Stochastic Thermodynamic Interpretation of Information Geometry

Sosuke Ito

RIES, Hokkaido University, N20 W10, Kita-ku, Sapporo, Hokkaido 001-0020, Japan

In recent years, the unified theory of information and thermodynamics has been intensively discussed in the context of stochastic thermodynamics. The unified theory reveals that information theory would be useful to understand nonstationary dynamics of systems far from equilibrium. In this Letter, we have found a new link between stochastic thermodynamics and information theory well-known as information geometry. By applying this link, an information geometric inequality can be interpreted as a thermodynamic uncertainty relationship between speed and thermodynamic cost. We have numerically applied an information geometric inequality to a thermodynamic model of a biochemical enzyme reaction.

The crucial relationship between thermodynamics and information theory has been well studied in the last decades [1]. Historically, thermodynamic-informational links had been discussed in the context of the second law of thermodynamics and the paradox of Maxwell's demon [2]. Recently, several studies have newly revealed thermodynamic interpretations of informational quantities such as the Kullback-Leibler divergence [3], mutual information [4–6], the transfer entropy and information flow [7–19]. The above interpretations of informational quantities are based on the theory of stochastic thermodynamics [20,21], which mainly focus on the entropy production in stochastic dynamics of small systems far from equilibrium.

The information thermodynamic relationship has been attractive not only in terms of Maxwell's demon, but also in terms of geometry [22–28]. Indeed, differential geometric interpretations of thermodynamics have been discussed especially in a near-equilibrium system [22,29–34]. Moreover, the technique of differential geometry in information theory, well-known as information geometry [35], has received remarkable attention in the fields of neuroscience, signal processing, quantum mechanics, and machine learning [36–38]. In spite of the deep link between information and thermodynamics, the direct connection between thermodynamics and information geometry has been elusive, especially for nonstationary and nonequilibrium dynamics. For example, G. E. Crooks discovered a link between thermodynamics and information geometry [22,32] based on the Gibbs ensemble, and then his discussion is only valid for a near-equilibrium system.

In this Letter, we discover a fundamental link between information geometry and thermodynamics based on stochastic thermodynamics for the master equation. We mainly report two inequalities, derived thanks to information geometry, and we interpret them within the theory of stochastic thermodynamics. The first inequality connects the environmental entropy change rate to the mean change of the local thermodynamic force rate. The second inequality can be interpreted as a kind of thermodynamic uncertainty relationship, or thermodynamic trade-off relationship [39–51], between the speed of a transition from one state to another and the thermodynamic cost related to the entropy change of thermal baths in a near-equilibrium system. We numerically illustrate these two inequalities on a model of a biochemical enzyme reaction.

Stochastic thermodynamics.—To clarify a link between stochastic thermodynamics and information geometry, we start with the formalism of stochastic thermodynamics for the master equation [20,21], which is also known as the Schnakenberg network theory [52,53].

Here, we consider a $(n + 1)$-states system. We assume that transitions between states are induced by n_bath-multiple thermal baths. The master equation for the probability p_x (≥ 0, $\sum_{x=0}^n p_x = 1$) to find the state at $x = \{0, 1, \ldots, n\}$ is given by

$$\frac{d}{dt} p_x = \sum_{\nu=1}^{n_\text{bath}} \sum_{x'=0}^{n} W_{x' \to x}^{(\nu)} p_{x'}, \qquad (1)$$

where $W_{x' \to x}^{(\nu)}$ is the transition rate from x' to x induced by νth thermal bath. We assume a nonzero value of the transition rate $W_{x' \to x}^{(\nu)} > 0$ for any $x \neq x'$. We also assume the condition

$$\sum_{x=0}^{n} W_{x' \to x}^{(\nu)} = 0, \qquad (2)$$

or, equivalently, $W_{x' \to x'}^{(\nu)} = -\sum_{x \neq x'} W_{x' \to x}^{(\nu)} < 0$, which leads to the conservation of probability $d(\sum_{x=0}^n p_x)/dt = 0$. Equation (2) indicates that the master equation is then given by the thermodynamic flux from the state x' to x [52],

$$J_{x' \to x}^{(\nu)} := W_{x' \to x}^{(\nu)} p_{x'} - W_{x \to x'}^{(\nu)} p_x, \qquad (3)$$

$$\frac{d}{dt}p_x = \sum_{\nu=1}^{n_{\text{bath}}} \sum_{x'=0}^{n} J_{x'\to x}^{(\nu)}. \tag{4}$$

If dynamics are reversible (i.e., $J_{x'\to x}^{(\nu)} = 0$ for any x, x', and ν), the system is said to be in thermodynamic equilibrium. If we consider the conjugated thermodynamic force

$$F_{x'\to x}^{(\nu)} := \ln[W_{x'\to x}^{(\nu)} p_{x'}] - \ln[W_{x\to x'}^{(\nu)} p_x], \tag{5}$$

thermodynamic equilibrium is equivalently given by $F_{x'\to x}^{(\nu)} = 0$ for any x, x', and ν.

