Uncertainty relations in stochastic processes: An information inequality approach

Yoshihiko Hasegawa* and Tan Van Vu†
Department of Information and Communication Engineering, Graduate School of Information Science and Technology, The University of Tokyo, Tokyo 113-8656, Japan

(Received 1 October 2018; revised manuscript received 20 March 2019; published 24 June 2019)

The thermodynamic uncertainty relation is an inequality stating that it is impossible to attain higher precision than the bound defined by entropy production. In statistical inference theory, information inequalities assert that it is infeasible for any estimator to achieve an error smaller than the prescribed bound. Inspired by the similarity between the thermodynamic uncertainty relation and the information inequalities, we apply the latter to systems described by Langevin equations, and we derive the bound for the fluctuation of thermodynamic quantities. When applying the Cramér-Rao inequality, the obtained inequality reduces to the fluctuation-response inequality. We find that the thermodynamic uncertainty relation is a particular case of the Cramér-Rao inequality, in which the Fisher information is the total entropy production. Using the equality condition of the Cramér-Rao inequality, we find that the stochastic total entropy production is the only quantity that can attain equality in the thermodynamic uncertainty relation. Furthermore, we apply the Chapman-Robbins inequality and obtain a relation for the lower bound of the ratio between the variance and the sensitivity of systems in response to arbitrary perturbations.

DOI: 10.1103/PhysRevE.99.062126

I. INTRODUCTION

Over the past two decades, substantial progress has been made in terms of universal relations among thermodynamic quantities, such as fluctuation theorems and generalized second laws [1–4]. One of the key achievements in this area is the thermodynamic uncertainty relation [5–14], which states that fluctuations in thermodynamic quantities are bounded from below by the reciprocal of entropy production. The thermodynamic uncertainty relation provides a theoretical justification for our intuition that higher precision is inevitably accompanied by larger energy consumption. Universal relations between "cost" and "quality" also exist in fields other than thermodynamics. It is an empirical truism that as the amount of available data increases, inferences on parameters become more precise. Information inequalities provide theoretical support for this intuition by giving the lower bounds for estimators. Information inequalities are known to be the basis for inequalities in other fields; for instance, Heisenberg's uncertainty principle can be derived through these inequalities [15,16]. The universality of information inequalities leads us to posit that they may play an important role in stochastic thermodynamic systems.

Herein, we regard fluctuations of thermodynamic quantities as errors in statistical estimators, thereby obtaining inequality relations for quantities of stochastic processes. In particular, we obtain the Cramér-Rao inequality for systems described by Langevin equations, which relates the fluctuation of thermodynamic quantities to the Fisher information. This relation reduces to a recently discovered fluctuation-response inequality [17]. We show that the thermodynamic uncertainty relation is a particular case of the Cramér-Rao inequality in which the Fisher information is the total entropy production. Using the equality condition of the Cramér-Rao inequality, we find that the stochastic total entropy production is the only quantity that can attain equality in the thermodynamic uncertainty relation. Furthermore, we apply the Chapman-Robbins inequality, which is a generalization of the Cramér-Rao inequality, to the systems to show that the ratio between the variance and the sensitivity is bounded from below by the reciprocal of the Pearson divergence for any perturbation. As an application of the Chapman-Robbins inequality, we obtain an explicit inequality between the phase variance and the phase sensitivity of stochastic limit cycle oscillators.

II. MODEL

We consider the following N-dimensional Ito Langevin equation for $\boldsymbol{x} \equiv [x_1, x_2, \dots, x_N]^\top$:

$$\dot{\boldsymbol{x}} = \boldsymbol{A}_\theta(\boldsymbol{x}, t) + \sqrt{2}\boldsymbol{C}(\boldsymbol{x}, t)\boldsymbol{\xi}(t), \tag{1}$$

where $\boldsymbol{\xi}(t) \equiv [\xi_1(t), \dots, \xi_M(t)]^\top$ is white Gaussian noise with $\langle \xi_i(t)\xi_j(t')\rangle = \delta_{ij}\delta(t-t')$ (M is the number of noise terms), $\boldsymbol{A}_\theta(\boldsymbol{x}, t) \equiv [A_{\theta,1}(\boldsymbol{x}, t), \dots, A_{\theta,N}(\boldsymbol{x}, t)]^\top$ is a drift vector with a real parameter θ, and $\boldsymbol{C}(\boldsymbol{x}, t) \equiv [C_{ij}(\boldsymbol{x}, t)]$ is an $N \times M$ noise matrix. θ is a parameter to be estimated with predefined estimators. For simplicity, we assume that θ is a scalar, but the calculation can be easily generalized to a multidimensional $\boldsymbol{\theta}$. Let $P_\theta(\boldsymbol{x}, t)$ be the probability density function of $\boldsymbol{x}$ at time t. Defining $[B_{ij}(\boldsymbol{x}, t)] = \boldsymbol{B}(\boldsymbol{x}, t) \equiv \boldsymbol{C}(\boldsymbol{x}, t)\boldsymbol{C}(\boldsymbol{x}, t)^\top$, the Fokker-Planck equation (FPE) of Eq. (1) is [18,19]

$$\partial_t P_\theta(\boldsymbol{x}, t) = \widehat{\mathcal{L}}_\theta(\boldsymbol{x}, t)P_\theta(\boldsymbol{x}, t), \tag{2}$$

*hasegawa@biom.t.u-tokyo.ac.jp
†tan@biom.t.u-tokyo.ac.jp

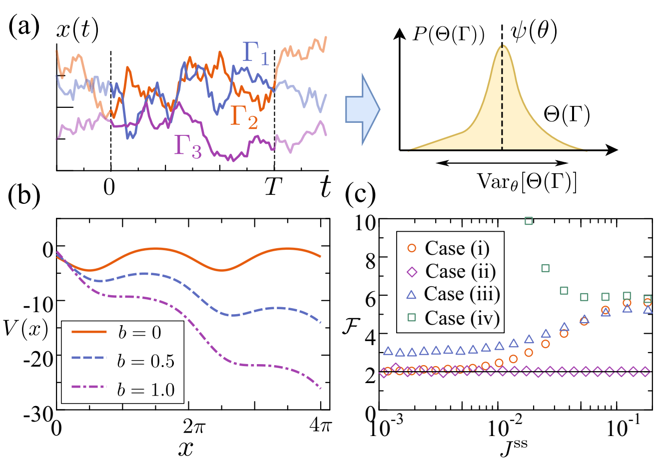

FIG. 1. (a) Relation between trajectory Γ and its estimator $\Theta(\Gamma)$. Given a trajectory $\Gamma = [x(t)]_{t=0}^{t=T}$, $\Theta(\Gamma)$ is an unbiased estimator for $\psi(\theta)$. Repeating the measurement and estimation, we obtain the probability density of $\Theta(\Gamma)$. The variance of the probability density corresponds to $\mathrm{var}_\theta[\Theta(\Gamma)]$, where $\langle\Theta(\Gamma)\rangle_\theta = \psi(\theta)$. (b) Effective potential function considered in the numerical verification. The effective potential $V(x)$ is plotted for $b = 0$ (solid line), 0.5 (dashed line), and 1.0 (dot-dashed line) with $a = 2$. (c) Numerical verification of the equality condition of the thermodynamic uncertainty relation. $\mathcal{F}$ are plotted as a function of J^{ss}, where $\mathcal{F} = 2$ for the equality of the thermodynamic uncertainty relation. Cases (i) (circles) and (ii) (diamonds) satisfy the equality condition, while cases (iii) (triangles) and (iv) (squares) do not. The parameter settings are $D = 1.0$, $a = 2$, $b \in [0.001, 1.0]$, and $T = 4.0$.

where $\widehat{\mathcal{L}}_\theta(\boldsymbol{x}, t) \equiv -\sum_i \partial_{x_i} A_{\theta,i}(\boldsymbol{x}, t) + \sum_{i,j} \partial_{x_i}\partial_{x_j} B_{ij}(\boldsymbol{x}, t)$ is an FPE operator. The probability current is

$$J_{\theta,i}(\boldsymbol{x}, t) \equiv \left\{A_{\theta,i}(\boldsymbol{x}, t) - \sum_j \partial_{x_j} B_{ij}(\boldsymbol{x}, t)\right\} P_\theta(\boldsymbol{x}, t). \tag{3}$$

Now we consider the estimation of the parameter θ from the measurement of a stochastic trajectory generated by Eq. (1) over an interval from $t = 0$ to T. Let $\boldsymbol{\Gamma} \equiv [\boldsymbol{x}(t)]_{t=0}^{t=T}$ be the measured trajectory and $\mathcal{P}_\theta(\boldsymbol{\Gamma})$ be the probability of $\boldsymbol{\Gamma}$ [Fig. 1(a)]. For an arbitrary function $f(\boldsymbol{\Gamma})$, we define its expectation as $\langle f(\boldsymbol{\Gamma})\rangle_\theta \equiv \int \mathcal{D}\boldsymbol{\Gamma}\, f(\boldsymbol{\Gamma})\mathcal{P}_\theta(\boldsymbol{\Gamma})$. We consider an estimator $\Theta(\boldsymbol{\Gamma})$, which is an unbiased estimator for $\psi(\theta)$, and thus we have $\langle\Theta(\boldsymbol{\Gamma})\rangle_\theta = \psi(\theta)$. Since $\boldsymbol{\Gamma}$ is a stochastic trajectory, $\Theta(\boldsymbol{\Gamma})$ is a random variable. Therefore, if we repeat the measurement and the estimation, $\Theta(\boldsymbol{\Gamma})$ is distributed around $\psi(\theta)$, whose variance is identified as the variance of the estimator $\Theta(\boldsymbol{\Gamma})$ [Fig. 1(a)].

