

α -divergence improves the entropy production estimation via machine learningEuijoon Kwon and Yongjoo Baek ^{*}*Department of Physics and Astronomy & Center for Theoretical Physics, Seoul National University, Seoul 08826, Republic of Korea*

(Received 4 March 2023; revised 27 November 2023; accepted 5 January 2024; published 30 January 2024)

Recent years have seen a surge of interest in the algorithmic estimation of stochastic entropy production (EP) from trajectory data via machine learning. A crucial element of such algorithms is the identification of a loss function whose minimization guarantees the accurate EP estimation. In this study we show that there exists a host of loss functions, namely, those implementing a variational representation of the α -divergence, which can be used for the EP estimation. By fixing α to a value between -1 and 0 , the α -NEEP (Neural Estimator for Entropy Production) exhibits a much more robust performance against strong nonequilibrium driving or slow dynamics, which adversely affects the existing method based on the Kullback-Leibler divergence ($\alpha = 0$). In particular, the choice of $\alpha = -0.5$ tends to yield the optimal results. To corroborate our findings, we present an exactly solvable simplification of the EP estimation problem, whose loss function landscape and stochastic properties give deeper intuition into the robustness of the α -NEEP.

DOI: [10.1103/PhysRevE.109.014143](https://doi.org/10.1103/PhysRevE.109.014143)**I. INTRODUCTION**

How irreversible does a process look? One may pose this question for two distinct reasons. First, whether a biological process requires energy dissipation is often a subject of much debate [1,2]. To resolve this issue, it is useful to note that irreversibility suggests energy dissipation. Various hallmarks of irreversibility, such as the breaking of the fluctuation-dissipation theorem [3] and the presence of nonequilibrium probability currents in the phase space [4,5], have been used to determine whether energy is dissipated. Second, whether a nonequilibrium system allows for an effective equilibrium description is an important issue. For instance, in active matter, despite the energy dissipation at the microscopic level, it has been argued that the large-scale phenomena allow for an effective equilibrium description [6–10]. If we can quantify the irreversibility of an empirical process at various levels of coarse graining [11,12], it will provide us with helpful clues as to whether we should look for an effective equilibrium theory for the process.

Based on the framework of stochastic thermodynamics, modern thermodynamics assigns entropy production (EP) to each stochastic trajectory based on its irreversibility [13]. Thus, empirically measuring the irreversibility of a process is closely tied to the problem of estimating EP from sampled trajectories [14–21]. A straightforward approach to the problem is to evaluate the relevant transition probabilities by directly counting the number of trajectory segments, which is called the *plug-in method* [14,15]. The method, readily applicable to discrete systems, can also be applied to continuous systems through the use of kernel functions [16]. However, while this method is simple and intuitive, it requires a huge ensemble of lengthy trajectories for accurate estimations (*curse of*

dimensionality). More recent studies proposed methods based on universal lower bounds of the average EP, such as the thermodynamic uncertainty relations [16–19] and the entropic bound [20]. While these methods do not suffer from the curse of dimensionality and are applicable even to nonstationary processes [19,20], their accuracy is impaired when the underlying bounds are not tight. Moreover, these methods are applicable only to the estimation of the average EP, not the EP of each trajectory.

Meanwhile, with the advent of machine learning techniques in physics, a novel method for EP estimation using artificial neural networks has been developed [21]. This method, called the Neural Estimator for Entropy Production (NEEP), minimizes the loss function based on a variational representation of the Kullback-Leibler (KL) divergence. Without any presupposed discretization of the phase space and using the rich expressivity of neural networks, the NEEP suffers far less from the complications of the sampling issues and is applicable to a diverse range of stochastic processes [19].

Still, the NEEP has its limits. Its accuracy deteriorates when the nonequilibrium driving is strong or when the dynamics slows down so that the phase space is poorly sampled. In this study we show that the NEEP can be significantly improved by changing the loss function. Toward this purpose, we propose the α -NEEP, which generalizes the NEEP. Instead of the KL divergence, the α -NEEP utilizes the α -divergence, which has been mainly used in the machine learning community [22–25]. We demonstrate that the α -NEEP with nonzero values of α shows much more robust performance for a broader range of nonequilibrium driving and sampling quality, with $\alpha = -0.5$ showing the optimal performance overall. This is corroborated by an analytically tractable simplification of the α -NEEP that shows the optimality of $\alpha = -0.5$.

The rest of this paper is organized as follows. After reviewing the original NEEP and its limitations (Sec. II), we introduce the α -NEEP (Sec. III) and demonstrate its enhanced

^{*}y.baek@snu.ac.kr

performance for three different examples of nonequilibrium systems (Sec. IV). Then we investigate the rationale behind the observed results using a simplified model describing how the α -NEEP works (Sec. V). Finally, we sum up the results and discuss their implications (Sec. VI).

II. OVERVIEW OF THE ORIGINAL NEEP

We first give a brief overview of how the original NEEP [21] estimates EP at the trajectory level. Suppose our goal is to estimate EP of a Markov process in discretized time, \mathbf{x}_t , in a d -dimensional space. For every ordered pair of states, denoted by $\mathbf{x} \equiv (\mathbf{x}_t, \mathbf{x}_{t+1})$, there is EP associated with the transition between them, which is given by the ratio between the forward and the backward path probabilities

$$\Delta S(\mathbf{x}) = \log \frac{p(\mathbf{x})}{p(\tilde{\mathbf{x}})}, \quad (1)$$

where $\tilde{\mathbf{x}} \equiv (\mathbf{x}_{t+1}, \mathbf{x}_t)$. Note that, throughout this study, we use the unit system in which the Boltzmann constant can be set to unity ($k_B = 1$). Then it follows that the ensemble average of this EP is equivalent to the KL divergence, which satisfies the inequality

$$\begin{aligned} \langle \Delta S(\mathbf{x}) \rangle &= D_{\text{KL}}[p(\mathbf{x}), p(\tilde{\mathbf{x}})] \\ &\geq \left\langle \log r(\mathbf{x}) - \frac{p(\tilde{\mathbf{x}})}{p(\mathbf{x})} r(\mathbf{x}) + 1 \right\rangle_{p(\mathbf{x})} \end{aligned} \quad (2)$$

for any positive function $r(\mathbf{x})$, given that $\langle \cdot \rangle_{p(\mathbf{x})}$ denotes the average with respect to the distribution $p(\mathbf{x})$. This inequality can be proven as follows: since \log is a concave function, the line tangent to any point never falls below the function. Thus, $\frac{1}{r_0}(r - r_0) + \log r_0 \geq \log r$ for any r and r_0 . By putting $r_0 = r_0(\mathbf{x}) = p(\mathbf{x})/p(\tilde{\mathbf{x}})$ and taking the average with respect to $p(\mathbf{x})$, we get the inequality. In this derivation, we immediately note that the equality condition is satisfied if and only if $r(\mathbf{x}) = r_0(\mathbf{x}) = p(\mathbf{x})/p(\tilde{\mathbf{x}})$. Hence, by varying $r(\mathbf{x})$ to maximize the right-hand side of Eq. (2), we accurately estimate the average EP $\langle \Delta S(\mathbf{x}) \rangle$. For this reason, Eq. (2) is called the *variational representation* of the KL divergence. Moreover, as a byproduct, we also obtain the function $r_0(\mathbf{x})$, which yields an accurate estimate for trajectory-level EP by $\Delta S(\mathbf{x}) = \log r_0(\mathbf{x})$.

