Thermodynamics of collective enhancement of precision

Yoshihiko Hasegawa*

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

(Received 8 February 2018; revised manuscript received 15 April 2018; published 7 September 2018)

The circadian oscillator exhibits remarkably high temporal precision, despite its exposure to several fluctuations. The central mechanism that protects the oscillator from fluctuations is a collective enhancement of precision where a population of coupled oscillators displays higher temporal precision than that achieved without coupling. Since coupling is essentially information exchange between oscillators, we herein investigate the relation between the temporal precision and the information flow between oscillators in the linearized Kuramoto model by using stochastic thermodynamics. For general coupling, we find that the temporal precision is bounded from below by the information flow. We generalize the model to incorporate a time-delayed coupling and demonstrate that the same relation also holds for the time-delayed case. Furthermore, the temporal precision is demonstrated to be improved in the presence of the time delay, and we show that the increased information flow is responsible for the time-delay-induced precision improvement.

DOI: 10.1103/PhysRevE.98.032405

I. INTRODUCTION

The circadian oscillator is a biological clock that is prevalent in biological organisms ranging from bacteria to humans [1,2]. Since the circadian oscillator is induced by biochemical reactions on a cellular level, it is subject to several fluctuations. Despite such stochasticity, the circadian oscillator is known to exhibit incredibly high temporal precision [3] where the temporal standard deviation is approximately 3–5 min in 24 h [4]. In a single oscillator level, a phase response curve, which quantifies the dynamics of the oscillator, is optimized so as to achieve high temporal precision [5,6]. Still, this single oscillator level improvement does not seem to fully explain the above mentioned temporal precision. Another significant precision improvement arises at the population level [7]. A previous study [8] experimentally showed that when oscillators are coupled, each oscillator exhibits higher precision than that realizable without coupling. This phenomenon is referred to as *collective enhancement of precision* (CEP) [7,9,10]. Since mammalian circadian oscillation occurs primarily in the suprachiasmatic nucleus (SCN), which is a collection of 10^4 cells, CEP is considered to be responsible for the high temporal precision in the circadian oscillator [7].

CEP is induced by coupling, which is essentially the exchange of information. Information enables the violation of the second law of thermodynamics through the Maxwell demon. High precision is achieved at the cost of energy consumption, which is known as the thermodynamic uncertainty relation [11–15]. When we ignore the coupling between oscillators, CEP appears to achieve thermodynamically impossible precision, motivating us to analyze CEP from the viewpoint of information analogous to the Maxwell demon (Fig. 1). Using stochastic thermodynamics [16–18], we find that the temporal variance of coupled oscillators is bounded from below by the information flow conferred by coupling. Furthermore, we generalize the obtained relation to incorporate time-delayed coupling and show that the same relation also holds for the time-delayed case. Intriguingly, the temporal precision in the time-delayed case is improved as compared to the nondelayed case, and we can ascribe this time-delay-induced improvement to the increase in the information flow.

II. MODEL

We consider N coupled identical oscillators, which are entrained by an external periodic signal with the angular frequency Ω and are subject to noise [Fig. 2(a)]. Let ϕ_i be the phase of the ith oscillator, and let ω be the angular frequency of oscillators. Then, the phase dynamics of each oscillator can be described by the following forced Kuramoto model [9,19,20]:

$$\dot{\phi}_i = \omega + L\,\sin(\Omega t - \phi_i) + f_i(t) + \xi_i(t), \tag{1}$$

$$f_i(t) = \sum_{j\in V, j\neq i} K_{ij}\sin(\phi_j - \phi_i), \tag{2}$$

where $V = \{1, 2, \ldots, N\}$ and we aggregate all coupling effects, which act on ϕ_i, into a single coupling variable $f_i(t)$. The sinusoidal term in Eq. (1) represents the periodic signal, and $L > 0$ denotes its strength. A sinusoidal function in Eq. (2) is a coupling function that represents the interaction between oscillators, where K_{ij} is the coupling strength between the ith and the jth oscillators. Note that we allow both symmetric ($K_{ij} = K_{ji}$) and asymmetric ($K_{ij} \neq K_{ji}$) couplings. The phase reduction with which the phase equation of Eq. (1) is derived assumes Langevin equations to be interpreted in the Stratonovich sense. Moreover, $\xi_i(t)$ is zero-mean white Gaussian noise with the correlation $\langle \xi_i(t)\xi_{i'}(t')\rangle =$

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

 ©2018 American Physical Society

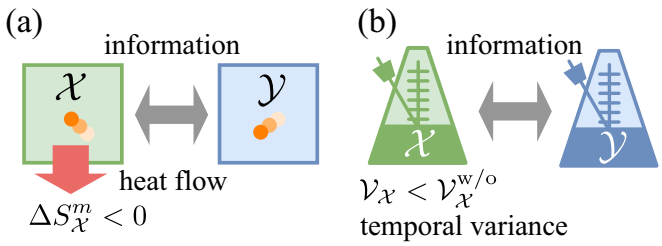

FIG. 1. Analogy between (a) the Maxwell demon and (b) the CEP. (a) When a system $\mathcal{X}$ is coupled to another system $\mathcal{Y}$, $\mathcal{X}$'s medium entropy $\Delta S^m_{\mathcal{X}}$, which is the heat from $\mathcal{X}$ to the medium (divided by the temperature), can be negative. (b) When an oscillator $\mathcal{X}$ is coupled to another oscillator $\mathcal{Y}$, $\mathcal{V}_{\mathcal{X}}$ (the temporal variance with coupling) can be smaller than $\mathcal{V}^{\mathrm{w/o}}_{\mathcal{X}}$(the temporal variance without coupling).

$2D\delta_{ii'}\delta(t-t')$ (D is the noise strength). The derivation of Eq. (1) can be found in Refs. [21,22]. We hereinafter assume that all of the oscillators are synchronized to the periodic signal. Although CEP is often studied for $L = 0$, we assume $L > 0$, which enables an approximation of the synchronized behavior as a nonequilibrium steady state.

Since there are efferent projections from the SCN to other neurons, the output of the SCN has been suggested to be generated by the average of a subset of oscillators and not by the average of all oscillators [10]. Therefore, we observe the average phase $\Phi(t) = (1/N_O)\sum_{i\in V_O}\phi_i(t)$, where V_O is a set of observable oscillators and N_O is its element number [Fig. 2(b)] [10]. We also define V_L and N_L to be a set of latent variables (the output of which is not observed) and its element number, respectively. Every oscillator belongs to the whole set $V = V_O \cup V_L$, and V_O and V_L are disjoint, i.e., $V_O \cap V_L = \varnothing$. See Fig. 2(b). Defining a relative phase x_i observed from a rotating frame as $x_i = \phi_i - \Omega t$, the synchronizing assumption is that x_i is well concentrated around 0. This condition can usually be satisfied when $|\omega - \Omega|/L$ is sufficiently small and D is sufficiently weak. Let $\mu_A = \langle A\rangle$, $\mathcal{V}_A = \langle A^2\rangle - \langle A\rangle^2$, and $\mathcal{C}_{AB} = \langle AB\rangle - \langle A\rangle\langle B\rangle$, where A and B are arbitrary random variables, and let $X(t) = \Phi(t) - \Omega t = (1/N_O)\sum_{i\in V_O} x_i(t)$, which is the average phase