In stochastic thermodynamics [21], we treat the entropy change of the thermal bath and the system in a stochastic way. In the transition from x' to x, the stochastic entropy change of the νth thermal bath is defined as

$$\Delta\sigma_{x'\to x}^{\text{bath}(\nu)} := \ln\frac{W_{x'\to x}^{(\nu)}}{W_{x\to x'}^{(\nu)}}, \tag{6}$$

and the stochastic entropy change of the system is defined as the stochastic Shannon entropy change

$$\Delta\sigma_{x'\to x}^{\text{sys}} := \ln p_{x'} - \ln p_x, \tag{7}$$

respectively. The thermodynamic force is then given by the sum of entropy changes in the transition from x' to x, induced by the νth thermal bath $F_{x'\to x}^{(\nu)} = \Delta\sigma_{x'\to x}^{\text{bath}(\nu)} + \Delta\sigma_{x'\to x}^{\text{sys}}$. This fact implies that the system is in equilibrium if the sum of entropy changes is zero for any transitions.

The total entropy production rate $\dot{\Sigma}^{\text{tot}}$ is given by the sum of the products of thermodynamic forces and fluxes over possible transitions. To simplify notations, we introduce the set of directed edges $E = \{(x'\to x, \nu) | 0 \le x' < x \le n, 1 \le \nu \le n_{\text{bath}}\}$, which denotes the set of all possible transitions between two states. The total entropy production rate is then given by

$$\dot{\Sigma}^{\text{tot}} := \sum_{(x'\to x,\nu)\in E} J_{x'\to x}^{(\nu)} F_{x'\to x}^{(\nu)} = \langle F \rangle, \tag{8}$$

where a parenthesis $\langle \cdots \rangle$ is defined as $\langle A \rangle := \sum_{(x'\to x,\nu)\in E} J_{x'\to x}^{(\nu)} A_{x'\to x}^{(\nu)}$ for any function of edge $A_{x'\to x}^{(\nu)}$. Because signs of the thermodynamic force $F_{x'\to x}^{(\nu)}$ and the flux $J_{x'\to x}^{(\nu)}$ are the same, the total entropy production rate is non-negative

$$\langle F \rangle = \langle \Delta\sigma^{\text{bath}} \rangle + \langle \Delta\sigma^{\text{sys}} \rangle \ge 0, \tag{9}$$

that is well-known as the second law of thermodynamics.

Information geometry.—Next, we introduce information theory, which is well-known as information geometry [35].

In this Letter, we only consider the discrete distribution group $\mathbf{p} = (p_0, p_1, ..., p_n)$, $p_x \ge 0$, and $\sum_{x=0}^{n} p_x = 1$. This discrete distribution group gives the n-dimensional manifold S_n, because the discrete distribution is given by $n+1$ parameters $(p_0, p_1, ..., p_n)$ under the constraint $\sum_{x=0}^{n} p_x = 1$. To introduce a geometry on the manifold S_n, we conventionally consider the Kullback-Leibler divergence [54] between two distributions $\mathbf{p}$ and $\mathbf{p}' = (p_0', p_1', ..., p_n')$, defined as

$$D_{\text{KL}}(\mathbf{p}||\mathbf{p}') := \sum_{x=0}^{n} p_x \ln\frac{p_x}{p_x'}. \tag{10}$$

The square of the line element ds is defined as the second-order Taylor series of the Kullback-Leibler divergence

$$ds^2 := \sum_{x=0}^{n} \frac{(dp_x)^2}{p_x} = 2D_{\text{KL}}(\mathbf{p}||\mathbf{p}+d\mathbf{p}), \tag{11}$$

where $d\mathbf{p} = (dp_0, dp_1, ..., dp_n)$ is the infinitesimal displacement that satisfies $\sum_{x=0}^{n} dp_x = 0$. This square of the line element is directly related to the Fisher information metric [55] [see also Supplemental Material (SM) [56]].

The manifold S_n leads to the geometry of the n-sphere surface of radius 2 (see also Fig. 1), because the square of the line element is also given by $ds^2 = \sum_{x=0}^{n} (2dr_x)^2$ under the constraint $\mathbf{r}\cdot\mathbf{r} = \sum_x (\sqrt{p_x})^2 = 1$, where $\mathbf{r}$ is the unit vector defined as $\mathbf{r} = (r_0, r_1, ..., r_n) := (\sqrt{p_0}, \sqrt{p_1}, ..., \sqrt{p_n})$ and $\cdot$ denotes the inner product. The statistical length $\mathcal{L}$ [57,58]

$$\mathcal{L} := \int ds = \int \frac{ds}{dt} dt, \tag{12}$$

from the initial state $\mathbf{r}_{\text{ini}}$ to the final state $\mathbf{r}_{\text{fin}}$ is then bounded by
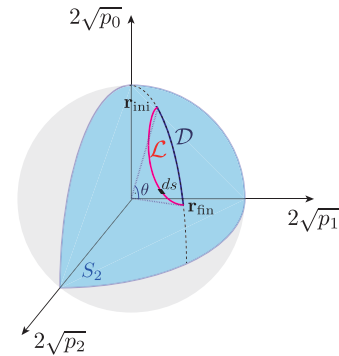