III. CRAMÉR-RAO INEQUALITY

A. Derivation of the uncertainty relation

The Cramér-Rao inequality provides the lower bound for the variance of estimators. Applying the Cramér-Rao inequality [20–22] to $\Theta(\boldsymbol{\Gamma})$, the following relation holds (Appendix A):

$$\frac{\mathrm{var}_\theta[\Theta(\boldsymbol{\Gamma})]}{[\partial_\theta\langle\Theta(\boldsymbol{\Gamma})\rangle_\theta]^2} \geqslant \frac{1}{\mathcal{I}(\theta)}, \tag{4}$$

where $\mathrm{var}_\theta[f(\boldsymbol{\Gamma})] \equiv \langle\{f(\boldsymbol{\Gamma}) - \langle f(\boldsymbol{\Gamma})\rangle_\theta\}^2\rangle_\theta$ and $\mathcal{I}(\theta)$ is the Fisher information:

$$\mathcal{I}(\theta) \equiv \left\langle\left(\frac{\partial}{\partial\theta}\ln\mathcal{P}_\theta(\boldsymbol{\Gamma})\right)^2\right\rangle_\theta = -\left\langle\frac{\partial^2}{\partial\theta^2}\ln\mathcal{P}_\theta(\boldsymbol{\Gamma})\right\rangle_\theta. \tag{5}$$

Let $\mathcal{P}_\theta(\boldsymbol{\Gamma}|\boldsymbol{x}^0)$ be the probability of $\boldsymbol{\Gamma}$ given $\boldsymbol{x}^0$ at $t = 0$. By using a path integral [14,23–25], we obtain (Appendix B)

$$\mathcal{P}_\theta(\boldsymbol{\Gamma}|\boldsymbol{x}^0) = \mathscr{N}\exp\left[-\int_0^T dt\, \mathscr{A}_\theta(\boldsymbol{x}(t), t)\right], \tag{6}$$

$$\mathscr{A}_\theta(\boldsymbol{x}, t) \equiv \frac{1}{4}\{(\dot{\boldsymbol{x}} - \boldsymbol{A}_\theta)^\top \boldsymbol{B}^{-1}(\dot{\boldsymbol{x}} - \boldsymbol{A}_\theta)\}, \tag{7}$$

where $\mathscr{N}$ is a term that does not depend on θ. In Eq. (6), we employ a prepoint discretization. Due to the prepoint discretization, cross terms such as $A_\theta(x, t)B(x, t)^{-1}\dot{x}$ should be interpreted as $A_\theta(x, t)B(x, t)^{-1} \bullet \dot{x}$, where $\bullet$ denotes the Ito product. Although $\mathscr{A}_\theta(x, t)$ is different for a midpoint discretization [26], both discretizations reduce to the same expression for additive noise systems [27]. We have $\mathcal{P}_\theta(\boldsymbol{\Gamma}) = \mathcal{P}_\theta(\boldsymbol{\Gamma}|\boldsymbol{x}^0)P_\theta(\boldsymbol{x}^0)$, where $P_\theta(\boldsymbol{x}^0)$ is the initial probability density of $\boldsymbol{x}^0$ at $t = 0$ [$\int \mathcal{D}\boldsymbol{\Gamma}\mathcal{P}_\theta(\boldsymbol{\Gamma}) = 1$]. The log-probability is calculated as

$$\ln\mathcal{P}_\theta(\boldsymbol{\Gamma}) = \ln\mathscr{N} + \ln P_\theta(\boldsymbol{x}^0) - \frac{1}{4}\int_0^T dt(\dot{\boldsymbol{x}} - \boldsymbol{A}_\theta)^\top \boldsymbol{B}^{-1}(\dot{\boldsymbol{x}} - \boldsymbol{A}_\theta). \tag{8}$$

From Eq. (5), we need to calculate the second derivative of Eq. (8) with respect to θ:

$$\frac{\partial^2}{\partial\theta^2}\ln\mathcal{P}_\theta(\boldsymbol{\Gamma}) = \frac{\partial^2}{\partial\theta^2}\ln P_\theta(\boldsymbol{x}^0) - \frac{1}{2}\int_0^T dt\left(\frac{\partial}{\partial\theta}\boldsymbol{A}_\theta\right)^\top \boldsymbol{B}^{-1}\left(\frac{\partial}{\partial\theta}\boldsymbol{A}_\theta\right) + \frac{1}{2}\int_0^T dt\,(\dot{\boldsymbol{x}} - \boldsymbol{A}_\theta)^\top \bullet \boldsymbol{B}^{-1}\left(\frac{\partial^2}{\partial\theta^2}\boldsymbol{A}_\theta\right). \tag{9}$$

When applying the expectation $\langle\cdots\rangle_\theta$ to Eq. (9), the last term disappears [cf. Eq. (B8) in Appendix B]. Therefore, from Eq. (5), the Fisher information is given by

$$\mathcal{I}(\theta) = -\left\langle\frac{\partial^2}{\partial\theta^2}\ln P_\theta(\boldsymbol{x}^0)\right\rangle_\theta + \frac{1}{2}\left\langle\int_0^T dt\left(\frac{\partial}{\partial\theta}\boldsymbol{A}_\theta^\top\right)\boldsymbol{B}^{-1}\left(\frac{\partial}{\partial\theta}\boldsymbol{A}_\theta\right)\right\rangle_\theta. \tag{10}$$

Equation (4) reduces to the recently proposed fluctuation-response inequality [17]. Suppose $\boldsymbol{A}_\theta(\boldsymbol{x}, t) = \boldsymbol{\mathcal{A}}(\boldsymbol{x}, t) + \theta\boldsymbol{\mathcal{Z}}(\boldsymbol{x}, t)$, where θ is a sufficiently small real parameter and $\boldsymbol{\mathcal{Z}}(\boldsymbol{x}, t)$ is an arbitrary perturbation. Since θ is sufficiently small, $\partial_\theta\langle\Theta(\boldsymbol{\Gamma})\rangle_\theta$ can be approximated by $\partial_\theta\langle\Theta(\boldsymbol{\Gamma})\rangle_\theta|_{\theta=0} \simeq [\langle\Theta(\boldsymbol{\Gamma})\rangle_\theta - \langle\Theta(\boldsymbol{\Gamma})\rangle_{\theta=0}]/\theta$. Entering these expressions into Eq. (4), we obtain the fluctuation-response inequality:

$$\frac{\mathrm{var}_{\theta=0}[\Theta(\boldsymbol{\Gamma})]}{[\langle\Theta(\boldsymbol{\Gamma})\rangle_\theta - \langle\Theta(\boldsymbol{\Gamma})\rangle_{\theta=0}]^2} \geqslant \frac{1}{\mathcal{C}} = \frac{1}{\theta^2\mathcal{I}(0)}, \tag{11}$$

where $\mathcal{C} \equiv \theta^2\mathcal{I}(0)$ and $\mathcal{I}(0) = \frac{1}{2}\langle\int_0^T dt\,\boldsymbol{\mathcal{Z}}^\top\boldsymbol{B}^{-1}\boldsymbol{\mathcal{Z}}\rangle_{\theta=0}$. Note that the boundary term is ignored in Eq. (11). The boundary term is zero when the probability density function remains unchanged upon perturbation, which is the case for the thermodynamic uncertainty relation. From Eq. (11), $1/[\theta^2\mathcal{I}(0)]$ is the lower bound of the fluctuation-response inequality. Because the fluctuation-response inequality holds only for sufficiently small θ, the perturbation should be sufficiently weak.

We explore the thermodynamic uncertainty relation from a statistical inference perspective. Reference [17] provides an alternative derivation of the finite-time current thermodynamic uncertainty relation [10] with the fluctuation-response inequality using the notion of a virtual perturbation. Let us consider a system $\boldsymbol{A}(\boldsymbol{x},t)=\boldsymbol{A}(\boldsymbol{x})$ and $\boldsymbol{C}(\boldsymbol{x},t)=\boldsymbol{C}(\boldsymbol{x})$ in Eq. (1) at a steady state. The thermodynamic uncertainty relation considers the generalized current

$$\Theta_{\mathrm{cur}}(\boldsymbol{\Gamma})\equiv\int_0^T\boldsymbol{\Lambda}(\boldsymbol{x})^\top\circ\dot{\boldsymbol{x}}\,dt,\tag{12}$$

where $\boldsymbol{\Lambda}(\boldsymbol{x})\equiv[\Lambda_1(\boldsymbol{x}),\dots,\Lambda_N(\boldsymbol{x})]^\top$ is an arbitrary projection function and $\circ$ denotes the Stratonovich product. Using the virtual perturbation technique [17], we define

$$A_{\theta,i}(\boldsymbol{x})\equiv(\theta+1)A_i(\boldsymbol{x})-\frac{\theta}{P^{\mathrm{ss}}(\boldsymbol{x})}\sum_j\partial_{x_j}B_{ij}(\boldsymbol{x})P^{\mathrm{ss}}(\boldsymbol{x}),\tag{13}$$

where $P^{\mathrm{ss}}(\boldsymbol{x})$ is the steady-state distribution of the unperturbed dynamics (i.e., the dynamics for the case in which $\theta=0$). Note that the steady-state distribution corresponding to $\boldsymbol{A}_\theta(\boldsymbol{x})$ of Eq. (13) does not depend on θ [17]. Using Eq. (13), we find

$$\begin{aligned}\langle\Theta_{\mathrm{cur}}(\boldsymbol{\Gamma})\rangle_\theta&=\left\langle\int_0^T\boldsymbol{\Lambda}(\boldsymbol{x})^\top\circ\dot{\boldsymbol{x}}\,dt\right\rangle_\theta\\&=T\int d\boldsymbol{x}\,\boldsymbol{\Lambda}(\boldsymbol{x})^\top\boldsymbol{J}_\theta^{\mathrm{ss}}(\boldsymbol{x})\\&=T\int d\boldsymbol{x}\,\boldsymbol{\Lambda}(\boldsymbol{x})^\top(1+\theta)\boldsymbol{J}^{\mathrm{ss}}(\boldsymbol{x})\\&=(\theta+1)\jmath,\end{aligned}\tag{14}$$

where $\boldsymbol{J}^{\mathrm{ss}}(\boldsymbol{x})\equiv[J_1^{\mathrm{ss}}(\boldsymbol{x}),\dots,J_N^{\mathrm{ss}}(\boldsymbol{x})]^\top$ and $\boldsymbol{J}_\theta^{\mathrm{ss}}(\boldsymbol{x})\equiv[J_{\theta,1}^{\mathrm{ss}}(\boldsymbol{x}),\dots,J_{\theta,N}^{\mathrm{ss}}(\boldsymbol{x})]^\top$ are the steady-state probability currents of unperturbed and perturbed dynamics, respectively, and $\jmath\equiv\langle\Theta_{\mathrm{cur}}(\boldsymbol{\Gamma})\rangle_{\theta=0}=T\int d\boldsymbol{x}\,\boldsymbol{\Lambda}(\boldsymbol{x})^\top\boldsymbol{J}^{\mathrm{ss}}(\boldsymbol{x})$. $\jmath$ corresponds to the averaged current and $\partial_\theta\langle\Theta_{\mathrm{cur}}(\boldsymbol{\Gamma})\rangle_\theta=\jmath$ from Eq. (14). From Eq. (10), the Fisher information is

$$\begin{aligned}\mathcal{I}(0)&=\frac{1}{2}\left\langle\int_0^T dt\left(\frac{\boldsymbol{J}^{\mathrm{ss}}(\boldsymbol{x})^\top}{P^{\mathrm{ss}}(\boldsymbol{x})}\right)\boldsymbol{B}(\boldsymbol{x})^{-1}\left(\frac{\boldsymbol{J}^{\mathrm{ss}}(\boldsymbol{x})}{P^{\mathrm{ss}}(\boldsymbol{x})}\right)\right\rangle_{\theta=0}\\&=\frac{T}{2}\int d\boldsymbol{x}\frac{\boldsymbol{J}^{\mathrm{ss}}(\boldsymbol{x})^\top\boldsymbol{B}(\boldsymbol{x})^{-1}\boldsymbol{J}^{\mathrm{ss}}(\boldsymbol{x})}{P^{\mathrm{ss}}(\boldsymbol{x})}.\end{aligned}\tag{15}$$

By using Eqs. (4) and (15), the thermodynamic uncertainty relation is obtained as follows:

$$\frac{\mathrm{var}_{\theta=0}[\Theta_{\mathrm{cur}}(\boldsymbol{\Gamma})]}{\jmath^2}\geqslant\frac{2}{\Delta S_{\mathrm{tot}}},\tag{16}$$

where ΔS_{tot} is the total entropy production:

$$\Delta S_{\mathrm{tot}}\equiv T\int d\boldsymbol{x}\,\frac{\boldsymbol{J}^{\mathrm{ss}}(\boldsymbol{x})^\top\boldsymbol{B}(\boldsymbol{x})^{-1}\boldsymbol{J}^{\mathrm{ss}}(\boldsymbol{x})}{P^{\mathrm{ss}}(\boldsymbol{x})}.\tag{17}$$

Equation (17) is the total entropy production assuming that all variables are even under time reversal. When systems include odd variables (e.g., underdamped systems), the total entropy production is expressed differently. In particular, the thermodynamic uncertainty relation including only the total entropy production term is violated in underdamped systems [28,29]. The calculations above show that, from the perspective of statistical inference, the current $\Theta_{\mathrm{cur}}(\boldsymbol{\Gamma})$ is an estimator that infers θ and the total entropy production corresponds to the Fisher information in θ-space. The Fisher information describes the log-likelihood change when varying a parameter θ. If the change is large, the curvature of the log-likelihood becomes steeper, which results in a more accurate parameter inference.