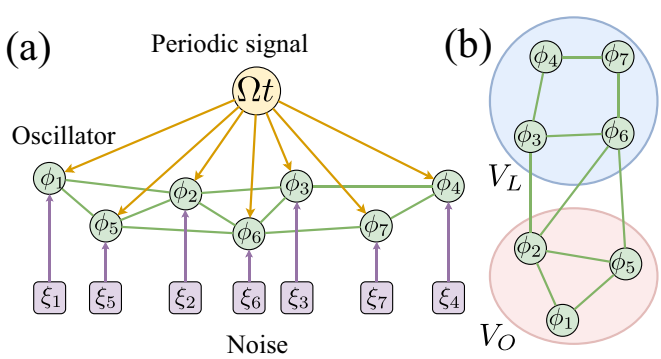

FIG. 2. Coupled oscillators model. (a) Adopted coupled oscillator model for $N = 7$ ($V = \{1, 2, \ldots, 7\}$). Each oscillator is entrained by a periodic signal with the angular frequency Ω and is subject to noise $\xi_i(t)$. (b) Observable set V_O and latent set V_L, where $V_O = \{1, 2, 5\}$ and $V_L = \{3, 4, 6, 7\}$ in this example.

seen from a rotating frame. The temporal variance of the average phase $\Phi(t)$ is given by the variance $\mathcal{V}_X$. Since each phase x_i is well concentrated around 0 (we assumed that all of the oscillators are synchronized with the periodic signal), we apply a Taylor series expansion to Eqs. (1) and (2) to obtain

$$\dot{X} = -L(X-c) + F(t) + \Xi_O(t), \tag{3}$$

$$\dot{x}_i = -L(x_i - c) + \sum_{j\in V,\, j\neq i} K_{ij}(x_j - x_i) + \xi_i(t)(i \in V), \tag{4}$$

where $F(t)$ is defined by

$$F(t) = \frac{1}{N_O}\sum_{i\in V_O} f_i(t) = \frac{1}{N_O}\sum_{i\in V_O}\sum_{j\in V,\, j\neq i} K_{ij}(x_j - x_i), \tag{5}$$

and $c = (\omega - \Omega)/L$ and $\Xi_O(t) = (1/N_O)\sum_{i\in V_O}\xi_i(t)$ [$\langle\Xi_O(t)\Xi_O(t')\rangle = (2D/N_O)\delta(t-t')$, $\langle\Xi_O(t)\xi_i(t')\rangle = (2D/N_O)\delta(t-t')$ for $i \in V_O$, and $\langle\Xi_O(t)\xi_i(t')\rangle = 0$ for $i \in V_L$]. Here, $F(t)$ includes all of the coupling effects that act on x_i for $i \in V_O$. Therefore, the information flow between X and F is of interest. We calculate the information flow between X and F based on Refs. [23,24], which is defined as

$$\begin{aligned}\dot{\mathcal{I}}_X(X;F) &= -\int J_X(X,F)\frac{\partial}{\partial X}\ln P(F|X)dX\,dF \\ &= -\frac{1}{dt}\left\langle \ln\left(\frac{P[F(t)|X(t+dt)]}{P[F(t)|X(t)]}\right)\right\rangle,\end{aligned} \tag{6}$$

where the sign is reversed compared to the original definition. This quantity is equivalent to the learning rate [25–27]. In order to calculate the information flow, we introduce a hypothetical time delay $h > 0$ in Eq. (5), which yields $F(t) = (1/N_O)\sum_{i\in V_O}\sum_{j\in V,\, j\neq i} K_{ij}[x_j(t-h) - x_i(t-h)]$. The introduction of time-delay h is a mathematical trick, and, when calculating several quantities, we set $h \to 0$ afterwards. Using a Taylor expansion for a sufficiently small h, we obtain

$$\dot{F} = h^{-1}\left(-F + \frac{1}{N_O}\sum_{i\in V_O}\sum_{j\in V,\, j\neq i} K_{ij}(x_j - x_i)\right). \tag{7}$$

Then, we can treat Eqs. (3), (4), and (7) as coupled Langevin equations. Let $P(X, \boldsymbol{x}, F)$ be a time-dependent probability density function of X, $\boldsymbol{x} = [x_1, \ldots, x_N]$, and F [for notational convenience, we dropped time t in the argument in $P(X, \boldsymbol{x}, F)$]. Due to the linearity of Eqs. (3), (4), and (7), the steady-state distribution is a Gaussian distribution with mean $[\mu_X, \mu_{x_1}, \ldots, \mu_{x_N}, \mu_F] = [c, c, \ldots, c, 0]$. The corresponding Fokker-Planck equation (FPE) with respect to $P(X, \boldsymbol{x}, F)$ can be represented as

$$\partial_t P(X, \boldsymbol{x}, F) = \mathbb{L}(X, \boldsymbol{x}, F)P(X, \boldsymbol{x}, F), \tag{8}$$

where $\mathbb{L}(X, \boldsymbol{x}, F)$ is a FPE operator, which is defined by

$$\begin{aligned}\mathbb{L}(X, \boldsymbol{x}, F) = &-\frac{\partial}{\partial X}[-L(X-c) + F] + \frac{D}{N_O}\frac{\partial^2}{\partial X^2} \\ &-\frac{\partial}{\partial F}\frac{1}{h}\left\{-F + \frac{1}{N_O}\sum_{i\in V_O}\sum_{j\in V,\, j\neq i} K_{ij}(x_j - x_i)\right\}\end{aligned}$$

$$-\sum_{i\in V}\frac{\partial}{\partial x_i}\left\{-L(x_i-c)+\sum_{j\in V,\,j\neq i}K_{ij}(x_j-x_i)\right\}$$
$$+D\sum_{i\in V}\frac{\partial^2}{\partial x_i^2}+\frac{2D}{N_O}\sum_{i\in V_O}\frac{\partial^2}{\partial x_i\partial X}. \tag{9}$$

We integrate out $\boldsymbol{x}$ in Eqs. (8) and (9) to obtain

$$\partial_t P(X,F)=-\partial_X J_X(X,F)-\partial_F J_F(X,F), \tag{10}$$

where $J_X(X,F)$ and $J_F(X,F)$ are probability currents, which are defined as

$$J_X(X,F)=\{-L(X-c)+F\}P-\frac{D}{N_O}\frac{\partial P}{\partial X}, \tag{11}$$

$$J_F(X,F)=h^{-1}\{-F+\mathcal{H}(X,F)\}P. \tag{12}$$

in which $P=P(X,F)$ and $\mathcal{H}(X,F)=\int(1/N_O)\sum_{i\in V_O}\sum_{j\in V,j\neq i}K_{ij}(x_j-x_i)P(\boldsymbol{x}|X,F)d\boldsymbol{x}$. Although the FPE operator $\mathbb{L}(X,\boldsymbol{x},F)$ of Eq. (9) contains cross terms, such as $\partial_X\partial_{x_i}$, due to nonvanishing correlation $\langle\Xi_O(t)\xi_i(t)\rangle\neq 0$ for $i\in V_O$, they disappear when integrating out $\boldsymbol{x}$. From Eq. (10), we can obtain the generalized second law following the conventional procedure in stochastic thermodynamics [17,18,23,24]. We first consider the Shannon entropy $S(X,F)=-\int P(X,F)\ln P(X,F)dX\,dF$. The time derivative is $\dot{S}(X,F)=\dot{S}_X(X,F)+\dot{S}_F(X,F)$, where $\dot{S}_X(X,F)=-\int J_X(X,F)\partial_X\ln P(X,F)dX\,dF$, and $\dot{S}_F(X,F)=-\int J_F(X,F)\partial_F\ln P(X,F)dX\,dF$. We can divide $J_X(X,F)$ of Eq. (11) into two contributions $J_X^R(X,F)$ and $J_X^I(X,F)$, which are defined as $J_X^I(X,F)=-L(X-c)P-(D/N_O)\partial_X P$ and $J_X^R=FP$, respectively. In feedback cooling of Brownian particles, J_X^I and J_X^R correspond to the irreversible and reversible portions of the probability current [23,28]. Then, $\dot{S}_X(X,F)$ is calculated as

