Information geometric inequalities of chemical thermodynamics

Kohei Yoshimura [1] and Sosuke Ito [1,2]

[1]*Department of Physics, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0031, Japan*
[2]*JST, PRESTO, 4-1-8 Honcho, Kawaguchi, Saitama, 332-0012, Japan*

We study a connection between chemical thermodynamics and information geometry. We clarify a relation between the Gibbs free energy of an ideal dilute solution and an information geometric quantity called an f divergence. From this relation, we derive information geometric inequalities that give a speed limit for the changing rate of the Gibbs free energy and a general bound of chemical fluctuations. These information geometric inequalities can be regarded as generalizations of the Cramér–Rao inequality for chemical reaction networks described by rate equations, where un-normalized concentration distributions are of importance rather than probability distributions. They hold true for damped oscillatory reaction networks and systems where the total concentration is not conserved, so that the distribution cannot be normalized. We also formulate a trade-off relation between speed and time on a manifold of concentration distribution by using the geometrical structure induced by the f divergence. Our results apply to both closed and open chemical reaction networks; thus they are widely useful for thermodynamic analysis of chemical systems from the viewpoint of information geometry.

I. INTRODUCTION

The history of chemical thermodynamics originates around the middle of the nineteenth century [1,2]. The chemical reaction in an ideal dilute solution is one of the main subjects in chemical thermodynamics. For example, the static nature of an ideal dilute solution under near-equilibrium conditions has been well studied since then. After the invention of an area of mathematics called chemical reaction network theory (CRNT) around the 1970s [3,4], its dynamic properties have also been well investigated [5]. One of the most important results of CRNT is that a class of chemical reaction networks called a complex balanced network has a Lyapunov function, which can be associated with thermodynamic quantities such as the Gibbs free energy [3,6,7]. This Lyapunov function is called the pseudo-Helmholtz function, and its time derivative is connected to the entropy production rate [8,9].

On the other hand, thermodynamics for stochastic processes have been well studied as stochastic thermodynamics [10–13]. In stochastic thermodynamics, physical quantities are given by probabilities, and we can discuss relations between thermodynamics and information theory [14–26] because probability plays a crucial role in information theory [27]. For example, in recent years, stochastic thermodynamics met a branch of information theory called information geometry [28–34], and its importance has been verified in recent studies of thermodynamic uncertainty relations [35–43]. An

information geometric quantity called the Fisher information gives several geometric bounds such as the Cramér–Rao bound [27–29], and these bounds indicate thermodynamic uncertainty relations in stochastic thermodynamics [33,44–48].

Although the analogy between stochastic thermodynamics and chemical thermodynamics has been studied [8,9,11,49–57], connections between chemical thermodynamics and information theory are still vague because rate equations which govern chemical reactions are based on un-normalized concentration distributions rather than probability distributions. Nevertheless, a few research efforts have been conducted from the perspective of a connection between chemical thermodynamics and information theory [9,58], focusing on the fact that the pseudo-Helmholtz function has a similar form to the Kullback–Leibler divergence, which plays a fundamental role in information theory.

In this paper, we clarify a connection between chemical thermodynamics and information theory from the viewpoint of information geometry. In information geometry, f divergence is well studied as a measure of the difference between two positive measures. The two positive measures do not have to be necessarily normalized like probability distributions, and the Kullback–Leibler divergence may not be well defined for them. We show that the pseudo-Helmholtz function is given not by the Kullback–Leibler divergence, but by an f divergence. Because the pseudo-Helmholtz function is a representative quantity of a chemical reaction system, this connection reveals how an information geometric concept plays a fundamental role in chemical thermodynamics. Introducing a generalization of the Fisher information from an f divergence and an averagelike quantity which is more suitable to chemical reaction networks than the ordinary average, we obtain a generalization of the Cramér–Rao inequality to

CRNT. This generalized Cramér–Rao inequality indicates the speed limit for the changing rate of the Gibbs free energy in terms of the fluctuation of the chemical potential. We numerically confirm the inequalities for a damped oscillatory reaction network, specifically the Brusselator model [59], and a system where the sum of concentrations is not conserved so that the distribution cannot be normalized. We also examine the geometry of concentration distributions. We formulate trade-off relations between time and speed in terms of information geometry. The trade-off relations are numerically illustrated by solving an association reaction.

This paper is organized as follows. In Sec. II, we formulate chemical reaction networks and introduce the pseudo-Helmholtz function. Section III extends the formulation to open chemical reaction networks (CRNs). Section IV is an introduction to information geometry of both probability spaces and positive measure spaces. We discuss mathematical properties of an f divergence and the connection between the Fisher information and an f divergence. Section V gives information geometric inequalities, which are the main results of this paper. We show that the speed limits for the changing rate of the Gibbs free energy and more general observables are given by the intrinsic speed, the Fisher information, for both closed and open cases. Also, we indicate that the speed limit is regarded as a generalization of the Cramér–Rao inequality. In addition, we examine the geometrical structure of concentrations and reveal trade-off relations between speed and time. In Sec. VI, we confirm our main results through three characteristic models of chemical reaction networks. A conclusion and a further perspective of research are given in Sec. VII.

II. THERMODYNAMICS OF CLOSED CHEMICAL REACTION NETWORKS

A. Kinetics of chemical reaction networks

In this paper, we consider the thermodynamics of a dilute solution with the temperature and pressure kept constant. Since the solvent is dominant, the volume is regarded as a constant. In this chapter, we focus on closed systems.

We consider a CRN consisting of N species of molecules $\{X_i\}_{i=1,2,\ldots,N}$ in a closed vessel. A CRN is defined as a set of M reactions

$$\sum_{i=1}^{N} \nu_{i\rho} X_i \underset{k_\rho^-}{\overset{k_\rho^+}{\rightleftharpoons}} \sum_{i=1}^{N} \kappa_{i\rho} X_i, \tag{1}$$
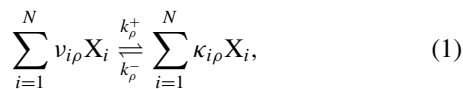
where reactions are labeled with $\rho = 1, 2, \ldots, M$, stoichiometric coefficients $\nu_{i\rho}$ and $\kappa_{i\rho}$ are non-negative integers, and $k_\rho^\pm$ are rate constants. The reactions are assumed to be reversible. In the closed CRN, the time evolution of X_i's concentration $[X_i]$ is described by the following rate equation:

$$\frac{d[X_i]}{dt} = \sum_{\rho=1}^{M} (\kappa_{i\rho} - \nu_{i\rho}) J_\rho, \tag{2}$$

where J_ρ is the reaction rate of the ρth reaction. According to the Waage–Guldberg law of mass action [60], the reaction rate J_ρ is given by

$$J_\rho = J_\rho^+ - J_\rho^-, \tag{3}$$

with

$$J_\rho^+ = k_\rho^+ \prod_{i=1}^{N} [X_i]^{\nu_{i\rho}}, \quad J_\rho^- = k_\rho^- \prod_{i=1}^{N} [X_i]^{\kappa_{i\rho}}, \tag{4}$$

where J_ρ^+ and J_ρ^- are the forward and reverse reaction rates, respectively. Using the coefficients in Eq. (2), we define the stoichiometric matrix of the CRN as a $N \times M$ matrix with its (i, ρ) element $S_{i\rho} := \kappa_{i\rho} - \nu_{i\rho}$, which corresponds to the change in $[X_i]$ when one unit of the ρth reaction proceeds. The rate equation can be written briefly in vector notation as

$$\frac{d[X]}{dt} = SJ. \tag{5}$$

The rate equation (2) has *a priori* conserved quantities. If $\ell \in \mathbb{R}^N$ satisfies

$$\ell^T S = 0^T, \tag{6}$$

i.e., $\ell \in \ker S^T = \{v \mid S^T v = 0\}$, the time derivative of $\ell \cdot [X]$ is zero

$$\frac{d}{dt}(\ell \cdot [X]) = \ell^T SJ = 0, \tag{7}$$

where the superscript T means transposition. Thus $\ell \cdot [X]$ is conserved. We call a left null vector of a stoichiometric matrix a conservation law. Note that a closed CRN has at least one conservation law, corresponding to the conservation of the total mass.

B. Equilibrium conditions

In thermodynamics, it is postulated that a closed system relaxes to equilibrium, at which a certain free energy is minimized depending on the condition. In the present case, the function to be minimized is the Gibbs free energy; thus the equilibrium distribution $[X]^{eq}$ is defined as a distribution that minimizes the Gibbs free energy.

Since the solution is supposed to be dilute, the Gibbs free energy per unit volume G and the chemical potentials μ_i are expressed as

$$G = \sum_{i=1}^{N} \mu_i [X_i] - RT \sum_{i=1}^{N} [X_i] + G_0, \tag{8}$$

$$\mu_i = \frac{\partial G}{\partial [X_i]} = \mu_i^\circ(T) + RT \ln[X_i], \tag{9}$$

where G_0 is a constant, R is the gas constant, and μ_i° are the standard chemical potentials, which are independent of the concentration [2,8,9]. It is plausible to call G the Gibbs free energy because the volume does not change.

Possible values of the concentration are restricted because the concentration changes obeying the rate equation. This can be seen by integrating the rate equation. Let s_ρ be the ρth column vector of the stoichiometric matrix, $S = (s_1, \ldots, s_M)$. Then $[X]$ at t is obtained as

$$[X] = [X]_0 + S \int_0^t dt\, J \tag{10}$$

$$= [X]_0 + \sum_{\rho=1}^{M} \left(\int_0^t dt\, J_\rho \right) s_\rho, \tag{11}$$

where $[\mathbf{X}]_0$ is the initial concentration. Thus a change in concentration $[\mathbf{X}] - [\mathbf{X}]_0$ must be a linear combination of $\{s_\rho\}_{\rho=1,2,...,M}$. So a set of concentrations that $[\mathbf{X}]$ may reach is given by

$$\mathcal{S}([\mathbf{X}]_0) := \{[\mathbf{X}]_0 + \mathsf{S}\boldsymbol{\xi} \mid \boldsymbol{\xi} \in \mathbb{R}^M\} \cap \mathbb{R}_{\geqslant 0}^N, \qquad (12)$$

where $\mathbb{R}_{\geqslant 0}^N$ is the set of N-dimensional vectors with nonnegative elements. This set is called the stoichiometric compatibility class [5], and $\boldsymbol{\xi}$ is the extent of reaction.

A necessary condition for equilibrium can be obtained as the derivative of the Gibbs free energy with respect to ξ_ρ vanishes for all ρ,

$$\frac{\partial G}{\partial \xi_\rho} = \sum_{i=1}^N \frac{\partial [X_i]}{\partial \xi_\rho} \frac{\partial G}{\partial [X_i]} = \sum_{i=1}^N \mathsf{S}_{i\rho} \mu_i^{\mathrm{eq}} = [(\boldsymbol{\mu}^{\mathrm{eq}})^\mathsf{T}\mathsf{S}]_\rho = 0. \qquad (13)$$

So the equilibrium distribution $[\mathbf{X}]^{\mathrm{eq}}$ satisfies

$$\sum_{i=1}^N (\mu_i^\circ + RT \ln[X_i]^{\mathrm{eq}})\mathsf{S}_{i\rho} = 0. \qquad (14)$$

Note that this means that the equilibrium chemical potential $\boldsymbol{\mu}^{\mathrm{eq}}$ is a conservation law.

On the other hand, the equilibrium state is often characterized by the detailed balance, which is based on kinetics,

$$J_\rho^+ = J_\rho^-. \qquad (15)$$

The consistency between the thermodynamic condition of equilibrium (13) and this detailed balance condition leads to a relation

$$\frac{k_\rho^+}{k_\rho^-} = \exp\left(-\frac{(\boldsymbol{\mu}^{\circ\mathsf{T}}\mathsf{S})_\rho}{RT}\right), \qquad (16)$$

which is called the local detailed balance property. It can be regarded as a bridge between thermodynamics and kinetics.