B. Equality condition near equilibrium

Identifying the thermodynamic uncertainty relation as the Cramér-Rao inequality, we can obtain the equality condition of the thermodynamic uncertainty relation, which was not reported in the approach based on the fluctuation-response inequality. From the equality condition of the Cramér-Rao inequality, Eq. (4) is satisfied with equality if and only if the following relation holds:

$$\frac{\partial}{\partial\theta}\ln\mathcal{P}_\theta(\boldsymbol{\Gamma})=\mu(\theta)[\Theta(\boldsymbol{\Gamma})-\psi(\theta)],\tag{18}$$

where $\mu(\theta)$ is a scaling function [Eq. (A3) in Appendix A]. When this relation holds, the thermodynamic uncertainty relation also holds with equality. For simplicity, we first consider a one-dimensional system with periodic boundary conditions. Converting from Ito to Stratonovich-type currents [cf. Eq. (C14) in Appendix C], the left-hand side of Eq. (18) is

$$\begin{aligned}&\frac{\partial}{\partial\theta}\ln\mathcal{P}_\theta(\Gamma)\\&=\frac{1}{2}\int_0^T dt\left[\frac{J^{\mathrm{ss}}}{P^{\mathrm{ss}}(x)B(x)}\bullet\dot{x}-\frac{J^{\mathrm{ss}}A_\theta(x)}{P^{\mathrm{ss}}(x)B(x)}\right]\\&=\frac{1}{2}\int_0^T dt\,\frac{J^{\mathrm{ss}}}{P^{\mathrm{ss}}(x)B(x)}\circ\dot{x}-\frac{1+\theta}{2}\int_0^T dt\,\frac{[J^{\mathrm{ss}}]^2}{P^{\mathrm{ss}}(x)^2B(x)}.\end{aligned}\tag{19}$$

Accordingly, the right-hand side of the same equation becomes

$$\begin{aligned}&\mu(\theta)[\Theta_{\mathrm{cur}}(\Gamma)-\psi(\theta)]\\&=\mu(\theta)\left[\int_0^T dt\,\Lambda(x)\circ\dot{x}-(1+\theta)\jmath\right],\end{aligned}\tag{20}$$

where we used Eq. (14) in the last line [$\psi(\theta)=\langle\Theta_{\mathrm{cur}}(\Gamma)\rangle_\theta=(1+\theta)\jmath$]. Without loss of generality, we set $\mu(\theta)=J^{\mathrm{ss}}/2$ because multiplying $\Lambda(x)$ with a constant results in the same bound. Correspondence between Eqs. (19) and (20) should hold for an arbitrary trajectory Γ to attain equality in the thermodynamic uncertainty relation. From Eqs. (18)–(20), we determine that for an arbitrary Γ, the following relation should be satisfied:

$$\Xi(\Gamma)-(1+\theta)J^{\mathrm{ss}}\Psi(\Gamma)=0,\tag{21}$$

where

$$\Xi(\Gamma) \equiv \int_0^T dt \left[\Lambda(x) - \frac{1}{P^{\rm ss}(x)B(x)}\right] \circ \dot{x}, \quad (22)$$

$$\Psi(\Gamma) \equiv T \int dx\, \Lambda(x) - \int_0^T \frac{1}{P^{\rm ss}(x)^2 B(x)} dt. \quad (23)$$

Equation (21) is a necessary and sufficient condition for the thermodynamic uncertainty relation to hold with equality. To satisfy Eq. (21), $\Xi(\Gamma)$ and $(1+\theta)J^{\rm ss}\Psi(\Gamma)$ should individually vanish [30]. From the condition $\Xi(\Gamma) = 0$, we find that the only quantity that satisfies the equality is the current:

$$\Theta_{\rm tot}(\Gamma) \equiv \int_0^T dt \frac{1}{P^{\rm ss}(x)B(x)} \circ \dot{x}. \quad (24)$$

For $B(x) = D$ (additive noise), $\Theta_{\rm tot}(\Gamma)$ is $\Theta_{\rm tot}(\Gamma) \propto \int_0^T dt\, \dot{s}_{\rm tot}(t)$, where $\dot{s}_{\rm tot}$ is the stochastic total entropy production rate:

$$\begin{aligned}\dot{s}_{\rm tot} &\equiv \frac{\dot{q}}{D} - \frac{d}{dt}\ln P^{\rm ss}(x)\\ &= \frac{A(x)\circ\dot{x}}{D} - \frac{\partial_x P^{\rm ss}(x)}{P^{\rm ss}(x)}\circ\dot{x} = \frac{J^{\rm ss}}{P^{\rm ss}(x)D}\circ\dot{x}. \quad (25)\end{aligned}$$

Here, $\dot{q} \equiv A(x)\circ\dot{x}$ is the stochastic heat dissipation rate [2]. Although the first term $\Xi(\Gamma)$ in Eq. (21) vanishes with $\Theta_{\rm tot}(\Gamma)$, the second term does not. Still, when we consider a near-equilibrium condition, $J^{\rm ss} \to 0$, the second term in Eq. (21) converges to 0. Our result shows that $\Theta_{\rm tot}(\Gamma)$ (and its multiples) is the only quantity that can attain equality near equilibrium. This also holds for multidimensional systems with $\boldsymbol{B}(\boldsymbol{x}) = \boldsymbol{D}$ (additive noise). Particularly, converting from Ito to Stratonovich-type current [Eq. (C15) in Appendix C], we repeat the same calculations as for the one-dimensional case to obtain

$$\begin{aligned}\ln \mathcal{P}_\theta(\boldsymbol{\Gamma}) &= \frac{1}{2}\int_0^T dt \frac{\boldsymbol{J}^{\rm ss}(\boldsymbol{x})^\top \boldsymbol{D}^{-1}}{P^{\rm ss}(\boldsymbol{x})}\circ\dot{\boldsymbol{x}}\\ &\quad - \frac{1+\theta}{2}\int_0^T dt \frac{\boldsymbol{J}^{\rm ss}(\boldsymbol{x})^\top \boldsymbol{D}^{-1}\boldsymbol{J}^{\rm ss}(\boldsymbol{x})}{P^{\rm ss}(\boldsymbol{x})^2}. \quad (26)\end{aligned}$$

Equation (26) is the multidimensional generalization of Eq. (19). Again, we define the following current:

$$\Theta_{\rm tot}(\boldsymbol{\Gamma}) \equiv \int_0^T \frac{\boldsymbol{J}^{\rm ss}(\boldsymbol{x})^\top \boldsymbol{D}^{-1}}{P^{\rm ss}(\boldsymbol{x})}\circ\dot{\boldsymbol{x}}\, dt. \quad (27)$$

Repeating the analysis of the one-dimensional case, we find that the equality of the thermodynamic uncertainty relation is satisfied near equilibrium if and only if the current is Eq. (27) (and its multiples). $\Theta_{\rm tot}(\boldsymbol{\Gamma})$ of Eq. (27) satisfies $\Theta_{\rm tot}(\boldsymbol{\Gamma}) = \int_0^T \dot{s}_{\rm tot} dt$, where $\dot{s}_{\rm tot}$ is the stochastic total entropy production rate for multidimensional systems:

$$\dot{s}_{\rm tot} \equiv \boldsymbol{A}(\boldsymbol{x})^\top \boldsymbol{D}^{-1}\circ\dot{\boldsymbol{x}} - \frac{d}{dt}\ln P^{\rm ss}(\boldsymbol{x}) = \frac{\boldsymbol{J}^{\rm ss}(\boldsymbol{x})^\top \boldsymbol{D}^{-1}}{P^{\rm ss}(\boldsymbol{x})}\circ\dot{\boldsymbol{x}}. \quad (28)$$

In Eq. (28), $\boldsymbol{A}(\boldsymbol{x})^\top \boldsymbol{D}^{-1}\circ\dot{\boldsymbol{x}}$ corresponds to the stochastic medium entropy rate. The stochastic total entropy production has been shown to attain equality near equilibrium in Ref. [13], but it was not shown that it is the only quantity that attains equality. This current was shown to satisfy the equality of the thermodynamic uncertainty relation using the linear response [31], and it was demonstrated to provide the tightest quadratic bound for the rate function [7].

C. Exact equality condition

As the equality condition discussed above is asymptotic with respect to $J^{\rm ss} \to 0$, the equality is not met exactly. Next, we seek an exact equality condition for the one-dimensional case. With $\Theta_{\rm tot}(\Gamma)$ [Eq. (24)], the first term $\Xi(\Gamma)$ in Eq. (21) vanishes, while the second term $(1+\theta)J^{\rm ss}\Psi(\Gamma)$ does not. Therefore, $\Psi(\Gamma) = 0$ should hold for an arbitrary Γ to attain equality in the thermodynamic uncertainty relation. In Eq. (23), the first term is a constant due to the integration with respect to x, but the second term $\int_0^T 1/[P^{\rm ss}(x)^2 B(x)]dt$ depends on Γ. Therefore, to satisfy $\Psi(\Gamma) = 0$ for an arbitrary Γ, the integrand $1/[P^{\rm ss}(x)^2 B(x)]$ should be constant, which yields

$$P^{\rm ss}(x) = \frac{c}{\sqrt{B(x)}}, \quad (29)$$

with $c > 0$ being a normalization constant. Indeed, substituting Eq. (29) into $\Psi(\Gamma)$ [Eq. (23)], we find $\Psi(\Gamma) = 0$. From $J^{\rm ss} = A(x)P^{\rm ss}(x) - \partial_x B(x)P^{\rm ss}(x)$, we obtain $A(x)$ as follows:

$$A(x) = \frac{J^{\rm ss} + \partial_x B(x)P^{\rm ss}(x)}{P^{\rm ss}(x)} = \kappa C(x) + C(x)\frac{d}{dx}C(x), \quad (30)$$

where κ is an arbitrary parameter. $\Lambda(x)$ is given by

$$\Lambda(x) \propto \frac{1}{P^{\rm ss}(x)B(x)} \propto \frac{1}{\sqrt{B(x)}} = \frac{1}{C(x)}. \quad (31)$$

Regardless of the system being near equilibrium, equality in the thermodynamic uncertainty relation is attained if and only if Eqs. (30) and (31) hold, which has not been demonstrated in the literature.