$$\dot{S}_X(X,F)=\int\left(\frac{N_O\left(J_X^I\right)^2}{DP}+\frac{N_O J_X^I L(X-c)}{D}\right)dX\,dF, \tag{13}$$

where $J_X^I=J_X^I(X,F)$ and $P=P(X,F)$. In Eq. (13), the first term is non-negative, and the second term is calculated as $(1/D)\int N_O J_X^I(X,F)L(X-c)dX\,dF=-(N_O L^2/D)\langle(X-c)^2\rangle+L$. According to Ref. [24], we have $\dot{\mathcal{I}}_X(X;F)=-\dot{S}(X)+\dot{S}_X(X,F)$, where $S(X)=-\int P(X)\ln P(X)dX$. Under the synchronization assumption, X is in a steady state, and thus $\dot{S}(X)=0$. Combining these representations, the variance $\mathcal{V}_X=\langle(X-c)^2\rangle$ satisfies

$$\mathcal{V}_X\geqslant\mathcal{V}_X^{\mathrm{LB}}=\frac{D}{L^2N_O}[L-\dot{\mathcal{I}}_X(X;F)], \tag{14}$$

where $\mathcal{V}_X^{\mathrm{LB}}$ is the lower bound of the variance $\mathcal{V}_X$. Equation (14) is our main result. When there is no coupling, $\mathcal{V}_X=D/(LN_O)$. Therefore, without information flow, the variance does not improve beyond $D/(LN_O)$, and Eq. (14) shows that the CEP is induced by the information flow due to coupling between oscillators. Equation (14) indicates that the CEP has the same mathematical property as the feedback cooling by the Maxwell demon [29,30] where the feedback reduces the variance. The angular frequency Ω affects the inequality of Eq. (14) through $c=(\omega-\Omega)/L$, showing that the inequality does not have Ω dependence when $\omega=\Omega$. Note that we cannot reach the desired inequality of Eq. (14) by using the positivity of $\int J_X(X,F)^2/P(X,F)dX\,dF$ in the conventional total entropy production rate [31]. The transfer entropy rate [32] is a similar quantity to the information flow. It has been shown that the transfer entropy rate from X to F is greater than or equal to the information flow [24,33,34]. Therefore, the same inequality as Eq. (14) holds for the transfer entropy rate but the bound becomes weaker than the information flow case.

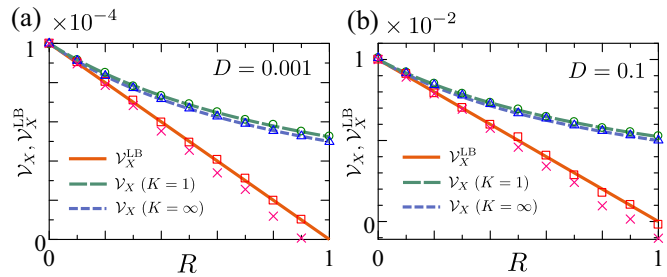

FIG. 3. Variance $\mathcal{V}_X$ and its lower bound $\mathcal{V}_X^{\mathrm{LB}}$ as a function of $R=N_L/N_O$ for (a) $D=0.001$ and (b) $D=0.1$. The solid, long-dashed, and dashed lines denote $\mathcal{V}_X^{\mathrm{LB}}$, $\mathcal{V}_X$ with $K=1$, and $\mathcal{V}_X$ with $K=\infty$, respectively, obtained analytically. The squares, crosses, circles, and triangles denote $\mathcal{V}_X^{\mathrm{LB}}$ with $K=1$, $\mathcal{V}_X^{\mathrm{LB}}$ with $K=10$, $\mathcal{V}_X$ with $K=1$, and $\mathcal{V}_X$ with $K=10$, respectively, obtained by Monte Carlo simulation. The other parameters are $N_O=10$, $L=1$, $\omega=1$, and $\Omega=1$.

III. EXAMPLES

A. Globally coupled model

We calculate $\mathcal{V}_X$ and $\mathcal{V}_X^{\mathrm{LB}}$ analytically for a global uniform coupling case $K_{ij}=K$. Defining $Y=(1/N_L)\sum_{i\in V_L}x_i$, we have

$$\dot{Y}=-L(Y-c)-F/R+\Xi_L(t), \tag{15}$$

$$\dot{F}=h^{-1}\{-F+KN_L(Y-X)\}, \tag{16}$$

where $R=N_L/N_O$ and $\Xi_L(t)=(1/N_L)\sum_{i\in V_L}\xi_i(t)$ [$\langle\Xi_L(t)\Xi_L(t')\rangle=(2D/N_L)\delta(t-t')$ and $\langle\Xi_L(t)\Xi_O(t')\rangle=0$]. Then, Eqs. (3), (15), and (16) constitute coupled Langevin equations of the global uniform coupling model. Taking a limit $h\to 0$ for which the hypothetical time delay in F vanishes, we can calculate $\mathcal{V}_X$, $\mathcal{V}_F$, and $\mathcal{C}_{XF}$ analytically due to the linearity of the coupled equations. Then, the information flow $\dot{\mathcal{I}}_X(X;F)$ can be calculated as follows (Appendix A):

$$\dot{\mathcal{I}}_X(X;F)=LR. \tag{17}$$

We plot $\mathcal{V}_X^{\mathrm{LB}}$ and $\mathcal{V}_X$ of the global uniform coupling case for $D=0.001$ [Fig. 3(a)] and $D=0.1$ [Fig. 3(b)]. The other parameters are shown in the caption of Fig. 3. In Fig. 3, the solid line denotes $\mathcal{V}_X^{\mathrm{LB}}$, and the long-dashed and dashed lines show $\mathcal{V}_X$ for $K=1$ and $K=\infty$, respectively. We also carried out a Monte Carlo simulation (for details, please see

Appendix D), where $\mathcal{V}_X$'s are denoted by circles ($K = 1$) and triangles ($K = 10$, which is intended to simulate $K = \infty$), and $\mathcal{V}_X^{\rm LB}$'s are denoted by squares ($K = 1$) and crosses ($K = 10$). Note that the Monte Carlo simulation is carried out for equations that do not use the Taylor approximation. Apparently, $\mathcal{V}_X \geqslant \mathcal{V}_X^{\rm LB}$ for $R \geqslant 0$, which verifies the inequality relation. For $R > 1$, $\mathcal{V}_X^{\rm LB}$ becomes negative, and the lower bound does not have a practical meaning. The difference is $\mathcal{V}_X - \mathcal{V}_X^{\rm LB} = L^{-2}\int J_X^I(X,F)^2/P(X,F)dX\,dF$, and the probability current $|J_X^I|$ becomes larger for larger R, which is responsible for the gap between $\mathcal{V}_X^{\rm LB}$ and $\mathcal{V}_X$. The inequality saturates when $F \to 0$, i.e., there is no coupling. Figure 3 shows that the Monte Carlo results agree with the analytical calculations, even for a relatively large noise intensity of $D = 0.1$ [Fig. 3(b)], and this demonstrates the validity of the synchronization assumption and the Taylor approximation. When R is sufficiently small, the variance $\mathcal{V}_X$ is close to $\mathcal{V}_X^{\rm LB}$, which shows that oscillators maximally exploit the information flow to improve the variance. As shown in the next model, higher temporal precision can be achieved if the information flow $\dot{\mathcal{I}}_X(X;F)$ can be increased.