Kim *et al.* [21] used these properties to construct the loss function of the NEEP. More specifically, they introduce $s_\theta(\mathbf{x})$, an estimator for trajectory-level EP parametrized by θ , and put $r(\mathbf{x}) = e^{s_\theta(\mathbf{x})}$. Then Eq. (2) can be rewritten as

$$\langle \Delta S(\mathbf{x}) \rangle \geq \langle s_\theta(\mathbf{x}) - e^{s_\theta(\tilde{\mathbf{x}})} + 1 \rangle_{p(\mathbf{x})}, \quad (3)$$

where $\langle r(\mathbf{x}) p(\tilde{\mathbf{x}})/p(\mathbf{x}) \rangle_{p(\mathbf{x})} = \langle r(\mathbf{x}) \rangle_{p(\tilde{\mathbf{x}})} = \langle r(\tilde{\mathbf{x}}) \rangle_{p(\mathbf{x})}$ has been used based on the one-to-one correspondence between \mathbf{x} and $\tilde{\mathbf{x}}$. Furthermore, since EP is odd under time reversal, i.e., $\Delta S(\mathbf{x}) = -\Delta S(\tilde{\mathbf{x}})$, it is natural to impose the same condition on s_θ . This leads to the inequality

$$\langle \Delta S(\mathbf{x}) \rangle \geq \langle s_\theta(\mathbf{x}) - e^{-s_\theta(\mathbf{x})} + 1 \rangle_{p(\mathbf{x})}, \quad (4)$$

which motivates the loss function

$$\mathcal{L}(\theta) = \langle -s_\theta(\mathbf{x}) + e^{-s_\theta(\mathbf{x})} - 1 \rangle_{p(\mathbf{x})}, \quad (5)$$

so that the minimization of $\mathcal{L}(\theta)$ ensures the accurate EP estimation $s_\theta(\mathbf{x}) = \Delta S(\mathbf{x})$.

It is notable that $\mathcal{L}(\theta)$ defined above is a convex functional of s_θ . Thus, as long as the θ -dependence of s_θ is well behaved, any gradient-descent algorithm can reach the global minimum of $\mathcal{L}(\theta)$ without getting trapped in a local minimum. In this regard, the rugged loss function landscape is not a major issue of the NEEP.

However, the performance of the NEEP strongly depends on how well $p(\mathbf{x})$ is sampled. Since the second term of $\mathcal{L}(\theta)$ depends exponentially on $s_\theta(\mathbf{x})$, rare transitions with minute $p(\mathbf{x})$ can make nonnegligible contributions to $\mathcal{L}(\theta)$ when $e^{-s_\theta(\mathbf{x})}$ is extremely large. Since the frequency of rare events is subject to considerable sampling noise, the performance of the original NEEP deteriorates in the presence of a strong nonequilibrium driving, which induces rare transitions with large negative EP. In the following section, we propose a loss function that remedies this weakness of the NEEP.

III. FORMULATION OF THE α -NEEP

Here we formulate a generalization of the NEEP loss function with the goal of mitigating its strong sampling-noise dependence. We note that the loss function needs not be an estimator of average EP $\langle \Delta S(\mathbf{x}) \rangle$, for our goal is to estimate $\Delta S(\mathbf{x})$ at the level of each trajectory. Thus, while the original NEEP uses the variational representation of the KL divergence corresponding to $\langle \Delta S(\mathbf{x}) \rangle$, we propose a different approach based on the variational representation of the α -divergence, which quantifies the difference between a pair of probability distributions $p(\mathbf{x})$ and $q(\mathbf{x})$ as

$$D_\alpha[p; q] \equiv \left\langle \frac{[p(\mathbf{x})/q(\mathbf{x})]^\alpha - 1}{\alpha(1 + \alpha)} \right\rangle_{p(\mathbf{x})}. \quad (6)$$

Since this reduces to the KL divergence in the limit $\alpha \rightarrow 0$, our approach generalizes the NEEP by introducing an extra parameter α . To emphasize this aspect, we term our method the α -NEEP.

The goal of the α -NEEP is to find $r(\mathbf{x})$ that minimizes the loss function

$$\mathcal{L}_\alpha[r] \equiv \left\langle -\frac{r(\mathbf{x})^\alpha - 1}{\alpha} + \frac{q(\mathbf{x}) r(\mathbf{x})^{1+\alpha} - 1}{p(\mathbf{x}) (1 + \alpha)} \right\rangle_{p(\mathbf{x})}, \quad (7)$$

where $p(\mathbf{x})$ and $q(\mathbf{x})$ are probability density functions, and α is a real number other than 0 and -1 . See Appendix B for discussion of these two exceptional cases. It can be rigorously shown (see Appendix B) that $\mathcal{L}_\alpha[r]$ satisfies the inequality

$$\mathcal{L}_\alpha[r] \geq -D_\alpha[p; q], \quad (8)$$

where the equality is achieved if and only if $r(\mathbf{x}) = p(\mathbf{x})/q(\mathbf{x})$ for all \mathbf{x} . In other words, by minimizing $\mathcal{L}_\alpha[r]$ to find $D_\alpha[p; q]$, we also obtain an estimate for the ratio $p(\mathbf{x})/q(\mathbf{x})$. We note that the properties of $\mathcal{L}_\alpha[r]$ used here are also valid for a much more general class of loss functions, as discussed in [22,23] (also see Appendix B).

Based on Eq. (8), we can construct a loss function

$$\mathcal{L}_\alpha(\theta) = \left\langle -\frac{e^{\alpha s_\theta(\mathbf{x})} - 1}{\alpha} + \frac{e^{-(1+\alpha)s_\theta(\mathbf{x})} - 1}{1 + \alpha} \right\rangle. \quad (9)$$

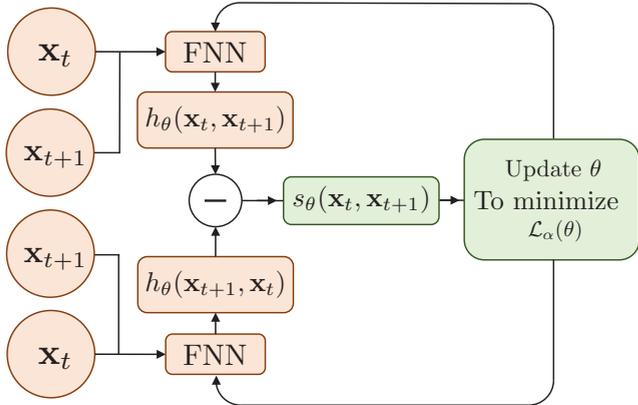


FIG. 1. Schematic illustration of the neural-network implementation of the α -NEEP.

Note that this reduces to the loss function of the original NEEP shown in Eq. (5) in the limit $\alpha \rightarrow 0$. If s_θ is sufficiently well behaved, the minimization of $\mathcal{L}_\alpha(\theta)$ yields the minimizer θ^* , which satisfies $\mathcal{L}_\alpha(\theta^*) = -D_\alpha[p(\mathbf{x}), p(\tilde{\mathbf{x}})]$ and $\Delta S(\mathbf{x}) = s_{\theta^*}(\mathbf{x})$. The former is generally not equal to average EP $\langle \Delta S(\mathbf{x}) \rangle$ (unless $\alpha \rightarrow 0$), but the latter ensures the accurate estimation of trajectory-level EP $\Delta S(\mathbf{x})$.

Comparing Eqs. (5) and (9), one readily observes that the exponential dependence on $s_\theta(\mathbf{x})$ can be made much weaker in $\mathcal{L}_\alpha(\theta)$ by choosing the value of α between -1 and 0 . Since this mitigates the detrimental effects of the sampling error associated with rare trajectories with large negative $s_\theta(\mathbf{x})$, one can naturally expect that the performance of the α -NEEP is much more robust against strong nonequilibrium driving. This is confirmed in the following sections.

Before proceeding, a few remarks are in order:

(1) The loss function satisfies $\mathcal{L}_{-(1+\alpha)}(\theta) = \mathcal{L}_\alpha(\theta)$, so the α -NEEP is symmetric under the exchange $\alpha \leftrightarrow -(1 + \alpha)$. For this reason, in the rest of this paper, we focus on the regime $-0.5 \leq \alpha \leq 0$ (the regime $\alpha > 0$ leads to very poor performance and is left out).

(2) From the antisymmetry $\Delta S(\mathbf{x}) = -\Delta S(\tilde{\mathbf{x}})$, we may set the estimator s_θ to be related to the feed-forward neural network (FNN) output h_θ as

$$s_\theta(\mathbf{x}) = h_\theta(\mathbf{x}) - h_\theta(\tilde{\mathbf{x}}), \quad (10)$$

so that the neural network focuses on the estimators that satisfy the antisymmetry of EP for more efficient training. The method described so far is schematically illustrated in Fig. 1.

(3) We emphasize that the minimized \mathcal{L}_α is not directly related to average EP. In all cases we compute the average EP by averaging s_θ over the sampled transitions.

IV. EXAMPLES

To assess the performance of the α -NEEP for various values of α , we apply the method to toy models of nonequilibrium systems, namely, the two-bead model, the Brownian gyrotator, and the driven Brownian particle.