For the multidimensional case with $\boldsymbol{B}(\boldsymbol{x}) = \boldsymbol{D}$, from Eq. (26), the exact equality of the thermodynamic uncertainty relation is satisfied if and only if the current is proportional to $\Theta_{\rm tot}(\boldsymbol{\Gamma})$ [Eq. (27)] and $\boldsymbol{J}^{\rm ss}(\boldsymbol{x})\boldsymbol{D}^{-1}\boldsymbol{J}^{\rm ss}(\boldsymbol{x})/P^{\rm ss}(\boldsymbol{x})^2$ is constant for any $\boldsymbol{x}$. However, it is difficult to specify systems satisfying the latter condition.

D. Example: Particle in a periodic potential

We numerically confirm these equality conditions by considering a particle on a periodic potential with a period of 2π subject to an external force. We consider the following periodic drift:

$$A(x) = [a + \sin(x)][b + \cos(x)], \quad (32)$$

where $a > 1$ and $b \geqslant 0$ are model parameters. The drift defined by Eq. (32) is the sum of the periodic potential and the external force. Let $V(x)$ be an effective potential function of Eq. (32):

$$\begin{aligned}V(x) &\equiv -\int A(x)dx\\ &= -\frac{1}{2}[a + \sin(x)]^2 - b[ax - \cos(x)]. \quad (33)\end{aligned}$$

In Fig. 1(b), $V(x)$ is plotted for $b = 0$ (solid line), 0.5 (dashed line), and 1.0 (dot-dashed line) with $a = 2$. When $b = 0$, because $V(x)$ does not have a tilt, the system is in equilibrium at a steady state. We employ the drift of Eq. (32) since $C(x)$ satisfying Eq. (30) can be expressed analytically. From Eq. (16), we define

$$\mathcal{F} \equiv \frac{\mathrm{var}_{\theta=0}[\Theta(\Gamma)]\Delta S_{\mathrm{tot}}}{J^2} \geqslant 2, \tag{34}$$

which is $\mathcal{F} = 2$ for the equality of the thermodynamic uncertainty relation. We numerically calculate $\mathcal{F}$ by repeating simulations $N_S = 5.0 \times 10^6$ times with temporal resolution $\Delta t = 0.0002$ [parameters are shown in the caption of Fig. 1(c)]. We consider the following four cases: (i) $\Lambda(x) = 1/[P^{\mathrm{ss}}(x)B(x)]$, $C(x) = \sqrt{D}$; (ii) $\Lambda(x) = 1/C(x) \propto 1/[P^{\mathrm{ss}}(x)B(x)]$, $C(x) = a + \sin(x)$; (iii) $\Lambda(x) = 1, C(x) = \sqrt{D}$; and (iv) $\Lambda(x) = A(x)$, $C(x) = \sqrt{D}$ (see Appendix D). Cases (i) and (ii) satisfy the near-equilibrium equality condition and (ii) further satisfies Eq. (30), while (iii) and (iv) do not. The current of (iv) depicts the stochastic heat dissipation. Figure 1(c) shows $\mathcal{F}$ as a function of J^{ss} for (i), (ii), (iii), and (iv), which are described by circles, diamonds, triangles, and squares, respectively. For $J^{\mathrm{ss}} \to 0$, only (i) and (ii) show $\mathcal{F} \to 2$. Case (ii) exhibits $\mathcal{F} \simeq 2$ for all J^{ss}, indicating that it satisfies the equality even when the system is far from equilibrium, as expected. Cases (iii) and (iv) are $\mathcal{F} > 2$; thus, they do not satisfy the equality of the thermodynamic uncertainty relation, which agrees with our theoretical result.

IV. CHAPMAN-ROBBINS INEQUALITY

A. Derivation of uncertainty relation

A different information inequality can be applied to the system from the identification of thermodynamic systems in terms of statistical inference. By applying the Chapman-Robbins inequality [20–22] (Appendix A), which is a generalization of the Cramér-Rao inequality, to Eq. (1), the following relation holds:

$$\frac{\mathrm{var}_{\theta=0}[\Theta(\mathbf{\Gamma})]}{[\langle\Theta(\mathbf{\Gamma})\rangle_\theta - \langle\Theta(\mathbf{\Gamma})\rangle_{\theta=0}]^2} \geqslant \frac{1}{\mathbb{D}_{\mathrm{PE}}[\mathcal{P}_\theta||\mathcal{P}_{\theta=0}]}, \tag{35}$$

where $\mathbb{D}_{\mathrm{PE}}[\mathcal{P}_\theta||\mathcal{P}_{\theta=0}]$ is the Pearson divergence between $\mathcal{P}_\theta$ and $\mathcal{P}_{\theta=0}$ defined by

$$\mathbb{D}_{\mathrm{PE}}[\mathcal{P}_\theta||\mathcal{P}_{\theta=0}] \equiv \int \mathcal{D}\mathbf{\Gamma}\left(\frac{\mathcal{P}_\theta(\mathbf{\Gamma})}{\mathcal{P}_{\theta=0}(\mathbf{\Gamma})} - 1\right)^2 \mathcal{P}_{\theta=0}(\mathbf{\Gamma}). \tag{36}$$

In Eq. (36), $[\langle\Theta(\mathbf{\Gamma})\rangle_\theta - \langle\Theta(\mathbf{\Gamma})\rangle_{\theta=0}]^2$ describes the difference between two dynamics characterized by $\theta = 0$ and $\theta \neq 0$, which represents the sensitivity of the system. Thus, the ratio between the variance of the unperturbed dynamics and the sensitivity is bounded from below by the reciprocal of the Pearson divergence between the two dynamics. For $\theta \to 0$, Eq. (35) reduces to the Cramér-Rao inequality [Eq. (4)] and the fluctuation-response inequality. Although the fluctuation-response inequality only holds for sufficiently weak perturbations, as it holds locally around $\theta = 0$, Eq. (35) is satisfied for an arbitrary $\theta \neq 0$, indicating that Eq. (35) can be used beyond a linear-response regime. In stochastic thermodynamics, by using thermodynamic inequalities, several measures of efficiency have been calculated to evaluate the performance of systems [32–34]. Similarly, we can evaluate the efficiency in terms of sensitivity and precision with Eq. (35). As Eq. (35) holds for an arbitrary θ, the Chapman-Robbins inequality is often stated to provide the lower bound, which is at least as tight as the Cramér-Rao inequality:

$$\mathrm{var}_{\theta=0}[\Theta(\mathbf{\Gamma})] \geqslant \sup_\theta \frac{[\langle\Theta(\mathbf{\Gamma})\rangle_\theta - \langle\Theta(\mathbf{\Gamma})\rangle_{\theta=0}]^2}{\mathbb{D}_{\mathrm{PE}}[\mathcal{P}_\theta||\mathcal{P}_{\theta=0}]} \geqslant \frac{[\partial_\theta\langle\Theta(\mathbf{\Gamma})\rangle_\theta]^2_{\theta=0}}{\mathcal{I}(0)}. \tag{37}$$

However, in the present paper we only focus on the relation of Eq. (35).

B. Example 1: Linear Langevin equation

First, we study Eq. (35) in a linear Langevin equation because the Pearson divergence can be obtained analytically. We consider the following equation in Eq. (1):

$$A_\theta(x,t) = -\alpha x + \theta u(t), \quad C(x,t) = \sqrt{D}, \tag{38}$$

where $\alpha > 0$, $u(t)$ is an arbitrary input function, and D is the noise intensity. The initial condition is $x = 0$ at $t = 0$. By following the calculation of the path integral, the Pearson divergence is represented analytically by (Appendix E)

$$\mathbb{D}_{\mathrm{PE}}[\mathcal{P}_\theta||\mathcal{P}_{\theta=0}] = -1 + \exp\left(\frac{\theta^2}{2D}\int_0^T u(t)^2 dt\right). \tag{39}$$

When we define $\Theta_x(\Gamma) \equiv \int_0^T \dot{x}\,dt$, $\Theta_x(\Gamma)$ is the position $x(T)$. Therefore, the Chapman-Robbins inequality in Eq. (35) is

$$\frac{\mathrm{var}_{\theta=0}[x(T)]}{[\langle x(T)\rangle_\theta - \langle x(T)\rangle_{\theta=0}]^2} \geqslant \frac{1}{\mathbb{D}_{\mathrm{PE}}[\mathcal{P}_\theta||\mathcal{P}_{\theta=0}]}. \tag{40}$$

We also consider the lower bound of the fluctuation-response inequality [Eq. (11)]: $1/[\theta^2\mathcal{I}(0)] = 2D/[\theta^2\int_0^T u(t)^2 dt]$. Using $\exp(x) \geqslant 1 + x$ in Eq. (39), we can easily show

$$\mathbb{D}_{\mathrm{PE}}[\mathcal{P}_\theta||\mathcal{P}_{\theta=0}] \geqslant \frac{\theta^2}{2D}\int_0^T u(t)^2 dt = \theta^2\mathcal{I}(0). \tag{41}$$

When $x = 0$ at time $t = 0$, the variance and the mean of Eq. (38) are given by

$$\langle x(T)\rangle_\theta = \left[\theta\int_0^T u(t)e^{\alpha t}dt\right]e^{-\alpha T}, \tag{42}$$

$$\mathrm{var}_{\theta=0}[x(T)] = \frac{D}{\alpha}[1 - e^{-2\alpha T}]. \tag{43}$$

For any $u(t) \geqslant 0$, using the Cauchy-Schwarz inequality, we have

$$\left[\int_0^T u(t)e^{\alpha t}dt\right]^2 \leqslant \int_0^T u(t)^2 dt \int_0^T e^{2\alpha t}dt = \frac{e^{2\alpha T} - 1}{2\alpha}\int_0^T u(t)^2 dt, \tag{44}$$

which yields the following relation:

$$\frac{\mathrm{var}_{\theta=0}[x(T)]}{[\langle x(T)\rangle_\theta - \langle x(T)\rangle_{\theta=0}]^2} \geqslant \frac{1}{\theta^2\mathcal{I}(0)} \geqslant \frac{1}{\mathbb{D}_{\mathrm{PE}}[\mathcal{P}_\theta||\mathcal{P}_{\theta=0}]}. \tag{45}$$

Equation (45) indicates that the bound of the fluctuation-response inequality is always tighter than that of the

Chapman-Robbins inequality. This relation seems to be inconsistent with the fact that the Chapman-Robbins inequality is at least as tight as the Cramér-Rao inequality [cf. Eq. (37)]. However, Eq. (37) does not contradict Eq. (45). Although the fluctuation-response inequality always holds for the linear system, it is violated in nonlinear systems, as shown in the next example.