C. Gibbs free energy with divergence

From Eq. (13), $\boldsymbol{\mu}^{\mathrm{eq}}$ is a conservation law; therefore $\boldsymbol{\mu}^{\mathrm{eq}} \cdot [\mathbf{X}]$ is time invariant

$$\frac{d}{dt}(\boldsymbol{\mu}^{\mathrm{eq}} \cdot [\mathbf{X}]) = (\boldsymbol{\mu}^{\mathrm{eq}})^\mathsf{T}\mathsf{S}\boldsymbol{J} = 0. \qquad (17)$$

So the Gibbs free energy at equilibrium G^{eq} can be expressed as

$$G^{\mathrm{eq}} = \boldsymbol{\mu}^{\mathrm{eq}} \cdot [\mathbf{X}]^{\mathrm{eq}} - RT \sum_{i=1}^N [X_i]^{\mathrm{eq}} + G_0 \qquad (18)$$

$$= \boldsymbol{\mu}^{\mathrm{eq}} \cdot [\mathbf{X}] - RT \sum_{i=1}^N [X_i]^{\mathrm{eq}} + G_0 \qquad (19)$$

for a concentration $[\mathbf{X}](\neq [\mathbf{X}]^{\mathrm{eq}})$ at an arbitrary time, namely, in the same stoichiometric compatibility class as $[\mathbf{X}]^{\mathrm{eq}}$. As a result, we obtain the expression

$$G - G^{\mathrm{eq}}$$

$$= (\boldsymbol{\mu} - \boldsymbol{\mu}^{\mathrm{eq}}) \cdot [\mathbf{X}] - RT \sum_{i=1}^N [X_i] + RT \sum_{i=1}^N [X_i]^{\mathrm{eq}} \qquad (20)$$
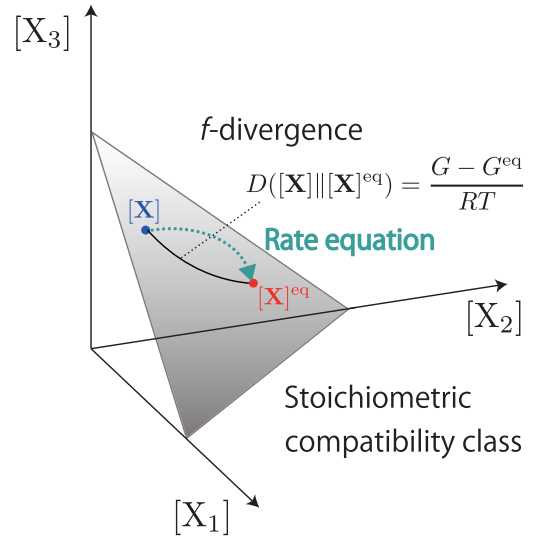

FIG. 1. Schematic of a stoichiometric compatibility class and the f divergence for $N = 3$. This f divergence gives the difference of the Gibbs free energy $(G - G^{\mathrm{eq}})/RT$.

$$= RT \sum_{i=1}^N \left([X_i] \ln \frac{[X_i]}{[X_i]^{\mathrm{eq}}} - [X_i] + [X_i]^{\mathrm{eq}}\right), \qquad (21)$$

by which a function is usually defined as

$$D([\mathbf{X}]\|[\mathbf{X}]^{\mathrm{eq}}) := \sum_{i=1}^N \left([X_i] \ln \frac{[X_i]}{[X_i]^{\mathrm{eq}}} - [X_i] + [X_i]^{\mathrm{eq}}\right). \qquad (22)$$

This suggestive form has been known for more than a half century, and $D([\mathbf{X}]\|[\mathbf{X}]^{\mathrm{eq}})$ is called the pseudo-Helmholtz function [3], Shear's Lyapunov function [6,7], or the relative entropy [9] (see Fig. 1). In terms of information geometry [28], it is regarded as an f divergence of a positive measure space. As we will see in Sec. IV, an f divergence is nonnegative and equal to zero if and only if the two arguments coincide. Therefore G is greater than or equal to G^{eq}, and $G = G^{\mathrm{eq}}$ only in equilibrium, $[\mathbf{X}] = [\mathbf{X}]^{\mathrm{eq}}$.

It is also known that the Gibbs free energy of a closed system never increases under the mass-action kinetics; that is, it is a Lyapunov function of a closed CRN. To show this fact, we calculate the time derivative of the Gibbs free energy

$$\frac{dG}{dt} = \boldsymbol{\mu} \cdot \frac{d[\mathbf{X}]}{dt} = \boldsymbol{\mu}^\mathsf{T}\mathsf{S}\boldsymbol{J}. \qquad (23)$$

From the local detailed balance property [Eq. (16)], $\boldsymbol{\mu}^\mathsf{T}\mathsf{S}$ can be transformed as

$$(\boldsymbol{\mu}^\mathsf{T}\mathsf{S})_\rho = (\boldsymbol{\mu}^{\circ\mathsf{T}}\mathsf{S})_\rho + RT \ln \prod_{i=1}^N [X_i]^{\mathsf{S}_{i\rho}}$$

$$= -RT \ln \frac{k_\rho^+ \prod_{i=1}^N [X_i]^{\nu_{i\rho}}}{k_\rho^- \prod_{i=1}^N [X_i]^{\kappa_{i\rho}}}$$

$$= -RT \ln \frac{J_\rho^+}{J_\rho^-}. \qquad (24)$$

By substituting this into Eq. (23), we have

$$\frac{dG}{dt} = -RT \sum_{\rho=1}^{M} (J_\rho^+ - J_\rho^-) \ln \frac{J_\rho^+}{J_\rho^-} \leqslant 0. \tag{25}$$

The last inequality follows from the fact that the signs of $J_\rho^+ - J_\rho^-$ and $\ln(J_\rho^+/J_\rho^-)$ are always the same for all ρ. Since dG/dt is negative unless the detailed balance is satisfied, and G is always not less than G^{eq}, hence G decreases to G^{eq} monotonically. Here, we point out two facts. First, the left-hand side of the inequality in Eq. (25) coincides with the opposite sign of the entropy production rate, which we define in the next section. Thus the inequality expresses the second law of thermodynamics. Second, the left-hand side is also represented by the f divergence D between $\boldsymbol{J}^+$ and $\boldsymbol{J}^-$

$$\frac{dG}{dt} = -RT[D(\boldsymbol{J}^+\|\boldsymbol{J}^-) + D(\boldsymbol{J}^-\|\boldsymbol{J}^+)]. \tag{26}$$

D. Entropy production rate and affinity

We can formulate the second law of thermodynamics in a CRN [9]. Here, we review a few important points. The entropy production rate σ due to chemical reactions is given by

$$\sigma = R \sum_{\rho=1}^{M} (J_\rho^+ - J_\rho^-) \ln \frac{J_\rho^+}{J_\rho^-}. \tag{27}$$

So it coincides with $-(1/T)dG/dt$ in closed CRNs. As we explained in the preceding section, the entropy production rate is always positive except for the equilibrium state, and that indicates the second law of thermodynamics.

The affinity F_ρ of a reaction is defined as follows [61]:

$$F_\rho := \sum_{i=1}^{N} v_{i\rho}\mu_i - \sum_{i=1}^{N} \kappa_{i\rho}\mu_i = -(\boldsymbol{\mu}^\mathsf{T}\mathsf{S})_\rho. \tag{28}$$

The affinity corresponds to the energy difference between the reactant and product of the ρth reaction. From the local detailed balance property [Eq. (16)], one can rewrite it as

$$F_\rho = RT \ln \frac{J_\rho^+}{J_\rho^-}. \tag{29}$$

Thus the entropy production is expressed as the sum of the product between the reaction rate and the affinity

$$T\sigma = \sum_{\rho=1}^{M} J_\rho F_\rho. \tag{30}$$

This expression shows that there is an analogy with stochastic thermodynamics of master equations [62].

III. THERMODYNAMICS OF OPEN CHEMICAL REACTION NETWORKS

We have already seen the connection between the Gibbs free energy G and the f divergence D in closed CRNs. On the other hand, in open CRNs, the Gibbs free energy is not directly related to the f divergence, even in a CRN called a complex balanced network where the f divergence becomes a Lyapunov function. This is because the chemical potential at a steady state is not a conservation law in general. However, we

can associate the Gibbs free energy with the f divergence. In this section, we formulate open CRNs and associate the Gibbs free energy with the f divergence in two ways: the method already known and the one we newly propose, respectively.

A. Setup for open CRNs

Let $\{Y_j\}_{j=1,2,\dots,N'}$ be the chemical species that are exchanged with the environment and $\{X_i\}_{i=1,2,\dots,N}$ be the other internal species. The former are assumed to be chemostatted; that is, their concentrations are constant.

The CRN consisting of these species is expressed as

$$\sum_{i=1}^{N} v_{i\rho}X_i + \sum_{j=1}^{N'} v_{(N+j)\rho}Y_j$$

$$\underset{k_\rho^-}{\overset{k_\rho^+}{\rightleftharpoons}} \sum_{i=1}^{N} \kappa_{i\rho}X_i + \sum_{j=1}^{N'} \kappa_{(N+j)\rho}Y_j. \tag{31}$$

The stoichiometric matrix $\mathsf{S} = (\kappa_{i\rho} - v_{i\rho})_{i=1,\dots,N+N'}^{\rho=1,\dots,M}$ can be decomposed into N rows of X part and N' rows of Y part

$$\mathsf{S} = \begin{pmatrix} \mathsf{S}^\mathsf{X} \\ \mathsf{S}^\mathsf{Y} \end{pmatrix}, \tag{32}$$

where $\mathsf{S}_{i\rho}^\mathsf{X} = \kappa_{i\rho} - v_{i\rho}$ $(i = 1, 2, \dots, N)$ and $\mathsf{S}_{j\rho}^\mathsf{Y} = \kappa_{(j+N)\rho} - v_{(j+N)\rho}$ $(j = 1, 2, \dots, N')$. Hereafter, for an $N + N'$ row quantity Q, let Q^X be the first N rows and Q^Y be the remainder as in the above case. Since the concentrations of the chemostatted species are assumed to be constant, the dynamics are expressed by the following rate equation:

$$\frac{d[\boldsymbol{Z}]}{dt} := \frac{d}{dt}\begin{pmatrix} [\boldsymbol{X}] \\ [\boldsymbol{Y}] \end{pmatrix} = \begin{pmatrix} \mathsf{S}^\mathsf{X}\boldsymbol{J} \\ \boldsymbol{0} \end{pmatrix}, \tag{33}$$

where $[\boldsymbol{Z}]^\mathsf{T} = ([\boldsymbol{X}]^\mathsf{T}\ \ [\boldsymbol{Y}]^\mathsf{T})$, $J_\rho = J_\rho^+ - J_\rho^-$, and

$$J_\rho^+ = k_\rho^+ \prod_{i=1}^{N}[X_i]^{v_{i\rho}} \prod_{j=1}^{M}[Y_j]^{v_{(j+N)\rho}}, \tag{34}$$

$$J_\rho^- = k_\rho^- \prod_{i=1}^{N}[X_i]^{\kappa_{i\rho}} \prod_{j=1}^{M}[Y_j]^{\kappa_{(j+N)\rho}}. \tag{35}$$

A steady state is defined as a state at which the concentration does not change in time, $\mathsf{S}^\mathsf{X}\boldsymbol{J} = \boldsymbol{0}$.

In open CRNs, $\boldsymbol{\ell} \cdot [\boldsymbol{Z}]$ is not necessarily conserved even if $\boldsymbol{\ell}$ is a conservation law, i.e., $\boldsymbol{\ell}^\mathsf{T}\mathsf{S} = \boldsymbol{0}^\mathsf{T}$ holds. Since we have

$$\frac{d}{dt}(\boldsymbol{\ell} \cdot [\boldsymbol{Z}]) = (\boldsymbol{\ell}^\mathsf{X})^\mathsf{T}\mathsf{S}^\mathsf{X}\boldsymbol{J}, \tag{36}$$

$\boldsymbol{\ell} \cdot [\boldsymbol{Z}]$ is conserved if $\boldsymbol{\ell}^\mathsf{X}$ belongs to $\ker(\mathsf{S}^\mathsf{X})^\mathsf{T}$. Thus a conservation law $\boldsymbol{\ell}$ leads to a conserved quantity if $\boldsymbol{\ell}$ is an element of the linear space $(\ker(\mathsf{S}^\mathsf{X})^\mathsf{T} \times \mathbb{R}^{N'}) \cap \ker \mathsf{S}^\mathsf{T} =: L_\mathsf{X}$, where $\times$ means the direct product between two linear spaces.

B. One way to associate G with D

One way to associate the Gibbs free energy with the f divergence is to decompose a basis of the space of conservation laws $\ker \mathsf{S}^\mathsf{T}$. This formalism is based on Rao and Esposito's paper [9]. It can be used only when the steady state is detailed balanced.

