic potentials. We also present numerical experiments on three multimodal distributions and a practical example in which the proposed method outperforms the conventional method, and we provide theoretical explanations.

I. INTRODUCTION

The development of nonequilibrium statistical mechanics has brought about a great many novel algorithms in computational physics and information science [1–5]. In particular, in recent years substantial progress has been made in the estimation of partition functions [1,4–7]. A partition function Z is the normalizing constant of a probability density function $f(x)$ described as a Gibbs distribution:

$$f(x) = \frac{1}{Z} e^{-\beta E(x)}, \tag{1}$$

$$Z := \int e^{-\beta E(x)} dx, \tag{2}$$

where $E(x)$ is the energy of the state x, and β is the inverse temperature. This quantity is fundamental in both natural and information sciences because it characterizes a system in equilibrium: in statistical mechanics, it corresponds to free energy that describes the stability of the system [8], and in machine learning, it is known as the model evidence, which gives an indicator to quantify the likelihood of models for observed data [9,10]. Moreover, the partition function in Eq. (2) can be regarded as a moment-generating function, which gives moments of a system. Therefore, it is important to efficiently calculate the partition function of a system.

However, it is challenging to obtain the partition function or the free energy of a given system. In terms of thermodynamics, the free-energy difference between the initial states and the final states is determined by the work exerted on the system during a quasistatic process, which is infeasible in finite time [11]. In terms of numerical computation, calculating

the partition function in Eq. (2) requires numerical integration over the entire conformation space, which is intractable, especially when the space has many dimensions.

The remarkable equality demonstrated by Jarzynski [1] makes it possible to obtain partition functions efficiently. The Jarzynski equality relates the free-energy difference ΔF to work W defined by the trajectories of states in a nonequilibrium process, which is performed by changing the system's configuration in finite time. The equality is given by

$$\langle e^{-\beta W} \rangle = e^{-\beta \Delta F} = \frac{Z}{Z_0}, \tag{3}$$

where Z_0 is the partition function of the initial distribution, and the notation $\langle \cdot \rangle$ is the expectation taken over all possible trajectories during the process. Equation (3) holds for arbitrary dynamics as long as the initial states are in equilibrium. From Eq. (3), we can estimate Z/Z_0 as an ensemble average of the exerted work. Using this equality, for example, the free-energy profile for a molecule was experimentally obtained by repeatedly pulling the molecule with laser optical tweezers [12]. Note that since Z_0 is rarely accessible in physical systems, we often obtain a relative value Z/Z_0 to some reference value Z_0, rather than Z.

An estimator of Z or Z/Z_0 can be calculated by simulating the nonequilibrium process [13,14], and we focus on a process modeled by the overdamped Langevin dynamics in this work because the dynamics is commonly used for systems in which the energy function $E(x)$ of the target distribution is determined only by position, which is the case in molecular-dynamics simulation or machine learning. Hence, the estimator using the Jarzynski equality for Langevin dynamics is described as the Jarzynski estimator throughout this paper.

In practice, however, convergence of the naive Jarzynski estimator is notoriously slow [15,16]. That is because work W

[*]ishida@biom.t.u-tokyo.ac.jp
[†]hasegawa@biom.t.u-tokyo.ac.jp

has a large variance, and the rare negative value of W, which significantly contributes to the estimator, cannot be efficiently sampled. To overcome such a problem, many methods have been proposed. For example, Refs. [17,18] devised initial states to alleviate sample bias by allowing almost any distribution for initial states using linear equations. References [5,19] used a backward process in which the work distribution can be used to estimate the partition function through Crook's fluctuation theorem [6]. Despite such important advances, however, adopting the Langevin dynamics still has a convergence problem: The variance of the estimator remains even if the system is controlled by the optimal protocol unless the distributions of the initial and final states are identical [20,21]. To overcome such inevitable variance, the authors in Ref. [7] introduced an additional flow field and transformed dynamics. They showed that the method achieves zero variance estimation, which means one can obtain the true partition function with only one trajectory, given an ideal flow field, although such a flow field seems impossible to obtain in practical cases.

In this paper, we address the problem of the inevitable error of Jarzynski estimators and significantly improve the convergence in fundamental cases. The key idea of our work is converting the stochastic process described by the Langevin dynamics to a deterministic virtual one. Specifically, we introduce auxiliary momentum and employ the Hamiltonian dynamics to simulate a target system initially connected to a heat bath. In contrast to simulating all particles constituting a system and a heat bath with the Hamiltonian dynamics, the dynamics employed in our method is virtual, viz. not physical, because we transform a system initially coupled to a heat bath to an isolated one and treat particles only in the system.

Our theoretical analyses reveal that our method achieves second-order acceleration with respect to the duration of simulated dynamics, and furthermore it achieves zero variance estimation at some conditions with harmonic potentials. This property is realized when we employ parallel transport and scaling of harmonic potentials, which are models of a moving laser trap and the time-dependent strength of the trap, respectively [20]. Moreover, the proposed method mitigates variance when the peaks of the target distribution are far from those of the initial distribution. We conduct numerical experiments on four model systems for which the proposed method outperforms the conventional one, and we discuss the results theoretically.

II. METHOD

Given a potential energy $U_{\mathrm{end}}(q)$, we aim to calculate the partition function Z defined by

$$Z := \int e^{-\beta U_{\mathrm{end}}(q)} dq, \qquad (4)$$

where q is a position. Hereafter, we assume $\beta = 1$ without loss of generality. The equilibrium probability distribution corresponding to $U_{\mathrm{end}}(q)$ is $f_{\mathrm{end}}^{\mathrm{eq}}(q) = \exp(-U_{\mathrm{end}}(q))/Z$. First, we select a probability distribution $f_{\mathrm{init}}^{\mathrm{eq}}(q)$ which is arbitrary. Then, to connect these distributions, let us consider a nonequilibrium process during time $t \in [0, \tau]$ characterized with a time-varying potential function $U(q; t)$ with boundary conditions $U(q; 0) = U_{\mathrm{init}}(q)$ and $U(q; \tau) = U_{\mathrm{end}}(q)$. The

process is also arbitrary as long as the boundary conditions are satisfied.

In the naive Jarzynski estimator based on the Langevin dynamics [13], which we call the Langevin Jarzynski estimator (LJE) throughout this paper, we consider a trajectory $q(t)$ governed by the overdamped Langevin dynamics:

$$\frac{dq}{dt} = -\frac{\partial U(q; t)}{\partial q} + \sqrt{2}\xi(t), \qquad (5)$$

where $\xi(t)$ is Gaussian noise such that its autocorrelation function is a δ function $\langle \xi(t)\xi(0) \rangle = \delta(t)$. Then, we define the work exerted during a process as

$$W := \int_0^\tau \frac{\partial U(q; t)}{\partial t} dt. \qquad (6)$$

The Jarzynski equality in Eq. (3) holds as an exponential ensemble average of the work of trajectories generated using Eq. (5). We obtain an estimator of Z/Z_0 by approximating the left-hand side of the Jarzynski equality with a finite number of samples. However, we need a significant number of trajectories to obtain an accurate result because work W has a large variance.

In the proposed method, which we call the Hamiltonian Jarzynski estimator (HJE), we augment the state space with virtual momentum p, which has the same number of dimensions with q. Along with p, we introduce a kinetic energy with a virtual time-dependent mass $m(t)$:

$$K(p; t) = \frac{\|p\|_2^2}{2m(t)}, \qquad (7)$$

where $\|\cdot\|_2$ is the Euclidean norm. Then, we define the Hamiltonian of the system as

$$H(q, p; t) := U(q; t) + K(p; t). \qquad (8)$$

Here the schedules of the Hamiltonian H and the mass $m(t)$ are arbitrary as long as the boundary conditions of $U(q; t)$ are satisfied as well as the LJE.

Then, we consider a trajectory $(q(t), p(t))$ under the Hamiltonian dynamics

$$\frac{dq}{dt} = \frac{\partial H(q, p; t)}{\partial p}, \qquad (9)$$

$$\frac{dp}{dt} = -\frac{\partial H(q, p; t)}{\partial q}, \qquad (10)$$

and we define work consistent with Eq. (6) [1]:

$$W := \int_0^\tau \frac{\partial H(q, p; t)}{\partial t} dt$$
$$= H(q(\tau), p(\tau); \tau) - H(q(0), p(0); 0). \qquad (11)$$