C. Example 2: Limit cycle oscillator

Equation (35) bounds the sensitivity and the precision, which is of particular interest in limit cycle oscillators. Circadian clocks are biological limit cycle oscillators in organisms that orchestrate the activities of several organs. Their temporal precision is incredibly high (the standard deviation of the period is 3–5 min over 24 h) [35] and several mechanisms have been proposed for such precision [36–39]. Simultaneously, circadian clocks have to synchronize to sunlight cycles such that biological activities are operational at specific times. As oscillators with higher sensitivity are vulnerable to periodic signals as well as noise, precision and sensitivity appear to be tradeoff factors, which is an uncertainty relation in stochastic oscillators [40–42].

We consider a deterministic limit cycle oscillator defined by $\dot{\boldsymbol{x}} = \boldsymbol{A}(\boldsymbol{x})$. We can define the phase ϕ on a closed orbit of the deterministic oscillation by $\dot{\phi} = \Omega$, where $\Omega \equiv 2\pi/\tau$ is the angular frequency of the oscillation (τ is the period of the deterministic oscillation). In the presence of an external signal, the dynamics obeys Eq. (1) with

$$\boldsymbol{A}_\theta(\boldsymbol{x}, t) = \boldsymbol{A}(\boldsymbol{x}) + \theta \boldsymbol{u}(t), \tag{46}$$

where $\boldsymbol{u}(t) = [u_1(t), \dots, u_N(t)]^\top$ depicts the signal. Although ϕ is defined only on the deterministic closed orbit, we can expand the definition of the phase over the entire $\boldsymbol{x}$ space, which is denoted by $\phi(\boldsymbol{x})$ [43]. $\phi(\boldsymbol{x})$ can be calculated by directly solving the ordinary differential equation (Appendix F). The integrated phase from $t = 0$ to $t = T$ is given by $\int_0^T \dot{\phi}(\boldsymbol{x}(t))dt$. Since the time derivative of ϕ is $\dot{\phi} = \sum_{i=1}^N (\partial_{x_i}\phi(\boldsymbol{x})) \circ \dot{x}_i$, we define a current

$$\Theta_\phi(\boldsymbol{\Gamma}) \equiv \int_0^T \sum_{i=1}^N \left(\frac{\partial}{\partial x_i}\phi(\boldsymbol{x})\right) \circ \dot{x}_i dt, \tag{47}$$

which is the integrated phase calculated from a trajectory $\boldsymbol{\Gamma}$. The temporal precision of the oscillator is quantified by $\mathrm{var}_{\theta=0}[\Theta_\phi(\boldsymbol{\Gamma})]$, which is the variance of the phase of unperturbed dynamics (lower variance corresponds to higher precision). The sensitivity of the oscillator can be quantified by the phase difference between perturbed and unperturbed dynamics. Therefore, we define the phase sensitivity as $[\langle\Theta_\phi(\boldsymbol{\Gamma})\rangle_\theta - \langle\Theta_\phi(\boldsymbol{\Gamma})\rangle_{\theta=0}]^2$. Figure 2(a) illustrates the phase $\phi(\boldsymbol{x})$ (the dotted line shows the isochron) and the phase difference between unperturbed and perturbed dynamics. From Eq. (35), the phase variance and the sensitivity satisfy the relation

$$\frac{\mathrm{var}_{\theta=0}[\Theta_\phi(\boldsymbol{\Gamma})]}{[\langle\Theta_\phi(\boldsymbol{\Gamma})\rangle_\theta - \langle\Theta_\phi(\boldsymbol{\Gamma})\rangle_{\theta=0}]^2} \geqslant \frac{1}{\mathbb{D}_{\mathrm{PE}}[\mathcal{P}_\theta||\mathcal{P}_{\theta=0}]}, \tag{48}$$

which shows that as $\mathbb{D}_{\mathrm{PE}}[\mathcal{P}_\theta||\mathcal{P}_{\theta=0}]$ increases, both precision and sensitivity are improved simultaneously. In limit cycle oscillators, perturbations are often applied to

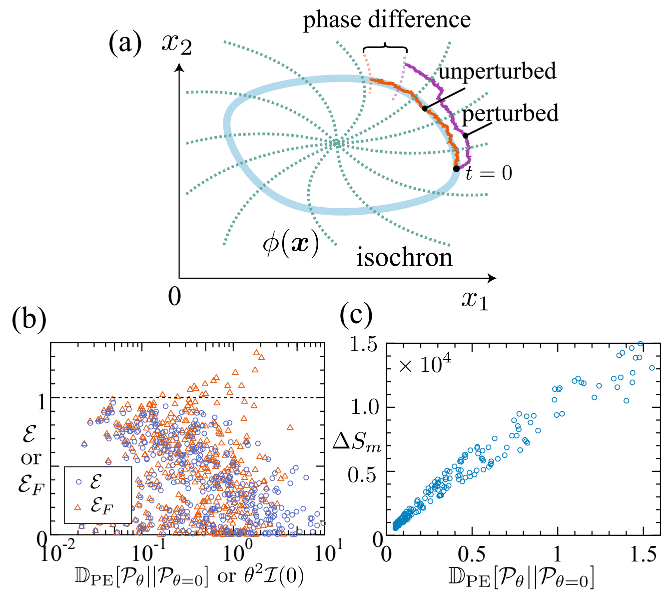

FIG. 2. (a) The phase $\phi(\boldsymbol{x})$ is defined on the coordinate space, and the isochron (green dotted lines) denotes the equiphase surface. The phase sensitivity is quantified by the phase difference between perturbed (purple line) and unperturbed (orange line) dynamics, with the deterministic oscillation shown by the light blue solid line. (b) Numerical verification of the Chapman-Robbins inequality in the stochastic oscillator. For random realizations, $\mathcal{E}$ (blue circles) and $\mathcal{E}_F$ (orange triangles) are plotted as a function of $\mathbb{D}_{\mathrm{PE}}[\mathcal{P}_\theta||\mathcal{P}_{\theta=0}]$ and $\theta^2\mathcal{I}(0)$, respectively. When the Chapman-Robbins inequality and the fluctuation-response inequality are satisfied, $\mathcal{E} \leqslant 1$ and $\mathcal{E}_F \leqslant 1$, respectively. Parameters are $D \in [0.05, 0.2]$, $\theta \in [0.1, 0.5]$, $T \in \{\tau/4, \tau/8, \tau/16\}$. (c) Medium entropy ΔS_m as a function of $\mathbb{D}_{\mathrm{PE}}[\mathcal{P}_\theta||\mathcal{P}_{\theta=0}]$, where circles denote random realizations. Parameters are $D \in [0.01, 0.2]$, $\theta = 0.05$, $T = \tau$ (τ is the period of the deterministic oscillation).

experimentally observe the properties of oscillators. Particularly, the sensitivity $[\langle\Theta_\phi(\boldsymbol{\Gamma})\rangle_\theta - \langle\Theta_\phi(\boldsymbol{\Gamma})\rangle_{\theta=0}]^2$ corresponds to the square of the phase response curve [44]. The phase variance and the sensitivity are common measures and were used in Refs. [40–42] to study their tradeoff relation. However, an explicit inequality between the two quantities has not been reported yet. We determine that their ratio is lower bounded by the Pearson divergence, which is an information quantity. Such information quantities play important roles in information thermodynamics [4]. The Pearson divergence between the original and perturbed trajectories is experimentally measurable because trajectory-based quantities were previously measured [45–49].

We numerically confirm the inequality relation of Eq. (48). We employ the Van der Pol oscillator [50], which is representative of many limit cycle oscillators, including circadian oscillators. This oscillator has been extensively employed in the literature. The noisy Van der Pol oscillator is defined by

$$\boldsymbol{A}_\theta(\boldsymbol{x}, t) = \begin{bmatrix} x_2 + \theta u(t) \\ \zeta(1 - x_1^2)x_2 - x_1 \end{bmatrix}, \; \boldsymbol{B} = \begin{bmatrix} D & 0 \\ 0 & D \end{bmatrix}, \tag{49}$$

where ζ is a model parameter ($\zeta = 2.5$ throughout), D is noise intensity, and $u(t) = 1$ for $t > 0$ and $u(t) = 0$ for $t \leqslant 0$. Using Monte Carlo simulations, we solve the Langevin equation (49) with time resolution $\Delta t = 0.0002$ and evaluate Pearson divergence, the sensitivity, and the phase variance (Appendix F). We randomly select D, θ, and T, and we repeat simulations $N_S = 5.0 \times 10^5$ times at each of the selected parameter settings [ranges of the random parameters are shown in the caption of Fig. 2(b)]. For initial values, we randomly select a point on the closed orbit of the deterministic oscillation [the light blue line in Fig. 2(a)]. We calculate

$$\mathcal{E} \equiv \frac{[\langle \Theta_\phi(\mathbf{\Gamma})\rangle_\theta - \langle \Theta_\phi(\mathbf{\Gamma})\rangle_{\theta=0}]^2}{\mathrm{var}_{\theta=0}[\Theta_\phi(\mathbf{\Gamma})]\mathbb{D}_{\mathrm{PE}}[\mathcal{P}_\theta||\mathcal{P}_{\theta=0}]}, \tag{50}$$

which should be $\mathcal{E} \leqslant 1$ according to Eq. (48). Identifying $\mathbb{D}_{\mathrm{PE}}[\mathcal{P}_\theta||\mathcal{P}_{\theta=0}]$ as the cost, we can regard $\mathcal{E}$ as the efficiency of oscillators (larger $\mathcal{E}$ corresponds to higher efficiency). In Fig. 2(b), we plot the random realizations of $\mathcal{E}$ (circles) as a function of $\mathbb{D}_{\mathrm{PE}}[\mathcal{P}_\theta||\mathcal{P}_{\theta=0}]$. For comparison, we define

$$\mathcal{E}_F \equiv \frac{[\langle \Theta_\phi(\mathbf{\Gamma})\rangle_\theta - \langle \Theta_\phi(\mathbf{\Gamma})\rangle_{\theta=0}]^2}{\mathrm{var}_{\theta=0}[\Theta_\phi(\mathbf{\Gamma})]\theta^2\mathcal{I}(0)}, \tag{51}$$

which is based on the fluctuation-response inequality, and we plot $\mathcal{E}_F$ (triangles) as a function of $\theta^2\mathcal{I}(0)$. When the fluctuation-response inequality is satisfied, $\mathcal{E}_F \leqslant 1$. We observe that all circles are located below 1, indicating that the Chapman-Robbins inequality is satisfied for all realizations. Conversely, some triangles are above 1, which suggests violation of the fluctuation-response inequality. Although the linear response provides an exact response for linear systems, it is accurate only for sufficiently weak perturbations in the case of nonlinear systems. Thus, the fluctuation-response inequality is violated for nonlinear cases.