To construct the desired basis, we exploit the linear space L_X. Let $\{\boldsymbol{\ell}^{\lambda_x}\}_{\lambda_x=1,2,\dots,\Lambda_x}$ be a basis of L_X; then we obtain a basis of ker S^T by adding some vectors $\{\boldsymbol{\ell}^{\lambda_{xy}}\}_{\lambda_{xy}=1,2,\dots,\Lambda_{xy}}$. The former vectors lead to quantities $\boldsymbol{\ell}^{\lambda_x} \cdot [\mathbf{Z}]$ that are conserved in an open CRN. We call them X-conservation laws and the latter XY-conservation laws. XY-conservation laws $\boldsymbol{\ell}^{\lambda_{xy}}$ are usually called broken laws because the remainder, X-conservation laws, are always true conservation laws. However, since an XY-conservation law might be a true conservation law, we do not use the conventional terminology.

From the local detailed balance condition [Eq. (16)], we have the following relation [see Eq. (24)]:

$$(\boldsymbol{\mu}^\mathsf{T}\mathsf{S})_\rho = -RT \ln \frac{J_\rho^+}{J_\rho^-}. \tag{37}$$

So if we suppose the steady state to be detailed balanced $J_\rho^+ = J_\rho^-$, i.e., be an equilibrium, the chemical potential at equilibrium $\boldsymbol{\mu}^{eq}$ becomes a conservation law. Then we can expand it with the prepared basis as

$$\boldsymbol{\mu}^{eq} = \sum_{\lambda_x=1}^{\Lambda_x} f_{\lambda_x} \boldsymbol{\ell}^{\lambda_x} + \sum_{\lambda_{xy}=1}^{\Lambda_{xy}} f_{\lambda_{xy}} \boldsymbol{\ell}^{\lambda_{xy}}. \tag{38}$$

A function $\mathcal{G}_1$ is defined by

$$\mathcal{G}_1 := G - \sum_{\lambda_{xy}=1}^{\Lambda_{xy}} f_{\lambda_{xy}} \boldsymbol{\ell}^{\lambda_{xy}} \cdot [\mathbf{Z}], \tag{39}$$

which is called the transformed Gibbs free energy in Ref. [9]. From Eq. (8) and the expansion of $\boldsymbol{\mu}^{eq}$ in Eq. (38), we see that

$$\mathcal{G}_1 = \sum_{i=1}^{N+N'} ([Z_i]\mu_i - RT[Z_i]) + G_0$$
$$- \left(\boldsymbol{\mu}^{eq} \cdot [\mathbf{Z}] - \sum_{\lambda_x=1}^{\Lambda_x} f_{\lambda_x} \boldsymbol{\ell}^{\lambda_x} \cdot [\mathbf{X}] \right). \tag{40}$$

Since the concentrations of the chemostatted species are constant, $[Y_j]$ coincide with $[Y_j]^{eq}$. Therefore we have

$$\mathcal{G}_1 = \sum_{i=1}^{N} \{[X_i](\mu_i - \mu_i^{eq}) - RT[X_i]\} + \text{const} \tag{41}$$
$$= RT \sum_{i=1}^{N} \left([X_i] \ln \frac{[X_i]}{[X_i]^{eq}} - [X_i] + [X_i]^{eq} \right) + \text{const}, \tag{42}$$

where we use the fact that $RT[\mathbf{X}]^{eq}$ is constant. Since the last constant term is equal to $\mathcal{G}_1$ at equilibrium, we write it as $\mathcal{G}_1^{eq}$. We finally obtain the following equation as in the closed CRN:

$$\mathcal{G}_1 = \mathcal{G}_1^{eq} + RT D([\mathbf{X}]\|[\mathbf{X}]^{eq}). \tag{43}$$

As shown in Ref. [9], $\mathcal{G}_1$ gives a bound to the irreversible work to manipulate nonequilibrium distributions. Its time derivative also provides the nonadiabatic entropy production rate as we will see later. The arbitrariness about the choice of the basis yields only a constant term [51]. However, we note that it is only defined for CRNs that relax to detailed balanced steady states. Therefore, if one uses $\mathcal{G}_1$, the number of systems that

can be examined with the f divergence, or in other words, the relative entropy, would be limited.

C. Another way to associate G with D

We newly propose another way of association, which is simpler and more widely applicable than the preceding one. Letting $[\mathbf{Z}]^{ss}$ be the concentration at a steady state and $\boldsymbol{\mu}^{ss} = (\mu_i^\circ + RT \ln[Z_i]^{ss})_{i=1,\dots,N+N'}$ be the chemical potential, we define

$$\mathcal{G}_2 := G - \boldsymbol{\mu}^{ss} \cdot [\mathbf{Z}]. \tag{44}$$

Since $[Y_j]$ coincide with $[Y_j]^{ss}$ as in the detailed balanced case, we have

$$\mathcal{G}_2 = \sum_{i=1}^{N} \{[X_i](\mu_i - \mu_i^{ss}) - RT[X_i]\} + \text{const} \tag{45}$$
$$= \sum_{i=1}^{N} \{[X_i](\mu_i - \mu_i^{ss}) - RT[X_i] + RT[X_i]^{ss}\} + \mathcal{G}_2^{ss} \tag{46}$$
$$= \mathcal{G}_2^{ss} + RT D([\mathbf{X}]\|[\mathbf{X}]^{ss}), \tag{47}$$

where $\mathcal{G}_2^{ss}$ is $\mathcal{G}_2$ at the steady state.

$\mathcal{G}_2$ is defined for general steady states and coincides with $\mathcal{G}_1$ up to a constant if the steady state is detailed balanced. In fact,

$$\mathcal{G}_2 = G - \boldsymbol{\mu}^{eq} \cdot [\mathbf{Z}] \tag{48}$$
$$= \mathcal{G}_1 - \sum_{\lambda_x=1}^{\Lambda_x} f_{\lambda_x} \boldsymbol{\ell}^{\lambda_x} \cdot [\mathbf{Z}] \tag{49}$$

and $\boldsymbol{\ell}^{\lambda_x} \cdot [\mathbf{Z}]$ are constants, so $\mathcal{G}_2 - \mathcal{G}_1 = \text{const}$. Thus $\mathcal{G}_2$ is a generalization of $\mathcal{G}_1$; then we denote $\mathcal{G}_2$ as $\mathcal{G}$ and call it the transformed Gibbs free energy.

The transformed Gibbs free energy we introduce here can be defined for open CRNs that do not satisfy detailed balance. Not all CRNs have steady-state solutions, but our definition in Eq. (44) enables us to study a much broader class of open CRNs than the previous one in Eq. (39) does. $\mathcal{G}_2$ also has the merit that it removes the arbitrariness in the choice of the basis as $\mathcal{G}_1$. For open CRNs with multiple steady states, one can define the transformed Gibbs free energy by choosing a steady state in Eq. (44). Remarkably, the following results hold regardless of the choice of the steady state. We comment that CRNs that do not have steady-state solutions are outside our framework. They include open CRNs that sustain oscillations.

D. Entropy production rate and affinity

The entropy production rate σ and the affinity F_ρ of an open CRN have the same form as those of a closed one. The entropy production rate of an open CRN can be decomposed into the adiabatic and nonadiabatic parts σ_a and σ_{na} [8,9], which are defined as

$$\sigma = \sigma_a + \sigma_{na}, \tag{50}$$
$$\sigma_a := R \sum_{\rho=1}^{M} J_\rho \ln \frac{J_\rho^{+,ss}}{J_\rho^{-,ss}}, \tag{51}$$

$$\sigma_{\mathrm{na}} := R \sum_{\rho=1}^{M} J_\rho \ln \frac{J_\rho^+ J_\rho^{-,\mathrm{ss}}}{J_\rho^- J_\rho^{+,\mathrm{ss}}}, \qquad (52)$$

where $J_\rho^{\pm,\mathrm{ss}}$ are the reaction rates at the steady state. The time derivative of the transformed Gibbs free energy gives the minus sign of the nonadiabatic entropy production rate $d\mathcal{G}/dt = -T\sigma_{\mathrm{na}}$. If the steady state is detailed balanced, or equivalently, $J_\rho^{+,\mathrm{ss}} = J_\rho^{-,\mathrm{ss}}$ holds, $d\mathcal{G}/dt$ can be written by the f divergence between the reaction rates as in the closed case

$$-\frac{d\mathcal{G}}{dt} = RT[D(\boldsymbol{J}^+\|\boldsymbol{J}^-) + D(\boldsymbol{J}^-\|\boldsymbol{J}^+)]. \qquad (53)$$

IV. INFORMATION GEOMETRY

Information geometry deals with a manifold of probability distributions $\boldsymbol{p} = (p_i)_{i=1,2,\dots,N} \in \mathbb{R}_{>0}^N$ that satisfy the normalization condition $\sum_{i=1}^{N} p_i = 1$, or a manifold of positive measures on a discrete set $\boldsymbol{m} = (m_i)_{i=1,2,\dots,N} \in \mathbb{R}_{>0}^N$, which does not have to be normalized [28]. The former manifold is called a probability simplex, and the latter is called a positive measure space. We use the term "distribution" for either a probability distribution or a positive measure in this section.

A. f divergence

A divergence $D(\cdot\|\cdot)$ is a measure of the separation between two distributions $\boldsymbol{m}$ and $\boldsymbol{n}$ that satisfies the following conditions [28]:

(i) $D(\boldsymbol{m}\|\boldsymbol{n}) \geqslant 0$;

(ii) $D(\boldsymbol{m}\|\boldsymbol{n}) = 0 \iff \boldsymbol{m} = \boldsymbol{n}$; and

(iii) $D(\boldsymbol{m}\|\boldsymbol{m} + d\boldsymbol{m}) = \dfrac{1}{2}\sum_{i,j=1}^{N} g_{ij} dm_i dm_j + o(dm^2)$, and

the matrix $(g_{ij})_{1\leqslant i,j\leqslant N}$ is positive definite.

Note that a divergence is similar to a distance function but it is not really a distance function because it is not symmetric: $D(\boldsymbol{m}\|\boldsymbol{n}) \neq D(\boldsymbol{n}\|\boldsymbol{m})$.

One of the well-known divergences is the f divergence [63], which has the following form:

$$D(\boldsymbol{m}\|\boldsymbol{n}) = \sum_{i=1}^{N} m_i f\left(\frac{n_i}{m_i}\right). \qquad (54)$$

The function f has to fulfill some conditions. For both kinds of manifold, f should be a convex differentiable function which satisfies $f(1) = 0$. If one chooses $f(x) = -\ln x$, D becomes the Kullback–Leibler divergence. For positive measure spaces, the condition $f'(1) = 0$ is imposed additionally. f is called a standard convex function when $f''(1) = 1$ holds.

The non-negativity of an f divergence is easily proved. Let $\boldsymbol{p}$ and $\boldsymbol{q}$ be probability distributions in a probability simplex. From Jensen's inequality, we have

$$\sum_{i=1}^{N} p_i f\left(\frac{q_i}{p_i}\right) \geqslant f\left(\sum_{i=1}^{N} p_i \frac{q_i}{p_i}\right) = f(1) = 0. \qquad (55)$$

Therefore an f divergence on a probability simplex is non-negative and equal to zero if and only if $\boldsymbol{p} = \boldsymbol{q}$. On the other hand, for a positive measure manifold, since f is convex and $f'(1) = 0$, f takes the minimum value 0 at $x = 1$. Thus f is non-negative, and so is an f divergence since all the coefficients of f are positive. An f divergence is zero if and only if $f(n_i/m_i) = 0$ for all i. This is equivalent to the two distributions being the same.

B. Fisher information

Letting θ be the parameter of distributions and D be a divergence, the Fisher information $\mathcal{I}(\theta)$ [28] is defined as

$$\mathcal{I}(\theta) := \sum_{i,j=1}^{N} g_{ij} \frac{dm_i}{d\theta} \frac{dm_j}{d\theta}, \qquad (56)$$

$$g_{ij} = \left.\frac{\partial}{\partial m_i} \frac{\partial}{\partial m_j} D(\boldsymbol{n}\|\boldsymbol{m})\right|_{\boldsymbol{n}=\boldsymbol{m}}. \qquad (57)$$

In information geometry,

$$ds^2 := \sum_{i,j=1}^{N} g_{ij} dm_i dm_j \simeq 2D(\boldsymbol{m}\|\boldsymbol{m} + d\boldsymbol{m}) \qquad (58)$$

is interpreted as the square of the line element between two close distributions. If the distributions are parametrized by the time t, the Fisher information becomes

$$\mathcal{I}(t) = \sum_{i,j=1}^{N} g_{ij} \frac{dm_i}{dt} \frac{dm_j}{dt} = \frac{ds^2}{dt^2}; \qquad (59)$$

so we define the intrinsic speed on the manifold ds/dt as

$$\frac{ds}{dt} := \sqrt{\mathcal{I}(t)}. \qquad (60)$$

An f divergence with a standard convex function leads to $g_{ij} = m_i^{-1}\delta_{ij}$, where δ_{ij} is Kronecker's delta; then the Fisher information always has the unique form

$$\mathcal{I}(\theta) = \sum_{i=1}^{N} \frac{1}{m_i}\left(\frac{dm_i}{d\theta}\right)^2. \qquad (61)$$

A significant fact related to the Fisher information is the Cramér–Rao inequality for a probability distribution $\boldsymbol{p}(\theta)$ [29]

$$\mathrm{Var}(\hat{\theta}) \geqslant \frac{1}{\mathcal{I}(\theta)}, \qquad (62)$$

where $\hat{\theta}$ is an unbiased estimator of θ (that is, $\langle\hat{\theta}\rangle_\theta := \sum_{i=1}^{N} p_i(\theta)\hat{\theta}_i = \theta$ holds) and $\mathrm{Var}(\hat{\theta})$ is the variance of $\hat{\theta}$, $\langle(\hat{\theta} - \langle\hat{\theta}\rangle_\theta)^2\rangle_\theta$.