FIG. 1. Schematic of information geometry on the manifold S_2. The manifold S_2 leads to the sphere surface of radius 2 (see also SM [56]). The statistical length $\mathcal{L}$ is bounded by the shortest length $\mathcal{D} = 2\theta = 2\cos^{-1}(\mathbf{r}_{\text{ini}}\cdot\mathbf{r}_{\text{fin}})$.

$$\mathcal{L} \geq 2\cos^{-1}(\mathbf{r}_{\text{ini}} \cdot \mathbf{r}_{\text{fin}}) := \mathcal{D}(\mathbf{r}_{\text{ini}}; \mathbf{r}_{\text{fin}}), \qquad (13)$$

because $\mathcal{D}(\mathbf{r}_{\text{ini}}; \mathbf{r}_{\text{fin}}) = 2\theta$ is the shortest length between $\mathbf{r}_{\text{ini}}$ and $\mathbf{r}_{\text{fin}}$ on the n-sphere surface of radius 2, where θ is the angle between $\mathbf{r}_{\text{ini}}$ and $\mathbf{r}_{\text{fin}}$ given by the inner product $\mathbf{r}_{\text{ini}} \cdot \mathbf{r}_{\text{fin}} = \cos\theta$.

Stochastic thermodynamics of information geometry.— Here, we discuss a relationship between the line element and conventional observables of stochastic thermodynamics, which give a stochastic thermodynamic interpretation of information geometric quantities.

By using the master equation (1) and definitions of the line element and thermodynamic quantities Eqs. (5), (6), and (11), we obtain stochastic thermodynamic expressions of ds^2/dt^2 (see also SM [56]),

$$\frac{ds^2}{dt^2} = \sum_{x=0}^{n} p_x \frac{d}{dt}\left(-\frac{1}{p_x}\frac{dp_x}{dt}\right) \qquad (14)$$

$$= -\sum_{x=0}^{n} p_x \frac{d}{dt}\left(\sum_{\nu=1}^{n_{\text{bath}}}\sum_{x'=0}^{n} W_{x \to x'}^{(\nu)} e^{-F_{x \to x'}^{(\nu)}}\right) \qquad (15)$$

$$= \left\langle\frac{d\Delta\sigma^{\text{bath}}}{dt}\right\rangle - \left\langle\frac{dF}{dt}\right\rangle. \qquad (16)$$

Equation (15) implies that geometric dynamics are driven by the thermodynamic factor $\exp[-F_{x \to x'}^{(\nu)}]$, that is well discussed in the context of stochastic thermodynamics (especially in the context of the fluctuation theorem [59–64]). The time evolution of the line element ds^2/dt^2 is directly related to the expected value of the time derivative of the rate-weighted thermodynamic factor $W_{x \to x'}^{(\nu)} e^{-F_{x \to x'}^{(\nu)}}$.

Another expression Eq. (16) gives a stochastic thermodynamic interpretation of information geometry, especially in case of a near-equilibrium system. The condition of an equilibrium system is given by $F_{x' \to x}^{(\nu)} = 0$ for any x', x, and ν. Then, the square of the line element is given by the entropy change in thermal baths $ds^2 \simeq \langle d\Delta\sigma^{\text{bath}}\rangle dt$ in a near-equilibrium system.

For example, in a near-equilibrium system, the probability distribution is assumed to be the canonical distribution $p_x = \exp(\beta(\phi - H_x))$, where $\phi := -\beta^{-1}\ln[\sum_{x=0}^{n}\exp(-\beta H_x)]$ is the Helmholtz free energy, β is the inverse temperature, and H_x is the Hamiltonian of the system in the state x. To consider a near-equilibrium transition, we assume that β and H_x can depend on time. From $ds^2 = [\langle d\Delta\sigma^{\text{bath}}\rangle - \langle dF\rangle]dt = -\langle d\Delta\sigma^{\text{sys}}\rangle dt$, we obtain $ds^2 = -\langle d\Delta\sigma^{\text{sys}}\rangle dt = -\langle d(\beta\Delta H)\rangle dt$ in a near-equilibrium system, where $\Delta H_{x' \to x} := H_x - H_{x'}$ is the Hamiltonian change from the state x' to x. Because $-\beta\Delta H$ can be considered as the entropy change of the thermal bath $\Delta\sigma^{\text{bath}}$, an expression $ds^2 = -\langle d(\beta\Delta H)\rangle dt$ for

the canonical distribution is consistent with a near equilibrium expression $ds^2 \simeq \langle d\Delta\sigma^{\text{bath}}\rangle dt$.