When the stochastic oscillator is approximated linearly around the deterministic orbit, Eq. (39) shows that the Pearson divergence increases exponentially as the noise intensity decreases. It is known that entropy production increases when noise intensity D decreases [51]. Therefore, lower noise intensity increases both entropy production and Pearson divergence. We numerically demonstrate a relation between $\mathbb{D}_{\mathrm{PE}}[\mathcal{P}_\theta||\mathcal{P}_{\theta=0}]$ and entropy production. Let ΔS_m be the medium entropy defined by

$$\Delta S_m \equiv \left\langle \frac{1}{D}\sum_{i=1}^{N}\int_0^T A_i(\boldsymbol{x})\circ \dot{x}_i dt\right\rangle. \tag{52}$$

When T is sufficiently large, the boundary term can be ignored and $\Delta S_m \simeq \Delta S_{\mathrm{tot}}$. Following the foregoing simulation procedure, we calculate ΔS_m and $\mathbb{D}_{\mathrm{PE}}[\mathcal{P}_\theta||\mathcal{P}_{\theta=0}]$ [parameter settings are shown in the caption for Fig. 2(c)]. In Fig. 2(c), we plot ΔS_m as a function of $\mathbb{D}_{\mathrm{PE}}[\mathcal{P}_\theta||\mathcal{P}_{\theta=0}]$ for fixed θ and T. We observe that ΔS_m increases when $\mathbb{D}_{\mathrm{PE}}[\mathcal{P}_\theta||\mathcal{P}_{\theta=0}]$ increases, showing that a larger Pearson divergence can be achieved for larger entropy production. Using simulations, Ref. [42] showed that both higher precision and higher sensitivity are achieved with higher entropy production, which is consistent with our results.

V. CONCLUSION

In this paper, we have applied information inequalities to systems described by Langevin equations to obtain inequalities in stochastic processes. We have identified that the thermodynamic uncertainty relation is a particular case of the Cramér-Rao inequality. Furthermore, we have applied the Chapman-Robbins inequality to the systems to show that the ratio between the variance and the sensitivity is bounded from below by the Pearson divergence. By bridging statistical inference theory and stochastic thermodynamic systems, this study can provide a useful basis for further developments with respect to thermodynamic bounds.

ACKNOWLEDGMENTS

This work was supported by Ministry of Education, Culture, Sports, Science and Technology (MEXT) KAKENHI Grants No. JP16K00325 and No. JP19K12153.

APPENDIX A: INFORMATION INEQUALITIES

Although information inequalities depict fundamental relations in statistics and machine learning, they are less known in physics. Therefore, we show their derivations for the readers' convenience [20–22].

Let X be a random variable, let $P_\theta(X)$ be a probability density function where θ is an arbitrary parameter, and let $\Theta(X)$ be an unbiased estimator of $\psi(\theta)$, which indicates $\psi(\theta) = \langle\Theta(X)\rangle_\theta$. Then the following relation holds:

$$\begin{aligned}\left\langle [\Theta(X)-\psi(\theta)]\left(\frac{\partial}{\partial\theta}\ln P_\theta(X)\right)\right\rangle_\theta &= \int dX[\Theta(X)-\psi(\theta)]\left(\frac{\partial}{\partial\theta}\ln P_\theta(X)\right)P_\theta(X)\\ &= \frac{\partial}{\partial\theta}\langle\Theta(X)\rangle_\theta.\end{aligned} \tag{A1}$$

Applying the Cauchy-Schwarz inequality to Eq. (A1), we obtain the Cramér-Rao inequality:

$$\begin{aligned}\mathrm{var}_\theta[\Theta(\Gamma)] \geqslant \frac{[\partial_\theta\langle\Theta(X)\rangle_\theta]^2}{\langle[\partial_\theta \ln P_\theta(X)]^2\rangle_\theta} &= \frac{[\partial_\theta\langle\Theta(X)\rangle_\theta]^2}{\left\langle -\partial_\theta^2 \ln P_\theta(X)\right\rangle_\theta}\\ &= \frac{[\partial_\theta\psi(\theta)]^2}{\mathcal{I}(\theta)},\end{aligned} \tag{A2}$$

where $\mathcal{I}(\theta)$ is the Fisher information. Its equality condition is obtained from that of the Cauchy-Schwarz inequality. From the left-hand side of Eq. (A1), if and only if the following condition is satisfied, the Cramér-Rao inequality holds with equality:

$$\frac{\partial}{\partial\theta}\ln P_\theta(X) = \mu[\Theta(X)-\psi(\theta)], \tag{A3}$$

where μ is a scaling parameter, which may depend on θ [i.e., $\mu = \mu(\theta)$].

The Chapman-Robbins inequality is a generalization of the Cramér-Rao inequality. For $\vartheta \neq \theta$, we notice that

$$\left\langle \frac{P_\vartheta(X)}{P_\theta(X)} - 1\right\rangle_\theta = \int dX[P_\vartheta(X) - P_\theta(X)] = 0. \tag{A4}$$

Then the following relation holds:

$$\left\langle [\Theta(X) - \psi(\theta)]\left(\frac{P_\vartheta(X)}{P_\theta(X)} - 1\right)\right\rangle_\theta$$
$$= \int dX[\Theta(X) - \psi(\theta)]\left(\frac{P_\vartheta(X)}{P_\theta(X)} - 1\right)P_\theta(X)$$
$$= \langle\Theta(X)\rangle_\vartheta - \langle\Theta(X)\rangle_\theta. \tag{A5}$$

Applying the Cauchy-Schwarz inequality to Eq. (A5), we obtain the Chapman-Robbins inequality:

$$\mathrm{var}_\theta[\Theta(X)] \geqslant \frac{[\langle\Theta(X)\rangle_\vartheta - \langle\Theta(X)\rangle_\theta]^2}{\left\langle\left(\frac{P_\vartheta(X)}{P_\theta(X)} - 1\right)^2\right\rangle_\theta}$$
$$= \frac{[\langle\Theta(X)\rangle_\vartheta - \langle\Theta(X)\rangle_\theta]^2}{\mathbb{D}_{\mathrm{PE}}[P_\vartheta||P_\theta]}, \tag{A6}$$

where $\mathbb{D}_{\mathrm{PE}}[P_\vartheta||P_\theta]$ is the Pearson divergence:

$$\mathbb{D}_{\mathrm{PE}}[P_\vartheta||P_\theta] \equiv \int dX\left(\frac{P_\vartheta(X)}{P_\theta(X)} - 1\right)^2 P_\theta(X)$$
$$= \int dX\left(\frac{P_\vartheta(X)}{P_\theta(X)}\right)^2 P_\theta(X) - 1. \tag{A7}$$

APPENDIX B: PATH INTEGRAL

Here we introduce the prepoint discretization procedure of the path integral according to Refs. [24,25]. We focus on the one-dimensional case because the calculations for the multidimensional case are laborious.

We consider the following Langevin equation (Ito interpretation):

$$\dot{x} = A_\theta(x,t) + \sqrt{2}C(x,t)\xi(t). \tag{B1}$$

We discretize time by dividing the interval $[0, T]$ into K equipartitioned intervals with time resolution Δt, where $T = K\Delta t$, $t^k \equiv k\Delta t$, and $x^k \equiv x(t^k)$ (superscripts denote points in a temporal sequence). Discretization of Eq. (B1) yields

$$x^{k+1} - x^k = \Delta t A_\theta(x^k, t^k) + \sqrt{2}C(x^k, t^k)\Delta w^k, \tag{B2}$$

where $\Delta w^k \equiv w^{k+1} - w^k = w(t^{k+1}) - w(t^k)$, with $w(t)$ depicting the Wiener process. Δw^k has the following properties:

$$\langle\Delta w^k\rangle = 0, \quad \langle\Delta w^k \Delta w^{k'}\rangle = \Delta t\delta_{kk'}. \tag{B3}$$

A stochastic trajectory $\mathcal{X} \equiv [x^1, x^2, \dots, x^K]$ is specified, given $\mathcal{W} \equiv [\Delta w^0, \Delta w^1, \dots, \Delta w^{K-1}]$ and x^0. The Wiener process Δw^k has the following probability density function:

$$P(\mathcal{W}) = \prod_{k=0}^{K-1} P(\Delta w^k) = \prod_{k=0}^{K-1} \frac{1}{\sqrt{2\pi\,\Delta t}} \exp\left[-\frac{(\Delta w^k)^2}{2\Delta t}\right]. \tag{B4}$$

Let us change variables in Eq. (B4) from $\mathcal{W} = [\Delta w^0, \Delta w^1, \dots, \Delta w^{K-1}]$ to $\mathcal{X} = [x^1, x^2, \dots, x^K]$. From Eq. (B2), the determinant of a Jacobian matrix is

$$\left|\frac{\partial(x^1, \dots, x^K)}{\partial(\Delta w^0, \dots, \Delta w^{K-1})}\right| = \prod_{k=0}^{K-1} \sqrt{2B(x^k, t^k)}, \tag{B5}$$

given that the determinant of a triangular matrix is a product of its diagonal elements, where we used $B(x,t) \equiv C(x,t)^2$. Using Eqs. (B2), (B4), and (B5), we obtain [52]

$$P_\theta(\mathcal{X}|x^0) = \left(\prod_{k=0}^{K-1} \frac{1}{\sqrt{4\pi\,\Delta t B(x^k, t^k)}}\right) \exp\left[-\frac{1}{4}\sum_{k=0}^{K-1} \Delta t\left\{\left(\frac{x^{k+1} - x^k}{\Delta t} - A_\theta(x^k, t^k)\right)^2 B(x^k, t^k)^{-1}\right\}\right]. \tag{B6}$$

In the limit $K \to \infty$, $\mathcal{X} \to \Gamma \equiv [x(t)]_{t=0}^{t=T}$, and we can write

$$\mathcal{P}_\theta(\Gamma|x^0) = \mathcal{N} \exp\left[-\frac{1}{4}\int_0^T dt[\dot{x} - A_\theta(x,t)]^2 B(x,t)^{-1}\right]. \tag{B7}$$

For an arbitrary function $g(x,t)$, the following relation holds:

$$\left\langle\int_0^T dt[\dot{x} - A_\theta(x,t)] \bullet g(x,t)\right\rangle_\theta$$
$$= \left\langle\sum_{k=0}^{K-1} \Delta t\left\{\left(\frac{x^{k+1} - x^k}{\Delta t} - A_\theta(x^k, t^k)\right)g(x^k, t^k)\right\}\right\rangle_\theta$$
$$= \left\langle\sum_{k=0}^{K-1} \sqrt{2}\Delta w^k C(x^k, t^k)g(x^k, t^k)\right\rangle_\theta = 0, \tag{B8}$$

where we used the property that x^k does not depend on Δw^k.