The second equation holds because the entire system is considered isolated throughout the process; that is, there is no heat dissipation. Finally, we can apply the Jarzynski equality for the set of trajectories and obtain the partition function Z with respect to $U_{\mathrm{end}}(q)$:

$$\langle e^{-W} \rangle = \frac{\int e^{-\beta H(q,p;\tau)} dq \, dp}{\int e^{-\beta H(q,p;0)} dq \, dp} = \frac{m(\tau)^{\frac{N}{2}} Z}{m(0)^{\frac{N}{2}} Z_0}, \qquad (12)$$
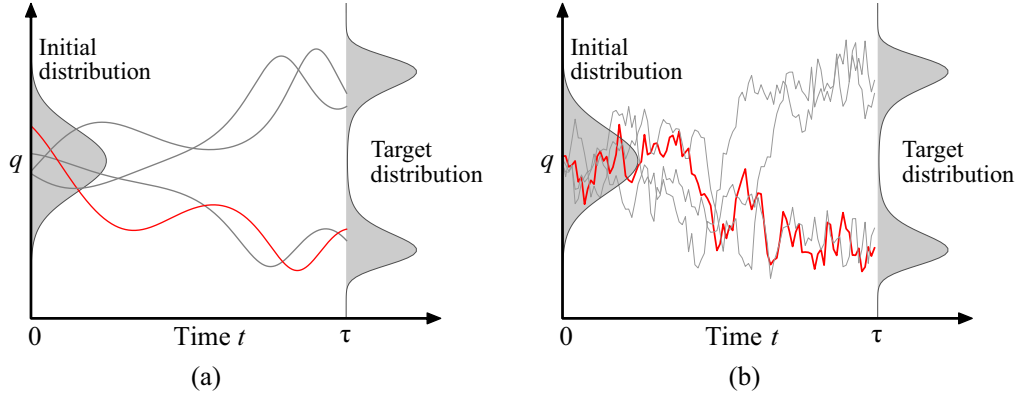
FIG. 1. Illustrations of (a) the proposed HJE method and (b) the conventional LJE method for estimating the partition function of a target distribution. Both methods generate trajectories, which are depicted by solid curves, whose initial positions are given by some initial distribution. A single trajectory for each method is colored in red (dark gray) for visibility. A protocol for each process is defined during time $t \in [0, \tau]$, and boundary values of the protocol are determined by the initial distribution and the target distribution. The main feature of the HJE is that its trajectories are generated by the Hamiltonian dynamics with virtual momentum, which is deterministic given the initial states, while the LJE adopts the Langevin dynamics, which is stochastic.

where N is the dimension of p. The second equation holds because of Eq. (8). In addition to the LJE, we can estimate Z by approximating the expected value on W in Eq. (12) with an ensemble average of finite samples.

Because the HJE adopts deterministic trajectories described by ordinary differential equations (ODEs), we can take advantage of efficient and accurate ODE solvers. This is superior to the case of the LJE, where stochastic differential equations have to be solved. Illustrations of the HJE and the LJE are shown in Figs. 1(a) and 1(b), respectively.

Our approach is inspired by the Hamiltonian Monte Carlo method [22]. This is an efficient sampling algorithm for accelerating transitions in state space by introducing auxiliary momentum and the Hamiltonian dynamics to replace the Langevin dynamics in the Metropolis-Hastings method [23]. In related work, another method introduced auxiliary momentum and the underdamped Langevin dynamics for estimating partition functions [24], but the trajectories in that method are still stochastic and therefore cannot take advantage of deterministic trajectories, in contrast to ours.

We note that the formation of the HJE is different from that based on the Hamiltonian dynamics, which traces all particles contained in the system and the connected heat bath. The HJE is also different from the methods that accurately simulate systems connected to thermostats, such as the Nosé-Hoover method [25].

III. THEORETICAL ANALYSIS

Before testing the proposed method in practical examples, we analytically calculated the properties of the HJE in linear systems, where the partition functions are known and the dynamics is analytically solvable. In particular, we focused on parallel transport, rotation, and scaling of a potential because linear systems are mainly composed of such transformations.

We compare the LJE and HJE in terms of their variance. To analyze the variance of the estimators, we evaluate the variance of work as a proxy using dissipative work W_{diss} [16],

which is more feasible. W_{diss} is defined by

$$W_{\text{diss}} := \langle W \rangle - \Delta F, \tag{13}$$

where $\langle W \rangle$ is the average work over all possible trajectories and ΔF is the difference of free energy between boundary distributions $f_{\text{init}}^{\text{eq}}$ and $f_{\text{end}}^{\text{eq}}$ satisfying $\Delta F = -\ln(Z/Z_0)$. Then we use the following equation as the first-order approximation of the variance of W, denoted by $\langle\langle W \rangle\rangle$ [16]:

$$\langle\langle W \rangle\rangle \approx 2 W_{\text{diss}}. \tag{14}$$

Moreover, we approximate the distributions during time development with a Gaussian distribution when the initial and final distribution of a protocol are Gaussian. This approximation is valid when the system is near equilibrium during the process.

A. Parallel transport

First, we explore the parallel transport of a one-dimensional (1D) harmonic potential. This is a model of dragged Brownian motion, such as a molecule trapped and pulled by an optical laser [20,26]. We let the protocol during $t \in [0, \tau]$ be

$$U(q; t) = \frac{\left(q - \frac{\mu_\tau}{\tau} t\right)^2}{2\sigma^2}, \tag{15}$$

where μ_τ is the transportation distance and σ is the standard deviation. This protocol means the harmonic potential is dragged from the origin to $q = \mu_t$ at a constant speed. The corresponding partition functions are $Z = Z_0 = \sqrt{2\pi\sigma^2}$. Then, we let the mass during the protocol be constant,

$$m(t) = m_0. \tag{16}$$

With this setting, we can analytically obtain W_{diss} for the LJE and HJE, which are given, respectively, by

$$W_{\text{diss}} = \frac{\mu_\tau^2}{\tau} - \frac{\mu_\tau^2 \sigma^2}{\tau^2}\left(1 - e^{-\frac{\tau}{\sigma^2}}\right) \tag{17}$$
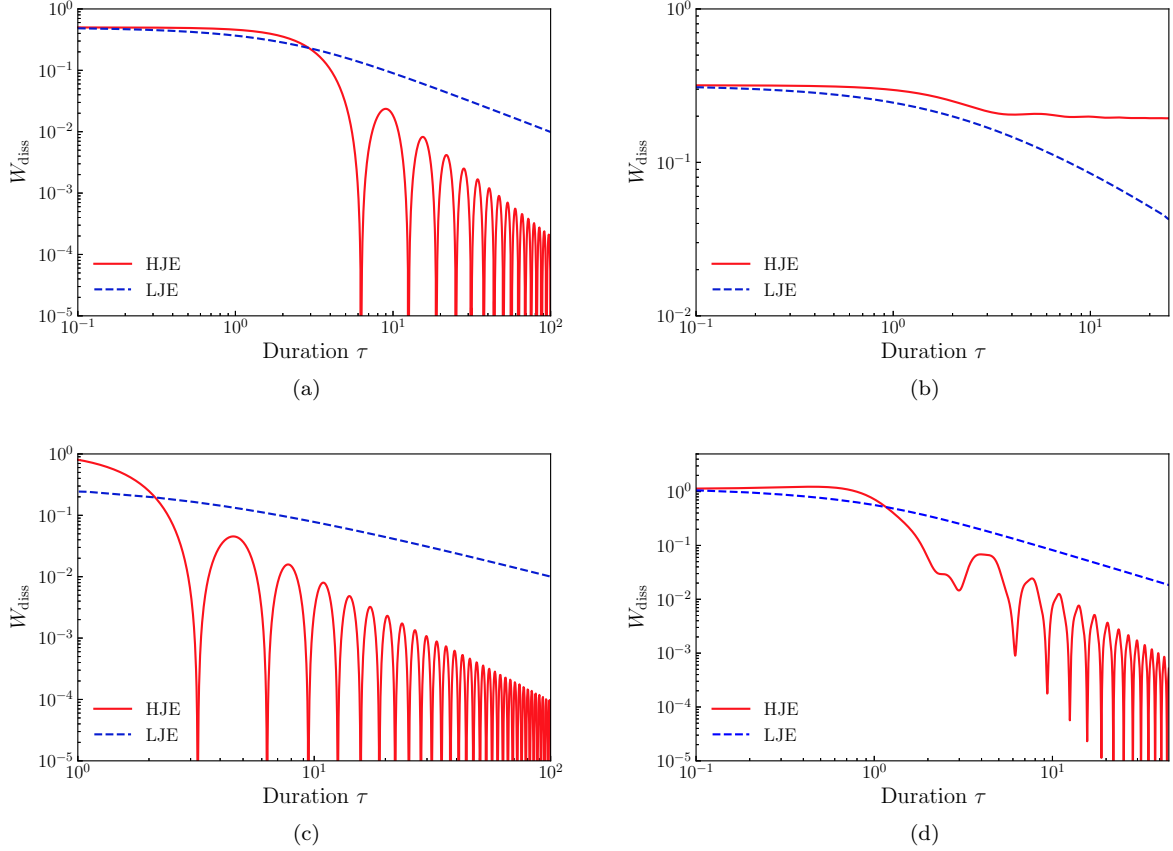
FIG. 2. Analytically obtained total dissipative work W_{diss} of the HJE and LJE as a function of duration τ in linear systems. W_{diss} of the HJE and LJE are depicted by a red solid curve and a blue dashed curve, respectively. (a) Parallel transport of a harmonic potential, where $\sigma^2 = 1$, $\mu_\tau = 1$, and $m_0 = 1$. (b) Scaling of a harmonic potential, where $\sigma_0^2 = 1$, $\sigma_\tau^2 = 4$, and $m(t) = 1$. (c) Scaling of a harmonic potential with time-dependent mass for the HJE, where $\sigma_0^2 = 1$, $\sigma_\tau^2 = 4$, and $\alpha = 1$. (d) Rotation of a harmonic potential, where $a = 8$, $b = 2$, and $m_0 = 1$. As a whole, the specific parameters do not change the results qualitatively.

and

$$W_{\text{diss}} = \frac{m_0 \mu_\tau^2}{\tau^2}\left(1 - \cos\frac{\tau}{\sqrt{m_0}\sigma}\right). \qquad (18)$$

Derivations of Eqs. (17) and (18) are shown in Appendixes A and B, respectively. Note that the result of Eq. (17) is consistent with a previous analysis of dissipative work in Ref. [27].