We also discuss the second order expansion of ds^2/dt^2 for the thermodynamic force in the SM [56], based on the linear irreversible thermodynamics [52]. Our discussion implies that the square of the line element (or the Fisher information metric) for the thermodynamic forces is related to the Onsager coefficients. Because of the Cramér-Rao bound [54,55], the Onsager coefficients are directly connected to a lower bound of the variance of unbiased estimator for parameters driven by the thermodynamic force.

Because of the non-negativity of the square of line element $ds^2/dt^2 \geq 0$, we have a thermodynamic inequality

$$\left\langle\frac{d\Delta\sigma^{\text{bath}}}{dt}\right\rangle \geq \left\langle\frac{dF}{dt}\right\rangle. \qquad (17)$$

The equality holds if the system is in a stationary state, i.e., $dp_x/dt = 0$ for any x. This result (17) implies that the change of the thermodynamic force rate is transferred to the environmental entropy change rate. The difference $\langle d\Delta\sigma^{\text{bath}}/dt\rangle - \langle dF/dt\rangle \geq 0$ can be interpreted as loss in the entropy change rate transfer due to the nonstationarity. If the environmental entropy change does not change in time (i.e., $d\Delta\sigma_{x' \to x}^{\text{bath}(\nu)}/dt = 0$ for any x' and x), the thermodynamic force change tends to decrease (i.e., $\langle dF/dt\rangle \leq 0$) in a transition. We stress that a mathematical property of the thermodynamic force in this result is different from the second law of thermodynamics $\langle F\rangle \geq 0$.

From Eq. (16), the statistical length $\mathcal{L} = \int_0^{\tau} dt(ds/dt)$ from time $t = 0$ to $t = \tau$ is given by

$$\mathcal{L} = \int_{t=0}^{t=\tau} dt \sqrt{\left\langle\frac{d\Delta\sigma^{\text{bath}}}{dt}\right\rangle - \left\langle\frac{dF}{dt}\right\rangle}. \qquad (18)$$

We then obtain the following thermodynamic inequality from Eqs. (13) and (18),

$$\int_{t=0}^{t=\tau} dt \sqrt{\left\langle\frac{d\Delta\sigma^{\text{bath}}}{dt}\right\rangle - \left\langle\frac{dF}{dt}\right\rangle} \geq \mathcal{D}(\mathbf{r}(0); \mathbf{r}(\tau)). \qquad (19)$$

The equality holds if the path of transient dynamics is a geodesic line on the manifold S_n. This inequality gives a geometric constraint of the entropy change rate transfer in a transition between two probability distributions $\mathbf{p}(0)$ and $\mathbf{p}(\tau)$.

Thermodynamic uncertainty.—We finally reach to a thermodynamic uncertainty relationship between speed and thermodynamic cost. Here, we consider the action $\mathcal{C} := (1/2)\int_{t=0}^{t=\tau} dt(ds^2/dt^2)$ from time $t = 0$ to $t = \tau$. From Eq. (16), the action $\mathcal{C}$ is given by

$$\mathcal{C} = \frac{1}{2} \int_{t=0}^{t=\tau} dt \left(\left\langle \frac{d\Delta\sigma^{\mathrm{bath}}}{dt} \right\rangle - \left\langle \frac{dF}{dt} \right\rangle \right). \qquad (20)$$

Especially in the case of a near-equilibrium system, the action $\mathcal{C}$ is given by $\mathcal{C} \simeq \int \langle d\Delta\sigma^{\mathrm{bath}} \rangle / 2$. If we assume the canonical distribution, we have $\mathcal{C} = -\int \langle d(\beta\Delta H) \rangle / 2$. Even for a system far from equilibrium, we can consider the action as a total amount of loss in the entropy change rate transfer. Therefore, the action can be interpreted as thermodynamic cost.

Because of the Cauchy-Schwarz inequality $\int_0^\tau dt \int_0^\tau (ds/dt)^2 dt \geq (\int_0^\tau (ds/dt) dt)^2$ [22], we obtain a thermodynamic uncertainty relationship between speed τ and thermodynamic cost $\mathcal{C}$,

$$\tau \geq \frac{\mathcal{L}^2}{2\mathcal{C}}. \qquad (21)$$

The equality holds if the speed of dynamics ds^2/dt^2 does not depend on time. By using the inequality (13), we also have a weaker bound

$$\tau \geq \frac{[\mathcal{D}(\mathbf{r}(0); \mathbf{r}(\tau))]^2}{2\mathcal{C}}. \qquad (22)$$

In a transition from $\mathbf{r}(0)$ to $\mathbf{r}(\tau)[\neq \mathbf{r}(0)]$, the thermodynamic cost $\mathcal{C}$ should be large if the transition time τ is small. In the case of a near-equilibrium system, we have $2\mathcal{C} = \int \langle d\Delta\sigma^{\mathrm{bath}} \rangle$ (or $2\mathcal{C} = -\int \langle d(\beta\Delta H) \rangle$), and then the inequality is similar to the quantum speed limit that is discussed in quantum mechanics [37]. We stress that this result is based on stochastic thermodynamics, not on quantum mechanics.