V. INFORMATION GEOMETRY IN CHEMICAL THERMODYNAMICS

A. Geometrical structure of chemical thermodynamics

A set of concentrations can be interpreted as a positive measure space in both closed and open CRNs. The measures are concentrations $[\boldsymbol{Z}]$, and the measurable set is the index set of species. This space is thought to have the *a priori* f divergence, with its standard convex function

$$f(x) = -\ln x + x - 1,$$

$$D([\mathbf{X}]\|[\mathbf{X}']) = \sum_{i=1}^{N} [X_i] f\left(\frac{[X_i']}{[X_i]}\right)$$

$$= \sum_{i=1}^{N} \left([X_i]\ln\frac{[X_i]}{[X_i']} - [X_i] + [X_i']\right), \quad (63)$$

because G of a closed CRN and $\mathcal{G}$ of an open CRN are described by the f divergence as

$$G = G^{\text{eq}} + RT D([\mathbf{X}]\|[\mathbf{X}]^{\text{eq}}) \quad (64)$$

and

$$\mathcal{G} = \mathcal{G}^{\text{ss}} + RT D([\mathbf{X}]\|[\mathbf{X}]^{\text{ss}}). \quad (65)$$

Let us confirm that f is a standard convex function. It is obviously smooth, and it is convex because $f''(x) = 1/x^2 > 0$. It is readily seen that the values at $x = 1$ are $f(1) = f'(1) = 0$ and $f''(1) = 1$. Hence f is a standard convex function.

As we pointed out in Sec. IV, a divergence gives a geometrical structure to a manifold, namely, a metric g_{ij}. An f divergence always provides the metric of the form $g_{ij} = \delta_{ij} m_i^{-1}$. Therefore manifolds of concentration distributions are considered to be equipped with the metric $g_{ij} = \delta_{ij}[X_i]^{-1}$. The metric is given naturally with respect to only internal species. This is because the concentrations of the chemostatted species are kept constant and do not appear in the divergence in Eq. (65). It is possible to extend the metric as $g_{ij} = \delta_{ij}[Z_i]^{-1}$ for $i, j \in \{1, 2, \ldots, N + N'\}$, but we just neglect the chemostatted species in this paper when we consider the geometrical structure of CRNs. Hence we only say here that the concentration distributions of closed and open CRNs have completely the same metric $g_{ij} = \delta_{ij}[X_i]^{-1}$, where $i, j \in \{1, 2, \ldots, N\}$.

Before stating our main results, we denote a few direct consequences of the geometrical structure. First, the square of the line element between the equilibrium concentration and a concentration close to it is related to the Gibbs free energy. If a concentration is close to equilibrium $[\mathbf{X}] = [\mathbf{X}]^{\text{eq}} + \delta[\mathbf{X}]$, the Gibbs free energy is expressed as $G = G^{\text{eq}} + \delta G$ with a small deviation δG. From Eq. (64), we obtain the expression of the square of the line element $ds^2 = 2D([\mathbf{X}]\|[\mathbf{X}]^{\text{eq}})$ by the Gibbs free energy

$$ds^2 = \frac{2}{RT}\delta G. \quad (66)$$

So the Gibbs free energy difference corresponds to the square of the distance from the equilibrium concentration under near-equilibrium conditions. This is true for open CRNs; that is, the fluctuation of the transformed Gibbs free energy from the steady-state value $\delta\mathcal{G}$ corresponds to ds^2 between the steady state and a concentration nearby $ds^2 = (2/RT)\delta\mathcal{G}$. Note that although the relations hold only under near-equilibrium or near-steady-state conditions, our results in the following sections apply to far-from-equilibrium systems except for Eqs. (74)–(76).

Second, we can consider the Fisher information of chemical reaction networks. Since the metric is $\delta_{ij}/[X_i]$ for closed and open CRNs and $d[Y_j]/dt = 0$, the Fisher information

$\mathcal{I}(t)$ is defined for both types of CRN as

$$\mathcal{I}(t) = \sum_{i=1}^{N} \frac{1}{[X_i]}\left(\frac{d[X_i]}{dt}\right)^2. \quad (67)$$

It is also represented as

$$\mathcal{I}(t) = -\frac{1}{RT}\sum_{\rho=1}^{M} J_\rho \frac{dF_\rho}{dt}. \quad (68)$$

We prove that this formula is true in both cases. Because $(d/dt)\ln[X_i] = (1/RT)(d\mu_i/dt)$, $\mathcal{I}(t)$ is transformed as follows:

$$\mathcal{I}(t) = \frac{1}{RT}\sum_{i=1}^{N} \frac{d[X_i]}{dt}\frac{d\mu_i}{dt} \quad (69)$$

$$= \frac{1}{RT}\sum_{i=1}^{N}\sum_{\rho=1}^{M} \frac{d\mu_i}{dt}\mathsf{S}_{i\rho} J_\rho. \quad (70)$$

If the CRN is closed, that ends the proof since $F_\rho = -(\boldsymbol{\mu}^\mathsf{T}\mathsf{S})_\rho$. On the other hand, if it is open,

$$\mathcal{I}(t) = \frac{1}{RT}\sum_{i=1}^{N+N'}\sum_{\rho=1}^{M} \frac{d\mu_i}{dt}\mathsf{S}_{i\rho} J_\rho \quad (71)$$

$$= -\frac{1}{RT}\sum_{\rho=1}^{M} \frac{dF_\rho}{dt} J_\rho, \quad (72)$$

where we used the fact that the chemical potentials of the chemostatted species do not change in time.

From the expression of the entropy production rate [Eq. (27)], which is valid for closed and open CRNs, Eq. (68) is also written as

$$\mathcal{I}(t) = -\frac{1}{R}\frac{d\sigma}{dt} + \frac{1}{RT}\sum_{\rho=1}^{M} \frac{dJ_\rho}{dt}F_\rho. \quad (73)$$

Therefore the Fisher information is associated with the entropy production rate with the additional term. Furthermore, under near-equilibrium conditions, $\mathcal{I}(t)$ is directly given by $d\sigma/dt$ with an error of higher order in δJ. That is,

$$\mathcal{I}(t) = -\frac{1}{2R}\frac{d\sigma}{dt} + o(\delta J^2) \quad (74)$$

is obtained. This is because the affinity is proportional to the reaction rate under near-equilibrium conditions. It can be proved as follows. $J_\rho^\pm$ can be decomposed into the equilibrium values $J_\rho^{+,\text{eq}} = J_\rho^{-,\text{eq}} =: J_\rho^{\text{eq}}$ and the fluctuations $\delta J_\rho^\pm$ as $J_\rho^\pm = J_\rho^{\text{eq}} + \delta J_\rho^\pm$. Then the affinities are obtained as $F_\rho = RT J_\rho/J_\rho^{\text{eq}} + o(J_\rho)$, where $J_\rho = \delta J_\rho^+ - \delta J_\rho^-$. So, we have

$$\sum_{\rho=1}^{M} \frac{dJ_\rho}{dt}F_\rho \simeq \sum_{\rho=1}^{M} RT \frac{J_\rho}{J_\rho^{\text{eq}}}\frac{dJ_\rho}{dt} \quad (75)$$

$$\simeq \frac{1}{2}\frac{d}{dt}\left(\sum_{\rho=1}^{M} J_\rho F_\rho\right) = \frac{T}{2}\frac{d\sigma}{dt}. \quad (76)$$

By substituting this into Eq. (73), we obtain the equality in Eq. (74).

B. Speed limit on the Gibbs free energy

We describe the main result in this section as follows. We state, discuss, and prove our assertion for closed CRNs. Then we obtain the same result in subsystems and open CRNs. This is because the discussion becomes concise without loss of generality in closed CRNs. This method of description is followed in the next section.

To state our main results, we define the *concentration integral* $\langle\!\langle \cdot \rangle\!\rangle$ of a quantity $\boldsymbol{q} = (q_1, q_2, \ldots, q_N)^\mathsf{T}$ as

$$\langle\!\langle \boldsymbol{q} \rangle\!\rangle := \sum_{i=1}^{N} q_i [\mathrm{X}_i]. \tag{77}$$

A concentration integral coincides with an average if the weight is normalized. It might be possible to use not the concentration but the normalized concentration $[\mathrm{X}_i]/\sum_{i=1}^{N}[\mathrm{X}_i]$ or the mole fraction as the weight. However, usually neither of them obeys any tractable differential equations such as the rate equation. Since we would like to consider time evolution, we use a concentration integral.

We further define the chemical variance of a chemical potential $\langle\!\langle \Delta \boldsymbol{\mu}^2 \rangle\!\rangle$ as

$$\langle\!\langle \Delta \boldsymbol{\mu}^2 \rangle\!\rangle := \langle\!\langle (\boldsymbol{\mu} - \boldsymbol{\mu}^{\mathrm{eq}})^2 \rangle\!\rangle = \sum_{i=1}^{N} \left(\mu_i - \mu_i^{\mathrm{eq}} \right)^2 [\mathrm{X}_i], \tag{78}$$

where $(\boldsymbol{\mu} - \boldsymbol{\mu}^{\mathrm{eq}})^2$ means the vector $[(\mu_i - \mu_i^{\mathrm{eq}})^2]_{i=1,2,\ldots,N}$. It is a variancelike quantity but differs from a variance in two aspects. One is that the weight is not normalized; namely, it is defined by the concentration integral. The other is that what is subtracted from the chemical potential $\boldsymbol{\mu}$ is not an average but the equilibrium chemical potential $\boldsymbol{\mu}^{\mathrm{eq}}$.

One of our main results is the fact that the time derivative of the Gibbs free energy is bounded above by the product between the Fisher information and the chemical variance of a chemical potential. First, for closed CRNs, it is the inequality

$$\left| \frac{dG}{dt} \right| \leqslant \sqrt{\mathcal{I}(t)} \sqrt{\langle\!\langle \Delta \boldsymbol{\mu}^2 \rangle\!\rangle}, \tag{79}$$

where $|\cdot|$ means the absolute value. This inequality gives an upper bound of the speed at which the Gibbs free energy decreases with the information geometric quantity, the Fisher information $\mathcal{I}(t)$. The inequality is rewritten equivalently as

$$\frac{1}{\mathcal{I}(t)} \left(\frac{dG}{dt} \right)^2 \leqslant \langle\!\langle \Delta \boldsymbol{\mu}^2 \rangle\!\rangle, \tag{80}$$

which is similar to the Cramér–Rao inequality [Eq. (62)]. We define a function of a quantity $\boldsymbol{q}$, a reference value $\bar{\boldsymbol{q}}$, and the time t as

$$v_{\boldsymbol{q}}(t, \bar{\boldsymbol{q}}) := \sqrt{\mathcal{I}(t)} \sqrt{\langle\!\langle (\boldsymbol{q} - \bar{\boldsymbol{q}})^2 \rangle\!\rangle}, \tag{81}$$

where $(\boldsymbol{q} - \bar{\boldsymbol{q}})^2 = [(q_i - \bar{q}_i)^2]_{i=1,2,\ldots,N}$; then Eq. (79) can be rewritten as

$$\left| \frac{dG}{dt} \right| \leqslant v_{\boldsymbol{\mu}}(t, \boldsymbol{\mu}^{\mathrm{eq}}). \tag{82}$$

The proof of Eq. (79) is straightforward. From Eq. (64), we have

$$\frac{dG}{dt} = RT \sum_{i=1}^{N} \frac{d[\mathrm{X}_i]}{dt} \ln \frac{[\mathrm{X}_i]}{[\mathrm{X}_i]^{\mathrm{eq}}} \tag{83}$$

$$= \sum_{i=1}^{N} \frac{d[\mathrm{X}_i]}{dt} \left(\mu_i - \mu_i^{\mathrm{eq}} \right); \tag{84}$$

then, using the Cauchy–Schwarz inequality, the inequality is obtained as follows:

$$\left| \frac{dG}{dt} \right| = \left| \sum_{i=1}^{N} \frac{1}{\sqrt{[\mathrm{X}_i]}} \frac{d[\mathrm{X}_i]}{dt} \sqrt{[\mathrm{X}_i]} \left(\mu_i - \mu_i^{\mathrm{eq}} \right) \right| \tag{85}$$

$$\leqslant \sqrt{\sum_{i=1}^{N} \frac{1}{[\mathrm{X}_i]} \left(\frac{d[\mathrm{X}_i]}{dt} \right)^2} \sqrt{\sum_{i=1}^{N} [\mathrm{X}_i] \left(\mu_i - \mu_i^{\mathrm{eq}} \right)^2}. \tag{86}$$

Here, we do not use any approximation or assumption of near-equilibrium; thus the speed limit holds far from equilibrium.