(i) *The two-bead model.* This model has been used in a number of previous studies as a benchmark for testing EP estimators [4,16,18,21]. The model consists of two

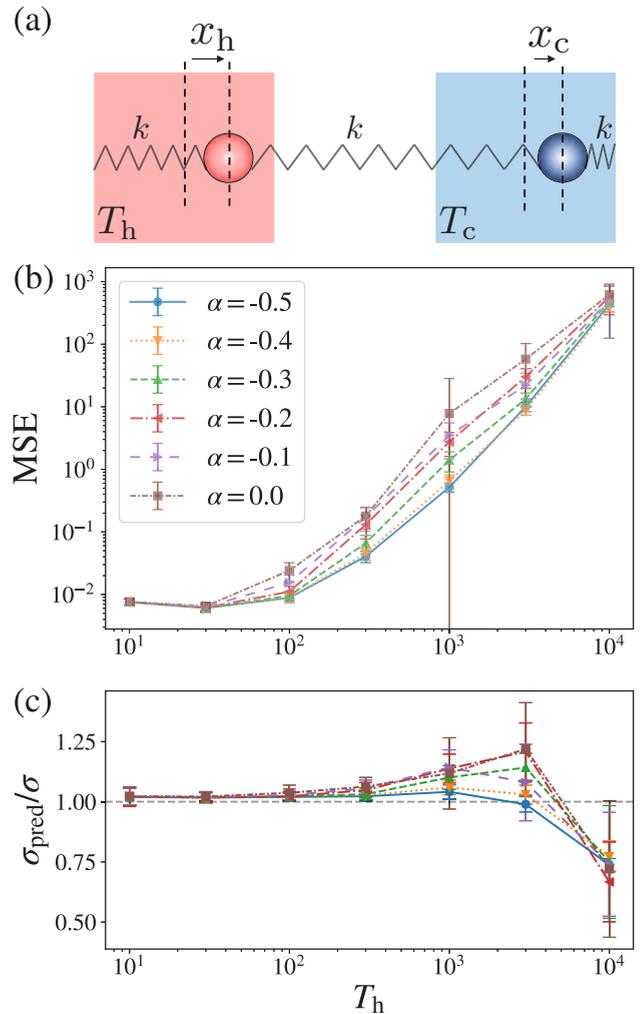


FIG. 2. (a) Illustration of the two-bead model. (b) Mean square error (MSE) of the EP estimate for various temperature differences. (c) Ratio between the estimated value σ_{pred} and the true value σ of average EP for the two-bead model. Temperature of the cold bath is fixed at $T_c = 1$. Each data point and error bar are obtained from 40 independent trainings.

one-dimensional (1D) overdamped beads which are connected to each other and to the walls on both sides by identical springs; see Fig. 2(a). The beads are in contact with heat baths at temperatures T_h and T_c with $T_h > T_c$. Denoting by x_h (x_c) the bead in contact with the hot (cold) bath, the stochastic equations of motion are given by

$$\gamma \dot{x}_h = k(-2x_h + x_c) + \sqrt{2\gamma T_h} \xi_h(t), \quad (11a)$$

$$\gamma \dot{x}_c = k(-2x_c + x_h) + \sqrt{2\gamma T_c} \xi_c(t). \quad (11b)$$

Here k is the spring constant, γ the friction coefficient, and $\xi_{h,c}$ the Gaussian thermal noise with zero means and $\langle \xi_h(t) \xi_h(t') \rangle = \langle \xi_c(t) \xi_c(t') \rangle = \delta(t - t')$. For infinitesimal displacements (dx_h, dx_c) , the associated EP is given by

$$\Delta S = \frac{k}{\gamma} \left[\frac{2x_h - x_c}{T_h} \circ dx_h + \frac{2x_c - x_h}{T_c} \circ dx_c \right] + \Delta S_{\text{sys}}, \quad (12)$$

where \circ denotes the Stratonovich product and ΔS_{sys} the change of the system's Shannon entropy, namely,

$$\Delta S_{\text{sys}} = -\ln \frac{p_s(x_h + dx_h, x_c + dx_c)}{p_s(x_h, x_c)}, \quad (13)$$

for the steady-state distribution $p_s(x_h, x_c)$. Since the system is fully linear, $p_s(x_h, x_c)$ can be calculated analytically. Thus the EP of this model can be calculated exactly using Eq. (12) and compared with the α -NEEP result.

To see how the predicted EP differs from the true EP, we observe the behavior of the mean square error (MSE) $\langle (s_\theta - \Delta S)^2 \rangle$. In Fig. 2(b) we observe that strengthening the nonequilibrium driving (by increasing T_h while keeping $T_c = 1$) tends to impair the EP estimation. This is because a stronger driving makes the reverse trajectories of typical trajectories rarer, lowering the sample quality. The adverse effects of the nonequilibrium driving are the strongest for the original NEEP ($\alpha = 0$), which are mitigated by choosing different values of α . Remarkably, choosing $\alpha = -0.5$ leads to the most robust performance against the driving.

As an alternative measure of the estimator's performance, we also observe the ratio between the predicted average EP σ_{pred} and the exact average EP σ . The results are shown in Fig. 2(c), which exhibit two different regimes. As T_h increases, there is a regime where the estimator overestimates average EP, which is followed by an underestimation regime. A detailed explanation for this behavior will be given in Sec. V using a simplified model. At the moment, we note that $\sigma_{\text{pred}}/\sigma$ tends to deviate away from 1 most strongly for the original NEEP ($\alpha = 0$), while choosing different values of α makes the ratio stay closer to 1. Again, the optimal value of α seems to be -0.5 .

(ii) *The Brownian gyrotor.* This simple model of a single-particle heat engine allows us to check the effects of a nonequilibrium driving apart from the temperature difference $T_h > T_c$. The dynamics of the model is governed by

$$\gamma \dot{x}_h = -\partial_{x_h} U(x_h, x_c) + \varepsilon x_c + \sqrt{2\gamma T_h} \xi_h(t), \quad (14a)$$

$$\gamma \dot{x}_c = -\partial_{x_c} U(x_h, x_c) + \delta x_h + \sqrt{2\gamma T_c} \xi_c(t), \quad (14b)$$

where $U(x_h, x_c) = \frac{1}{2}k(x_h^2 + x_c^2)$ is the harmonic potential, and $(\varepsilon x_c, \delta x_h)$ is a nonconservative force that drives the system out of equilibrium and enables work extraction. See Fig. 3(a) for an illustration of this system. For infinitesimal displacements (dx_h, dx_c) , the associated EP is given by

$$\Delta S = -\frac{Q_h}{T_h} - \frac{Q_c}{T_c} + \Delta S_{\text{sys}}, \quad (15)$$

where

$$Q_h = (\partial_{x_h} U - \varepsilon x_c) \circ dx_h, \quad (16a)$$

$$Q_c = (\partial_{x_c} U - \delta x_h) \circ dx_c, \quad (16b)$$

and ΔS_{sys} the change of the system entropy. Again, the system is fully linear, and the steady-state distribution can be calculated analytically, allowing exact calculations of EP at the trajectory level.