The inequality (22) gives the ratio between time-averaged thermodynamic cost $2\mathcal{C}/\tau$ and square of the velocity on manifold $([\mathcal{D}(\mathbf{r}(0); \mathbf{r}(\tau))]/\tau)^2$. Then, this ratio

$$\eta := \frac{[\mathcal{D}(\mathbf{r}(0); \mathbf{r}(\tau))]^2}{2\tau\mathcal{C}}, \qquad (23)$$

quantifies an efficiency for the conversion of power to speed. Because of the inequality (22) and its non-negativity, the efficiency η satisfies $0 \leq \eta \leq 1$, where $\eta = 1$ ($\eta = 0$) implies a high (low) efficiency.

Three states model of enzyme reaction.—We numerically illustrate thermodynamic inequalities of information geometry by using a thermodynamic model of a biochemical reaction. Here, we consider a three states model (see also SM [56]) that represents a chemical reaction $A + B \rightleftharpoons AB$ with enzyme X,

$$A + X \rightleftharpoons AX, \qquad (24)$$

$$A + B \rightleftharpoons AB, \qquad (25)$$

$$AX + B \rightleftharpoons AB + X. \qquad (26)$$

Here, we consider the probability distribution of states $x = A, AX, AB$. We assume that the system is attached to a single heat bath with an inverse temperature β (β_{AX}, β_{AB})

for each transition. The master equation is given by Eq. (1), where the transition rates are supposed to be

$$W_{A \to AX}^{(1)} = k_{AX+}[X], \qquad W_{AX \to A}^{(1)} = k_{AX+} e^{-\beta_{AX}\Delta\mu_{AX}},$$

$$W_{A \to AB}^{(1)} = k_{AB+}[B], \qquad W_{AB \to A}^{(1)} = k_{AB+} e^{-\beta_{AB}\Delta\mu_{AB}},$$

$$W_{AX \to AB}^{(1)} = k_+[B], \qquad W_{AB \to AX}^{(1)} = k_+ e^{-\beta\Delta\mu}[X], \qquad (27)$$

$[X]$ ($[B]$) is the concentration of X (B), k_{AX+}, k_{AB+}, and k_+ are the reaction rate constants, and $\Delta\mu_{AX}$, $\Delta\mu_{AB}$, and $\Delta\mu$ are the chemical potential differences. In this model, the entropy change of bath $\Delta\sigma_{x' \to x}^{\mathrm{bath}(\nu)}$ is given by this chemical potential difference (see also SM [56]) [65].

In a numerical simulation, we set $k_{AX+} = k_{AB+} = k_+ = 1$, $\beta_{AX}\Delta\mu_{AX} = 1$, $\beta_{AB}\Delta\mu_{AB} = 0.5$, and $\beta\Delta\mu = 2$. We assume that the time evolution of the concentrations is given by $[X] = \tan^{-1}(\omega_X t)$, $[B] = \tan^{-1}(\omega_B t)$ with $\omega_X = 1$ and $\omega_B = 2$, which means that the concentrations $[X]$ and $[B]$ perform as control parameters. At time $t = 0$, we set the initial probability distribution as $(p_A, p_{AX}, p_{AB}) = (0.9998, 0.0001, 0.0001)$.

In Fig. 2, we numerically show the inequality $\langle d\Delta\sigma^{\mathrm{bath}}/dt \rangle \geq \langle dF/dt \rangle$. We check that this inequality does not coincide with the second law of thermodynamics $\langle F \rangle \geq 0$. We also check the thermodynamic uncertainty relationship $\tau \geq \mathcal{L}^2/(2\mathcal{C})$ in Fig. 3. Because the path from the initial distribution $(p_A, p_{AX}, p_{AB}) = (0.9998, 0.0001, 0.0001)$ to the final distribution is close to the geodesic line, the thermodynamic uncertainty relationship gives a tight bound of the transition time τ.
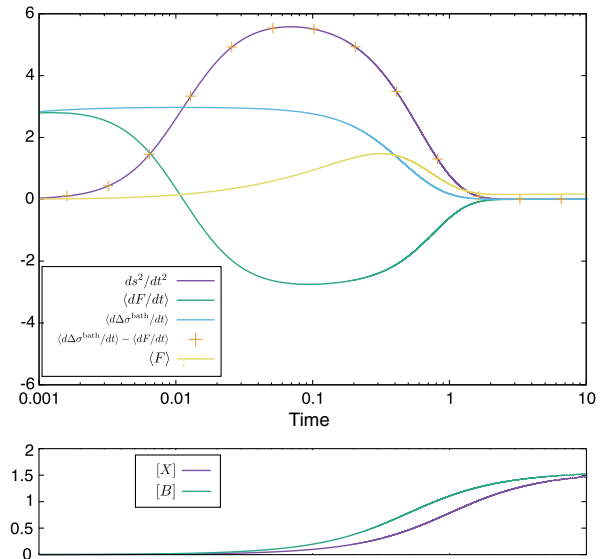