APPENDIX C: ITO AND STRATONOVICH CURRENTS

We present a relation between Ito and Stratonovich currents [cf. Eq. (C7)], both of which appear in the main text. Here, we explain this relation for the one-dimensional case, with the multidimensional generalization presented later.

Ito and its equivalent Stratonovich-Langevin equations are given by

$$dx = A(x)dt + \sqrt{2}C(x) \bullet dw, \tag{C1}$$

$$dx = [A(x) - C(x)C'(x)]dt + \sqrt{2}C(x) \circ dw, \tag{C2}$$

respectively. Let $\eta(x)$ be an arbitrary function of x. We are concerned with the relation between the following two terms:

$$U_I \equiv \int_0^T \eta(x) \bullet dw, \quad U_S \equiv \int_0^T \eta(x) \circ dw. \tag{C3}$$

Their discretized representations are

$$U_I = \sum_{k=0}^{K-1} \eta(x^k)\Delta w^k, \tag{C4}$$

$$U_S = \sum_{k=0}^{K-1} \eta\left(\frac{x^{k+1}+x^k}{2}\right)\Delta w^k. \tag{C5}$$

Applying a Taylor series expansion to Eq. (C5) and dropping terms whose orders are higher than $O(\Delta t)$, we obtain the following well-known relation [19]:

$$\begin{aligned} U_S &= \sum_{k=0}^{K-1} \left[\eta(x^k)\Delta w^k + \frac{\sqrt{2}}{2} C(x^k)\eta'(x^k)(\Delta w^k)^2 \right] \\ &= \int_0^T \eta(x) \bullet dw + \frac{\sqrt{2}}{2} \int_0^T C(x)\eta'(x) \bullet dw^2 \\ &= U_I + \frac{\sqrt{2}}{2} \int_0^T C(x)\eta'(x)dt, \end{aligned} \tag{C6}$$

where we used the relation $dw^2 = dt$ in the last line, which is valid for any nonanticipating function (see Chap. 4 in [19] for details).

Next, we consider Ito and Stratonovich currents of the following forms:

$$J_I \equiv \int_0^T \Lambda(x) \bullet \dot{x}\, dt, \quad J_S \equiv \int_0^T \Lambda(x) \circ \dot{x}\, dt, \tag{C7}$$

where $\Lambda(x)$ is a projection function. Their discretized representations are

$$J_I = \sum_{k=0}^{K-1} \Lambda(x^k)(x^{k+1} - x^k), \tag{C8}$$

$$J_S = \sum_{k=0}^{K-1} \Lambda\left(\frac{x^{k+1}+x^k}{2}\right)(x^{k+1} - x^k). \tag{C9}$$

Substituting Eqs. (C1) and (C2) into Eqs. (C8) and (C9), respectively, we obtain

$$J_I = \int_0^T \Lambda(x)A(x)dt + \sqrt{2}\int_0^T \Lambda(x)C(x) \bullet dw, \tag{C10}$$

$$\begin{aligned} J_S = &\int_0^T \Lambda(x)[A(x) - C(x)C'(x)]dt \\ &+ \sqrt{2}\int_0^T \Lambda(x)C(x) \circ dw. \end{aligned} \tag{C11}$$

By using Eq. (C6) [$\eta(x) = \Lambda(x)C(x)$], the following relation holds:

$$\int_0^T \Lambda(x)C(x) \circ dw = \int_0^T \Lambda(x)C(x) \bullet dw + \frac{\sqrt{2}}{2}\int_0^T C(x)\frac{d\Lambda(x)C(x)}{dx}dt. \tag{C12}$$

By substituting Eq. (C12) into Eq. (C11), a relation between the Stratonovich current J_S and the Ito current J_I is given by

$$\begin{aligned} J_S &= \int_0^T \Lambda(x)[A(x) - C(x)C'(x)]dt + \sqrt{2}\left[\int_0^T \Lambda(x)C(x)\bullet dw + \frac{\sqrt{2}}{2}\int_0^T C(x)\frac{d\Lambda(x)C(x)}{dx}dt\right] \\ &= \int_0^T \Lambda(x)A(x)dt + \sqrt{2}\int_0^T \Lambda(x)C(x) \bullet dw + \int_0^T \Lambda'(x)C(x)^2 dt \\ &= J_I + \int_0^T \Lambda'(x)C(x)^2 dt. \end{aligned} \tag{C13}$$

Therefore, we find the following relation:

$$\int_0^T \Lambda(x) \circ \dot{x}\, dt = \int_0^T \Lambda(x) \bullet \dot{x}\, dt + \int_0^T B(x)\frac{d\Lambda(x)}{dx}dt. \tag{C14}$$

For the multidimensional case $d\boldsymbol{x} = \boldsymbol{A}(\boldsymbol{x})dt + \sqrt{2}\boldsymbol{C}(\boldsymbol{x}) \bullet d\boldsymbol{w}$, we repeat the same calculations to obtain

$$\int_0^T \boldsymbol{\Lambda}(\boldsymbol{x})^\top \circ \dot{\boldsymbol{x}}\, dt = \int_0^T \boldsymbol{\Lambda}(\boldsymbol{x})^\top \bullet \dot{\boldsymbol{x}}\, dt + \int_0^T \mathrm{Tr}\left[\boldsymbol{B}(\boldsymbol{x})\frac{\partial \boldsymbol{\Lambda}(\boldsymbol{x})}{\partial \boldsymbol{x}}\right]dt, \tag{C15}$$

where $\boldsymbol{B}(\boldsymbol{x}) \equiv \boldsymbol{C}(\boldsymbol{x})\boldsymbol{C}(\boldsymbol{x})^\top$ and $[\partial\boldsymbol{\Lambda}(\boldsymbol{x})/\partial\boldsymbol{x}]_{ij} \equiv \partial\Lambda_i(\boldsymbol{x})/\partial x_j$ is a Jacobian matrix.

APPENDIX D: STEADY-STATE DISTRIBUTION OF PERIODIC SYSTEMS

Because $\Theta_{\mathrm{tot}}(\Gamma)$ is defined through the projection function $\Lambda(x) = 1/[P^{\mathrm{ss}}(x)B(x)]$ [Eq. (24)], we need to calculate the steady-state distribution $P^{\mathrm{ss}}(x)$, which can be found analytically as shown below.

Let $f(x)$ be a periodic potential $f(x) = f(x+2\pi)$ and let $\rho \geqslant 0$ be an external force. Suppose a system is given by

$$\dot{x} = \rho - f'(x) + \sqrt{2D}\xi(t), \tag{D1}$$

where $\rho = ab$ and $f(x) = -\frac{1}{2}[a+\sin(x)]^2 + b\cos(x)$ for Eq. (32). According to Ref. [18] (p. 287), the steady-state distribution of Eq. (D1) is

$$P^{\rm ss}(x) = \exp\left(-\frac{V(x)}{D}\right)\left[\mathfrak{N} - \frac{J^{\rm ss}}{D}\int_0^x \exp\left(\frac{V(x')}{D}dx'\right)\right], \tag{D2}$$

where $V(x) = f(x) - \rho x$, and $\mathfrak{N}$ is a normalization constant. $\mathfrak{N}$ and $J^{\rm ss}$ are determined by the two constraints

$$J^{\rm ss}\int_0^{2\pi} \exp\left(\frac{V(x)}{D}\right)dx = D\mathfrak{N}\left[1 - \exp\left(-\frac{2\pi\rho}{D}\right)\right], \tag{D3}$$

$$\int_0^{2\pi} P^{\rm ss}(x)dx = 1. \tag{D4}$$

Equations (D3) and (D4) are solved numerically to obtain $\mathfrak{N}$ and $J^{\rm ss}$.

APPENDIX E: PEARSON DIVERGENCE FOR THE LINEAR LANGEVIN PROCESS

The Pearson divergence is calculated analytically for a linear Langevin equation of Eq. (38) (the Ornstein-Uhlenbeck process). The discretized representation of Eq. (38) is

$$x^{k+1} - x^k = \Delta x^k = [-\alpha x^k + \theta u^k]\Delta t + \sqrt{2D}\Delta w^k, \tag{E1}$$

where $u^k \equiv u(t^k)$. The probability of the discretized trajectory $\mathcal{X} = [x^1, x^2, \dots, x^K]$ given x^0 is [cf. Eq. (B6)]

$$P_\theta(\mathcal{X}|x^0) = \frac{1}{(4\pi D\Delta t)^{K/2}}\exp\left[-\frac{1}{4D\Delta t}\sum_{k=0}^{K-1}(x^{k+1} - x^k - \{-\alpha x^k + \theta u^k\}\Delta t)^2\right]. \tag{E2}$$

The Pearson divergence between $P_\theta(\mathcal{X})$ and $P_{\theta=0}(\mathcal{X})$ is [cf. Eq. (A7)]

$$\mathbb{D}_{\rm PE}[P_\theta||P_{\theta=0}] = \int\prod_{k=0}^{K}dx^k\left[\frac{P_\theta(\mathcal{X}|x^0)P_\theta(x^0)}{P_{\theta=0}(\mathcal{X}|x^0)P_{\theta=0}(x^0)}\right]^2 P_{\theta=0}(\mathcal{X}|x^0)P_{\theta=0}(x^0) - 1. \tag{E3}$$

Let us introduce new variables y^k $(k = 1, 2, \dots, K)$, defined by

$$y^{k+1} \equiv x^{k+1} - x^k + \alpha x^k\Delta t. \tag{E4}$$

The determinant of a Jacobian is

$$\left|\frac{\partial(y^1, y^2, \dots, y^K)}{\partial(x^1, x^2, \dots, x^K)}\right| = 1. \tag{E5}$$

Using Eqs. (E4) and (E5), the probability density of $\mathcal{Y} \equiv [y^1, y^2, \dots, y^K]$ is

$$P_\theta(\mathcal{Y}|x^0) = \frac{1}{(4\pi D\Delta t)^{K/2}}\exp\left[-\frac{1}{4D\Delta t}\sum_{k=0}^{K-1}(y^{k+1} - \theta u^k\Delta t)^2\right]. \tag{E6}$$

Therefore, the Pearson divergence is given by

$$\begin{aligned}
\mathbb{D}_{\rm PE}[P_\theta||P_{\theta=0}] &= \int dx^0\int\prod_{k=1}^{K}dy^k\left[\frac{P_\theta(\mathcal{Y}|x^0)P_\theta(x^0)}{P_{\theta=0}(\mathcal{Y}|x^0)P_{\theta=0}(x^0)}\right]^2 P_{\theta=0}(\mathcal{Y}|x^0)P_{\theta=0}(x^0) - 1\\
&= -1 + \int dx^0\left(\frac{P_\theta(x^0)}{P_{\theta=0}(x^0)}\right)^2 P_{\theta=0}(x^0)\int\prod_{k=1}^{K}\frac{dy^k}{(4\pi D\Delta t)^{K/2}}\\
&\quad\times\exp\left[-\frac{1}{2D\Delta t}\sum_{k=0}^{K-1}(y^{k+1} - \theta u^k\Delta t)^2 + \frac{1}{2D\Delta t}\sum_{k=0}^{K-1}(y^{k+1})^2 - \frac{1}{4D\Delta t}\sum_{k=0}^{K-1}(y^{k+1})^2\right]\\
&= -1 + \exp\left[\sum_{k=0}^{K-1}\frac{(u^k)^2\Delta t}{2D}\theta^2\right]\int dx^0\left(\frac{P_\theta(x^0)}{P_{\theta=0}(x^0)}\right)^2 P_{\theta=0}(x^0).
\end{aligned} \tag{E7}$$

When the initial distributions are the same for $\theta \neq 0$ and $\theta = 0$, in the limit of $K \to \infty$, we obtain Eq. (39).