Equations (17) and (18) reveal prominent features of the HJE: its W_{diss} converges in the second order as a function of the reciprocal of τ, while that of the LJE converges in the first order, and furthermore the HJE achieves $W_{\text{diss}} = 0$ at $\tau = (2n+1)\pi\sqrt{m_0}\sigma$ for any $n \in \mathbb{N}$ and $\mu_\tau \in \mathbb{R}$, while the LJE never achieves $W_{\text{diss}} = 0$ with finite τ [20]. Therefore, W_{diss} of the HJE converges faster than that of the LJE when the duration of protocol τ is longer. We note that in the case of $\tau \to 0$, both Eqs. (17) and (18) reduce to $W_{\text{diss}} = \mu_\tau^2/(2\sigma^2)$, which is equal to the case of importance sampling [23]. Importance sampling is a widely used method to estimate an expected value. We plot Eqs. (17) and (18) in Fig. 2(a), which shows that while W_{diss} of the HJE is larger than that of the LJE when the duration τ is small, the HJE more significantly reduces its W_{diss} than the LJE as τ becomes large.

Additionally, we note that the HJE for parallel transport of multidimensional harmonic potentials works as well as it does

for 1D because in that case the dynamics can be decomposed into a set of independent 1D parallel transports.

B. Scaling

Next, we explore the scaling of a harmonic potential by changing the variance of the corresponding distribution from σ_0^2 to σ_τ^2 through a protocol. This is a model of traps in stochastic systems whose strength varies temporally [20]. The initial distribution and the target distribution are given, respectively, by

$$f_{\text{init}}^{\text{eq}}(q) = \frac{1}{Z_0}e^{-\frac{q^2}{2\sigma_0^2}}, \qquad (19)$$

$$f_{\text{end}}^{\text{eq}}(q) = \frac{1}{Z}e^{-\frac{q^2}{2\sigma_\tau^2}}, \qquad (20)$$

where $Z_0 = \sqrt{2\pi\sigma_0^2}$ and $Z = \sqrt{2\pi\sigma_\tau^2}$. Therefore, the corresponding potential functions are $U_{\text{init}}(q) = q^2/(2\sigma_0^2)$ and $U_{\text{end}}(q) = q^2/(2\sigma_\tau^2)$. Then we let a linear protocol during $t \in [0, \tau]$ be

$$U(q;t) = \left(1 - \frac{t}{\tau}\right)U_{\text{init}}(q) + \frac{t}{\tau}U_{\text{end}}(q) \qquad (21)$$

and let the mass be constant, $m(t) = 1$.

We plot W_{diss} of the HJE and LJE in Fig. 2(b). The details of the calculation are shown in Appendixes C and D, respectively. Figure 2(b) shows that W_{diss} of the HJE converges to some positive value while that of the LJE converges to zero as a function of τ. This result of the HJE is consistent with the well-known fact that an isolated system will generally not be in equilibrium after some protocol, even if the protocol takes infinite time.

However, we can overcome the limitation by recalling the fact that the mass in our method is virtual and arbitrary. To compensate for the difference of variance σ_0^2 and σ_τ^2, we adopt a time-dependent mass. While such a mass protocol is almost infeasible in reality, we can use it because our mass is virtual.

The initial and final distributions are the same as $f_{\text{init}}^{\text{eq}}(q)$ and $f_{\text{end}}^{\text{eq}}(q)$, respectively, in the above analysis. Then, we let the protocol during $t \in [0, \tau]$ be

$$U(q;t) = \frac{q^2}{2\sigma(t)^2}, \tag{22}$$

$$\sigma(t)^2 = \sigma_0^2 e^{\frac{\gamma}{\tau}t}, \tag{23}$$

where $\gamma = \ln\sigma_\tau^2/\sigma_0^2$. For the protocol of a kinetic energy, let the mass be time-dependent:

$$m(t) = \frac{\alpha}{\sigma(t)^2}, \tag{24}$$

where α is a positive constant. In this case, we obtain W_{diss} of the HJE as

$$W_{\text{diss}} = -1 + \cos^2 \omega\tau + \left(\frac{\alpha\omega^2}{2} + \frac{1}{2\alpha\omega^2} + \frac{\alpha\gamma^4}{32\omega^2\tau^4} \right.$$
$$\left. + \frac{\alpha\gamma^2}{4\tau^2} + \frac{\gamma^2}{4\omega^2\tau^2} \right) \sin^2 \omega\tau, \tag{25}$$

$$\omega = \frac{\sqrt{4\tau^2/\alpha - \gamma^2}}{2\tau}, \tag{26}$$

where $\tau > \sqrt{\alpha}|\gamma|/2$. Details of the derivation are shown in Appendix D. Note that W_{diss} of the LJE needs numerical integration, whose details are shown in Appendix C. We plot W_{diss} of the HJE and LJE in Fig. 2(c). As well as the case for parallel transport, the HJE achieves $W_{\text{diss}} = 0$ with $\omega\tau = \pi n$ for any $n \in \mathbb{N}$, that is, $\tau = \sqrt{\frac{\gamma^2}{4} + \pi^2 n^2}$, and it converges faster than the LJE thanks to the virtual mass.

C. Parallel transport and scaling

Next, we investigate protocols that include both parallel transport and scaling. In particular, we compare W_{diss} of the HJE and LJE as the distance of parallel transport changes and the magnitude of scaling increases. The initial and final distributions are $\mathcal{N}(0, \sigma^2)$ and $\mathcal{N}(\mu, \sigma^2)$, respectively, where $\mathcal{N}(\mu, \sigma^2)$ denotes a Gaussian distribution with mean μ and variance σ^2. Then we let the protocol during $t \in [0, \tau]$ be

$$U(q;t) = \frac{1 + a\frac{t}{\tau}}{2\sigma^2}\left(q - \mu\frac{t}{\tau}\right)^2. \tag{27}$$

This protocol describes that the potential is dragged from the origin to $x = a$ and its strength is varied at a constant ratio over time. Here, we adopt a constant mass of $m(t) = m_0$.
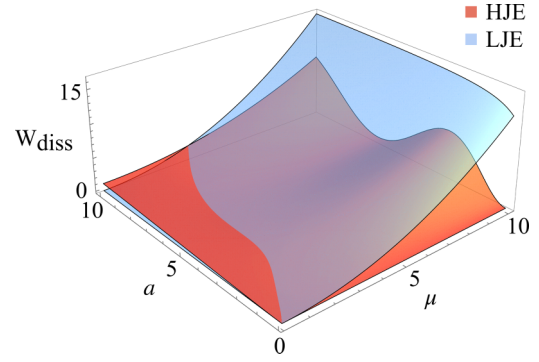

FIG. 3. Total dissipative work W_{diss} exerted from a harmonic potential that is dragged and scaled taking duration $\tau = 2\pi$, where the HJE and LJE are depicted by red (dark gray) and blue (light gray) surfaces, respectively. μ is the dragged distance and a characterizes how much the potential is scaled.

We plot W_{diss} as functions of a and μ with fixed $\tau = 2\pi$ in Fig. 3. The derivation of W_{diss} for the protocol is shown in Appendixes E and F. Figure 3 shows that W_{diss} of the LJE is smaller than that of the HJE when μ is small, which is consistent with the analysis in Sec. III B for the case of $\mu = 0$ in Eq. (27). However, W_{diss} of the HJE is significantly smaller than that of the LJE when μ is large. In addition, as described in Sec. III A, W_{diss} is zero for the HJE when $a = 0$. This result suggests that as the dragged distance μ becomes larger, the superiority of the HJE increases, even when the initial distribution and the final distribution cannot be connected only by parallel transport.

D. Rotation

Finally, we employ a model of rotation protocol. The initial distribution $f_{\text{init}}^{\text{eq}}$ and the target distribution $f_{\text{end}}^{\text{eq}}$ are

$$f_{\text{init}}^{\text{eq}}(x, y) = \frac{1}{Z_0}e^{-(ax^2+by^2)}, \tag{28}$$

$$f_{\text{end}}^{\text{eq}}(x, y) = \frac{1}{Z}e^{-(bx^2+ay^2)}, \tag{29}$$

where a and b are positive values characterizing the variance of x and y, respectively. These distributions are equal to each other through $\pi/2$ rotation. The corresponding partition functions are $Z = Z_0 = \pi/\sqrt{ab}$.

Then, we let a potential function during $t \in [0, \tau]$ be

$$U(x, y;t) = ax'(t)^2 + by'(t)^2, \tag{30}$$

where (x', y') is the rotation of (x, y):

$$\begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} = \begin{pmatrix} \cos\theta(t) & -\sin\theta(t) \\ \sin\theta(t) & \cos\theta(t) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{31}$$

with $\theta(t) = \pi t/(2\tau)$.

Under this configuration, we can calculate W_{diss} analytically, and its derivation is shown in Appendixes G and H. We plot W_{diss} as a function of τ in in Fig. 2(d). From Fig. 2(d), we find that although W_{diss} of the LJE decreases monotonically, that of the HJE oscillates and its local minima are much smaller than W_{diss} of the LJE at the same τ. Therefore, if we can choose a τ value where W_{diss} of the HJE is minimal, then

it converges significantly faster than the LJE as the number of trajectories increases.

IV. NUMERICAL EXPERIMENTS

In this section, we demonstrate the effectiveness of the HJE through four numerical experiments on model systems where the LJE works poorly. We begin with a toy model to clarify the efficacy of the proposed method. Then, we show the robustness of our method under conditions where the conventional method fails, and we verify that the HJE outperforms the LJE for a distribution with a large number of peaks. In addition, we conduct a polymer stretching experiment to clarify the meaning of the virtual trajectories in our method. The result of each experiment is explained based on the theoretical analysis given in the previous section.