B. Time-delayed model

Next, we generalize the model to include time delay in the coupling, which exists in reality and has been reported to be able to cause different dynamics in the Kuramoto model [35]. In the context of feedback cooling, a time-delayed case was investigated in Refs. [36,37]. Let $\tau \geqslant 0$ be the time delay of the coupling. Following Ref. [35], we incorporate the time delay into Eq. (1) using the following $f_i(t)$ instead of Eq. (2):

$$f_i(t) = \sum_{j\in V,\, j\neq i} K_{ij}\sin(\phi_j^\tau - \phi_i), \tag{18}$$

where superscript τ hereinafter denotes a time-delayed variable, i.e., $\phi_j^\tau(t) = \phi_j(t-\tau)$. Similar to the nondelayed case, we introduce a relative phase $x_i = \phi_i - \Omega t$. For the nondelayed case, $\mu_X = 0$ when $\omega = \Omega$. However, for the time-delayed case, μ_X deviates from 0, even when $\omega = \Omega$. Therefore, in order for the Taylor approximation to yield reliable results (i.e., x_i should be well concentrated around 0), $\Omega\tau$ should be sufficiently small. In such cases, we can use the Taylor approximation to obtain $F = (1/N_O)\sum_{i\in V_O} f_i = (1/N_O)\sum_{i\in V_O}\sum_{j\in V,\, j\neq i} K_{ij}(x_j^\tau - x_i - \Omega\tau)$, and the introduction of the hypothetical delay h leads us to $\dot{F} = (1/h)\{-F + (1/N_O)\sum_{i\in V_O}\sum_{j\in V,\, j\neq i} K_{ij}(x_j^\tau - x_i - \Omega\tau)\}$. Following Ref. [38], an effective FPE, which does not include time-delayed variables, can be derived for the time-delayed case. Let $\boldsymbol{x}^\tau = [x_1^\tau, \ldots, x_N^\tau]$ and $P(X,\boldsymbol{x},F;\boldsymbol{x}^\tau)$ be the probability density of X, $\boldsymbol{x}$, and F at time t, and $\boldsymbol{x}^\tau$ at time $t-\tau$. Calculating the marginal distribution $P(X,F) = \int P(X,\boldsymbol{x},F;\boldsymbol{x}^\tau)d\boldsymbol{x}\,d\boldsymbol{x}^\tau$, the FPE for $P(X,F)$ is similar to that given by Eqs. (10)–(12). Therefore, the inequality as in Eq. (14) holds for the time-delayed case (Appendix B).

C. Globally coupled time-delayed model

Next, we calculate $\mathcal{V}_X$ and $\mathcal{V}_X^{\rm LB}$ for a globally coupled time-delayed case. For this case, we obtained the following equations:

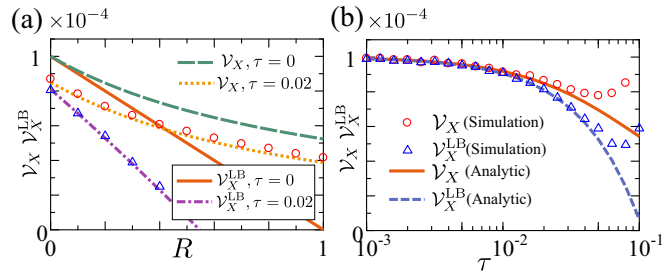

FIG. 4. Variance $\mathcal{V}_X$ and its lower bound $\mathcal{V}_X^{\rm LB}$ for a time-delayed case. (a) $\mathcal{V}_X$ and $\mathcal{V}_X^{\rm LB}$ as a function of $R = N_L/N_O$ for the time-delayed case ($\tau = 0.02$) and $K = 1$. The dotted and dot-dashed lines denote $\mathcal{V}_X$ and $\mathcal{V}_X^{\rm LB}$ for $\tau = 0.02$, respectively, obtained analytically. The solid and long-dashed lines denote $\mathcal{V}_X^{\rm LB}$ and $\mathcal{V}_X$, respectively, for the nondelayed case ($\tau = 0$) and are identical to those in Fig. 3(a). The circles and triangles denote $\mathcal{V}_X$ and $\mathcal{V}_X^{\rm LB}$, respectively, for $\tau = 0.02$ obtained by Monte Carlo simulations. (b) Variance $\mathcal{V}_X$ and its lower bound $\mathcal{V}_X^{\rm LB}$ as a function of the time-delay τ for $K = 1$ and $N = N_O = 10$. The solid and dashed lines denote $\mathcal{V}_X$ and $\mathcal{V}_X^{\rm LB}$, respectively, obtained analytically. The circles and triangles denote $\mathcal{V}_X$ and $\mathcal{V}_X^{\rm LB}$, respectively, obtained by Monte Carlo simulations. In (a) and (b), the other unspecified parameters are the same as those in Fig. 3(a).

$$\dot{X} = -LX + F + \Xi_O(t), \tag{19}$$

$$\dot{Y} = K\{(N_L - 1)Y^\tau - (N-1)Y + N_O X^\tau\} - LY + \Xi_L(t), \tag{20}$$

$$F = K\{(N_O - 1)X^\tau - (N-1)X + N_L Y^\tau\}, \tag{21}$$

where we have applied a parallel translation so that $\mu_X = \mu_Y = \mu_F = 0$ holds for the steady state. Since the system is the steady state under the synchronization assumption, we can use the Fourier transform to calculate the covariance matrix [38,39]. Let $\mathcal{F}[\cdots]$ be the Fourier transform operator, and let ν be the Fourier variable. We define $\mathcal{S}_F(\nu) = \mathcal{F}[\langle F(t)F(0)\rangle]$, $\mathcal{S}_X(\nu) = \mathcal{F}[\langle X(t)X(0)\rangle]$, and $\mathcal{S}_{XF}(\nu) = \mathcal{F}[\langle X(t)F(0)\rangle]$. By virtue of the Wiener-Khinchin theorem, $\mathcal{V}_X$, $\mathcal{V}_F$, and $\mathcal{C}_{XF}$ can be obtained by $\mathcal{V}_X = (2\pi)^{-1}\int_{-\infty}^{\infty}\mathcal{S}_X(\nu)d\nu$, $\mathcal{V}_F = (2\pi)^{-1}\int_{-\infty}^{\infty}\mathcal{S}_F(\nu)d\nu$, and $\mathcal{C}_{XF} = (2\pi)^{-1}\int_{-\infty}^{\infty}\mathcal{S}_{XF}(\nu)d\nu$ (Appendix C). Since the integration is complicated, we evaluate these integrals numerically.