FIG. 2. Numerical calculation of thermodynamic quantities in the three states model of enzyme reaction. We numerically show the non-negativity of $ds^2/dt^2 \geq 0$ and $ds^2/dt^2 = -\langle dF/dt \rangle + \langle d\Delta\sigma^{\mathrm{bath}}/dt \rangle$ in the graph. We also show the total entropy change rate $\langle F \rangle \geq 0$. We note that $d\langle F \rangle/dt$ is not equal to $\langle dF/dt \rangle$.
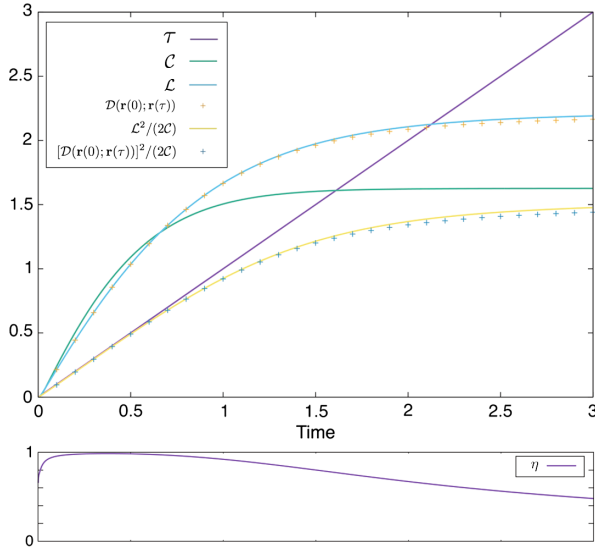
FIG. 3. Numerical calculation of the thermodynamic uncertainty relationship in the three states model of enzyme reaction. We numerically show the geometric inequality $\mathcal{L} \geq \mathcal{D}(\mathbf{r}(0); \mathbf{r}(\tau))$, the thermodynamic uncertainty relationship $\tau \geq \mathcal{L}^2/(2\mathcal{C}) \geq [\mathcal{D}(\mathbf{r}(0); \mathbf{r}(\tau))]^2/(2\mathcal{C})$, and the efficiency η in the graph.

Conclusion.—In this Letter, we reveal a link between stochastic thermodynamic quantities (J, F, $\Delta\sigma^{\text{sys}}$, $\Delta\sigma^{\text{bath}}$) and information geometric quantities (ds^2, $\mathcal{L}$, $\mathcal{D}$, $\mathcal{C}$). Because the theory of information geometry is applicable to various fields of science, such as neuroscience, signal processing, machine learning, and quantum mechanics, this link would help us to understand a thermodynamic aspect of such a topic. The trade-off relationship between speed and thermodynamic cost Eq. (22) would be helpful to understand biochemical reactions, and it gives a new insight into recent studies of the relationship between information and thermodynamics in biochemical processes [7,42,66–70].

Note added.—Recently, I heard that Schuyler B. Nicholson independently discovered a similar result such as Eq. (16) [71].