The above discussion can be extended to subsystems. If we are concerned with some specific species $S = \{\mathrm{X}_i\}_{i \in A_S} \subset \{\mathrm{X}_1, \mathrm{X}_2, \ldots, \mathrm{X}_N\}$, where A_S is the index set of S, we can define the partial Gibbs free energy G_S as

$$G_S := RT \sum_{i \in A_S} \left([\mathrm{X}_i] \ln \frac{[\mathrm{X}_i]}{[\mathrm{X}_i]^{\mathrm{eq}}} - [\mathrm{X}_i] + [\mathrm{X}_i]^{\mathrm{eq}} \right). \tag{87}$$

The inequality in Eq. (79) also holds for this partial Gibbs free energy

$$\left| \frac{dG_S}{dt} \right| \leqslant \sqrt{\mathcal{I}_S(t)} \sqrt{\langle\!\langle \Delta \boldsymbol{\mu}^2 \rangle\!\rangle_S} =: v_{\boldsymbol{\mu}, S}(t, \boldsymbol{\mu}^{\mathrm{eq}}), \tag{88}$$

where

$$\mathcal{I}_S(t) := \sum_{i \in A_S} \frac{1}{[\mathrm{X}_i]} \left(\frac{d[\mathrm{X}_i]}{dt} \right)^2, \tag{89}$$

$$\langle\!\langle \Delta \boldsymbol{\mu}^2 \rangle\!\rangle_S := \sum_{i \in A_S} (\mu_i - \mu_i^{\mathrm{eq}})^2 [\mathrm{X}_i]. \tag{90}$$

The partial Gibbs free energy can show nontrivial behavior, e.g., oscillation, even in a closed CRN, while the Gibbs free energy of a total system decreases monotonically. Hence this bound [Eq. (88)] also becomes a nontrivial one. The proof is almost the same as that of Eq. (79).

Furthermore, the speed limit obtains in even open CRNs for the transformed Gibbs free energy $\mathcal{G}$. The time derivative of $\mathcal{G}$ is bounded above as the Gibbs free energy of closed systems was by the Fisher information and the deviation of chemical potential

$$\left| \frac{d\mathcal{G}}{dt} \right| \leqslant \sqrt{\mathcal{I}_{S_\mathrm{X}}(t)} \sqrt{\langle\!\langle (\boldsymbol{\mu} - \boldsymbol{\mu}^{\mathrm{ss}})^2 \rangle\!\rangle_{S_\mathrm{X}}}, \tag{91}$$

where $S_\mathrm{X} = \{\mathrm{X}_1, \ldots, \mathrm{X}_N\}$ and $(\boldsymbol{\mu} - \boldsymbol{\mu}^{\mathrm{ss}})^2$ is the vector $[(\mu_i - \mu_i^{\mathrm{ss}})^2]_{i=1,2,\ldots,N+N'}$. Since the transformed Gibbs free energy is given by the f divergence, one can prove the inequality in Eq. (91) easily. This inequality obtains in any open CRN which has a steady-state solution.

We comment on the experimental importance of the speed limit. The inequality in Eq. (79) can be written as follows:

$$\frac{1}{\sqrt{\langle\!\langle\Delta\boldsymbol{\mu}^2\rangle\!\rangle}}\left|\frac{dG}{dt}\right| \leqslant \frac{ds}{dt}. \tag{92}$$

The denominator of the left-hand side is seen as the entire driving force of the CRN, and the numerator is the corresponding changing rate. Thus the ratio itself can be interpreted as something like the transport coefficient in the linear response theory. The inequality in Eq. (92) shows that it is suppressed by the intrinsic speed on the stoichiometric compatibility class ds/dt. Hence, if ds/dt of a CRN is small, we can see that the CRN responds weakly to a change in chemical potentials. Since the intrinsic speed ds/dt needs only the current concentration and the concentration's time derivative, it is experimentally obtained more easily than the transport-coefficient-like quantity that needs a standard chemical potential or equilibrium concentration. In addition, it is remarkable that the information geometric speed ds/dt can be defined and considered in CRNs with multistability and sustained oscillation, though these systems are out of the range of our study.

C. Generalized Cramér–Rao inequality for chemical reaction networks

We consider the result in the previous section further. If we use the original definition of the Gibbs free energy [Eq. (8)], the time derivative of G is given by

$$\frac{dG}{dt} = \sum_{i=1}^{N}\frac{d[X_i]}{dt}\mu_i, \tag{93}$$

while in Eq. (84) there is an additional term $-\sum_{i=1}^{N}(d[X_i]/dt)\mu_i^{\mathrm{eq}}$. This is because $\boldsymbol{\mu}^{\mathrm{eq}}$ satisfies Eq. (13) and thus is orthogonal to $d[\mathbf{X}]/dt = \mathsf{S}\boldsymbol{J}$. Hence it is crucial for the speed limit [Eq. (79)] that $\boldsymbol{\mu}^{\mathrm{eq}}$ is a conservation law. This suggests that we can make use of conservation laws to evaluate the time derivative of a concentration integral.

For example, letting $\bar{\boldsymbol{q}}$ be a conservation law, we can add $0 = -\bar{\boldsymbol{q}}\cdot(d[\mathbf{X}]/dt)$ to the time derivative of $\langle\!\langle\boldsymbol{q}\rangle\!\rangle$

$$\frac{d}{dt}\langle\!\langle\boldsymbol{q}\rangle\!\rangle = \boldsymbol{q}\cdot\frac{d[\mathbf{X}]}{dt} = (\boldsymbol{q}-\bar{\boldsymbol{q}})\cdot\frac{d[\mathbf{X}]}{dt}. \tag{94}$$

Then, we obtain the following inequality in the same way as the proof of Eq. (79):

$$\left|\frac{d}{dt}\langle\!\langle\boldsymbol{q}\rangle\!\rangle\right| \leqslant \sqrt{\mathcal{I}(t)}\sqrt{\langle\!\langle(\boldsymbol{q}-\bar{\boldsymbol{q}})^2\rangle\!\rangle} = v_{\boldsymbol{q}}(t,\bar{\boldsymbol{q}}), \tag{95}$$

or equivalently,

$$\frac{1}{\mathcal{I}(t)}\left(\frac{d}{dt}\langle\!\langle\boldsymbol{q}\rangle\!\rangle\right)^2 \leqslant \langle\!\langle(\boldsymbol{q}-\bar{\boldsymbol{q}})^2\rangle\!\rangle. \tag{96}$$

We may call this inequality the generalized Cramér–Rao inequality for CRNs. Equation (95) reveals the fact that the Fisher information acts as a speed limit not only on the Gibbs free energy but also on general quantities.

We can construct the $\bar{\boldsymbol{q}}$ ($=:\bar{\boldsymbol{q}}^{\mathrm{min}}$) that minimizes $v_{\boldsymbol{q}}(t,\cdot)$. We introduce a diagonal matrix X
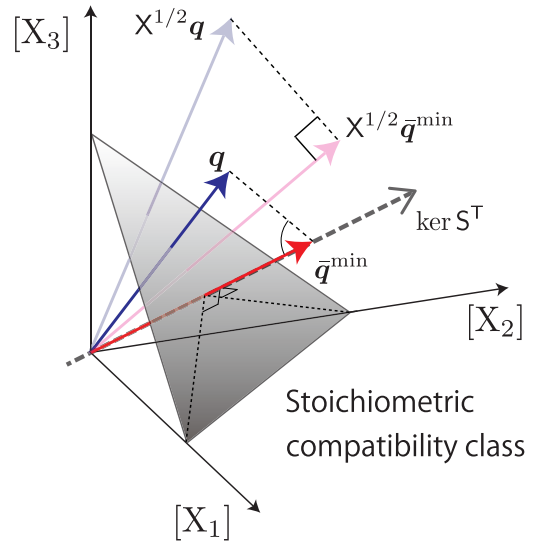
$$\mathsf{X}_{ij} := \delta_{ij}[X_i]; \tag{97}$$

FIG. 2. Geometric picture of $\boldsymbol{q}$ and other quantities appearing in Sec. V C for $N = 3$. $\bar{\boldsymbol{q}}^{\mathrm{min}}$, which makes the right-hand sides of the inequalities in Eqs. (95) and (96) smallest, is obtained as a projection of $\boldsymbol{q}$ onto $\ker\mathsf{S}^{\mathsf{T}}$, which is perpendicular to the stoichiometric compatibility class. It is not the orthogonal projection in the Euclidean space. Instead, $\mathsf{X}^{1/2}\boldsymbol{q}$ is orthogonally projected to $\mathsf{X}^{1/2}\bar{\boldsymbol{q}}^{\mathrm{min}}$.

then we can rewrite $\langle\!\langle(\boldsymbol{q}-\bar{\boldsymbol{q}})^2\rangle\!\rangle$ as

$$\sum_{i=1}^{N}(q_i-\bar{q}_i)^2[X_i] = \|\mathsf{X}^{1/2}(\boldsymbol{q}-\bar{\boldsymbol{q}})\|^2 \tag{98}$$

$$= \|\mathsf{X}^{1/2}\boldsymbol{q} - \mathsf{X}^{1/2}\bar{\boldsymbol{q}}\|^2, \tag{99}$$

where $\|\cdot\|$ is the Euclidean norm. Since the linear space that $\mathsf{X}^{1/2}\bar{\boldsymbol{q}}$ belongs to is

$$\mathsf{X}^{1/2}\ker\mathsf{S}^{\mathsf{T}} := \{\mathsf{X}^{1/2}\boldsymbol{\ell}\mid\boldsymbol{\ell}\in\ker\mathsf{S}^{\mathsf{T}}\}, \tag{100}$$

we see that Eq. (99) is smallest when $\mathsf{X}^{1/2}\bar{\boldsymbol{q}}$ is the orthogonal projection of $\mathsf{X}^{1/2}\boldsymbol{q}$ onto $\mathsf{X}^{1/2}\ker\mathsf{S}^{\mathsf{T}}$ (see Fig. 2). Therefore, letting $\{\mathsf{X}^{1/2}\boldsymbol{\ell}^\lambda\}_{\lambda=1,2,\dots,\Lambda}$ be an orthonormal basis of $\mathsf{X}^{1/2}\ker\mathsf{S}^{\mathsf{T}}$, $\bar{\boldsymbol{q}}^{\mathrm{min}}$ is given by

$$\bar{\boldsymbol{q}}^{\mathrm{min}} = \mathsf{X}^{-1/2}\sum_{\lambda=1}^{\Lambda}(\mathsf{X}^{1/2}\boldsymbol{\ell}^\lambda\cdot\mathsf{X}^{1/2}\boldsymbol{q})\mathsf{X}^{1/2}\boldsymbol{\ell}^\lambda \tag{101}$$

$$= \sum_{\lambda=1}^{\Lambda}[(\boldsymbol{\ell}^\lambda)^{\mathsf{T}}\mathsf{X}\boldsymbol{q}]\boldsymbol{\ell}^\lambda. \tag{102}$$

These results can be restricted to subsystems as in Eq. (88). For a subset of species S, we have

$$\left|\frac{d}{dt}\langle\!\langle\boldsymbol{q}\rangle\!\rangle_S\right| \leqslant v_{\boldsymbol{q},S}(t,\bar{\boldsymbol{q}}) \tag{103}$$

$$:= \sqrt{\mathcal{I}_S(t)}\sqrt{\sum_{i\in A_S}(q_i-\bar{q}_i)^2[X_i]}. \tag{104}$$

The right-hand side is minimized when $\bar{q}_i = \bar{q}_i^{\mathrm{min}}$ for all $i\in A_S$.