A. 1D double-well potential

In this experiment, we validate the efficiency of the HJE through a simple but suggestive model. The target distribution is characterized by a double-well potential function,

$$f_{\text{end}}^{\text{eq}}(q) = \frac{1}{Z} e^{-kq^4 + q^2}, \tag{32}$$

where k is a positive parameter. Double-well potentials are widely used to describe bistable states in stochastic systems [28,29]. We use the standard Gaussian distribution as an initial distribution $f_{\text{init}}^{\text{eq}}$. Then, we construct a protocol for the potential energy during $t \in [0, \tau]$ by linearly interpolating the boundaries:

$$U(q; t) = \left(1 - \frac{t}{\tau}\right)\frac{q^2}{2} + \frac{t}{\tau}(kq^4 - q^2). \tag{33}$$

We plot the error of the estimators of Z/Z_0 as a function of τ in Fig. 4, which shows that, although the LJE has an error smaller than that of the HJE when τ is small, the HJE has substantially lower error than that of the LJE as τ becomes large. Therefore, it is suggested that if we adopt a sufficiently large duration τ, we can estimate the partition function of the target distribution more efficiently by the HJE. Moreover, the error of the HJE oscillates moderately as a function of τ. These results can be explained by the theoretical analysis in the previous section. The protocol of this experiment is approximated with two independent parallel transports of a harmonic potential: one a harmonic potential moving from the origin to the right peak of the target distribution, and the other one moving from the origin to the left peak of the target distribution. Then, if the protocol is parallel transport, the error of the HJE oscillates and decreases faster than that of the LJE as a function of τ, which is shown in Sec. III A.

B. Gaussian mixture model

Next, we examine the robustness of the proposed method against the arrangement of peaks in a target distribution through a two-dimensional (2D) bimodal distribution. In particular, we explore a mixture of two Gaussian distributions with various arrangements. The partition functions for the Gaussian mixture models are $Z = 1$ regardless of specific
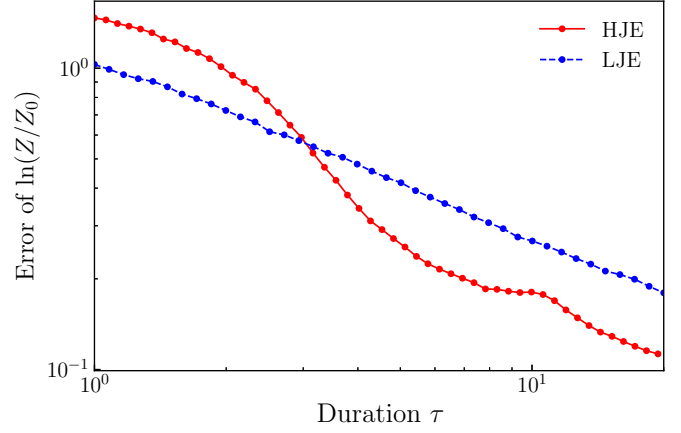

FIG. 4. Error of the estimated partition function for the HJE and LJE, which are depicted by a red solid curve with dots and a blue dashed curve with dots, respectively, as a function of τ for a 1D double-well potential. The target distribution is characterized by $k = 1/16$ and its partition function is $Z = 146.372$, which is numerically confirmed. We generated 10 trajectories at each τ to estimate the partition function and repeated the procedure 100 000 times for both the HJE and LJE. The adopted numeric schemes are the Runge-Kutta fourth-order method for the HJE and the Euler-Maruyama method [30] for the LJE. Applying a higher-order scheme for stochastic differential equations, the order 1.0 strong stochastic Runge-Kutta method [31], for the LJE makes no qualitative difference to the result (not shown). We set the time step for numerical integration to be $\Delta t = 10^{-3}$.

parameters, which is appropriate for comparing the results for target distributions with different parameters.

The target distribution is a mixture of Gaussian distributions $\mathcal{N}(\mu_P, \sigma^2)$ and $\mathcal{N}(\mu_Q, \sigma^2)$ in state space $q = (x, y)$, where $\mu_P = (a, s)$ and $\mu_Q = (a, -s)$. That is,

$$f_{\text{end}}^{\text{eq}}(x, y) = \frac{1}{Z} \exp\left(-U_{\text{end}}(x, y)\right), \tag{34}$$

$$U_{\text{end}}(x, y) = -\log\left(\frac{1}{4\pi\sigma^2} e^{-\frac{(x-a)^2 + (y-s)^2}{2\sigma^2}}\right.$$
$$\left. + \frac{1}{4\pi\sigma^2} e^{-\frac{(x-a)^2 + (y+s)^2}{2\sigma^2}}\right). \tag{35}$$

As shown in Fig. 5(a), the distribution given by Eq. (34) is bimodal, and its peaks are located at points $P = \mu_P$ and $Q = \mu_Q$. Hereafter, we define the displacement s, which describes how the two peaks are separated from each other. Then we adopt a Gaussian distribution at the origin O as the initial distribution, that is, $U_{\text{init}}(x, y) = (x^2 + y^2)/(2\sigma^2)$, and we let the protocol between $t \in [0, \tau]$ be

$$U(x, y; t) = \left(1 - \frac{t}{\tau}\right)U_{\text{init}}(x, y) + \frac{t}{\tau}U_{\text{end}}(x, y). \tag{36}$$

We sweep the displacement s to reveal how the arrangement of the peaks of the distributions affects the HJE and LJE. Moreover, we let $\tau = 2\pi\sqrt{\sigma^2}$, where W_{diss} of the HJE in the parallel transport protocol given by Eq. (18) vanishes.

We plot the error of the estimated Z/Z_0 as a function of s in Fig. 5(b), which shows that the HJE has an error
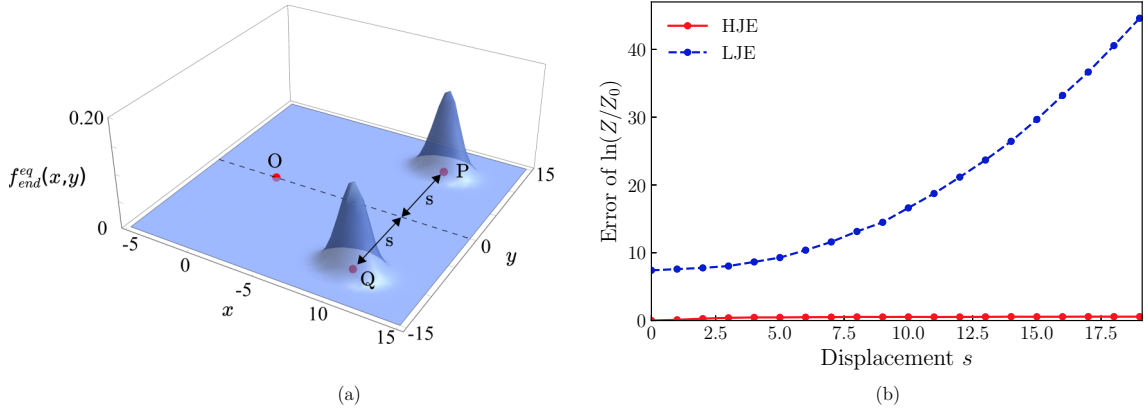
FIG. 5. Illustration of the target distribution and error of the estimated partition function. (a) Plot of the target distribution. The distribution is a mixture of Gaussian distributions, and the peaks are located at points $P = (a, s)$ and $Q = (a, -s)$. We define the displacement s given by the distance between the x-axis and the peaks of the target distribution so as to characterize their arrangement. (b) Error for the HJE and LJE, which are depicted by a red solid curve with dots and a blue dashed curve with dots, respectively, as a function of s, where $\tau = 2\pi$, $\sigma^2 = 1$, and $a = 10$. We generated 10 trajectories at each s to estimate the partition function and repeated the procedure 1000 times for the HJE and LJE. The settings for the numerical schemes are the same as in Sec. IV A.

significantly smaller than that of the LJE for any displacement s. Moreover, although the error of the LJE increases rapidly as a function of s, that of the HJE increases only slightly. These results can be explained by the analysis of parallel transport in Sec. III A. When $s = 0$, the target distribution is a Gaussian distribution, as explained by the analysis in Sec. III A, which demonstrates that the error of the HJE vanishes regardless of the distance between the peaks of the initial and final distributions. This suggests that the significantly small error of the HJE in this experiment is due to the protocol being approximately decomposable into two independent parallel transports: on a harmonic potential moving from O to P and the other one moving from O to Q. As displacement s grows, the peaks of the target distribution separate from each other and the distance becomes large between the points P or Q, which are the means of the peaks, and the origin O, which is the mean of the initial distribution. If the protocol can be considered as a parallel transport, the distance between O and P or Q has no effect on the performance of the HJE, while the error of the LJE increases as a function of the distance. These are the reasons why the qualitative difference appears in the experiment shown in Fig. 5(b).

C. Multimodal distribution

In this experiment, we employ a distribution with many peaks to demonstrate the effectiveness of the HJE for multimodal distributions. We let the dimension of the state space q be 16 and the target distribution defined over q is

$$f_{\text{end}}^{\text{eq}}(q) = \frac{1}{Z} \exp\left(-U(q;\tau)\right), \tag{37}$$

$$U_{\text{end}}(q) = -\ln\left\{\sum_{i=1}^{N} \exp(-\cosh\|q - \mu_i\|_2)\right\}, \tag{38}$$

where N is the number of peaks and μ_i is the location of the ith peak.