[1] J. M. Parrondo, J. M. Horowitz, and T. Sagawa, Thermodynamics of information, Nat. Phys. **11**, 131 (2015).
[2] *Maxwell's Demon: Entropy, Information, Computing*, edited by H. S. Leff and A. F. Rex (Princeton University Press, Princeton, NJ, 2014).
[3] R. Kawai, J. M. R. Parrondo, and C. Van den Broeck, Dissipation: The Phase-Space Perspective, Phys. Rev. Lett. **98**, 080602 (2007).
[4] T. Sagawa and M. Ueda, Generalized Jarzynski Equality under Nonequilibrium Feedback Control, Phys. Rev. Lett. **104**, 090602 (2010).
[5] S. Still, D. A. Sivak, A. J. Bell, and G. E. Crooks, Thermodynamics of Prediction, Phys. Rev. Lett. **109**, 120604 (2012).
[6] T. Sagawa and M. Ueda, Fluctuation Theorem with Information Exchange: Role of Correlations in Stochastic Thermodynamics, Phys. Rev. Lett. **109**, 180602 (2012).
[7] S. Ito and T. Sagawa, Information Thermodynamics on Causal Networks, Phys. Rev. Lett. **111**, 180603 (2013).
[8] D. Hartich, A. C. Barato, and U. Seifert, Stochastic thermodynamics of bipartite systems: Transfer entropy inequalities and a Maxwell's demon interpretation, J. Stat. Mech. (2014) P02016.
[9] D. Hartich, A. C. Barato, and U. Seifert, Sensory Capacity: An information theoretical measure of the performance of a sensor, Phys. Rev. E **93**, 022116 (2016).
[10] R. E. Spinney, J. T. Lizier, and M. Prokopenko, Transfer entropy in physical systems and the arrow of time, Phys. Rev. E **94**, 022135 (2016).
[11] S. Ito, Backward transfer entropy: Informational measure for detecting hidden Markov models and its interpretations in thermodynamics, gambling and causality, Sci. Rep. **6**, 36831 (2016).
[12] G. E. Crooks and S. E. Still, Marginal and conditional second laws of thermodynamics, arXiv:1611.04628.
[13] A. E. Allahverdyan, D. Janzing, and G. Mahler, Thermodynamic efficiency of information and heat flow, J. Stat. Mech. (2009) P09011.
[14] J. M. Horowitz and M. Esposito, Thermodynamics with Continuous Information Flow, Phys. Rev. X **4**, 031015 (2014).
[15] J. M. Horowitz and H. Sandberg, Second-law-like inequalities with information and their interpretations, New J. Phys. **16**, 125007 (2014).
[16] N. Shiraishi and T. Sagawa, Fluctuation theorem for partially masked nonequilibrium dynamics, Phys. Rev. E **91**, 012130 (2015).
[17] N. Shiraishi, S. Ito, K. Kawaguchi, and T. Sagawa, Role of measurement-feedback separation in autonomous Maxwell's demons, New J. Phys. **17**, 045012 (2015).
[18] S. Yamamoto, S. Ito, N. Shiraishi, and T. Sagawa, Linear irreversible thermodynamics and Onsager reciprocity for information-driven engines, Phys. Rev. E **94**, 052121 (2016).
[19] S. Goldt and U. Seifert, Stochastic Thermodynamics of Learning, Phys. Rev. Lett. **118**, 010601 (2017).
[20] K. Sekimoto, *Stochastic Energetics* (Springer, New York, 2010).
[21] U. Seifert, Stochastic thermodynamics, fluctuation theorems and molecular machines, Rep. Prog. Phys. **75**, 126001 (2012).

[22] G. E. Crooks, Measuring Thermodynamic Length, Phys. Rev. Lett. **99**, 100602 (2007).

[23] F. H. Edward and G. E. Crooks, Length of Time's Arrow, Phys. Rev. Lett. **101**, 090602 (2008).

[24] M. Polettini and M. Esposito, Nonconvexity of the relative entropy for Markov dynamics: A Fisher information approach, Phys. Rev. E **88**, 012112 (2013).

[25] H. Tajima and M. Hayashi, Finite-size effect on optimal efficiency of heat engines, Phys. Rev. E **96**, 012128 (2017).

[26] H. Shimazaki, Neural Engine Hypothesis, *Dynamic Neuroscience* (Springer, Cham, 2018) pp. 267–291; H. Shimazaki, Neurons as an information-theoretic engine, arXiv:1512.07855.

[27] S. B. Nicholson and E. J. Kim, Investigation of the statistical distance to reach stationary distributions, Phys. Lett. A **379**, 83 (2015).

[28] S. Lahiri, J. Sohl-Dickstein, and S. Ganguli, A universal tradeoff between power, precision and speed in physical communication, arXiv:1603.07758.

[29] F. Weinhold, Metric geometry of equilibrium thermodynamics, J. Chem. Phys. **63**, 2479 (1975).

[30] G. Ruppeiner, Thermodynamics: A Riemannian geometric model, Phys. Rev. A **20**, 1608 (1979).

[31] P. Salamon and R. S. Berry, Thermodynamic Length and Dissipated Availability, Phys. Rev. Lett. **51**, 1127 (1983).

[32] D. A. Sivak and G. E. Crooks, Thermodynamic Metrics and Optimal Paths, Phys. Rev. Lett. **108**, 190602 (2012).

[33] B. B. Machta, Dissipation Bound for Thermodynamic Control, Phys. Rev. Lett. **115**, 260603 (2015).

[34] G. M. Rotskoff, G. E. Crooks, and E. Vanden-Eijnden, Geometric approach to optimal nonequilibrium control: Minimizing dissipation in nanomagnetic spin systems, Phys. Rev. E **95**, 012148 (2017).

[35] S. I. Amari and H. Nagaoka, *Methods of Information Geometry*, Vol. 191 (Oxford University Press, New York, 2000).

[36] M. Oizumi, N. Tsuchiya, and S. I. Amari, Unified framework for information integration based on information geometry, Proc. Natl. Acad. Sci. U.S.A. **113**, 14817 (2016).

[37] D. P. Pires, M. Cianciaruso, L. C. Céleri, G. Adesso, and D. O. Soares-Pinto, Generalized Geometric Quantum Speed Limits, Phys. Rev. X **6**, 021031 (2016).

[38] S. I. Amari, *Information Geometry and its Applications* (Springer, Japan, 2016).