Setting $T_h/T_c = 10$ and $\varepsilon = -\delta$, we vary the magnitude of ε to assess the robustness of the α -NEEP in terms of the MSE and the ratio $\sigma_{\text{pred}}/\sigma$, as shown in Figs. 3(b) and 3(c), respectively. The results are qualitatively similar to the case of the

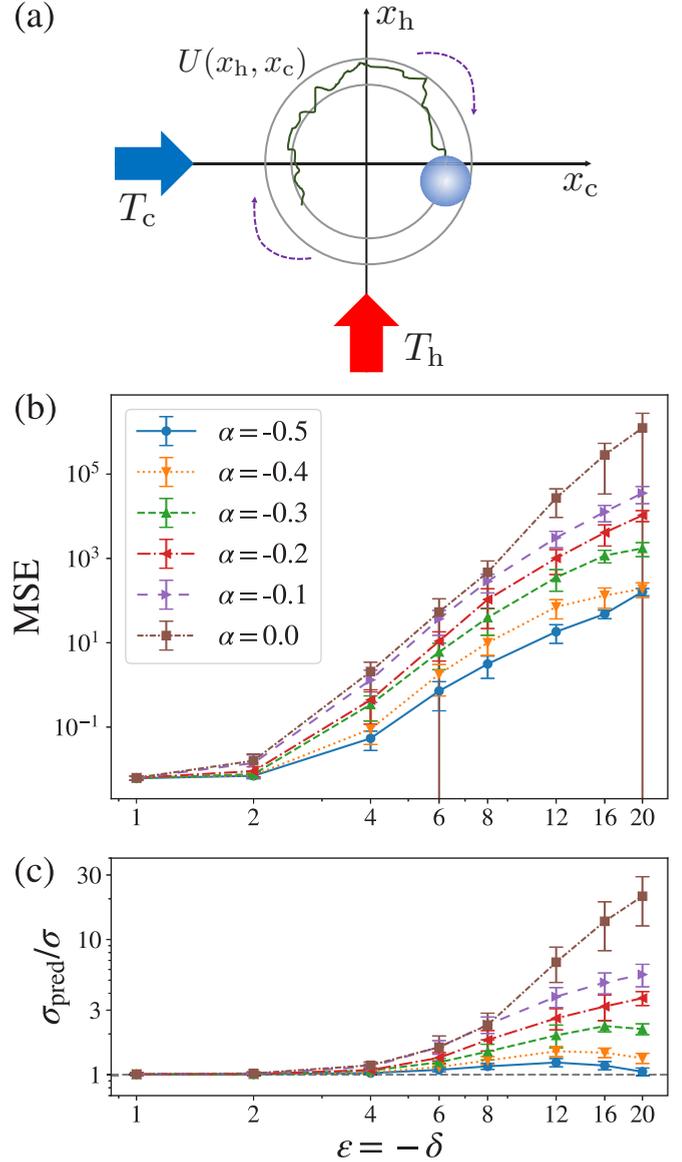


FIG. 3. (a) Illustration of the Brownian gyrotor. Circles represent the equipotential lines, and the dashed arrows indicate the directions of the nonconservative driving. (b) MSE of the EP estimate for the Brownian gyrotor model as the magnitude of nonconservative force, $\varepsilon = -\delta$, is varied. (c) Ratio between the estimated value σ_{pred} and the true value σ of average EP for the Brownian gyrotor. Temperatures are fixed at $T_h = 10$ and $T_c = 1$. Each data point and error bar are obtained from 40 independent trainings.

two-bead model: as the nonconservative driving gets stronger, the performance of the original NEEP ($\alpha = 0$) deteriorates the most, while other values of α yield more robust results. Again, $\alpha = -0.5$ seems to be the optimal choice.

(iii) *The driven Brownian particle.* While the two examples given above were both linear systems, we also consider a nonlinear system featuring a 1D overdamped Brownian particle in a periodic potential $U(x) = A \sin x$ driven by a constant force f . The motion of the particle is described by the Langevin equation

$$\gamma \dot{x} = f - U'(x) + \sqrt{2\gamma T} \xi(t), \quad (17)$$

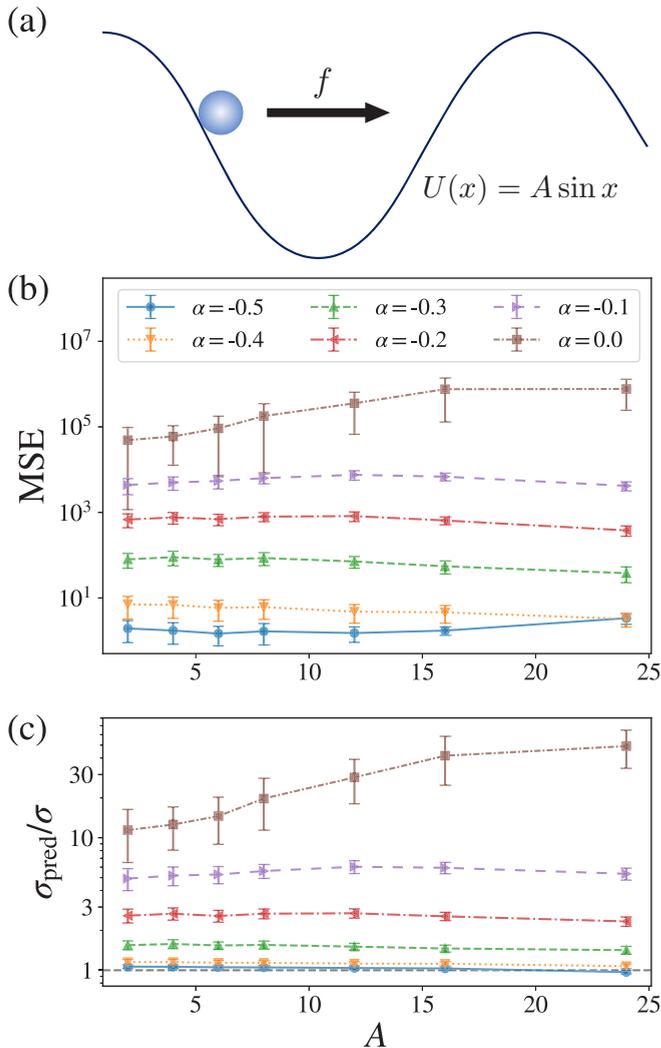


FIG. 4. (a) Illustration of the driven Brownian particle. (b) MSE of the EP estimate for the driven Brownian particle as the potential depth A is varied. (c) Ratio between the estimated value σ_{pred} and the true value σ of the average EP for the driven Brownian particle. Strength of the nonequilibrium driving is fixed at $f = 32$ and the temperature at $T = 1$. Each data point and error bar are obtained from 40 independent trainings.

where $\xi(t)$ is a Gaussian white noise with unit variance. See Fig. 4(a) for an illustration of the model. For sufficiently large A , this model can approximate the behaviors of the Markov jump process on a discrete chain. For this model, the EP associated with the infinitesimal displacement dx is given by

$$\Delta S = \frac{-f dx + U(x + dx) - U(x)}{T} + \Delta S_{\text{sys}}, \quad (18)$$

where $\Delta S_{\text{sys}} = -\ln p_s(x + dx)/p_s(x)$ again denotes the Shannon entropy change for the steady-state distribution $p_s(x)$. Since the system is 1D, it is straightforward to obtain $p_s(x)$ by numerical integration. Thus, the EP of this model can also be calculated exactly and compared to the α -NEEP result.

Fixing $f = 32$, the performance of the α -NEEP for this model is shown in Figs. 4(b) and 4(c) in terms of the MSE and

the ratio $\sigma_{\text{pred}}/\sigma$, respectively. Due to the presence of a strong background driving ($f = 32$), there are already considerable differences among different methods at $A = 0$. But it is worth noting that increasing the amplitude A of the periodic potential $U(x)$ clearly increases the MSE and makes $\sigma_{\text{pred}}/\sigma$ deviate farther away from 1 for the original NEEP ($\alpha = 0$). This may be the consequence of rarer movements (jumps from one potential well to the next) across the system as the potential well gets deeper, which means rare trajectories are even more poorly sampled. The α -NEEPs with nonzero values of α are much more robust against the increase of A , with $\alpha = -0.5$ showing the best performance overall.

V. SIMPLE GAUSSIAN MODEL

The results shown thus far clearly indicate that, by choosing a nonzero value of α , the α -NEEP can exhibit a much more robust performance against the adverse effects of the nonequilibrium driving. Moreover, $\alpha = -0.5$ seems to exhibit the best performance in many cases. To gain more intuition into these results, we simplify the EP estimation problem to the density-ratio estimation problem for a 1D random variable. To be specific, we estimate the log ratio $s(x) \equiv \log p(x)/p(-x)$ given samples drawn from the distribution $p(x)$. It is intuitively clear that this problem is structurally equivalent to EP estimation.

For further simplification, we set

$$p(x) = \begin{cases} \mathcal{N} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] & \text{for } |x - \mu| \leq k\sigma, \\ 0 & \text{otherwise.} \end{cases} \quad (19)$$

Here \mathcal{N} is a suitable normalization factor, μ the positive mean of the distribution, σ the width of the distribution, and k a positive number truncating the tails of the distribution. While $k = \infty$ corresponds to the perfect sampling of a Gaussian distribution, a finite k corresponds to the case where the tails of the distribution are poorly sampled.