Figure 4(a) shows $\mathcal{V}_X$ and $\mathcal{V}_X^{\rm LB}$ as a function of R for time delay $\tau = 0.02$ and $K = 1$. The other parameters are the same as in Fig. 3(a). In Fig. 4(a), the dotted and dot-dashed lines denote $\mathcal{V}_X$ and $\mathcal{V}_X^{\rm LB}$ for $\tau = 0.02$, respectively. The solid and long-dashed lines represent $\mathcal{V}_X^{\rm LB}$ and $\mathcal{V}_X$ for the nondelayed case and are identical to those in Fig. 3(a) (shown for comparison). We also carried out Monte Carlo simulations for $\tau = 0.02$, and these data are shown by circles and triangles, which correspond to $\mathcal{V}_X$ and $\mathcal{V}_X^{\rm LB}$, respectively. We again confirm that the inequality of Eq. (14) also holds for the delayed case. Figure 4(a) shows that $\mathcal{V}_X$ for $\tau = 0.02$ is lower than that of the nondelayed case. Furthermore, $\mathcal{V}_X^{\rm LB}$ is also lower for the delayed case. In order to clarify the effect of the time delay, we plot $\mathcal{V}_X$ and $\mathcal{V}_X^{\rm LB}$ as a function of τ for $N = N_O = 10$ and

$K = 1$ in Fig. 4(b). The solid and dashed lines denote $\mathcal{V}_X$ and $\mathcal{V}_X^{\mathrm{LB}}$, respectively, obtained analytically, and the circles and triangles denote $\mathcal{V}_X$ and $\mathcal{V}_X^{\mathrm{LB}}$, respectively, obtained by Monte Carlo simulations. For $\tau < 0.03$, the analytical and Monte Carlo results agree, whereas they do not agree for $\tau > 0.03$. This is because $\phi_j^\tau - \phi_i$ in the coupling deviates from 0 when τ increases, which degrades the reliability of the Taylor approximation. Intriguingly, Fig. 4(b) shows that when τ increases to $\tau \sim 0.04$, both $\mathcal{V}_X$ and $\mathcal{V}_X^{\mathrm{LB}}$ decrease. Since the lower bound $\mathcal{V}_X^{\mathrm{LB}}$ is smaller in the presence of a time delay, the precision improvement can be ascribed to the increase in the information flow $\dot{\mathcal{I}}_X(X;F)$. Although the time delay in the Kuramoto model is often studied from the viewpoint of coherence and incoherence [35], we show that the time delay is beneficial to temporal precision of oscillators.

IV. CONCLUSION

We obtained the inequality relating the temporal variance in the Kuramoto model with the information flow conferred by coupling between oscillators. The CEP is a universal phenomenon and is not limited to coupled oscillators. For instance, the precision of cellular concentration inference is substantially improved when the concentration is measured by a population of cells [40]. A similar calculation can be applied in such cases, which will be left for future studies.

ACKNOWLEDGMENT

This work was supported by MEXT KAKENHI Grant No. JP16K00325.

APPENDIX A: GLOBALLY COUPLED MODEL

We calculate $\mathcal{V}_X^{\mathrm{LB}}$ and $\mathcal{V}_X$ for a uniformly globally coupled model, where $K_{ij} = K$. The Langevin equations for X and Y are given by Eqs. (3) and (15), respectively. The coupling variable F is

$$F = KN_L(Y - X). \tag{A1}$$

Introducing the hypothetical delay h in Eq. (A1), we have

$$F(t) = KN_L[Y(t-h) - X(t-h)]. \tag{A2}$$

Assuming that h is sufficiently small, the coupling variable obeys Eq. (16). Equations (3), (15), and (16) constitute coupled Langevin equations. The corresponding FPE is

$$\begin{aligned}\frac{\partial}{\partial t}P(X,Y,F) = &-\frac{\partial}{\partial X}\{-L(X-c)+F\}P(X,Y,F) + \frac{D}{N_O}\frac{\partial^2}{\partial X^2}P(X,Y,F) - \frac{\partial}{\partial Y}\left\{-L(Y-c) - \frac{F}{R}\right\}P(X,Y,F)\\ &+ \frac{D}{N_L}\frac{\partial^2}{\partial Y^2}P(X,Y,F) - \frac{\partial}{\partial F}\frac{1}{h}\{-F + KN_L(Y-X)\}P(X,Y,F).\end{aligned} \tag{A3}$$

The steady-state distribution of Eq. (A3) is a Gaussian distribution,

$$P(X,Y,F) = \mathcal{N}(X,Y,F|\boldsymbol{\mu}_{XYF}, \boldsymbol{\Sigma}_{XYF}),$$

where $\mathcal{N}(\cdot)$ denotes a multivariate Gaussian distribution and $\boldsymbol{\mu}_{XYF}$ and $\boldsymbol{\Sigma}_{XYF}$ are its mean vector and covariance matrix, respectively. We can obtain the mean and covariance in the steady state. For $h \to 0$, the hypothetical delay vanishes, and F reduces to Eq. (A1). The mean vector is $\boldsymbol{\mu}_{XYF} = [c, c, 0]$, and the covariance matrix $\boldsymbol{\Sigma}_{XYF}$ is

$$\boldsymbol{\Sigma}_{XYF} = \begin{bmatrix} \frac{(KN_O+L)D}{N_O L[K(R+1)N_O+L]} & \frac{DK}{L[K(R+1)N_O+L]} & -\frac{KRD}{K(R+1)N_O+L} \\ \frac{DK}{L[K(R+1)N_O+L]} & \frac{(KN_O R+L)D}{RN_O L[K(R+1)N_O+L]} & \frac{DK}{K(R+1)N_O+L} \\ -\frac{KRD}{K(R+1)N_O+L} & \frac{DK}{K(R+1)N_O+L} & \frac{N_O(R+1)RK^2D}{K(R+1)N_O+L} \end{bmatrix}. \tag{A4}$$

With Eq. (6), the information flow is obtained as follows:

$$\dot{\mathcal{I}}_X(X;F) = -\frac{\mathcal{C}_{XF}\{\mathcal{C}_{XF}(D + N_O\mathcal{C}_{XF}) - N_O\mathcal{V}_F\mathcal{V}_X\}}{N_O\mathcal{V}_X\left(\mathcal{C}_{XF}^2 - \mathcal{V}_F\mathcal{V}_X\right)}. \tag{A5}$$

From Eqs. (A4) and (A5), we obtain Eq. (17).