Under near-equilibrium conditions, it can be proved that $\boldsymbol{\mu}^{\mathrm{eq}}$ coincides with $\bar{\boldsymbol{\mu}}^{\mathrm{min}}$ within an error of second order of the deviation $\Delta[\mathbf{X}] := [\mathbf{X}] - [\mathbf{X}]^{\mathrm{eq}}$. For any $\boldsymbol{\ell} \in \ker \mathsf{S}^{\mathsf{T}}$,

$$(\mathsf{X}^{1/2}(\boldsymbol{\mu} - \boldsymbol{\mu}^{\mathrm{eq}})) \ \cdot \ \mathsf{X}^{1/2}\boldsymbol{\ell} \tag{105}$$

$$= RT \sum_{i=1}^{N} [\mathrm{X}_i] \ln \frac{[\mathrm{X}_i]}{[\mathrm{X}_i]^{\mathrm{eq}}} \ell_i \tag{106}$$

$$= RT \sum_{i=1}^{N} \ell_i([\mathrm{X}_i] - [\mathrm{X}_i]^{\mathrm{eq}}) + \mathcal{O}(\Delta[\mathbf{X}]^2) \tag{107}$$

$$= RT \Delta(\boldsymbol{\ell} \cdot [\mathbf{X}]) + \mathcal{O}(\Delta[\mathbf{X}]^2), \tag{108}$$

and $\boldsymbol{\ell} \cdot [\mathbf{X}]$ does not change in time; therefore $\mathsf{X}^{1/2}(\boldsymbol{\mu} - \boldsymbol{\mu}^{\mathrm{eq}})$ is approximately orthogonal to $\mathsf{X}^{1/2} \ker \mathsf{S}^{\mathsf{T}}$. This means that $\mathsf{X}^{1/2}\boldsymbol{\mu}$ is orthogonally projected to $\mathsf{X}^{1/2}\boldsymbol{\mu}^{\mathrm{eq}} \in \mathsf{X}^{1/2} \ker \mathsf{S}^{\mathsf{T}}$; therefore we see that $\boldsymbol{\mu}^{\mathrm{eq}} \simeq \bar{\boldsymbol{\mu}}^{\mathrm{min}}$. Then the speed limit [Eq. (79)] is the tightest Cramér–Rao bound under near-equilibrium conditions with an error of second order of the concentration deviation.

Finally, let us extend the results to open CRNs. For a quantity $\boldsymbol{q} \in \mathbb{R}^{N+N'}$, a speed limit on the change in $\langle\langle\boldsymbol{q}\rangle\rangle$ is given by

$$\frac{d}{dt}\langle\langle\boldsymbol{q}\rangle\rangle = \sum_{i=1}^{N} q_i \sum_{\rho=1}^{M} \mathsf{S}_{i\rho} J_\rho \tag{109}$$

because $d[\mathrm{Y}_j]/dt = 0$. Thus it is bounded by quantities related to the internal species

$$\left| \frac{d}{dt}\langle\langle\boldsymbol{q}\rangle\rangle \right| \leqslant \sqrt{\mathcal{I}_{S_{\mathrm{X}}}(t)} \sqrt{\langle\langle(\boldsymbol{q} - \bar{\boldsymbol{q}})^2\rangle\rangle_{S_{\mathrm{X}}}}, \tag{110}$$

where $\bar{\boldsymbol{q}}$ is an element of $\ker(\mathsf{S}^{\mathsf{X}})^{\mathsf{T}} \times \mathbb{R}^{N'}$. The generalized Cramér–Rao bound for an open CRN is also obtained as

$$\frac{1}{\mathcal{I}_{S_{\mathrm{X}}}(t)} \left(\frac{d}{dt}\langle\langle\boldsymbol{q}\rangle\rangle \right)^2 \leqslant \langle\langle(\boldsymbol{q} - \bar{\boldsymbol{q}})^2\rangle\rangle_{S_{\mathrm{X}}}. \tag{111}$$

D. Geometry of stoichiometric compatibility class

In addition to the above results, we consider an information geometric aspect of stoichiometric compatibility classes. Both closed and open CRNs are considered in the same notation because the chemostatted species do not affect the geometrical structure. A concentration distribution is confined to a stoichiometric compatibility class. We can consider this confinement to be a consequence of constraints $\boldsymbol{\ell}^\lambda \cdot [\mathbf{X}] = L^\lambda$(const), besides the discussion in Sec. II. This is reminiscent of the situation of probability distributions where they are restricted by the normalization condition $\sum_i p_i = 1$. In fact, this constraint is a key to deduce basic results in information geometry, such as Čencov's theorem [28,64]. On the other hand, the more generic restrictions of chemical reactions need a nontrivial extension of information geometry.

We indicate an information geometric characterization of stoichiometric compatibility classes. Let $r_i := \sqrt{[\mathrm{X}_i]}$ and $a_i^\lambda := \sqrt{L^\lambda/|\ell_i^\lambda|}$. Note that L^λ can be always non-negative by choosing $\boldsymbol{\ell}^\lambda$ properly. Then the metric with respect to r_i is the Euclidean because

$$ds^2 = \sum_{i=1}^{N} \frac{1}{[\mathrm{X}_i]} d[\mathrm{X}_i]^2 = \sum_{i=1}^{N} (2dr_i)^2. \tag{112}$$

If $L^\lambda \neq 0$, the constraint on concentration distributions becomes

$$\sum_{i=1}^{N} \mathrm{sgn}(\ell_i^\lambda) \left(\frac{r_i}{a_i^\lambda} \right)^2 = 1, \tag{113}$$

which is an equation of a quadric surface, such as an ellipsoid or a hyperboloid. If $L^\lambda = 0$, the constraint represents a cone that contains the origin $[\mathbf{X}] = \mathbf{0}$ as a specific case of quadric surfaces. This consideration shows that the stoichiometric compatibility class can be considered as the intersection between such quadric surfaces and the positive orthant.

We also consider the geometry of reaction dynamics on a stoichiometric compatibility class. The length $\mathcal{L}$ of the path between the concentration at $t = 0$, $[\mathbf{X}]_{t=0}$, and that at $t = \tau$, $[\mathbf{X}]_{t=\tau}$, is defined as

$$\mathcal{L} := \int_0^\tau \frac{ds}{dt} dt = \int_0^\tau \sqrt{\sum_{i=1}^{N} \frac{1}{[\mathrm{X}_i]} \left(\frac{d[\mathrm{X}_i]}{dt} \right)^2} dt. \tag{114}$$

From the speed limit on the Gibbs free energy [Eq. (79)], the transport-coefficient-like quantity gives a lower bound to $\mathcal{L}$

$$\mathcal{L} \geqslant \int_0^\tau \frac{1}{\sqrt{\langle\langle\Delta\boldsymbol{\mu}^2\rangle\rangle}} \left| \frac{dG}{dt} \right| dt. \tag{115}$$

Moreover, applying the Cauchy–Schwarz inequality to $\mathcal{L}$, we obtain the inequality

$$\mathcal{L}^2 \leqslant 2\tau\mathcal{C}, \tag{116}$$

where $\mathcal{C}$ is called the thermodynamic cost and is defined as

$$\mathcal{C} := \frac{1}{2} \int_0^\tau dt\mathcal{I}(t). \tag{117}$$

Here, $1/2$ is a conventional coefficient like the $1/2$ in the action of a free particle in analytical mechanics. As shown in Sec. V A, under near-equilibrium conditions, the Fisher information of a closed CRN is expressed as

$$\mathcal{I}(t) = -\frac{1}{2R} \frac{d\sigma}{dt} + o(\Delta[\mathbf{X}]^2); \tag{118}$$

thus its integral is

$$\int_0^\tau dt\mathcal{I}(t) = \frac{\sigma(t=0) - \sigma(t=\tau)}{2R} + o(\Delta[\mathbf{X}]^2). \tag{119}$$

Since this quantity has the same dimension as reaction rates J_ρ and $d[\mathrm{X}_i]/dt$, it may represent how fast the system relaxes in the time interval τ even far from equilibrium. Hence we can interpret $\mathcal{C}$ as the mean relaxation rate. Then, we have a more informative expression of the inequality in Eq. (116),

$$\tau \geqslant \frac{\mathcal{L}^2}{2\mathcal{C}}. \tag{120}$$

It shows a trade-off relation between time and the mean relaxation rate in CRNs. One might think that $\mathcal{L}$ and $\mathcal{C}$ have the same information because they differ just by the powers of integrands except for the coefficients. However, we cannot derive such a connection between $\mathcal{L}$ and the entropy production rate, so the two quantities represent different aspects of a system.

Apart from the dynamics, one can define the shortest distance between two distributions $[\mathbf{X}]$ and $[\mathbf{X}']$ in the same stoichiometric compatibility class as

$$\mathcal{D} := \inf_{\gamma} \int_0^\tau \frac{ds}{dt} dt = \inf_{\gamma} \int_0^\tau \sqrt{\sum_{i=1}^N \frac{1}{\gamma_i} \left(\frac{d\gamma_i}{dt}\right)^2} dt, \quad (121)$$

where the infimum is taken over the paths of concentration distribution that satisfy the initial and final conditions and are contained in the same stoichiometric compatibility class as $[\mathbf{X}]$ and $[\mathbf{X}']$, namely, $\{\boldsymbol{\gamma} : [0, \tau] \to \mathcal{S}([\mathbf{X}]) \,|\, \boldsymbol{\gamma}(0) = [\mathbf{X}], \boldsymbol{\gamma}(\tau) = [\mathbf{X}']\}$. Since the shortest distance always serves as a lower bound of the length $\mathcal{L}$, another trade-off relation obtains

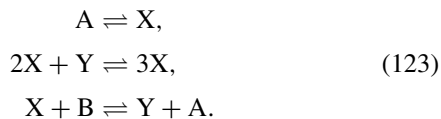$$\tau \geqslant \frac{\mathcal{D}^2}{2\mathcal{C}}. \quad (122)$$

Although this inequality is a weaker bound than Eq. (120), it gives a lower bound to the mean relaxation rate $\mathcal{C} \geqslant \mathcal{D}^2/2\tau$, which needs only information of the time interval and the initial and final distributions.

VI. EXAMPLES

Through three examples of CRNs, we check our results: the speed limits [Eqs. (79) and (88)], the generalized Cramér–Rao inequality [Eq. (95)], and the trade-off relations [Eqs. (120) and (122)].

A. Speed limit in damped Brusselator

The first example is the Brusselator [59,65], which is a notable model of oscillating reactions such as the Belousov–Zhabotinsky reaction. We consider the following CRN:

$$\begin{aligned} \mathrm{A} &\rightleftharpoons \mathrm{X}, \\ 2\mathrm{X} + \mathrm{Y} &\rightleftharpoons 3\mathrm{X}, \\ \mathrm{X} + \mathrm{B} &\rightleftharpoons \mathrm{Y} + \mathrm{A}. \end{aligned} \quad (123)$$

Then the concentrations obey the rate equations below:

$$\begin{aligned} \frac{d[\mathrm{X}]}{dt} &= J_1 - J_3, \\ \frac{d[\mathrm{Y}]}{dt} &= -J_2 + J_3, \\ \frac{d[\mathrm{A}]}{dt} &= -J_1 + J_3, \\ \frac{d[\mathrm{B}]}{dt} &= -J_3, \end{aligned} \quad (124)$$

where

$$J_1 = k_1^+[\mathrm{A}] - k_1^-[\mathrm{X}], \quad (125)$$
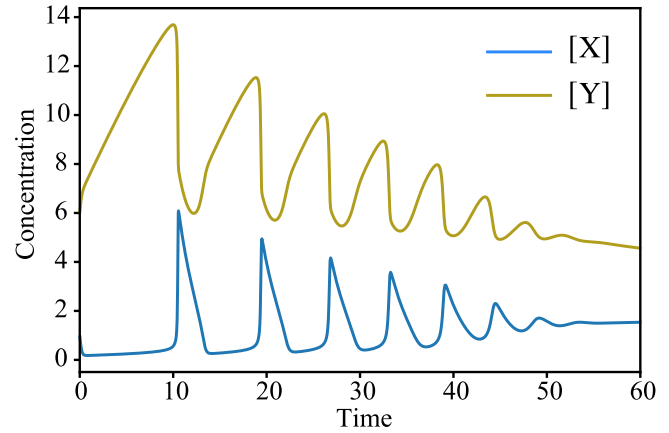

FIG. 3. The time evolutions of X's and Y's concentrations obtained by integrating Eq. (124) with the parameters $k_1^+ = 1 \times 10^{-3}$, $k_1^- = k_2^+ = k_2^- = 1$, $k_3^+ = 1 \times 10^{-2}$, $k_3^- = 1 \times 10^{-4}$, $[\mathrm{X}]_0 = 1$, $[\mathrm{Y}]_0 = 6$, and $[\mathrm{A}]_0 = [\mathrm{B}]_0 = 1 \times 10^3$. They oscillate and then relax to equilibrium.

$$J_2 = k_2^+[\mathrm{X}]^2[\mathrm{Y}] - k_2^-[\mathrm{X}]^3, \quad (126)$$

$$J_3 = k_3^+[\mathrm{X}][\mathrm{B}] - k_3^-[\mathrm{Y}][\mathrm{A}]. \quad (127)$$

In a usual Brusselator model, the change in two species, X and Y, is of interest, so the others are assumed to be constant because of the abundance. On the other hand, because we consider the system to be closed, we do not set [A] and [B] constant, but sufficiently large to observe the damped oscillation of [X] and [Y].