For $k = \infty$, the correct answer to the problem is a linear function $s(x) = \theta_0 x$, where $\theta_0 \equiv 2\mu/\sigma^2$. Thus, for further simplicity, we focus on the *one-parameter model* $s_\theta(x) = \theta x$, which estimates $s(x)$ using only a single parameter θ . For this problem, the suitable loss function is obtained as an analog of Eq. (9):

$$\mathcal{L}_\alpha(\theta) = \left\langle -\frac{(1 + \alpha)e^{\alpha\theta x} - 1}{\alpha} + e^{-(1+\alpha)\theta x} \right\rangle_{p(x)}. \quad (20)$$

If k is large but finite, the minimum of this loss function shifts to $\theta_0 + \Delta\theta$, where $\Delta\theta$ can be expanded to the leading orders in $1/k$:

$$\Delta\theta \sim \exp\left[-\frac{k^2}{2} + \frac{2\mu}{\sigma}\left(\left|\alpha + \frac{1}{2}\right| + \frac{1}{2}\right)k\right]. \quad (21)$$

This clearly shows that $\alpha = -0.5$ gives the least shift $\Delta\theta$, as also illustrated by various results shown in Fig. 5.

In Fig. 5(a) we show that the shift of the minimum $\Delta\theta$ tends to increase as the tail sampling becomes poorer (i.e., k decreases). The landscapes of the loss function $\mathcal{L}_\alpha(\theta)$, shown in the inset of Fig. 5(a), also confirm this observation. The increase of the error with the potential depth A in Figs. 1(d) and 2(b) may primarily be due to the same effect.

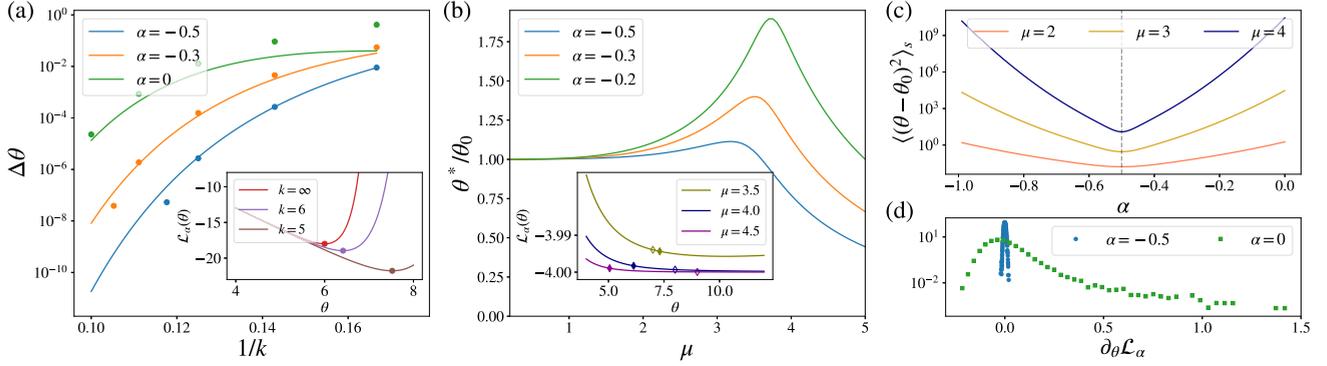


FIG. 5. Performance of the exactly solvable one-parameter model. (a) Shift $\Delta\theta$ of the loss function minimum as a function of the truncation parameter k . Circles are results obtained by numerical minimization, and solid lines are from the small $1/k$ expansion. (Inset) Loss function landscapes, with circles indicating the minima. We fixed $\mu = 3$, $\sigma = 1$, and $\alpha = 0$. (b) Ratio of the estimated minimum θ^* to the true minimum θ_0 as the bias μ is varied. The optimal points are calculated using the criterion that the loss function gradient satisfies $|\partial_\theta \mathcal{L}_\alpha(\theta)| < 10^{-3}$ for the first time as θ increases from 0. We fixed $k = 4$ and $\sigma = 1$. (Inset) Loss function landscape. Open diamonds indicate the true minima θ_0 , and filled diamonds represent the estimated minima θ^* . The parameters $\alpha = -0.5$ and $k = 4$ are fixed. (c) MSE of θ . The vertical dashed line shows that the error is minimized at $\alpha = -0.5$. (d) Distribution of the loss function gradient $\partial_\theta \mathcal{L}_\alpha$ at the minimum $\theta_0 = 2$ for $\mu = \sigma = 1$.

In Fig. 5(b) we plot the ratio between the estimated minimum θ^* and the true minimum θ_0 as a function of the mean μ , which is an analog of the nonequilibrium driving. We note that here θ^* is the lowest value of θ at which the slope of the loss function becomes less than 10^{-3} . We observe that an overestimation regime ($\theta^*/\theta_0 > 1$) crosses over to an underestimation regime ($\theta^*/\theta_0 < 1$) as μ grows. This is in striking agreement with the trends shown in Fig. 2(a). The reason why θ^* underestimates θ_0 for large μ can be understood by the flattened loss function landscapes shown in the inset of Fig. 5(b). In this regime the dynamics of θ (starting from $\theta = 0$) slows down, ending up at a value (filled diamonds) even lower than θ_0 (empty diamonds). This effect is due to the samples with $x < 0$ vanishing when μ is too large. We expect that a similar mechanism might be at play behind the observed behavior of $\sigma_{\text{pred}}/\sigma$ shown in Fig. 2(a). If we had used a broader range of nonequilibrium driving, the same behaviors might have been observed for other models as well, although this remains to be checked.

The one-parameter model also allows us to examine the effects of the finite minibatch size M . While the ideal loss function is given in Eq. (20), the loss function used in the actual training looks like

$$\mathcal{L}_\alpha(\theta; M) = \frac{1}{M} \sum_i \left[-\frac{(1 + \alpha)e^{\alpha\theta X_i} - 1}{\alpha} + e^{-(1 + \alpha)\theta X_i} \right], \quad (22)$$

where X_1, \dots, X_M are i.i.d. Gaussian random variables of mean μ and variance σ^2 . When M is large and finite, using the central limit theorem (CLT), the gradient of this loss function can be approximated as [26,27]

$$\left. \frac{\partial \mathcal{L}_\alpha}{\partial \theta} \right|_{\theta=\theta_t} = \bar{K}(\theta_t - \theta_0) + \sqrt{\frac{\Lambda}{M}} N_t + o(M^{-1/2}), \quad (23)$$

where $\theta_0 = \text{argmin}(\langle \mathcal{L}_\alpha(\theta) \rangle)$, $\bar{K} = \partial_\theta^2 \langle \mathcal{L}_\alpha \rangle|_{\theta=\theta_0}$, and $\Lambda = \text{Var}[\partial_\theta \mathcal{L}_\alpha|_{\theta=\theta_0, M=1}]$. When the stochastic gradient descent $\theta_{t+1} = \theta_t - \lambda(\partial \mathcal{L}_\alpha / \partial \theta)|_{\theta=\theta_t}$ reaches the steady state, the MSE

of θ is given by

$$\langle(\theta - \theta_0)^2\rangle_s = \frac{\lambda\Lambda}{M\bar{K}(2 - \lambda\bar{K})} + o(M^{-1}). \quad (24)$$

This leading-order behavior is shown in Fig. 5(c) for various values of μ . For all cases, the MSE of θ is minimized at $\alpha = -0.5$, which is consistent with the smallest error bars observed at $\alpha = -0.5$ in Figs. 1 and 2. Hence, $\alpha = -0.5$ yields the most consistent EP estimator.

Direct measurements of the loss function gradient at the minimum also confirm the above result. As shown in Fig. 5(d), the gradient $\partial_\theta \mathcal{L}_\alpha$ is far more broadly distributed for $\alpha = 0$ than for $\alpha = -0.5$. Moreover, due to the subleading effects (beyond the CLT) of finite M , the gradient for $\alpha = 0$ features a large skewness. These show that the training dynamics for the original NEEP ($\alpha = 0$) tends to be far more volatile and unstable than for the α -NEEP with $\alpha = -0.5$.