APPENDIX F: NUMERICAL SIMULATIONS

In the main text, we performed numerical simulations. In this Appendix, we explain how these simulations are implemented.

1. Monte Carlo simulations

To solve the Ito Langevin equations, we used Eq. (B2) (this method is known as the Euler-Maruyama scheme). Stratonovich-type currents are calculated by Eq. (C9).

We numerically calculate the Pearson divergence. We generate trajectories from the Langevin equations with parameter $\theta = 0$. Let N_S be the number of generated trajectories and $\boldsymbol{\mathcal{X}}_i$ be the ith realization of the trajectories. Let us consider $\boldsymbol{B}(\boldsymbol{x}, t) = \boldsymbol{D}$. The integral of the Pearson divergence is approximated by the following summation:

$$\mathbb{D}_{\mathrm{PE}}[\mathcal{P}_\theta || \mathcal{P}_{\theta=0}] \simeq \frac{1}{N_S} \sum_{i=1}^{N_S} \left(\frac{P_\theta(\boldsymbol{\mathcal{X}}_i)}{P_{\theta=0}(\boldsymbol{\mathcal{X}}_i)} - 1 \right)^2, \tag{F1}$$

where

$$P_\theta(\boldsymbol{\mathcal{X}}|\boldsymbol{x}_0) = \exp\left[-\frac{\Delta t}{4} \sum_{k=0}^{K-1} \sum_{i,j} \left(\frac{x_i^{k+1} - x_i^k}{\Delta t} - A_{\theta,i}(\boldsymbol{x}^k, t^k) \right) D_{ij}^{-1} \left(\frac{x_j^{k+1} - x_j^k}{\Delta t} - A_{\theta,j}(\boldsymbol{x}^k, t^k) \right) \right]. \tag{F2}$$

Here D_{ij}^{-1} is an i, jth element of $\boldsymbol{D}^{-1}$. We omitted $\mathcal{N}$ because it cancels out in Eq. (F1).

2. Phase definition

The phase for limit cycle oscillators can be defined [43]. For deterministic oscillators, we define phase ϕ on a closed orbit by

$$\frac{d\phi}{dt} = \Omega, \tag{F3}$$

where Ω is the angular frequency of the deterministic oscillation. We can expand the definition of the phase into an entire space $\boldsymbol{x}$, where $\boldsymbol{x}$ is an N-dimensional vector. Let $\boldsymbol{x}_a$ be a point on the closed orbit and $\boldsymbol{x}_b$ be a point that is *not* on the orbit. According to Eq. (F3), we can determine $\phi(\boldsymbol{x}_a)$. As the closed orbit is an attractor in limit cycle oscillators, $\boldsymbol{x}_b$ eventually converges to the closed orbit for $t \to \infty$. We let $\boldsymbol{x}_a$ and $\boldsymbol{x}_b$ time-evolve for the same duration. If the two points eventually converge to the same point on the closed orbit, then we assign $\phi(\boldsymbol{x}_b) = \phi(\boldsymbol{x}_a)$.

[1] F. Ritort, in *Advances in Chemical Physics*, edited by S. A. Rice (Wiley, Hoboken, NJ, 2008), Vol. 137, pp. 31–123.
[2] U. Seifert, Rep. Prog. Phys. **75**, 126001 (2012).
[3] C. Van den Broeck and M. Esposito, Physica A **418**, 6 (2015).
[4] J. M. R. Parrondo, J. M. Horowitz, and T. Sagawa, Nat. Phys. **11**, 131 (2015).
[5] A. C. Barato and U. Seifert, Phys. Rev. Lett. **114**, 158101 (2015).
[6] A. C. Barato and U. Seifert, J. Phys. Chem. B **119**, 6555 (2015).
[7] T. R. Gingrich, J. M. Horowitz, N. Perunov, and J. L. England, Phys. Rev. Lett. **116**, 120601 (2016).
[8] M. Polettini, A. Lazarescu, and M. Esposito, Phys. Rev. E **94**, 052104 (2016).
[9] P. Pietzonka, A. C. Barato, and U. Seifert, Phys. Rev. E **93**, 052145 (2016).
[10] J. M. Horowitz and T. R. Gingrich, Phys. Rev. E **96**, 020103(R) (2017).
[11] K. Proesmans and C. V. den Broeck, Europhys. Lett. **119**, 20001 (2017).
[12] P. Pietzonka, F. Ritort, and U. Seifert, Phys. Rev. E **96**, 012101 (2017).
[13] S. Pigolotti, I. Neri, E. Roldán, and F. Jülicher, Phys. Rev. Lett. **119**, 140604 (2017).
[14] A. Dechant and S.-i. Sasa, J. Stat. Mech: Theor. Expt. (2018) 063209.
[15] A. Dembo, Dept. Statistics, Stanford Univ., Stanford, CA, Tech. Rep **75** (1990).
[16] A. Dembo, T. M. Cover, and J. A. Thomas, IEEE Trans. Inf. Theor. **37**, 1501 (1991).
[17] A. Dechant and S.-i. Sasa, arXiv:1804.08250.
[18] H. Risken, *The Fokker-Planck Equation: Methods of Solution and Applications*, 2nd ed. (Springer, Berlin, Heidelberg, 1989).
[19] C. Gardiner, *Stochastic Methods: A Handbook for the Natural and Social Sciences* (Springer, Berlin, Heidelberg, 2009).
[20] A. Stuart, J. K. Ord, and S. Arnold, *Classical Inference and the Linear Model*, 6th ed., Kendall's Advanced Theory of Statistics Vol. 2A (Arnold, London, 1999).
[21] G. Casella and R. L. Berger, *Statistical Inference*, edited by C. Crockett (Duxbury, Belmont, California, 2001).
[22] E. L. Lehmann and G. Casella, *Theory of Point Estimation* (Springer, New York, 2003).
[23] V. Y. Chernyak, M. Chertkov, and C. Jarzynski, J. Stat. Mech: Theor. Expt. (2006) P08001.
[24] P. C. Bressloff, in *Stochastic Processes in Cell Biology*, edited by P. C. Bressloff (Springer International, Cham, Switzerland, 2014), pp. 577–617.
[25] H. S. Wio, *Path Integrals for Stochastic Processes: An Introduction* (World Scientific, Singapore, 2013).
[26] Y. Tang, R. Yuan, and P. Ao, J. Chem. Phys. **141**, 044125 (2014).

[27] A. B. Adib, J. Phys. Chem. B **112**, 5910 (2008).
[28] L. P. Fischer, P. Pietzonka, and U. Seifert, Phys. Rev. E **97**, 022143 (2018).
[29] T. V. Vu and Y. Hasegawa, arXiv:1901.05715.
[30] Terms with and without $\dot{x}$ cannot cancel each other. Suppose canceling both terms is possible, i.e., there exist functions $h_1(x)$ and $h_2(x)$ such that $\int_0^T h_1(x) \circ \dot{x}\, dt = \int_0^T h_2(x)dt$ for all Γ [we exclude a trivial case $h_1(x) = h_2(x) = 0$]. Such a condition yields $h_1(x) \circ \dot{x} - h_2(x) = 0$, indicating that we can determine $\dot{x}$ deterministically given x. However, due to the stochastic nature of Langevin equations, such a condition is not possible.
[31] K. Macieszczak, K. Brandner, and J. P. Garrahan, Phys. Rev. Lett. **121**, 130601 (2018).
[32] A. C. Barato, D. Hartich, and U. Seifert, New J. Phys. **16**, 103024 (2014).
[33] D. Hartich, A. C. Barato, and U. Seifert, Phys. Rev. E **93**, 022116 (2016).
[34] S. Goldt and U. Seifert, Phys. Rev. Lett. **118**, 010601 (2017).
[35] K. T. Moortgat, T. H. Bullock, and T. J. Sejnowski, J. Neurophysiol. **83**, 971 (2000).
[36] A. T. Winfree, *The Geometry of Biological Time*, 2nd ed. (Springer, New York, 2001).
[37] D. J. Needleman, P. H. E. Tiesinga, and T. J. Sejnowski, Physica D **155**, 324 (2001).
[38] H. Kori, Y. Kawamura, and N. Masuda, J. Theor. Biol. **297**, 61 (2012).
[39] Y. Hasegawa, Phys. Rev. E **98**, 032405 (2018).
[40] Y. Hasegawa and M. Arita, J. R. Soc. Interface **11**, 20131018 (2014).
[41] Y. Hasegawa and M. Arita, Phys. Rev. Lett. **113**, 108101 (2014).
[42] C. Fei, Y. Cao, Q. Ouyang, and Y. Tu, Nat. Commun. **9**, 1434 (2018).
[43] Y. Kuramoto, *Chemical Oscillations, Waves, and Turbulence* (Dover, Mineola, NY, 2003).
[44] M. A. Schwemmer and T. J. Lewis, in *Phase Response Curves in Neuroscience: Theory, Experiment, and Analysis*, edited by N. Schultheiss, A. Prinz, and R. Butera (Springer, New York, 2012).
[45] E. H. Trepagnier, C. Jarzynski, F. Ritort, G. E. Crooks, C. J. Bustamante, and J. Liphardt, Proc. Natl. Acad. Sci. (USA) **101**, 15038 (2004).
[46] D. Collin, F. Ritort, C. Jarzynski, S. B. Smith, I. Tinoco Jr., and C. Bustamante, Nature (London) **437**, 231 (2005).
[47] S. Schuler, T. Speck, C. Tietz, J. Wrachtrup, and U. Seifert, Phys. Rev. Lett. **94**, 180602 (2005).
[48] C. Tietz, S. Schuler, T. Speck, U. Seifert, and J. Wrachtrup, Phys. Rev. Lett. **97**, 050602 (2006).
[49] D. Andrieux, P. Gaspard, S. Ciliberto, N. Garnier, S. Joubaud, and A. Petrosyan, Phys. Rev. Lett. **98**, 150601 (2007).
[50] B. van der Pol, London, Edinburgh, Dublin Philos. Mag. J. Sci. **2**, 978 (1926).
[51] T. Tomé, Braz. J. Phys. **36**, 1285 (2006).
[52] Instead of using the Jacobian, Ref. [24] used a δ function to change the variables. The two approaches are equivalent.