We use a zero-mean Gaussian distribution with a variance σ^2 as the initial distribution:

$$f_{\text{init}}^{\text{eq}}(q) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{\|q\|_2^2}{2\sigma^2}\right). \tag{39}$$

Then we let the protocol during $t \in [0, \tau]$ be

$$U(q;t) = \left(1 - \frac{t}{\tau}\right)\frac{\|q\|_2^2}{2}$$
$$+ \frac{t}{\tau}\left(-\ln\left\{\sum_{i=1}^{N} \exp\left(-\cosh\left\|q - \mu_i\frac{t}{\tau}\right\|_2\right)\right\}\right). \tag{40}$$

We sweep the number of peaks N from 1 to 256 in order to reveal how the performance of the HJE scales as the target distribution has more peaks. For each N, we test 20 sets of $\{\mu_i\}_{i=1}^{N}$, which are randomly sampled from a Gaussian distribution $\mathcal{N}(0, 25)$.

We plot the error of the estimators of Z/Z_0 in Fig. 6 as a function of N, which shows that the error of the HJE is substantially smaller that that of the LJE for all N. Roughly speaking, the HJE accurately estimates the order of the ground truth, while the LJE differs by two digits. This result can be explained by the theoretical analysis in Sec. III and the previous experiment on a Gaussian mixture model in Sec. IV B. The protocol in this experiment can be approximated by independent parallel transport from the initial distribution to peaks in the target distribution, where the HJE has been demonstrated to perform well. On the other hand, the peak in the initial distribution and those in the target distribution have different shapes in terms of the second- or higher-order moments, which results in the increment of the variance of the estimator as shown in the theoretical analysis of the scaling protocol in Sec. III B. Then we conclude that the reason for the success of the HJE in this experiment is because the contribution of parallel transport is more substantial than deterioration by scaling, which is suggested by the theoretical analysis of
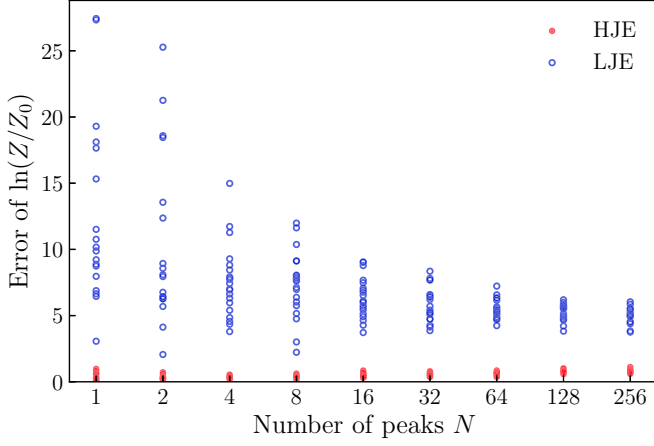
FIG. 6. Error of the estimated partition function for the HJE and LJE, which are depicted by red dots and blue circles, respectively, as a function of N characterizing multimodal distributions in 16-dimensional space. Points at each N correspond to different sets of $\{\mu_i\}$. The ground truth of Z is numerically obtained as in Appendix I. We tested 20 sets of $\{\mu_i\}$. Then, we generated 1000 trajectories for each set to estimate the partition function, and we repeated the procedure 10 times for the HJE and LJE. The duration of the process is $\tau = 4\pi$. The settings for the numerical schemes are the same as in Sec. IV A.

the protocol that includes both parallel transport and scaling in Sec. III C.

D. Stretching polymer chain in a solvent

Finally, we clarify the meaning of the virtual dynamics of our method and its efficacy through an experiment using the stretching Rouse model, which is an ideal polymer chain in a solvent [32,33]. In the model, a polymer chain is split into subsections that show rubberlike behavior. The subsections are considered beads with no volume and connected to adjacent elements with harmonic springs. The beads follow the Langevin dynamics. Note that each bead in the Rouse model corresponds not to a molecule but to a segment of a molecular chain.

We conduct a numerical experiment stretching the chain with two optical traps. One trap grabs one end point of the chain and is fixed at its initial position, and the other grabs the other end point of the chain and moves to the target position. The interaction between the traps and the grabbed beads is also modeled as harmonic springs [34]. To formulate the system, we consider a series of elements: an optical trap, N beads, and the other trap, each connected with springs. We represent the position of the ith element by the scalar value q_i: q_0 for the first trap, $q_1, \ldots, q_n$ for beads, and q_{N+1} for the other trap. The system is illustrated in Fig. 7(a).

Note that in the proposed method, the beads are considered to be free particles, which have no physical counterparts because the subsection-dividing coarse-graining in the Rouse model is valid only when the polymer is in a solvent or other similar circumstances, but not in a vacuum.
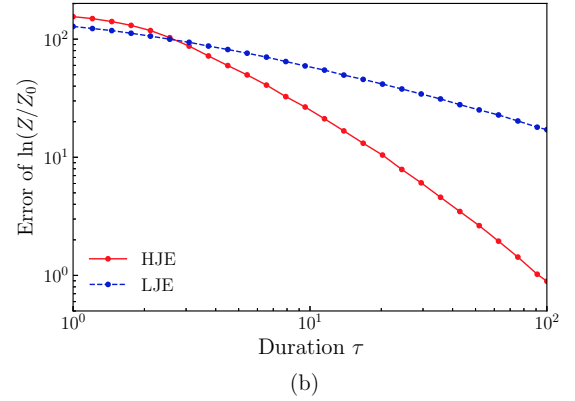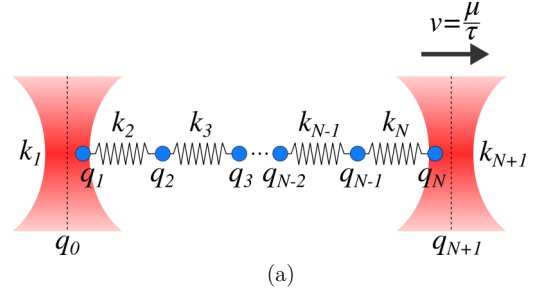

FIG. 7. Illustration of the stretched Rouse model and the error of the estimated partition function. (a) Model description. A Rouse model consists of N beads of positions q_1 to q_N and harmonic springs connecting them with spring constants k_2 to k_N. Each end point of the chain is held by optical laser traps with positions q_0 and q_{N+1} and spring constants k_0 and k_{N+1}, respectively. The right trap moves at a constant velocity $v = \mu/\tau$. (b) Error of the estimated partition function with the protocol of the stretching Rouse model for the HJE and LJE, depicted by a red solid curve with dots and a blue dashed curve with dots, respectively, as a function of τ. The model is characterized by $N = 100$, $\mu = 20$, and $k_n = 1$ for $n = 1, \ldots, N$. We generated 10 trajectories at each τ to estimate the partition function and repeated the procedure 2000 times for both the HJE and LJE. The settings for the numerical schemes are the same as in Sec. IV A.

The potential energy of the chain is given by

$$U(q;t) = \sum_{n=1}^{N+1} \frac{k_n}{2}(q_n - q_{n-1})^2, \qquad (41)$$

where k_n is the strength of the nth spring connecting elements n and $n + 1$.

We consider a linear protocol moving the trap position from 0 to μ at a constant velocity μ/τ. That is, the initial and final distributions are

$$f_{\text{init}}^{\text{eq}}(q) = \frac{1}{Z_0} \exp\left(-U(q;0)\right), \qquad (42)$$

$$f_{\text{end}}^{\text{eq}}(q) = \frac{1}{Z} \exp\left(-U(q;\tau)\right), \qquad (43)$$

and the position of the moving trap is described by $q_{N+1}(t) = t\mu/\tau$. The exact solution of Z for this model is characterized by the position of the moving trap q_{N+1}:

$$Z(q_{N+1}) = \sqrt{\frac{(2\pi)^n}{\det A}} \exp\left(-\frac{1}{2}\bar{k}q_{N+1}^2\right), \qquad (44)$$

where A is a trigonal matrix with elements $A_{n,n} = k_n + k_{n+1}$, $A_{n,n+1} = -k_{n+1}$, and $A_{n,n-1} = -k_n$, and $\bar{k}$ is the harmonic mean of k_n satisfying $1/\bar{k} = \sum_{n=1}^{N+1} 1/k_n$ [34]. The initial distribution $f_{\mathrm{init}}^{\mathrm{eq}}(q)$ is a Gaussian distribution with a mean of zero and a covariance matrix A^{-1}.

We sweep the duration of the process τ and plot the error of the estimated Z/Z_0 for $N = 100$ in Fig. 7(b), in which the error for the HJE decreases faster and becomes smaller than that for the LJE as τ increases. Therefore, for large enough τ, the HJE outperforms the LJE in terms of fast convergence.

We now provide a theoretical explanation for the result. The distribution of positions during the time development is Gaussian, and its mean and variance are given as follows [34]:

$$\langle q \rangle = A^{-1}h, \tag{45}$$

$$\langle (q - \langle q \rangle)(q^T - \langle q^T \rangle) \rangle = A^{-1}, \tag{46}$$

where $h^T = (0, 0, \dots, 0, k_{N+1}q_{N+1}(t))$ is time-dependent. Therefore, the protocol can be considered to represent parallel transport of the peak in a distribution of positional states. Then, as we showed theoretically in Sec. III A, the HJE achieves faster convergence than the LJE for parallel transport. Thus, the HJE works well for this experiment.

V. CONCLUDING REMARKS

In this paper, we have addressed the problem of slow convergence of Jarzynski estimators for partition functions by introducing deterministic virtual trajectories instead of the original stochastic trajectories. We proved that our method achieves second-order acceleration with respect to the duration of simulated dynamics and furthermore zero-variance estimation in the case of harmonic potentials. Then, we numerically showed that the HJE outperforms the LJE in four examples, and we provided theoretical explanations of why the HJE exhibits better performance in the examples. In short, it is suggested that if the protocol can be approximated by some independent parallel transport protocols, the error of the HJE decreases faster than that of the LJE as the duration τ grows. We suggest that probability distributions modeling multistable systems would satisfy this requirement.