While the Brusselator exhibits a damped oscillation of concentration as in Fig. 3, $|dG/dt|$ is suppressed by $v_\mu(t, \boldsymbol{\mu}^{\mathrm{eq}})$ as shown in Fig. 4. The speed limit [Eq. (79)] is verified for the damped oscillatory CRN. Also the inequality for the partial Gibbs free energy [Eq. (88)] can be confirmed. Focusing on $S = \{\mathrm{X}, \mathrm{Y}\}$, we show its appearance in Fig. 5.

There are situations where two curves are close, i.e., the inequality is tight. That occurs when the chemical potential changes exponentially. Because we used the Cauchy–Schwarz
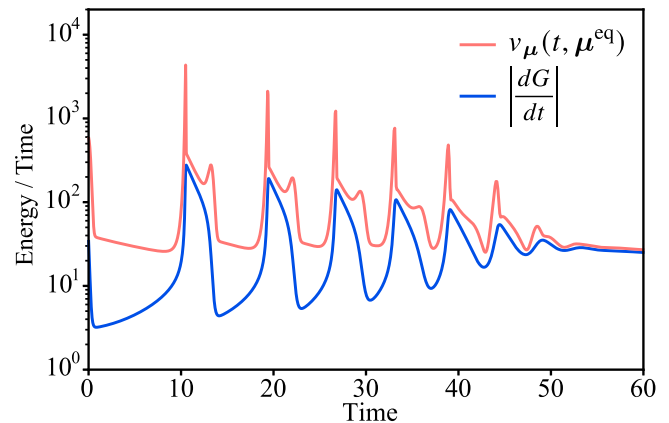

FIG. 4. The speed limit [Eq. (79)] with respect to the reaction system (123). The speed of the Gibbs free energy change cannot exceed $v_\mu(t, \boldsymbol{\mu}^{\mathrm{eq}})$.
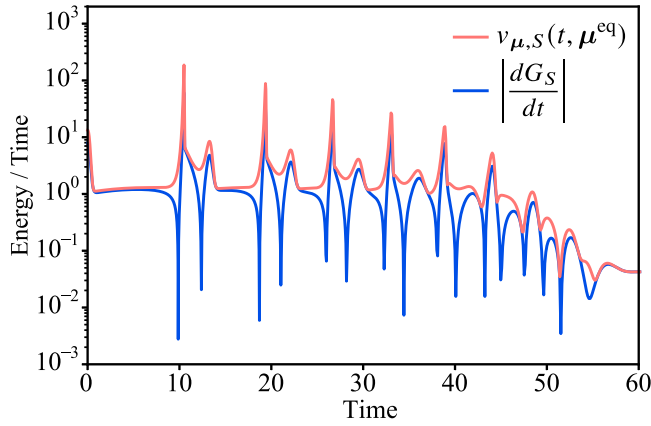
FIG. 5. The speed limit [Eq. (88)] for a subset $S = \{X, Y\}$. The speed of the partial Gibbs free energy change also does not exceed the corresponding function $v_{\mu,S}(t, \boldsymbol{\mu}^{\mathrm{eq}})$.

inequality to prove the speed limit in Eq. (86), equality holds when there exists a constant α such that for all i,

$$\frac{1}{\sqrt{[X_i]}} \frac{d[X_i]}{dt} = \alpha \sqrt{[X_i]} (\mu_i - \mu_i^{\mathrm{eq}}). \qquad (128)$$

This is equivalent to

$$\frac{d}{dt}(\mu_i - \mu_i^{\mathrm{eq}}) = \alpha'(\mu_i - \mu_i^{\mathrm{eq}}), \qquad (129)$$

where $\alpha' = RT\alpha$. Thus the equality holds when the deviation of the chemical potential $\mu_i - \mu_i^{\mathrm{eq}}$ is proportional to the exponential function $e^{\alpha' t}$. If the constant α' is negative, this condition means an exponential decay to equilibrium. Under near-equilibrium conditions, the system is assumed to relax to equilibrium exponentially. Hence $\boldsymbol{\mu}^{\mathrm{eq}}$ should be the minimizer of the speed limit. We proved that $\boldsymbol{\mu}^{\mathrm{eq}} \simeq \bar{\boldsymbol{\mu}}^{\mathrm{min}}$ under near-equilibrium conditions in Sec. V C. Therefore the equality condition is consistent with the previous discussion.

B. Generalized Cramér–Rao inequality in a model where the total concentration does not conserve

We confirm the validity of the generalized Cramér–Rao inequality [Eq. (95)] furthermore by observing the following CRN:

$$\begin{aligned} 2A &\rightleftharpoons B, \\ A + B &\rightleftharpoons B + C. \end{aligned} \qquad (130)$$
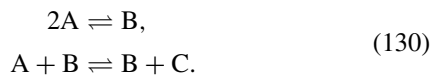
This CRN is simple but sufficient to break the conservation of total concentration as shown in Fig. 6. Note that the preceding Brusselator model looks complicated but preserves the total concentration, $[X] + [Y] + [A] + [B]$; thus the concentration can be normalized by dividing this constant.

The stoichiometric matrix is

$$\mathsf{S} = \begin{pmatrix} -2 & -1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad (131)$$
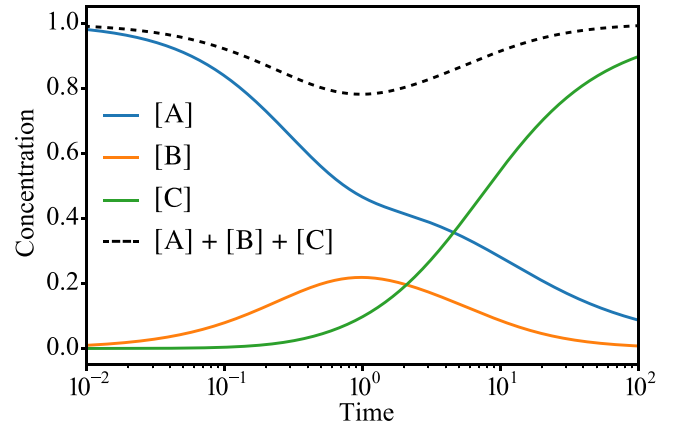

FIG. 6. The time evolution of the concentrations in the CRN in Eq. (130) calculated with the parameters $k_1^+ = k_1^- = k_2^+ = 1$, $k_2^- = 1 \times 10^{-5}$, $[A]_0 = 1$, and $[B]_0 = [C]_0 = 1 \times 10^{-5}$. The total concentration $[A] + [B] + [C]$ (dashed line) is not constant.

so that the conservation law is only

$$\boldsymbol{\ell} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \qquad (132)$$

up to a scale factor. Therefore $[A] + 2[B] + [C] =: L$ becomes a constant instead of the total concentration.

If we set $\boldsymbol{q} = (1, 1, 1)^{\mathsf{T}}$, $\langle\!\langle \boldsymbol{q} \rangle\!\rangle$ is the total concentration

$$\langle\!\langle \boldsymbol{q} \rangle\!\rangle = [A] + [B] + [C]. \qquad (133)$$

The projection $\bar{\boldsymbol{q}}^{\mathrm{min}}$ is given by

$$\bar{\boldsymbol{q}}^{\mathrm{min}} = \frac{\boldsymbol{\ell}^{\mathsf{T}} \mathsf{X} \boldsymbol{q}}{\boldsymbol{\ell}^{\mathsf{T}} \mathsf{X} \boldsymbol{\ell}} \boldsymbol{\ell} = \frac{L}{[A] + 4[B] + [C]} \boldsymbol{\ell}, \qquad (134)$$

where a denominator $\boldsymbol{\ell}^{\mathsf{T}} \mathsf{X} \boldsymbol{\ell}$ appears, unlike in Eq. (102), because $\mathsf{X}^{1/2} \boldsymbol{\ell}$ is not a unit vector here. Then the change in the total concentration is bounded as in Fig. 7. To compare with the tightest bound $v_q(t, \bar{\boldsymbol{q}}^{\mathrm{min}})$, the bound given by the trivial conservation law $\bar{\boldsymbol{q}} = \boldsymbol{0}$, $v_q(t, 0)$, is presented together.
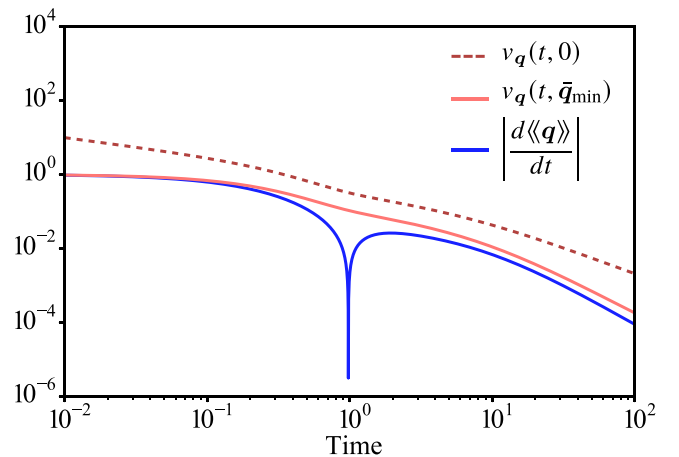

FIG. 7. Bounds on the changing rate of the total concentration $\langle\!\langle \boldsymbol{q} \rangle\!\rangle$. The bound given by $\bar{\boldsymbol{q}}^{\mathrm{min}}$ is much tighter than that given by $\bar{\boldsymbol{q}} = \boldsymbol{0}$.
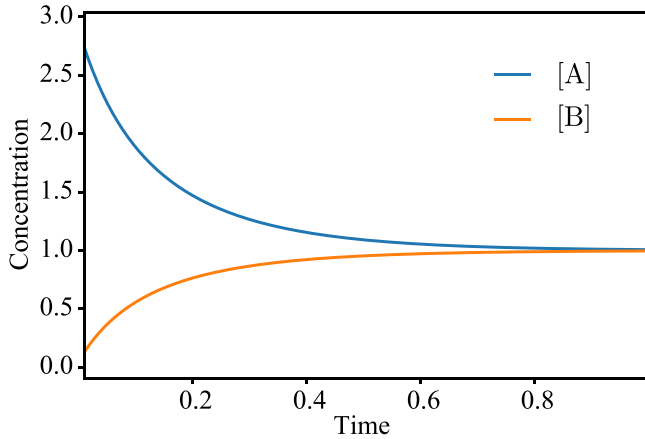
FIG. 8. The time evolution of the concentrations in the association reaction. The parameters are set to $k^+ = k^- = 1$, $[A]_0 = 2.9$, and $[B]_0 = 0.05$.