APPENDIX B: TIME-DELAYED MODEL

In the main text, we consider the time-delayed case. Assuming that $\Omega\tau$ is sufficiently small, we can apply the Taylor approximation to obtain

$$\frac{dX}{dt} = \omega - \Omega - LX + F + \Xi_O(t), \tag{B1}$$

$$\frac{dx_i}{dt} = \omega - \Omega - Lx_i + \sum_{j\in V,\, j\neq i} K_{ij}\left(x_j^\tau - x_i - \Omega\tau\right) + \xi_i(t) \quad (i \in V), \tag{B2}$$

$$\frac{dF}{dt} = \frac{1}{h}\left\{-F + \frac{1}{N_O}\sum_{i\in V_O}\sum_{j\in V,\, j\neq i} K_{ij}\left(x_j^\tau - x_i - \Omega\tau\right)\right\}. \tag{B3}$$

When deriving Eq. (B3), we again introduced the hypothetical time-delay h as in Eq. (A2). Suppose that the mean of $\boldsymbol{x}$ in the steady state is $[\mu_{x_1}, \mu_{x_2}, \ldots, \mu_{x_N}] = [c_{x_1}, c_{x_2}, \ldots, c_{x_N}]$, where $[c_{x_1}, c_{x_2}, \ldots, c_{x_N}]$ satisfies the following relation:

$$\omega - \Omega - Lc_{x_i} + \sum_{j\in V,\, j\neq i} K_{ij}\left(c_{x_j} - c_{x_i} - \Omega\tau\right) = 0 \quad (i \in V).$$

Equations (B1)–(B3) can be written as

$$\frac{dX}{dt} = -L(X - c_X) + F - c_F + \Xi_O(t), \tag{B4}$$

$$\frac{dx_i}{dt} = -L\left(x_i - c_{x_i}\right) + \sum_{j\in V,\, j\neq i} K_{ij}\left(x_j^\tau - x_i - \Omega\tau\right) - c_{f_i} + \xi_i(t) \quad (i \in V), \tag{B5}$$

$$\frac{dF}{dt} = \frac{1}{h}\left\{-F + \frac{1}{N_O}\sum_{i\in V_O}\sum_{j\in V,\, j\neq i} K_{ij}\left(x_j^\tau - x_i - \Omega\tau\right)\right\}, \tag{B6}$$

where

$$c_{f_i} = \sum_{j\in V,\, j\neq i} K_{ij}\left(c_{x_j} - c_{x_i} - \Omega\tau\right), \quad c_X = \frac{1}{N_O}\sum_{i\in V_O} c_{x_i}, \quad c_F = \frac{1}{N_O}\sum_{i\in V_O} c_{f_i}.$$

Following Ref. [38], we can obtain the time evolution of $P(X, \boldsymbol{x}, F) = \int P(X, \boldsymbol{x}, F; \boldsymbol{x}^\tau)d\boldsymbol{x}^\tau$ where $P(X, \boldsymbol{x}, F; \boldsymbol{x}^\tau)$ is the probability density of X, $\boldsymbol{x}$, and F at time t and of $\boldsymbol{x}^\tau$ at time $t - \tau$. Let $\mathcal{A}(X, \boldsymbol{x}, F)$ be an arbitrary function of X, $\boldsymbol{x}$, and F. The derivation in Ref. [38] is constructed on the Ito interpretation. If Langevin equations of interest obey the Stratonovich interpretation, they have to be converted to equivalent Ito equations [38]. Since Eqs. (B4)–(B6) include only additive noise terms, the calculations afterwards do not depend on the stochastic integral. Taking the expectation for $d\mathcal{A}(X, \boldsymbol{x}, F)/dt$, we apply the Ito calculus to obtain

$$\begin{aligned}
\left\langle\frac{d\mathcal{A}}{dt}\right\rangle &= \left\langle\frac{\partial\mathcal{A}}{\partial X}[-L(X - c_X) + F - c_F] + \frac{1}{h}\frac{\partial\mathcal{A}}{\partial F}\left[-F + \frac{1}{N_O}\sum_{i\in V_O}\sum_{j\in V,\, j\neq i} K_{ij}\left(x_j^\tau - x_i - \Omega\tau\right)\right]\right.\\
&\quad + \sum_{i\in V}\frac{\partial\mathcal{A}}{\partial x_i}\left[-L(x_i - c_{x_i}) + \sum_{j\in V,\, j\neq i} K_{ij}\left(x_j^\tau - x_i - \Omega\tau\right) - c_{f_i}\right]\\
&\quad \left. + \frac{D}{N_O}\frac{\partial^2\mathcal{A}}{\partial X^2} + \sum_{i\in V} D\frac{\partial^2\mathcal{A}}{\partial x_i^2} + \sum_{i\in V_O}\frac{2D}{N_O}\frac{\partial^2\mathcal{A}}{\partial X\,\partial x_i}\right\rangle_{P(X,\boldsymbol{x},F;\boldsymbol{x}^\tau)}\\
&= \left\langle\frac{\partial\mathcal{A}}{\partial X}[-L(X - c_X) + F - c_F] + \frac{1}{h}\frac{\partial\mathcal{A}}{\partial F}\left[-F + \left\langle\frac{1}{N_O}\sum_{i\in V_O}\sum_{j\in V,\, j\neq i} K_{ij}\left(x_j^\tau - x_i - \Omega\tau\right)\right\rangle_{P(\boldsymbol{x}^\tau|X,\boldsymbol{x},F)}\right]\right.\\
&\quad + \sum_{i\in V}\frac{\partial\mathcal{A}}{\partial x_i}\left[-L(x_i - c_{x_i}) + \left\langle\sum_{j\in V,\, j\neq i} K_{ij}\left(x_j^\tau - x_i - \Omega\tau\right)\right\rangle_{P(\boldsymbol{x}^\tau|X,\boldsymbol{x},F)} - c_{f_i}\right]\\
&\quad \left. + \frac{D}{N_O}\frac{\partial^2\mathcal{A}}{\partial X^2} + \sum_{i\in V} D\frac{\partial^2\mathcal{A}}{\partial x_i^2} + \sum_{i\in V_O}\frac{2D}{N_O}\frac{\partial^2\mathcal{A}}{\partial X\,\partial x_i}\right\rangle_{P(X,\boldsymbol{x},F)}.
\end{aligned}$$

Here, we explicitly write the probability density with which the expectation is calculated. Letting $P(B)$ be a probability density function of arbitrary random variable(s) B [$\int P(B)dB = 1$], we define $\langle\cdots\rangle_{P(B)} = \int\cdots P(B)dB$ [e.g., $\langle\cdots\rangle_{P(\boldsymbol{x}^\tau|X,F,\boldsymbol{x})} = \int\cdots P(\boldsymbol{x}^\tau|X, \boldsymbol{x}, F)d\boldsymbol{x}^\tau$]. Since $\mathcal{A}(X, \boldsymbol{x}, F)$ is an arbitrary function, using the integration by parts, we obtain

$$\frac{\partial}{\partial t}P(X, \boldsymbol{x}, F) = \mathbb{L}_\tau(X, \boldsymbol{x}, F)P(X, \boldsymbol{x}, F), \tag{B7}$$

where $\mathbb{L}_\tau(X, \boldsymbol{x}, F)$ is the following FPE operator:

$$\mathbb{L}_\tau(X, \boldsymbol{x}, F) = -\frac{\partial}{\partial X}[-L(X - c_X) + F - c_F] - \sum_{i\in V}\frac{\partial}{\partial x_i}\left[-L(x_i - c_{x_i}) + \left\langle \sum_{j\in V,\, j\neq i} K_{ij}\left(x_j^\tau - x_i - \Omega\tau\right)\right\rangle_{P(\boldsymbol{x}^\tau|X,\boldsymbol{x},F)} - c_{f_i}\right]$$
$$-\frac{\partial}{\partial F}\frac{1}{h}\left[-F + \left\langle \frac{1}{N_O}\sum_{i\in V_O}\sum_{j\in V,\, j\neq i} K_{ij}\left(x_j^\tau - x_i - \Omega\tau\right)\right\rangle_{P(\boldsymbol{x}^\tau|X,\boldsymbol{x},F)}\right] + \frac{D}{N_O}\frac{\partial^2}{\partial X^2} + \sum_{i\in V} D\frac{\partial^2}{\partial x_i^2} + \sum_{i\in V_O}\frac{2D}{N_O}\frac{\partial^2}{\partial X\,\partial x_i}. \tag{B8}$$