The HJE uses the Hamiltonian dynamics to make the simulated trajectories deterministic. Theoretical analysis of parallel transport, scaling with a constant mass, and rotation of potentials implies that if the eigenfrequency of the initial distribution is preserved through the process, as in the cases of parallel transport and rotation, each trajectory of the HJE is less dissipative than when the frequency varies, which is the case of scaling. Furthermore, the HJE can compensate for the variation of the eigenfrequency during a process if an appropriate time-dependent mass is employed, as is the case of scaling with a time-dependent mass.

In addition to the above advantages, our work enables the employment of efficient and accurate ODE solvers when simulating the nonequilibrium process. Considering that the LJE needs to solve stochastic differential equations, this property makes the present method easier to apply.

Despite the fact that there are cases where W_{diss} of the HJE does not converge to zero even in the case of $\tau \to \infty$, like

the scaling protocol shown in Sec. III B, our work has the potential to be a dramatically efficient estimator of partition functions, especially for multistable systems. The HJE is a fundamental improvement of the LJE, so various future works based on the HJE could be expected to address its limitations, as many efficient methods have been developed based on the LJE. We presume that time-dependent mass, which we used in the scaling protocol in Sec. III B, is particularly promising for the HJE.

APPENDIX A: W_{diss} OF THE LJE FOR PARALLEL TRANSPORT

We derive W_{diss} of the LJE for the parallel transport protocol. When a state q of a system satisfies the Langevin dynamics, the time development of the probability distribution of the system $f(q, t)$ is described by following the Fokker-Planck equation [35] under the condition of the inverse temperature $\beta = 1$:

$$\frac{\partial f(q, t)}{\partial t} = \frac{\partial}{\partial q}\left[\frac{\partial U(q; t)}{\partial q} + \frac{\partial}{\partial q}\right]f(q, t). \tag{A1}$$

Using Eq. (A1), we obtain the time derivative of the average and variance of position, which are denoted as $\langle q(t) \rangle$ and $\langle\langle q \rangle\rangle$ at time t:

$$\frac{d\langle q(t) \rangle}{dt} = -\int \frac{\partial U(q; t)}{\partial q} f(q, t)dq, \tag{A2}$$

$$\frac{d\langle\langle q(t) \rangle\rangle}{dt} = -2\int q\frac{\partial U(q; t)}{\partial q}f(q, t)dq + 2 - 2\mu(t)\frac{d\mu(t)}{dt}. \tag{A3}$$

With the setting in Sec. III A, by solving Eqs. (A2) and (A3), we obtain

$$\langle q(t) \rangle = \frac{\mu_\tau}{\tau}\left[t - \sigma^2\left(1 - e^{-\frac{t}{\sigma_0^2}}\right)\right], \tag{A4}$$

$$\langle\langle q(t) \rangle\rangle = \sigma_0^2. \tag{A5}$$

We can verify that the distribution of the states at time t is a Gaussian distribution with mean $\mu(t)$ and variance $\sigma^2(t)$ by substituting $f(q, t)$ into Eq. (A1). Finally, with the definitions of work W [Eq. (6)] and W_{diss} [Eq. (13)] we obtain Eq. (17).

APPENDIX B: W_{diss} OF THE HJE FOR PARALLEL TRANSPORT

We derive W_{diss} of the HJE for the parallel transport protocol. To obtain the dissipative work W_{diss}, we calculate the Kullback-Leibler (KL) divergence between the equilibrium distribution $f_{\mathrm{end}}^{\mathrm{eq}}$ corresponding to the final Hamiltonian $H(q, p; \tau)$ and the realized distribution f_{end} at the end of the

process because the following equality holds [36]:

$$
\begin{aligned}
W_{\mathrm{diss}} &= D_{\mathrm{KL}}\big(f_{\mathrm{end}}\big\|f_{\mathrm{end}}^{\mathrm{eq}}\big) \\
&:= \int f_{\mathrm{end}}(q(\tau), p(\tau))\ln\frac{f_{\mathrm{end}}(q(\tau), p(\tau))}{f_{\mathrm{end}}^{\mathrm{eq}}(q(\tau), p(\tau))}dq(\tau)dp(\tau).
\end{aligned}
$$

(B1)

With the setting in Sec. III A, by solving the Hamiltonian dynamics with an initial condition (q_0, p_0) we have the final state $(q(\tau), p(\tau))$ as

$$
q(\tau) = \left(\frac{p_0}{\sqrt{m_0}} - \frac{\mu_\tau}{\tau}\right)\sigma\,\sin\frac{\tau}{\sqrt{m_0}\sigma} + q_0\cos\frac{\tau}{\sqrt{m_0}\sigma} + \mu_\tau,
$$

(B2)

$$
\begin{aligned}
p(\tau) &= \left(\frac{p_0}{\sqrt{m_0}} - \frac{\mu_\tau}{\tau}\right)\cos\frac{\tau}{\sqrt{m_0}\sigma} \\
&\quad - \frac{\sqrt{m_0}q_0}{\sigma}\sin\frac{\tau}{\sqrt{m_0}\sigma} + \frac{m_0\mu_\tau}{\tau}.
\end{aligned}
$$

(B3)

Then, Liouville's theorem [37], which describes that the Hamiltonian dynamics preserves the density of states along each trajectory, is used to obtain the final distribution as

$$
f_{\mathrm{end}}(q(\tau), p(\tau)) = f_{\mathrm{init}}^{\mathrm{eq}}(q_0, p_0).
$$

(B4)

Finally, we can calculate the KL divergence as

$$
\begin{aligned}
&D_{\mathrm{KL}}\big(f_{\mathrm{end}}\big\|f_{\mathrm{end}}^{\mathrm{eq}}\big) \\
&= \int f_{\mathrm{init}}^{\mathrm{eq}}(q(\tau), p(\tau))\ln\frac{f_{\mathrm{init}}^{\mathrm{eq}}(q(\tau), p(\tau))}{f_{\mathrm{end}}^{\mathrm{eq}}(q(\tau), p(\tau))}dq(\tau)dp(\tau) \\
&= \int f_{\mathrm{init}}^{\mathrm{eq}}(q(\tau), p(\tau))\ln\frac{f_{\mathrm{init}}^{\mathrm{eq}}(q(\tau), p(\tau))}{f_{\mathrm{end}}^{\mathrm{eq}}(q(\tau), p(\tau))}dq_0 dp_0 \\
&= \frac{m_0\mu_\tau^2}{\tau^2}\left(1 - \cos\frac{\tau}{\sqrt{m_0}\sigma}\right),
\end{aligned}
$$

(B5)

where the second equality holds because the Jacobian for the variable transformation $q(\tau) \to q_0$ and $p(\tau) \to p_0$ is 1. With Eqs. (B1) and (B5) we obtain Eq. (18).

APPENDIX C: W_{diss} OF THE LJE FOR SCALING

We derive W_{diss} for the scaling protocol. We obtain a differential equation on the variance of position $\langle\!\langle q(t)\rangle\!\rangle$ by substituting $U(q; t)$ given by Eq. (21) for Eq. (A3):

$$
\frac{d\langle\!\langle q(t)\rangle\!\rangle}{dt} = \frac{\sigma_\tau^2(\tau - t) + \sigma_0^2 t}{\sigma_0^2\sigma_\tau^2\tau}\langle\!\langle q(t)\rangle\!\rangle + 2.
$$

(C1)

Then, the average work exerted during the process is

$$
\begin{aligned}
\langle W\rangle &= \left\langle\int_0^\tau \frac{\partial U(q; t)}{\partial t}dt\right\rangle \\
&= \int_0^\tau \frac{\sigma_0^2 - \sigma_\tau^2}{2\sigma_0\sigma_\tau^2\tau}\langle\!\langle q(t)\rangle\!\rangle dt.
\end{aligned}
$$

(C2)

We can calculate Eq. (C2) by solving Eq. (C1) numerically. Finally, we obtain W_{diss} by subtracting $\Delta F = -\ln[\sigma_\tau/\sigma_0]$.

Next, we derive W_{diss} for the scaling protocol given by Eq. (22). The derivative of the variance of position is

$$
\frac{d\langle\!\langle q(t)\rangle\!\rangle}{dt} = 2(1 - e^{-\frac{\gamma t}{\tau}})\langle\!\langle q(t)\rangle\!\rangle.
$$

(C3)

Then, the average work exerted during the process is

$$
\langle W\rangle = \int_0^\tau -\frac{\gamma}{2\sigma_0^2\tau}e^{-\frac{\gamma t}{\tau}}\langle\!\langle q(t)\rangle\!\rangle dt.
$$

(C4)

We can calculate Eq. (C4) by solving Eq. (C3) numerically. Finally, we obtain W_{diss} by subtracting $\Delta F = -\ln[\sigma_\tau^2/\sigma_0^2]$.

APPENDIX D: W_{diss} OF THE HJE FOR SCALING

We derive W_{diss} of the HJE for the scaling protocol. The distribution during the process is not Gaussian in this case, but we approximate the distribution with a Gaussian distribution. This approximation is valid when the system is near equilibrium during the process.

When a state (q, p) of a system satisfies the Hamiltonian dynamics, the time development of the probability distribution of the system $f(q, p, t)$ is described by Liouville's equation [37]:

$$
\frac{\partial f}{\partial t} + \sum_{i=1}^n\left(\frac{\partial f}{\partial q_i}\frac{dq_i}{dt} + \frac{\partial f}{\partial p_i}\frac{dp_i}{dt}\right) = 0,
$$

(D1)

where n is the dimension of q and p.