[39] J. Uffink and J. van Lith, Thermodynamic uncertainty relations, Found. Phys. **29**, 655 (1999).

[40] G. Lan, P. Sartori, S. Neumann, V. Sourjik, and Y. Tu, The energy-speed-accuracy trade-off in sensory adaptation, Nat. Phys. **8**, 422 (2012).

[41] C. C. Govern and P. R. ten Wolde, Optimal resource allocation in cellular sensing systems, Proc. Natl. Acad. Sci. U.S.A. **111**, 17486 (2014).

[42] S. Ito and T. Sagawa, Maxwell's demon in biochemical signal transduction with feedback loop, Nat. Commun. **6**, 7498 (2015).

[43] A. C. Barato and U. Seifert, Thermodynamic Uncertainty Relation for Biomolecular Processes, Phys. Rev. Lett. **114**, 158101 (2015).

[44] T. R. Gingrich, J. M. Horowitz, N. Perunov, and J. L. England, Dissipation Bounds All Steady-State Current Fluctuations, Phys. Rev. Lett. **116**, 120601 (2016).

[45] N. Shiraishi, K. Saito, and H. Tasaki, Universal Trade-Off Relation Between Power and Efficiency for Heat Engines, Phys. Rev. Lett. **117**, 190601 (2016).

[46] P. Pietzonka, A. C. Barato, and U. Seifert, Universal bounds on current fluctuations, Phys. Rev. E **93**, 052145 (2016).

[47] A. C. Barato and U. Seifert, Cost and Precision of Brownian Clocks, Phys. Rev. X **6**, 041053 (2016).

[48] J. M. Horowitz and T. R. Gingrich, Proof of the finite-time thermodynamic uncertainty relation for steady-state currents, Phys. Rev. E **96**, 020103 (2017).

[49] K. Proesmans and C. Van den Broeck, Discrete-time thermodynamic uncertainty relation, Europhys. Lett. **119**, 20001 (2017).

[50] C. Maes, Frenetic Bounds on the Entropy Production, Phys. Rev. Lett. **119**, 160601 (2017).

[51] A. Dechant and S. I. Sasa, Current fluctuations and transport efficiency for general Langevin systems, arXiv:1708.08653.

[52] J. Schnakenberg, Network theory of microscopic and macroscopic behavior of master equation systems, Rev. Mod. Phys. **48**, 571 (1976).

[53] D. Andrieux and P. Gaspard, Fluctuation theorem for currents and Schnakenberg network theory, J. Stat. Phys. **127**, 107 (2007).

[54] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (John Wiley & Sons, New York, 2012).

[55] C. R. Rao, Information and the Accuracy Attainable in the Estimation of Statistical Parameters, in *Breakthroughs in Statistics* (Springer, New York, 1992), pp. 235–247.

[56] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.121.030605, for detailed calculations and explanations.

[57] W. K. Wootters, Statistical distance and Hilbert space, Phys. Rev. D **23**, 357 (1981).

[58] S. L. Braunstein and C. M. Caves, Statistical Distance and the Geometry of Quantum States, Phys. Rev. Lett. **72**, 3439 (1994).

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

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

[61] D. J. Evans and D. J. Searles, The fluctuation theorem, Adv. Phys. **51**, 1529 (2002).

[62] U. Seifert, Entropy Production Along a Stochastic Trajectory and an Integral Fluctuation Theorem, Phys. Rev. Lett. **95**, 040602 (2005).

[63] M. Esposito and C. Van den Broeck, Three faces of the second law. I. Master equation formulation, Phys. Rev. E **82**, 011143 (2010).

[64] M. Esposito and C. Van den Broeck, Three Detailed Fluctuation Theorems, Phys. Rev. Lett. **104**, 090601 (2010).

[65] T. Schmiedl and U. Seifert, Stochastic thermodynamics of chemical reaction networks, J. Chem. Phys. **126**, 044101 (2007).

[66] A. C. Barato, D. Hartich, and U. Seifert, Efficiency of cellular information processing, New J. Phys. **16**, 103024 (2014).

[67] P. Sartori, L. Granger, C. F. Lee, and J. M. Horowitz, Thermodynamic costs of information processing in sensory adaptation, PLoS Comput. Biol. **10**, e1003974 (2014).

[68] S. Bo, M. Del Giudice, and A. Celani, Thermodynamic limits to information harvesting by sensory systems, J. Stat. Mech. (2015) P01014.

[69] T. E. Ouldridge, C. C. Govern, and P. R. ten Wolde, Thermodynamics of Computational Copying in Biochemical Systems, Phys. Rev. X **7**, 021004 (2017).

[70] T. McGrath, N. S. Jones, P. R. ten Wolde, and T. E. Ouldridge, Biochemical Machines for the Interconversion of Mutual Information and Work, Phys. Rev. Lett. **118**, 028101 (2017).

[71] S. B. Nicholson, Uncertainty scars and the distance from equilibrium, arXiv:1801.02242.