For details of the derivations, please see Ref. [38]. Integrating Eqs. (B7) and (B8) with respect to $\boldsymbol{x}$, we obtain

$$\frac{\partial P(X, F)}{\partial t} = -\frac{\partial}{\partial X}[-L(X - c_X) + F - c_F]P(X, F) + \frac{D}{N_O}\frac{\partial^2}{\partial X^2}P(X, F) - \frac{\partial}{\partial F}\frac{1}{h}\{-F + \mathcal{H}_\tau(X, F)\}P(X, F), \tag{B9}$$

where

$$\mathcal{H}_\tau(X, F) = \left\langle \frac{1}{N_O}\sum_{i\in V_O}\sum_{j\in V,\, j\neq i} K_{ij}\left(x_j^\tau - x_i - \Omega\tau\right)\right\rangle_{P(\boldsymbol{x},\boldsymbol{x}^\tau|X,F)}.$$

Since Eq. (B9) is identical to Eqs. (10)–(12) when replacing $\mathcal{H}(X, F)$ with $\mathcal{H}_\tau(X, F)$ and $F \to F - c_F$ in a drift term, we obtain

$$\langle (X - c_X)^2\rangle \geqslant \frac{D}{N_O L^2}[L - \dot{\mathcal{I}}_X(X; F)],$$

which is the same inequality as the nondelayed model [Eq. (14)].

APPENDIX C: GLOBALLY COUPLED TIME-DELAY MODEL

We consider a globally coupled case $K_{ij} = K$ to calculate $\mathcal{V}_X^{\mathrm{LB}}$ and $\mathcal{V}_X$. The calculations can be classified into two cases: $N_L \geqslant 1$ and $N_L = 0$.

For $N_L \geqslant 1$, we obtain Eqs. (19)–(21) for X, Y, and F. We need to calculate the covariance matrix of X and F. These quantities can be obtained via the Fourier transform [38]. The detailed procedures (known as the Rice method) are shown in Ref. [39]. For an arbitrary random variable $A(t)$, we can define the Fourier transform and its inverse as follows:

$$\widetilde{A}(\nu) = \mathcal{F}[A(t)] = \int_{-\infty}^{\infty} e^{-i\nu t}A(t)dt, \quad A(t) = \mathcal{F}^{-1}[\widetilde{A}(\nu)] = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\nu t}\widetilde{A}(\nu)d\nu,$$

where i denotes the imaginary unit. Then, applying the Fourier transform, we obtain

$$\begin{bmatrix}\widetilde{X}(\nu)\\ \widetilde{Y}(\nu)\\ \widetilde{F}(\nu)\end{bmatrix} = \mathcal{M}\begin{bmatrix}\widetilde{\Xi}_O(\nu)\\ \widetilde{\Xi}_L(\nu)\\ 0\end{bmatrix}, \tag{C1}$$

where $\mathcal{M}$ is a regular matrix defined by

$$\mathcal{M} = \begin{bmatrix} L + i\nu & 0 & -1\\ -e^{-i\nu\tau}KN_O & L + K\{-1 - e^{-i\nu\tau}(N_L - 1) + N_L + N_O\} + i\nu & 0\\ K\{-1 + N_L - e^{-i\nu\tau}(N_O - 1) + N_O\} & -e^{-i\nu\tau}KN_L & 1\end{bmatrix}^{-1}.$$

We can consider the Fourier transform of the correlation function: $\mathcal{S}_X(\nu) = \mathcal{F}[\langle X(t)X(0)\rangle]$, $\mathcal{S}_F(\nu) = \mathcal{F}[\langle F(t)F(0)\rangle]$, $\mathcal{S}_{XF}(\nu) = \mathcal{F}[\langle X(t)F(0)\rangle]$, $\mathcal{S}_O(\nu) = \mathcal{F}[\langle \Xi_O(t)\Xi_O(0)\rangle] = 2D/N_O$, and $\mathcal{S}_L(\nu) = \mathcal{F}[\langle \Xi_L(t)\Xi_L(0)\rangle] = 2D/N_L$. Let A and B be arbitrary random variables. According to the Wiener-Khinchin theorem, the following relation holds [39]:

$$\langle \widetilde{A}(\nu)\widetilde{B}^*(\nu')\rangle = 2\pi\delta(\nu - \nu')\mathcal{S}_{AB}(\nu), \tag{C2}$$

$$\mathcal{S}_{AB}(\nu) = \mathcal{F}[\langle A(t)B^*(0)\rangle], \tag{C3}$$

where the superscript $*$ denotes the complex conjugate. From Eqs. (C1)–(C3), we obtain

$$\mathcal{S}_X(\nu) = |\mathcal{M}_{11}|^2\mathcal{S}_O(\nu) + |\mathcal{M}_{12}|^2\mathcal{S}_L(\nu), \tag{C4}$$

$$\mathcal{S}_F(\nu) = |\mathcal{M}_{31}|^2\mathcal{S}_O(\nu) + |\mathcal{M}_{32}|^2\mathcal{S}_L(\nu), \tag{C5}$$

$$\mathcal{S}_{XF}(\nu) = \mathcal{M}_{11}\mathcal{M}_{31}^*\mathcal{S}_O(\nu) + \mathcal{M}_{12}\mathcal{M}_{32}^*\mathcal{S}_L(\nu), \tag{C6}$$

where $\mathcal{M}_{ij}$ is the i, jth element of $\mathcal{M}$. When we assume a Gaussian distribution for $P(X, F)$, from Eq. (A5), it is

sufficient to calculate $\mathcal{V}_X$, $\mathcal{V}_F$, and $\mathcal{C}_{XF}$ for information flow $\dot{\mathcal{I}}_X$. From Eq. (C3), the inverse Fourier transform yields variances $\mathcal{V}_X$ and $\mathcal{V}_F$ and covariance $\mathcal{C}_{XF}$,

$$\mathcal{V}_X = \langle X(t)^2 \rangle = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathcal{S}_X(\nu)d\nu, \tag{C7}$$

$$\mathcal{V}_F = \langle F(t)^2 \rangle = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathcal{S}_F(\nu)d\nu, \tag{C8}$$

$$\mathcal{C}_{XF} = \langle X(t)F(t) \rangle = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathcal{S}_{XF}(\nu)d\nu. \tag{C9}$$

Since it is unlikely that we can obtain the inverse Fourier transforms of Eqs. (C7)–(C9) analytically, we calculated the transforms numerically.