VI. SUMMARY AND OUTLOOK

We proposed the α -NEEP, a generalization of the NEEP for estimating steady-state EP at the trajectory level. By choosing a value of α between -1 and 0 , the α -NEEP weakens the exponential dependence of the loss function on the EP estimator, effectively mitigating the adverse effects induced by poor sampling of transitions associated with large negative EP in the presence of strong nonequilibrium driving and/or deep potential wells. We also observed that $\alpha = -0.5$ tends to exhibit the optimal performance, which can be understood via a simplification of the original EP estimation problem, whose loss function landscape and relaxation properties are analytically tractable. The α -NEEP thus provides a powerful method for estimating the EP for much broader range of the nonequilibrium driving force and the timescale of dynamics. Identification of even better loss functions and optimization of other hyperparameters (network size, number of iterations, etc.) are left as future work. It would also be interesting to apply the α -NEEP to estimations of the EP of the Brownian movies [28] and stochastic systems with odd-parity

variables [29], which have been studied using the original NEP method.

ACKNOWLEDGMENTS

This work was supported by the POSCO Science Fellowship of the POSCO TJ Park Foundation. E.K. and Y.B. also thank Junghyo Jo and Sangyun Lee for helpful comments.

APPENDIX A: TRAINING DETAILS

We always use the fully connected network (FCN) with three hidden layers, with each layer composed of 512 nodes. Each training dataset consists of 10^6 trajectories. The neural network parameters are updated using the ReLU activation function and the Adam optimizer. The learning rate is fixed to 10^{-5} , and the weight decay is fixed to 5×10^{-5} . We halt the training after 10000 iterations, except for the results shown in Figs. 8 and 9 (see Appendix C), where we continue the training for a longer time to check the overfitting effects. All trainings are done on PyTorch with NVIDIA GeForce RTX 3090.

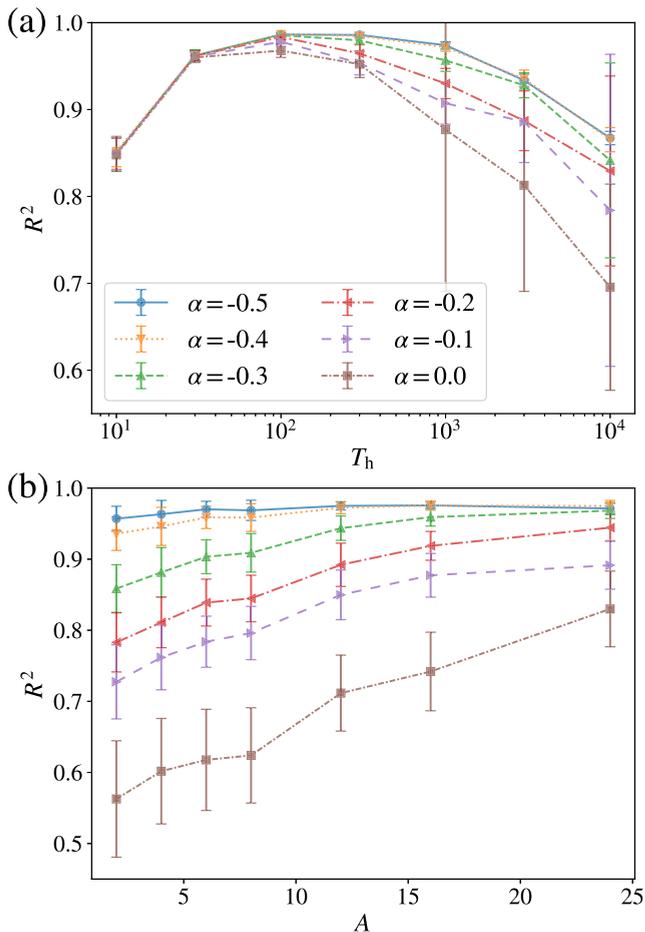


FIG. 6. Coefficient of determination R^2 for (a) the two-bead model with $T_c = 1$ and for (b) the driven Brownian particle with $f = 32$. 10^4 trajectories are used for each minibatch, and error bars indicate the standard deviations obtained from 40 independent trainings.

In subfigure (b) of Figs. 2–4, each minibatch consists of 10^4 trajectories. On the other hand, in subfigure (c) of Figs. 2–4, each minibatch consists of 10^5 trajectories.

APPENDIX B: DENSITY RATIO ESTIMATION VIA f -DIVERGENCE

Here we show that the the loss function $\mathcal{L}_\alpha[r]$ given in Eq. (7), whose minimization allows us to estimate the ratio between two probability density functions, can be generalized even further using the concept of f -divergence. Consider a convex, twice-differentiable real-valued function $f(u)$. Then the inequality

$$-pf'(u) + q[uf'(u) - f(u)] \geq -qf(p/q) \quad (\text{B1})$$

holds. We can verify this by differentiating the left-hand side (LHS) with respect to u , which yields $f''(u)(-p + qu)$. Thus, the LHS has a local minimum at $u = p/q$, and this is the only local minimum since f is convex. In addition, the second derivative of the LHS at $u = p/q$ equals $qf''(p/q)$, which is positive by the convexity. This proves the inequality (B1).

Using this result, we can design a loss function whose minimum is equal to the negative f -divergence between two probability distributions $p(\mathbf{x})$ and $q(\mathbf{x})$. To be specific, for any

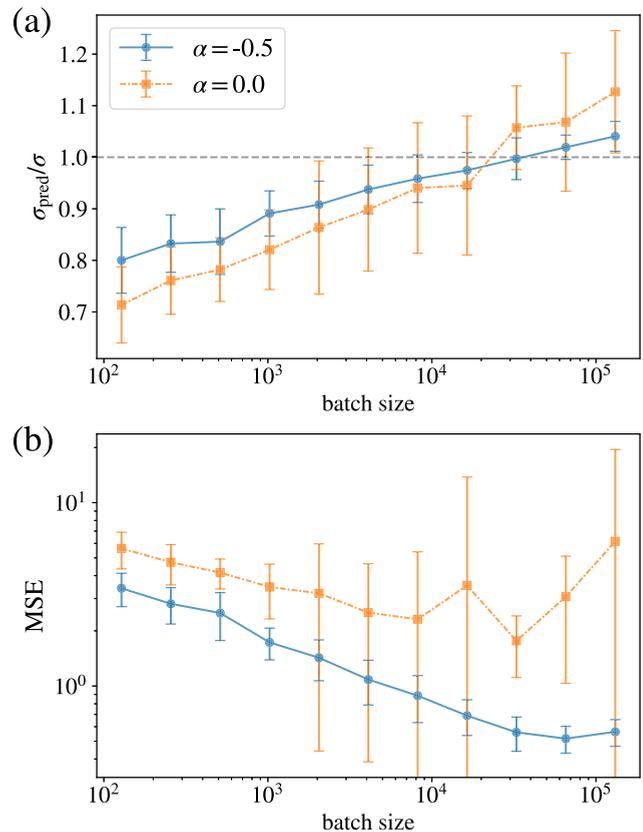


FIG. 7. Effects of the minibatch size on the performance of the α -NEP for the two-bead model with $T_h = 1000$ and $T_c = 1$. Error bars indicate the standard deviations obtained from 40 independent trainings.