Similar to the case of the Fokker-Planck equation in Appendix A, we use Eq. (D1) to obtain the time derivatives of states' variance $\langle\!\langle\cdot\rangle\!\rangle$ and covariance $\langle\!\langle\cdot, \cdot\rangle\!\rangle$ at time t under the protocol given by Eq. (21):

$$
\frac{d\langle\!\langle q(t)\rangle\!\rangle}{dt} = 2\langle\!\langle q(t), p(t)\rangle\!\rangle,
$$

(D2)

$$
\frac{d\langle\!\langle p(t)\rangle\!\rangle}{dt} = -2\langle\!\langle q(t), p(t)\rangle\!\rangle\frac{\sigma_\tau^2\tau + (\sigma_0^2 - \sigma_\tau^2)t}{\sigma_0^2\sigma_\tau^2\tau},
$$

(D3)

$$
\begin{aligned}
\frac{d\langle\!\langle q(t), p(t)\rangle\!\rangle}{dt} &= \frac{1}{\sigma_0^2\sigma_\tau^2\tau}\big([\langle\!\langle p(t)\rangle\!\rangle\sigma_0^2 - \langle\!\langle q(t)\rangle\!\rangle]\sigma_\tau^2\tau \\
&\quad - \langle\!\langle q(t)\rangle\!\rangle(\sigma_0^2 - \sigma_\tau^2)t\big).
\end{aligned}
$$

(D4)

By solving Eqs. (D2)–(D4), we can calculate $\langle W\rangle$ given by Eq. (C2).

Next, we derive W_{diss} of the HJE for the scaling protocol with time-dependent mass. With the setting in Sec. III B, by solving the Hamiltonian dynamics with an initial condition (q_0, p_0), we have the final state $(q(\tau), p(\tau))$ as

$$
q(\tau) = e^{\frac{\gamma}{2}}(c_1\cos\omega\tau + c_2\sin\omega\tau),
$$

(D5)

$$
\begin{aligned}
p(\tau) &= \frac{1}{\sigma_0^2}e^{-\frac{\gamma}{2}}\bigg\{\frac{\gamma}{2\tau}(c_1\cos\omega\tau + c_2\sin\omega\tau) \\
&\quad + (-c_1\omega\sin\omega\tau + c_2\omega\cos\omega\tau)\bigg\},
\end{aligned}
$$

(D6)

where

$$
\omega = \frac{\sqrt{4\tau^2/\alpha - \gamma^2}}{2\tau},
$$

(D7)

$$
c_1 = q_0,
$$

(D8)

$$
c_2 = \frac{1}{\omega}\left(\frac{\sigma_0^2}{\alpha}p_0 - \frac{\gamma}{2\tau}q_0\right)
$$

(D9)

for $\tau > \sqrt{\alpha}|\gamma|/2$. Then we have the distribution at $t = \tau$ and we obtain W_{diss} in Eq. (25) by calculating the KL divergence in the same way as in Appendix B.

APPENDIX E: W_{diss} OF THE LJE FOR PARALLEL TRANSPORT AND SCALING

We derive W_{diss} of the LJE for the protocol that includes both parallel transport and scaling. We obtain differential equations for the average and variance of position $\langle q(t) \rangle$ and $\langle\langle q(t) \rangle\rangle$ by substituting $U(q;t)$ in Sec. III C for Eqs. (A2) and (A3):

$$\frac{d\langle q(t)\rangle}{dt} = \frac{a\mu t^2 - [a\langle q(t)\rangle - \mu]\tau t - \tau^2\langle q(t)\rangle}{\tau^2\sigma^2}, \quad \text{(E1)}$$

$$\frac{d\langle\langle q(t)\rangle\rangle}{dt} = -2\frac{a\langle\langle q(t)\rangle\rangle t + [\langle\langle q(t)\rangle\rangle - \sigma^2]\tau}{\tau\sigma^2}. \quad \text{(E2)}$$

Then, the average work exerted during the process is

$$\langle W \rangle = \left\langle \int_0^\tau \frac{\partial U(q;t)}{\partial t} dt \right\rangle$$
$$= \int_0^\tau \left[\frac{a}{2\sigma\tau}\left\{ \langle\langle q(t)\rangle\rangle + \langle q(t)\rangle^2 - \frac{2\mu t}{\tau}\langle q(t)\rangle + \left(\frac{\mu t}{\tau}\right)^2 \right\} - \frac{1 + a\frac{t}{\tau}}{\sigma^2}\frac{\mu}{\tau}\left(\langle q(t)\rangle - \frac{\mu t}{\tau}\right) \right] dt. \quad \text{(E3)}$$

We can calculate Eq. (E3) by solving Eqs. (E1) and (E2) numerically. Finally, we obtain W_{diss} by subtracting $\Delta F = 1/2\ln(1 + a)$ from $\langle W \rangle$.

APPENDIX F: W_{diss} OF THE HJE FOR PARALLEL TRANSPORT AND SCALING

We derive W_{diss} of the HJE for the protocol that includes both parallel transport and scaling. In the same manner as in Appendix D, we obtain the time derivatives of states' average $\langle \cdot \rangle$, variance $\langle\langle \cdot \rangle\rangle$, and covariance $\langle\langle \cdot, \cdot \rangle\rangle$ at time t with the setting in Sec. III C:

$$\frac{d\langle q(t)\rangle}{dt} = \langle p(t)\rangle, \quad \text{(F1)}$$

$$\frac{d\langle p(t)\rangle}{dt} = \frac{a\mu t^2 - [a\langle q(t)\rangle - \mu]t\tau - \tau^2\langle q(t)\rangle}{\tau^2\sigma^2}, \quad \text{(F2)}$$

$$\frac{d\langle\langle q(t)\rangle\rangle}{dt} = 2\langle\langle q(t), p(t)\rangle\rangle, \quad \text{(F3)}$$

$$\frac{d\langle\langle p(t)\rangle\rangle}{dt} = -\frac{2\langle\langle q(t), p(t)\rangle\rangle(at + \tau)}{\tau\sigma^2}, \quad \text{(F4)}$$

$$\frac{d\langle\langle q(t), p(t)\rangle\rangle}{dt} = -\frac{at\langle\langle q(t)\rangle\rangle - [\sigma^2\langle\langle p(t)\rangle\rangle - \langle\langle q(t)\rangle\rangle]\tau}{\tau\sigma^2}. \quad \text{(F5)}$$

Then, the average work $\langle W \rangle$ is described by Eq. (E3), which is the same as that used for the LJE. This is because the work exerted during a process only depends on a potential energy function.

We can calculate $\langle W \rangle$ by solving Eqs. (F1)–(F5) numerically and finally obtain W_{diss} by subtracting $\Delta F = 1/2\ln(1 + a)$ from $\langle W \rangle$.

APPENDIX G: W_{diss} OF THE LJE FOR ROTATION

We derive W_{diss} of the LJE for the rotation protocol. In the same manner as in Appendix C, we obtain the time derivatives of states' average $\langle \cdot \rangle$, variance $\langle\langle \cdot \rangle\rangle$, and covariance $\langle\langle \cdot, \cdot \rangle\rangle$ at time t with the setting in Sec. III D:

$$\frac{d\langle\langle x(t)\rangle\rangle}{dt} = -4\langle\langle x(t)\rangle\rangle(a\cos\theta(t)^2 + b\sin\theta(t)^2) + 4\langle\langle x(t), y(t)\rangle\rangle(a - b)\cos\theta(t)\sin\theta(t) + 2, \quad \text{(G1)}$$

$$\frac{d\langle\langle y(t)\rangle\rangle}{dt} = -4\langle\langle y(t)\rangle\rangle(b\cos\theta(t)^2 + a\sin\theta(t)^2) + 4\langle\langle x(t), y(t)\rangle\rangle(a - b)\cos\theta(t)\sin\theta(t) + 2, \quad \text{(G2)}$$

$$\frac{d\langle\langle x(t), y(t)\rangle\rangle}{dt} = 2[\langle\langle x(t)\rangle\rangle + \langle\langle y(t)\rangle\rangle] \times (a - b) \times \cos\theta(t)\sin\theta(t) - 2\langle\langle x(t), y(t)\rangle\rangle(a + b). \quad \text{(G3)}$$

Then, the time derivative of $U(x, y; t)$ given by Eq. (30) is

$$\frac{\partial U(x, y; t)}{\partial t} = 2\frac{d\theta}{dt}\{(a - b)\sin\theta(t)\cos\theta(t)(-x^2 + y^2) - (a - b)(\cos^2\theta - \sin^2\theta)xy\}. \quad \text{(G4)}$$

We describe W_{diss} by x, y, and θ with the expected value of Eq. (G4):

$$W_{\text{diss}} = \left\langle \int_0^\tau \frac{\partial U(x, y; t)}{\partial t} dt \right\rangle$$
$$= \int_0^\tau \left\langle \frac{\partial U(x, y; t)}{\partial t} \right\rangle dt$$
$$= \int_0^\tau 2\frac{d\theta}{dt}\{(a - b)\sin\theta(t)\cos\theta(t)(-\langle x^2\rangle + \langle y^2\rangle) - (a - b)(\cos^2\theta(t) - \sin^2\theta(t))\langle xy\rangle dt\}. \quad \text{(G5)}$$

The expected values in Eq. (G5) can be obtained by numerically solving Eqs. (G1)–(G3).