We next consider the case in which $N_L = 0$. In a similar way, we have the following coupled equations:

$$\frac{dX}{dt} = -LX + F + \Xi_O(t), \quad F = K(N-1)(X^\tau - X).$$

The Fourier transform yields

$$\begin{bmatrix} \widetilde{X}(\nu) \\ \widetilde{F}(\nu) \end{bmatrix} = \mathcal{L}\begin{bmatrix} \widetilde{\Xi}_O(\nu) \\ 0 \end{bmatrix}, \tag{C10}$$

where $\mathcal{L}$ is a regular matrix defined as

$$\mathcal{L} = \begin{bmatrix} L + i\nu & -1 \\ K(N-1)(1-e^{-i\nu\tau}) & 1 \end{bmatrix}^{-1}.$$

Then, the Fourier transforms of the correlation functions are

$$\mathcal{S}_X(\nu) = |\mathcal{L}_{11}|^2 \mathcal{S}_O(\nu), \tag{C11}$$

$$\mathcal{S}_F(\nu) = |\mathcal{L}_{21}|^2 \mathcal{S}_O(\nu), \tag{C12}$$

$$\mathcal{S}_{XF}(\nu) = \mathcal{L}_{11}\mathcal{L}_{21}^* \mathcal{S}_O(\nu). \tag{C13}$$

Again, the variance and covariance can be obtained by the inverse Fourier transform of Eqs. (C7)–(C9).

APPENDIX D: MONTE CARLO SIMULATION

In order to solve Langevin equations numerically, we use the Euler method shown in Ref. [39] with time step $\epsilon = 0.001$. When calculating the information flow $\dot{\mathcal{I}}_X(X;F)$, we use the following approximation from Eq. (6):

$$\dot{\mathcal{I}}_X(X;F) \simeq -\frac{1}{\epsilon}\left\langle \ln\left(\frac{P[F(t)|X(t+\epsilon)]}{P[F(t)|X(t)]}\right)\right\rangle. \tag{D1}$$

Here, we use $F(t)$ based on Eq. (2) [nondelayed case] or Eq. (18) [time-delayed case]. In Eq. (D1), we need to calculate the probability density $P(X,F)$. Assuming a multivariate Gaussian distribution for $P(X,F)$, we empirically calculate the mean vector and covariance matrix for $P(X,F)$. Then, we can numerically calculate the information flow $\dot{\mathcal{I}}_X(X;F)$ as the average of 2×10^6 samples. Regarding the number of oscillators, we use $N_O = 10$ and $N_L = 0$–10 ($0 \leqslant R \leqslant 1$). Note that when we employ an excessively large N, the simulation suffers from an artifact due to the time discretization because discretized coupling terms $f_i(t)\epsilon$ are on the order of $O(N\epsilon)$.

[1] R. Refinetti, *Circadian Physiology*, 2nd ed. (CRC, Boca Raton, FL, 2005).
[2] H. Ukai and H. R. Ueda, Annu. Rev. Physiol. **72**, 579 (2010).
[3] K. T. Moortgat, T. H. Bullock, and T. J. Sejnowski, J. Neurophysiol. **83**, 971 (2000).
[4] J. T. Enright, Science **209**, 1542 (1980).
[5] Y. Hasegawa and M. Arita, J. R. Soc., Interface **11**, 20131018 (2014).
[6] Y. Hasegawa and M. Arita, Phys. Rev. Lett. **113**, 108101 (2014).
[7] A. T. Winfree, *The Geometry of Biological Time*, 2nd ed. (Springer, New York, 2001).
[8] J. Clay and R. L. DeHaan, Biophys. J. **28**, 377 (1979).
[9] D. J. Needleman, P. H. E. Tiesinga, and T. J. Sejnowski, Physica D **155**, 324 (2001).
[10] H. Kori, Y. Kawamura, and N. Masuda, J. Theor. Biol. **297**, 61 (2012).
[11] Y. Cao, H. Wang, Q. Ouyang, and Y. Tu, Nat. Phys. **11**, 772 (2015).
[12] A. C. Barato and U. Seifert, Phys. Rev. Lett. **114**, 158101 (2015).
[13] A. C. Barato and U. Seifert, Phys. Rev. X **6**, 041053 (2016).
[14] T. R. Gingrich, J. M. Horowitz, N. Perunov, and J. L. England, Phys. Rev. Lett. **116**, 120601 (2016).
[15] J. M. Horowitz and T. R. Gingrich, Phys. Rev. E **96**, 020103 (2017).
[16] F. Ritort, in *Advances in Chemical Physics*, edited by S. A. Rice (Wiley, Hoboken, NJ, 2008), Vol. 137, pp. 31–123.
[17] U. Seifert, Rep. Prog. Phys. **75**, 126001 (2012).
[18] C. Van den Broeck and M. Esposito, Physica A **418**, 6 (2015).
[19] H. Sakaguchi, Prog. Theor. Phys. **79**, 39 (1988).
[20] J. A. Acebrón, L. L. Bonilla, C. J. Pérez Vicente, F. Ritort, and R. Spigler, Rev. Mod. Phys. **77**, 137 (2005).
[21] A. Pikovsky, M. Rosenblem, and J. Kurths, *Synchronization: A Universal Concept in Nonlinear Sciences* (Cambridge University Press, Cambridge, U.K., 2001).
[22] Y. Kuramoto, *Chemical Oscillations, Waves, and Turbulence* (Dover, Mineola, NY, 2003).
[23] J. M. Horowitz and M. Esposito, Phys. Rev. X **4**, 031015 (2014).
[24] J. M. Horowitz, J. Stat. Mech.: Theory Exp. (2015) P03006.
[25] A. C. Barato, D. Hartich, and U. Seifert, New J. Phys. **16**, 103024 (2014).
[26] D. Hartich, A. C. Barato, and U. Seifert, Phys. Rev. E **93**, 022116 (2016).
[27] R. A. Brittain, N. S. Jones, and T. E. Ouldridge, J. Stat. Mech.: Theory Exp. (2017) 063502.
[28] T. Munakata and M. L. Rosinberg, J. Stat. Mech.: Theory Exp. (2013) P06014.
[29] J. M. Horowitz and H. Sandberg, New J. Phys. **16**, 125007 (2014).

[30] S. Ito and T. Sagawa, Nat. Commun. **6**, 7498 (2015).
[31] T. Tomé, Braz. J. Phys. **36**, 1285 (2006).
[32] T. Schreiber, Phys. Rev. Lett. **85**, 461 (2000).
[33] A. E. Allahverdyan, D. Janzing, and G. Mahler, J. Stat. Mech: Theory Exp. (2009) P09011.
[34] D. Hartich, A. C. Barato, and U. Seifert, J. Stat. Mech: Theory Exp. (2014) P02016.
[35] M. K. S. Yeung and S. H. Strogatz, Phys. Rev. Lett. **82**, 648 (1999).
[36] M. L. Rosinberg, T. Munakata, and G. Tarjus, Phys. Rev. E **91**, 042114 (2015).
[37] M. L. Rosinberg, G. Tarjus, and T. Munakata, Phys. Rev. E **95**, 022123 (2017).
[38] S. Guillouzic, I. L'Heureux, and A. Longtin, Phys. Rev. E **59**, 3970 (1999).
[39] H. Risken, *The Fokker-Planck Equation: Methods of Solution and Applications*, 2nd ed. (Springer, Berlin/Heidelberg, 1989).
[40] S. Fancher and A. Mugler, Phys. Rev. Lett. **118**, 078101 (2017).