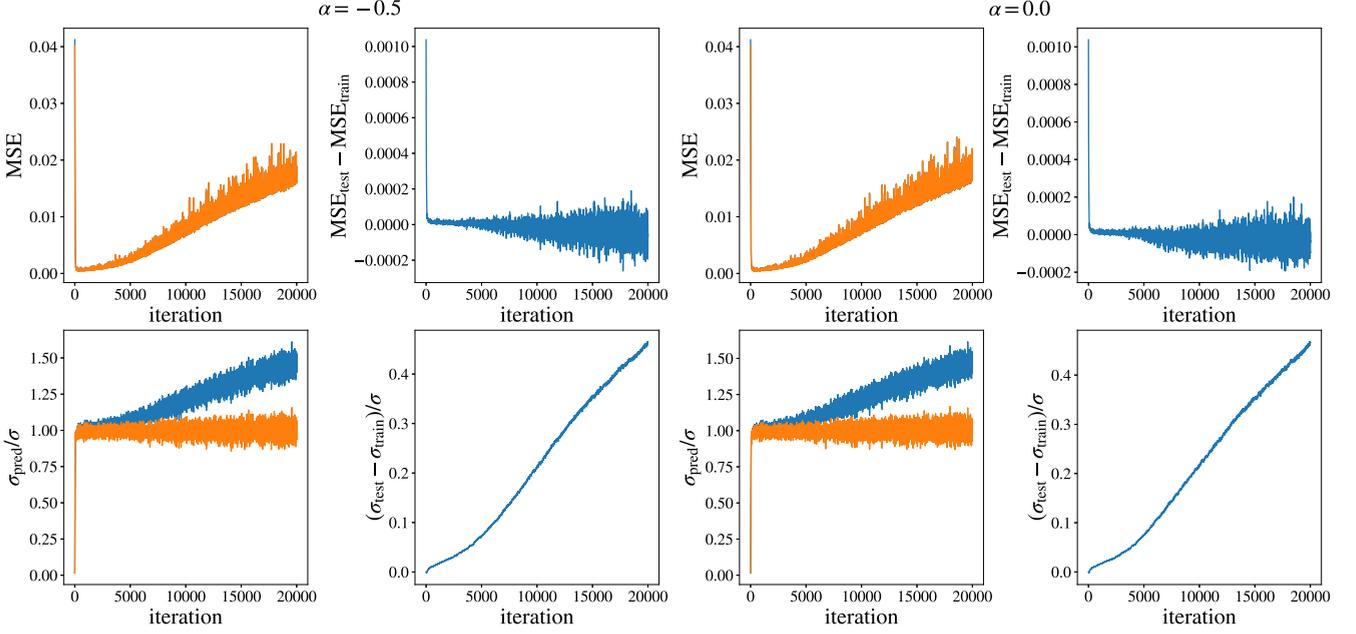


FIG. 8. Training dynamics of the α -NEEP for the two-bead model at $T_h = 10$. Each minibatch consists of 10^5 trajectories. The first and the third columns show the performance for the training set [blue (dark gray curve)] and the test set [orange (light gray curve)]. The second and the fourth columns show the difference in performance between the two datasets. The first (last) two columns correspond to $\alpha = -0.5$ ($\alpha = 0$).

function $r(\mathbf{x})$, we define

$$\begin{aligned} \mathcal{L}_f[r] &= \left\langle -f'(r(\mathbf{x})) + \frac{q(\mathbf{x})}{p(\mathbf{x})} [r(\mathbf{x})f'(r(\mathbf{x})) - f(r(\mathbf{x}))] \right\rangle_{p(\mathbf{x})} \\ &= \int d^{2d} \mathbf{x} \{ -p(\mathbf{x})f'(r(\mathbf{x})) \\ &\quad + q(\mathbf{x})[r(\mathbf{x})f'(r(\mathbf{x})) - f(r(\mathbf{x}))] \}. \end{aligned} \quad (\text{B2})$$

Using Eq. (B2), we conclude that

$$\mathcal{L}_f[r] \geq - \int d^{2d} \mathbf{x} q(\mathbf{x}) f\left(\frac{p(\mathbf{x})}{q(\mathbf{x})}\right) = -D_f[p:q], \quad (\text{B3})$$

where $D_f[p:q]$ is the f -divergence between the distributions $p(\mathbf{x})$ and $q(\mathbf{x})$, and the equality holds if and only if $r(\mathbf{x}) = p(\mathbf{x})/q(\mathbf{x})$ for all \mathbf{x} . By minimizing $\mathcal{L}_f[r]$, we can estimate $p(\mathbf{x})/q(\mathbf{x})$ as well as $D_f[p:q]$.

The loss function \mathcal{L}_α and the associated α -divergence discussed in the main text are obtained by choosing the function f to be

$$f_\alpha(u) = \begin{cases} \frac{u^{1+\alpha} - (1+\alpha)u + \alpha}{\alpha(1+\alpha)} & \text{for } \alpha \neq 0, -1, \\ u \log u & \text{for } \alpha = 0, \\ \log u + 1 - u & \text{for } \alpha = -1. \end{cases} \quad (\text{B4})$$

Note that $f_0(u) = \lim_{\alpha \rightarrow 0} f_\alpha(u)$ and $f_{-1}(u) = \lim_{\alpha \rightarrow -1} f_\alpha(u)$. It is straightforward to obtain Eq. (9) and its extensions to the cases $\alpha = 0$ and $\alpha = -1$ from this choice.

APPENDIX C: EXTRA NUMERICAL RESULTS

1. Coefficient of determination

In the literature the extent of agreement between a prediction and the true value is often expressed by the coefficient

of determination R^2 . Here we check how the behaviors of R^2 differ as the value of α changes for the cases of the two-bead model and the driven 1D Brownian particle.

For the two-bead model, as shown in Fig. 6(a), R^2 exhibits a nonmonotonic behavior as a function of T_h . The decrease of R^2 with increasing T_h reflects the detriment of the α -NEEP performance as the nonequilibrium driving gets stronger. Meanwhile, the decrease of R^2 as T_h decreases (getting closer to equilibrium $T_h = T_c = 1$) is due to the overfitting phenomenon discussed in the next section, which disrupts the linear relationship between the predicted EP and the true EP.

For the driven Brownian particle, as shown in Fig. 6(b), R^2 always increases with A . This may seem contradictory to how the MSE tends to increase or stay constant with increasing A in Fig. 4(b). Indeed, higher R^2 only means that there is a good linear relationship between the EP estimate s_θ and the true EP ΔS , not that s_θ and ΔS are close to each other. When A is increased, due to the slower dynamics, we may have $s_\theta > \Delta S$ for transitions with positive EP and $s_\theta < \Delta S$ for transitions with negative EP, which can make the linear relationship between s_θ and ΔS appear stronger. This example clearly shows that R^2 is not an adequate measure of the performance of EP estimators.

2. Effects of the minibatch size

The minibatch refers to the group of samples used for computing the gradient of the loss function. Smaller (larger) minibatches increase (decrease) the noisy component of the gradient, which in turn affects the performance of the α -NEEP.

We explicitly check the effects of the minibatch size using the two-bead model with $T_h = 1000$ and $T_c = 1$, as shown in Fig. 7. We use the ratio $\sigma_{\text{pred}}/\sigma$ and the MSE as two

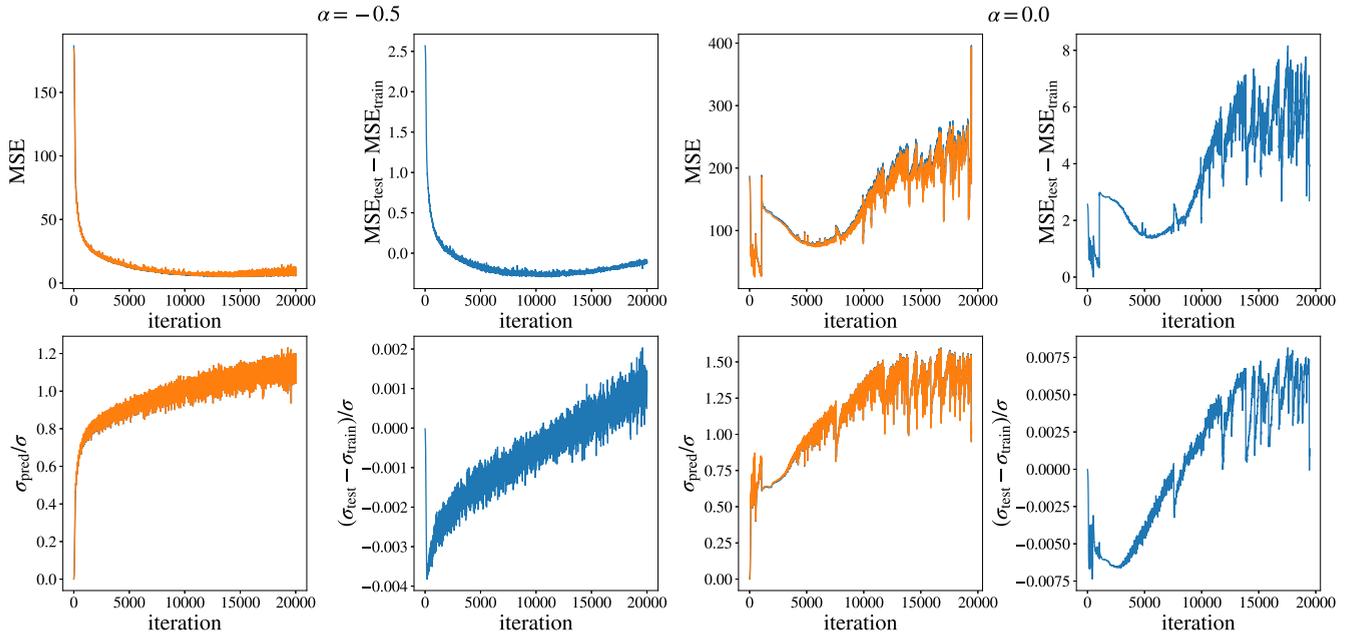


FIG. 9. Training dynamics of the α -NEEP for the two-bead model at $T_h = 3000$ and $T_c = 1$. Each minibatch consists of 10^5 trajectories. The first and the third columns show the performance for the training set [blue (dark gray curve)] and the test set [orange (light gray curve)]. The second and the fourth columns show the difference in performance between the two datasets. The first (last) two columns correspond to $\alpha = -0.5$ ($\alpha = 0$).