APPENDIX H: W_{diss} OF THE HJE FOR ROTATION

We derive W_{diss} of the HJE for the rotation protocol. In the same manner as in Appendix D, we obtain the time derivatives of states' average $\langle \cdot \rangle$ at time t under the setting in Sec. III D:

$$\frac{d\langle x^2\rangle}{dt} = 2\langle xp_x\rangle, \quad \text{(H1)}$$

$$\frac{d\langle y^2\rangle}{dt} = 2\langle yp_y\rangle, \quad \text{(H2)}$$

$$\frac{d\langle p_x^2\rangle}{dt} = -4[a\cos^2\theta(t) + b\sin^2\theta(t)]\langle xp_x\rangle + 4(a - b)\sin\theta(t)\cos\theta(t)\langle yp_x\rangle, \quad \text{(H3)}$$

$$\frac{d\langle p_y^2\rangle}{dt} = -4[b\cos^2\theta(t) + a\sin^2\theta(t)]\langle yp_y\rangle + 4(a - b)\sin\theta(t)\cos\theta(t)\langle xp_y\rangle, \quad \text{(H4)}$$

$$\frac{d\langle xp_x \rangle}{dt} = -2[a\cos^2\theta(t) + b\sin^2\theta(t)]\langle x^2 \rangle$$
$$+2(a-b)\cos\theta(t)\sin\theta(t)\langle xy \rangle + \langle p_x^2 \rangle, \quad \text{(H5)}$$

$$\frac{d\langle xy \rangle}{dt} = \langle xp_y \rangle + \langle yp_y \rangle, \quad \text{(H6)}$$

$$\frac{d\langle xp_y \rangle}{dt} = -2[b\cos^2\theta(t) + a\sin^2\theta(t)]\langle xy \rangle$$
$$+2(a-b)\sin\theta(t)\cos\theta(t)\langle x^2 \rangle + \langle p_x p_y \rangle, \quad \text{(H7)}$$

$$\frac{d\langle yp_x \rangle}{dt} = -2[a\cos^2\theta(t) + b\sin^2\theta(t)]\langle xy \rangle$$
$$+2(a-b)\sin\theta(t)\cos\theta(t)\langle y^2 \rangle + \langle p_x p_y \rangle, \quad \text{(H8)}$$

$$\frac{d\langle p_x p_y \rangle}{dt} = -2(a\langle xp_y \rangle + b\langle yp_y \rangle)\cos^2\theta(t)$$
$$+2(a-b)\sin\theta(t)\cos\theta(t)(\langle xp_x \rangle + \langle yp_y \rangle)$$
$$-2(a\langle yp_x \rangle + b\langle xp_x \rangle)\sin^2\theta(t), \quad \text{(H9)}$$

$$\frac{d\langle yp_y \rangle}{dt} = -2[b\cos^2\theta(t) + a\sin^2\theta(t)]\langle y^2 \rangle$$
$$+2(a-b)\sin\theta(t)\cos\theta(t)\langle xy \rangle + \langle p_y^2 \rangle. \quad \text{(H10)}$$

The expected values $\langle W \rangle$ given by Eq. (G5) can be obtained by numerically solving Eqs. (H1)–(H10).

APPENDIX I: Z FOR A MULTIMODAL DISTRIBUTION

We calculate the partition function Z of the target distribution given by Eq. (39) in Sec. IV C. Considering the superposition of distributions, Z is given by

$$Z = N\int_{-\infty}^{\infty} e^{-\cosh\|q\|_2}dq. \quad \text{(I1)}$$

Therefore, we focus on the integration on the right-hand side. Let S_{D-1} be the surface area of the $(D-1)$-dimensional unit sphere:

$$S_{D-1} = \frac{2\pi^{D/2}}{\Gamma(D/2)}, \quad \text{(I2)}$$

where $\Gamma(x)$ is the Gamma function. In D-dimensional polar coordinates, Eq. (I1) is

$$Z = NS_{D-1}\int_0^{\infty} e^{-\cosh r}r^{D-1}dr. \quad \text{(I3)}$$

We can obtain Z for $D = 16$ and arbitrary N by numerically calculating Eq. (I3).

[1] C. Jarzynski, Nonequilibrium Equality for Free Energy Differences, Phys. Rev. Lett. **78**, 2690 (1997).

[2] H. Suwa and S. Todo, Markov Chain Monte Carlo Method without Detailed Balance, Phys. Rev. Lett. **105**, 120603 (2010).

[3] M. Ohzeki, Quantum Annealing with the Jarzynski Equality, Phys. Rev. Lett. **105**, 050401 (2010).

[4] G. M. Rotskoff and E. Vanden-Eijnden, Dynamical Computation of the Density of States and Bayes Factors Using Nonequilibrium Importance Sampling, Phys. Rev. Lett. **122**, 150602 (2019).

[5] M. Habeck, Model evidence from nonequilibrium simulations, Adv. Neural Inf. Proc. Syst. 1753 (2017).

[6] G. E. Crooks, Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences, Phys. Rev. E **60**, 2721 (1999).

[7] S. Vaikuntanathan and C. Jarzynski, Escorted Free Energy Simulations: Improving Convergence by Reducing Dissipation, Phys. Rev. Lett. **100**, 190601 (2008).

[8] F. Reif, *Fundamentals of Statistical and Thermal Physics* (McGraw-Hill, New York, 1965).

[9] D. J. C. MacKay, *Information Theory, Inference, and Learning Algorithms* (Cambridge University Press, Cambridge, 2003).

[10] E. T. Jaynes, *Probability Theory* (Cambridge University Press, Cambridge, 2003).

[11] L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Pergamon, London, 1958), Vol. 5.

[12] G. Hummer and A. Szabo, Free energy profiles from single-molecule pulling experiments, Proc. Natl. Acad. Sci. (USA) **107**, 21441 (2010).

[13] F. Ritort, Nonequilibrium fluctuations in small systems: From physics to biology, Adv. Chem. Phys. **137**, 31 (2008).

[14] R. M. Neal, Annealed importance sampling, Statist. Comput. **11**, 125 (2001).

[15] C. Jarzynski, Rare events and the convergence of exponentially averaged work values, Phys. Rev. E **73**, 046105 (2006).

[16] J. Gore, F. Ritort, and C. Bustamante, Bias and error in estimates of equilibrium free-energy differences from nonequilibrium measurements., Proc. Natl. Acad. Sci. (USA) **100**, 12564 (2003).

[17] C. Yang, B. Wan, S. Xu, Y. Wang, and X. Zhou, Equilibrium sampling by reweighting nonequilibrium simulation trajectories, Phys. Rev. E **93**, 033309 (2016).

[18] B. Wan, C. Yang, Y. Wang, and X. Zhou, Jarzynski matrix equality: Calculating the free-energy difference by nonequilibrium simulations with an arbitrary initial distribution, Phys. Rev. E **93**, 043312 (2016).

[19] M. Arrar, F. M. Boubeta, M. E. Szretter, M. Sued, L. Boechi, and D. Rodriguez, On the accurate estimation of free energies using the Jarzynski equality, J. Comput. Chem. **40**, 688 (2019).

[20] T. Schmiedl and U. Seifert, Optimal Finite-Time Processes In Stochastic Thermodynamics, Phys. Rev. Lett. **98**, 108301 (2007).

[21] Y. Zhang, Work needed to drive a thermodynamic system between two distributions, Europhys. Lett. **128**, 30002 (2020).

[22] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, Hybrid Monte Carlo, Phys. Lett. B **195**, 216 (1987).

[23] C. Bishop, *Pattern Recognition and Machine Learning* (Springer, New York, 2006).

[24] J. Sohl-Dickstein and B. J. Culpepper, Hamiltonian annealed importance sampling for partition function estimation, arXiv:1205.1924.

[25] W. G. Hoover, Canonical dynamics: Equilibrium phase-space distributions, Phys. Rev. A **31**, 1695 (1985).

[26] C. Bustamante, J. Liphardt, and F. Ritort, The nonequilibrium thermodynamics of small systems, Phys. Today **58**(7), 43 (2005).

[27] J. Horowitz and C. Jarzynski, Illustrative example of the relationship between dissipation and relative entropy, Phys. Rev. E **79**, 021106 (2009).

[28] Y. Jun, M. Gavrilov, and J. Bechhoefer, High-Precision Test of Landauer's Principle in a Feedback Trap, Phys. Rev. Lett. **113**, 190601 (2014).

[29] A. Retzker, R. C. Thompson, D. M. Segal, and M. B. Plenio, Double Well Potentials and Quantum Phase Transitions in Ion Traps, Phys. Rev. Lett. **101**, 260504 (2008).

[30] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations* (Springer, New York, 1992).

[31] A. Rößler, Runge-Kutta methods for the strong approximation of solutions of stochastic differential equations, SIAM J. Numer. Anal. **48**, 922 (2010).

[32] P. E. Rouse, A theory of the linear viscoelastic properties of dilute solutions of coiling polymers, J. Chem. Phys. **21**, 1272 (1953).

[33] G. R. Strobl, *The Physics of Polymers* (Springer, New York, 1997).

[34] A. Dhar, Work distribution functions in polymer stretching experiments, Phys. Rev. E **71**, 036126 (2005).

[35] H. Risken and T. Frank, *The Fokker-Planck Equation* (Springer, New York, 1989).

[36] S. Vaikuntanathan and C. Jarzynski, Dissipation and lag in irreversible processes, Europhys. Lett. **87**, 60005 (2009).

[37] M. Tuckerman, *Statistical Mechanics: Theory and Molecular Simulation* (Oxford University Press, New York, 2010).