different measures of the α -NEEP performance. For small minibatches, the highly skewed distribution of the stochastic gradient shown in Fig. 5(d) causes underestimation of the EP. For large minibatches, the noisy component of the loss-function gradient decreases, revealing the properties of the loss function landscape of the training dataset. As discussed using the Gaussian model in Sec. V, the loss function landscape at a moderately strong nonequilibrium driving leads to the overestimation of the EP. Thus, as the minibatch size is increased, $\sigma_{\text{pred}}/\sigma$ grows beyond 1.

The nonmonotonic behaviors of the MSE also hint at the existence of an optimal minibatch size at the tradeoff between the skewed noise in the gradient (which drives the neural network towards underestimation) and the loss function landscape tilted towards overestimation. For both measures, the superiority of $\alpha = -0.5$ to $\alpha = 0$ is manifest.

3. Effects of overfitting

In many cases, when the training continues for too many iterations, artificial neural networks are known to exhibit overfitting behaviors. As shown in Figs. 8 and 9, we checked whether the α -NEEP is also subject to the same phenomena as the training continues up to 20 000 iterations. Towards this end, we created two independent datasets of trajectories exhibited by the two-bead model, namely, the training set and the test set. Only the former was used during the training of the α -NEEP, and we measured the MSE and the ratio $\sigma_{\text{pred}}/\sigma$ to assess the performance of the α -NEEP for each dataset.

In Fig. 8 we show the results for the weak nonequilibrium driving ($T_h = 10$ and $T_c = 1$). The first and the third columns show the two different measures of performance for the training dataset and the test dataset. Meanwhile, the second and the fourth columns show the difference between the corresponding measures obtained for two datasets. The overfitting phenomena are manifest from the increase of the MSE towards the end of the training. Interestingly, overfitting leads to an overestimation of the average EP only for the training dataset. We also note that the value of α is largely irrelevant to the extent of overfitting. This phenomenon can be explained as follows. Near equilibrium, the neural network swiftly reaches the loss function minimum. However, as the training continues, the neural network starts to see the detailed fluctuations of the training dataset. This makes the functional form of the estimator s_θ very rough, leading to the increase of the MSE for both datasets. But while the neural network now believes all trajectories in the training dataset to be highly irreversible and assigns high EP to them, the EP assigned to the trajectories in the test dataset stay unbiased. Thus, $\sigma_{\text{pred}}/\sigma$ grows larger only for the training dataset.

In Fig. 9 we show the results for the strong nonequilibrium driving ($T_h = 3000$ and $T_c = 1$). The subfigures are organized in exactly the same way as in Fig. 8. In this case the overfitting effects do exist. But they are not as pronounced as in the case of the weaker nonequilibrium driving, and the differences between the training and the test datasets stay small. Note that the curves for $\alpha = 0$ exhibit strong fluctuations, which is in agreement with the large fluctuations of the gradient shown in Fig. 5(d).

- [1] C. P. Brangwynne, G. H. Koenderink, F. C. MacKintosh, and D. A. Weitz, *J. Cell Biol.* **183**, 583 (2008).
- [2] S. C. Weber, A. J. Spakowitz, and J. A. Theriot, *Proc. Natl. Acad. Sci. USA* **109**, 7338 (2012).
- [3] D. Mizuno, C. Tardin, C. F. Schmidt, and F. C. MacKintosh, *Science* **315**, 370 (2007).
- [4] C. Battle, C. P. Broedersz, N. Fakhri, V. F. Geyer, J. Howard, C. F. Schmidt, and F. C. MacKintosh, *Science* **352**, 604 (2016).
- [5] J. Gladrow, N. Fakhri, F. C. MacKintosh, C. F. Schmidt, and C. P. Broedersz, *Phys. Rev. Lett.* **116**, 248301 (2016).
- [6] J. Tailleur and M. E. Cates, *Phys. Rev. Lett.* **100**, 218103 (2008).
- [7] T. Speck, J. Bialké, A. M. Menzel, and H. Löwen, *Phys. Rev. Lett.* **112**, 218304 (2014).
- [8] S. C. Takatori, W. Yan, and J. F. Brady, *Phys. Rev. Lett.* **113**, 028103 (2014).
- [9] T. F. F. Farage, P. Krinninger, and J. M. Brader, *Phys. Rev. E* **91**, 042310 (2015).
- [10] A. P. Solon, J. Stenhammar, M. E. Cates, Y. Kafri, and J. Tailleur, *Phys. Rev. E* **97**, 020602(R) (2018).
- [11] E. Fodor, C. Nardini, M. E. Cates, J. Tailleur, P. Visco, and F. van Wijland, *Phys. Rev. Lett.* **117**, 038103 (2016).
- [12] C. Nardini, E. Fodor, E. Tjhung, F. van Wijland, J. Tailleur, and M. E. Cates, *Phys. Rev. X* **7**, 021007 (2017).
- [13] U. Seifert, *Rep. Prog. Phys.* **75**, 126001 (2012).
- [14] E. Roldán and J. M. R. Parrondo, *Phys. Rev. Lett.* **105**, 150607 (2010).
- [15] E. Roldán and J. M. R. Parrondo, *Phys. Rev. E* **85**, 031129 (2012).
- [16] J. Li, J. M. Horowitz, T. R. Gingrich, and N. Fakhri, *Nat. Commun.* **10**, 1666 (2019).
- [17] T. Van Vu, V. T. Vo, and Y. Hasegawa, *Phys. Rev. E* **101**, 042138 (2020).
- [18] S. Otsubo, S. Ito, A. Dechant, and T. Sagawa, *Phys. Rev. E* **101**, 062106 (2020).
- [19] S. Otsubo, S. K. Manikandan, T. Sagawa, and S. Krishnamurthy, *Commun. Phys.* **5**, 11 (2022).
- [20] S. Lee, D.-K. Kim, J.-M. Park, W. K. Kim, H. Park, and J. S. Lee, *Phys. Rev. Res.* **5**, 013194 (2023).
- [21] D.-K. Kim, Y. Bae, S. Lee, and H. Jeong, *Phys. Rev. Lett.* **125**, 140604 (2020).
- [22] A. Basu, I. R. Harris, N. L. Hjort, and M. C. Jones, *Biometrika* **85**, 549 (1998).
- [23] M. Sugiyama, T. Suzuki, and T. Kanamori, *Density Ratio Estimation in Machine Learning* (Cambridge University Press, Cambridge, 2012).
- [24] S. Nowozin, B. Cseke, and R. Tomioka, *Advances in Neural Information Processing Systems*, edited by D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett (Curran Associates, Inc., 2016), Vol. 29, pp. 271–279.
- [25] M. I. Belghazi, A. Baratin, S. Rajeshwar, S. Ozair, Y. Bengio, A. Courville, and D. Hjelm, in *International Conference on Machine Learning* (PMLR, 2018), pp. 531–540.
- [26] Q. Li, C. Tai, and W. E, in *Proceedings of the 34th International Conference on Machine Learning*, Proceedings of Machine Learning Research, Vol. 70, edited by D. Precup and Y. W. Teh (PMLR, 2017), pp. 2101–2110.
- [27] P. Chaudhari and S. Soatto, in *2018 Information Theory and Applications Workshop (ITA)* (IEEE, 2018), pp. 1–10.
- [28] Y. Bae, D.-K. Kim, and H. Jeong, *Phys. Rev. Res.* **4**, 033094 (2022).
- [29] D.-K. Kim, S. Lee, and H. Jeong, *Phys. Rev. Res.* **4**, 023051 (2022).