C. Trade-off relations on stoichiometric compatibility class

We illustrate our discussion in Sec. V D by considering the following association reaction:

$$2A \rightleftharpoons B. \tag{135}$$

For this CRN, shown in Fig. 8, we can obtain the distance $\mathcal{D}$ analytically. We denote the conserved quantity $[A] + 2[B]$ as L. For new variables $r_1 = 2\sqrt{[A]}$ and $r_2 = 2\sqrt{[B]}$, we can write the metric and the constraint as

$$ds^2 = (dr_1)^2 + (dr_2)^2, \tag{136}$$

$$\left(\frac{r_1}{2\sqrt{L}}\right)^2 + \left(\frac{r_2}{\sqrt{2L}}\right)^2 = 1. \tag{137}$$

Because this constraint represents an elliptic, the coordinate can be parametrized by a parameter θ as $(r_1, r_2)_{t=0} = (2\sqrt{L}\cos\theta_1, \sqrt{2L}\sin\theta_1)$, $(r_1, r_2)_{t=\tau} = (2\sqrt{L}\cos\theta_2, \sqrt{2L}\sin\theta_2)$. Then the distance is obtained as

$$\mathcal{D} = \left| \int_{\theta_1}^{\theta_2} d\theta \sqrt{\left(\frac{dr_1}{d\theta}\right)^2 + \left(\frac{dr_2}{d\theta}\right)^2} \right| \tag{138}$$

$$= 2\sqrt{L} \left| \int_{\theta_1}^{\theta_2} d\theta \sqrt{1 - \frac{1}{2}\cos^2\theta} \right| \tag{139}$$

$$= 2\sqrt{L} \left| E\left(\theta_2; \frac{1}{\sqrt{2}}\right) - E\left(\theta_1; \frac{1}{\sqrt{2}}\right) \right|, \tag{140}$$

where $E(x; k) := \int_0^x d\theta \sqrt{1 - k^2 \cos^2\theta}$ is the incomplete elliptic integral of the second kind.

We introduce the following notation to check the trade-off relation for time t and time interval τ:

$$\mathcal{L}_\tau(t) := \int_t^{t+\tau} dt' \frac{ds}{dt'}, \tag{141}$$

$$\mathcal{C}_\tau(t) := \frac{1}{2} \int_t^{t+\tau} dt' \mathcal{I}(t'), \tag{142}$$

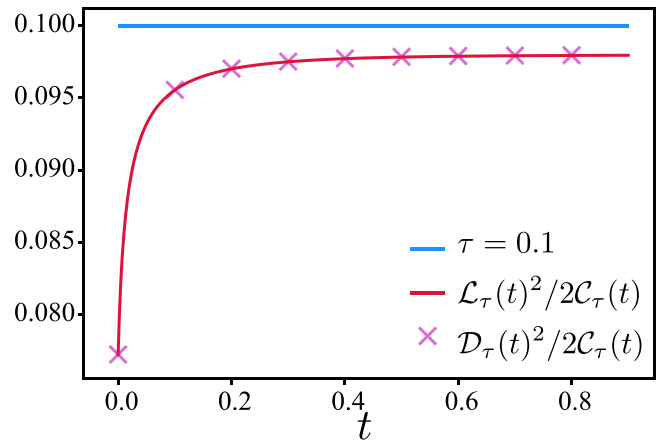$$\mathcal{D}_\tau(t) := \inf_\gamma \int_t^{t+\tau} dt' \frac{ds}{dt'}, \tag{143}$$

FIG. 9. The trade-off relations in Eqs. (120) and (122) are verified for the CRN in Eq. (135). For fixed $\tau = 0.1$, $\mathcal{L}_\tau^2/2\mathcal{C}_\tau$ and $\mathcal{D}_\tau^2/2\mathcal{C}_\tau$ give a lower bound. Because the concentration change is monotonic, $\mathcal{L}_\tau$ and $\mathcal{D}_\tau$ give the same bound.

where the infimum is taken over γ's that satisfy $\gamma(t) = [\mathbf{X}]_t$ and $\gamma(t + \tau) = [\mathbf{X}]_{t+\tau}$ and are contained in the same stoichiometric compatibility class as $[\mathbf{X}]_t$. For the fixed time interval $\tau = 0.1$, the trade-off relations are shown in Fig. 9. Since the relaxation is monotonic, the length $\mathcal{L}$ and the distance $\mathcal{D}$ should coincide with each other. In Fig. 9, $\mathcal{L}_\tau(t)^2/2\mathcal{C}_\tau(t)$ and $\mathcal{D}_\tau(t)^2/2\mathcal{C}_\tau(t)$ are actually the same.

VII. CONCLUSION

We have studied thermodynamics of chemical reaction networks in terms of information geometry. We have revealed that geometrical structure and the Fisher information can be obtained in CRNs by using information geometry. Then we have derived speed limits in CRNs, e.g., Eqs. (79), (91), and (95). Our results are not restricted to near-equilibrium conditions but hold even if the CRN is open or far from equilibrium. This broad range of application shows the universality of the speed limit. It has been shown that the speed limit can be interpreted as a generalization of the Cramér–Rao inequality [Eq. (96)] outside probability spaces. These results are mainly based on the form of the Gibbs free energy that includes the f divergence $D([\mathbf{X}]\|[\mathbf{X}]^{eq})$ and the conservation quantities dwelling in a CRN. We have further obtained a trade-off relation [Eq. (120)] between time and speed in CRNs examining the geometry of stoichiometric compatibility classes.

Our study provides a new perspective on chemical thermodynamics in terms of information geometry. It offers a framework to analyze the thermodynamic profile of biological systems. The use of information geometric measures such as the intrinsic speed ds/dt or the length of reactions $\mathcal{L}$ would bring a new perspective to informatic aspects of biology. Our results can be used if one can obtain the concentration distribution, so the range of application would be wide.

There is a more theoretical question. Though the information geometry of chemical thermodynamics is brought by the f divergence, one can ask whether there are more fundamental reasons why information geometry is applicable to CRNs. CRNs are just a kind of dynamical system and have nothing

to do with probability in its formulation. However, the fact that they can be studied by using information geometry, which is usually useful in probability theory, would be a clue to investigate the link between various dynamical systems in nature and information.

[1] J. W. Gibbs, On the equilibrium of heterogeneous substances, Am. J. Sci. Arts **s3-16**, 441 (1878).

[2] D. Kondepudi and I. Prigogine, *Modern Thermodynamics: From Heat Engines to Dissipative Structures* (Wiley, New York, 2014).

[3] F. Horn and R. Jackson, General mass action kinetics, Arch. Ration. Mech. Anal. **47**, 81 (1972).

[4] M. Feinberg, Complex balancing in general kinetic systems, Arch. Ration. Mech. Anal. **49**, 187 (1972).

[5] M. Feinberg, *Foundations of Chemical Reaction Network Theory* (Springer, New York, 2019).

[6] D. Shear, An analog of the Boltzmann H-theorem (a Liapunov function) for systems of coupled chemical reactions, J. Theor. Biol. **16**, 212 (1967).

[7] J. Higgins, Some remarks on Shear's Liapunov function for systems of chemical reactions, J. Theor. Biol. **21**, 293 (1968).

[8] H. Ge and H. Qian, Nonequilibrium thermodynamic formalism of nonlinear chemical reaction systems with Waage–Guldberg's law of mass action, Chem. Phys. **472**, 241 (2016).

[9] R. Rao and M. Esposito, Nonequilibrium Thermodynamics of Chemical Reaction Networks: Wisdom from Stochastic Thermodynamics, Phys. Rev. X **6**, 041064 (2016).

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

[11] T. Schmiedl and U. Seifert, Efficiency at maximum power: An analytically solvable model for stochastic heat engines, J. Chem. Phys. **126**, 044101 (2007).

[12] K. Sekimoto, *Stochastic Energetics* (Springer, New York, 2010).

[13] U. Seifert, Stochastic thermodynamics, fluctuation theorems and molecular machines, Rep. Prog. Phys. **75**, 126001 (2012).

[14] R. Kawai, J. M. R. Parrondo, and C. Van den Broeck, Dissipation: The Phase-Space Perspective, Phys. Rev. Lett. **98**, 080602 (2007).

[15] A. E. Allahverdyan, D. Janzing, and G. Mahler, Thermodynamic efficiency of information and heat flow, J. Stat. Mech.: Theory Exp. (2009) P09011.

[16] T. Sagawa and M. Ueda, Generalized Jarzynski Equality under Nonequilibrium Feedback Control, Phys. Rev. Lett. **104**, 090602 (2010).

[17] S. Toyabe, T. Sagawa, M. Ueda, E. Muneyuki, and M. Sano, Experimental demonstration of information-to-energy conversion and validation of the generalized Jarzynski equality, Nat. Phys. **6**, 988 (2010).

[18] S. Still, D. A. Sivak, A. J. Bell, and G. E. Crooks, Thermodynamics of Prediction, Phys. Rev. Lett. **109**, 120604 (2012).

[19] T. Sagawa and M. Ueda, Fluctuation Theorem with Information Exchange: Role of Correlations in Stochastic Thermodynamics, Phys. Rev. Lett. **109**, 180602 (2012).

[20] S. Ito and T. Sagawa, Information Thermodynamics on Causal Networks, Phys. Rev. Lett. **111**, 180603 (2013).

[21] J. M. Horowitz and M. Esposito, Thermodynamics with Continuous Information Flow, Phys. Rev. X **4**, 031015 (2014).

[22] 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.: Theory Exp. (2014) P02016.

[23] J. M. R. Parrondo, J. M. Horowitz, and T. Sagawa, Thermodynamics of information, Nat. Phys. **11**, 131 (2015).

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

[25] N. Shiraishi and T. Sagawa, Fluctuation theorem for partially masked nonequilibrium dynamics, Phys. Rev. E **91**, 012130 (2015).

[26] M. L. Rosinberg and J. M. Horowitz, Continuous information flow fluctuations, EPL **116**, 10007 (2016).

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

[28] S.-i. Amari and H. Nagaoka, *Methods of Information Geometry*, Translations of Mathematical Monographs, Vol. 191 (Oxford University Press, New York, 2000).

[29] C. R. Rao, Information and the accuracy attainable in the estimation of statistical parameters, Bull. Calcutta Math. Soc. **37**, 81 (1945), originally printed in 1945 and reprinted as *Breakthroughs in Statistics* (Springer, New York, 1992).

[30] G. Ruppeiner, Riemannian geometry in thermodynamic fluctuation theory, Rev. Mod. Phys. **67**, 605 (1995).

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

[32] G. M. Rotskoff and G. E. Crooks, Optimal control in nonequilibrium systems: Dynamic Riemannian geometry of the Ising model, Phys. Rev. E **92**, 060102(R) (2015).

[33] S. Ito, Stochastic Thermodynamic Interpretation of Information Geometry, Phys. Rev. Lett. **121**, 030605 (2018).

[34] S. Ito, M. Oizumi, and S.-i. Amari, Unified framework for the entropy production and the stochastic interaction based on information geometry, Phys. Rev. Res. **2**, 033048, (2020).

[35] J. M. Horowitz and T. R. Gingrich, Thermodynamic uncertainty relations constrain non-equilibrium fluctuations, Nat. Phys. **16**, 15 (2019).

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

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

[38] 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).

[39] M. Polettini, A. Lazarescu, and M. Esposito, Tightening the uncertainty principle for stochastic currents, Phys. Rev. E **94**, 052104 (2016).

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

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

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

[43] A. Dechant and S. I. Sasa, Current fluctuations and transport efficiency for general Langevin systems, J. Stat. Mech.: Theory Exp. (2018) 063209.

[44] A. Dechant, Multidimensional thermodynamic uncertainty relations, J. Phys. A: Math. Theor. **52**, 035001 (2018).

[45] Y. Hasegawa and T. Van Vu, Uncertainty relations in stochastic processes: An information inequality approach, Phys. Rev. E **99**, 062126 (2019).

[46] S. Ito and A. Dechant, Stochastic Time Evolution, Information Geometry, and the Cramér-Rao Bound, Phys. Rev. X **10**, 021056 (2020).

[47] S. Otsubo, S. Ito, A. Dechant, and T. Sagawa, Estimating entropy production by machine learning of short-time fluctuating currents, Phys. Rev. E **101**, 062106 (2020).

[48] S. Ito, Thermodynamics of information geometry as a generalization of the Glansdorff-Prigogine criterion for stability, arXiv:1908.09446.

[49] H. Ge and H. Qian, Physical origins of entropy production, free energy dissipation, and their mathematical representations, Phys. Rev. E **81**, 051133 (2010).

[50] H. Ge, M. Qian, and H. Qian, Stochastic theory of nonequilibrium steady states. Part II: Applications in chemical biophysics, Phys. Rep. **510**, 87 (2012).

[51] R. Rao and M. Esposito, Conservation laws and work fluctuation relations in chemical reaction networks, J. Chem. Phys. **149**, 245101 (2018).

[52] A. Wachtel, R. Rao, and M. Esposito, Thermodynamically consistent coarse graining of biocatalysts beyond Michaelis–Menten, New J. Phys. **20**, 042002 (2018).

[53] A. Lazarescu, T. Cossetto, G. Falasco, and M. Esposito, Large deviations and dynamical phase transitions in stochastic chemical networks, J. Chem. Phys. **151**, 064117 (2019).

[54] G. Galasco, T. Cossetto, E. Penocchio, and M. Esposito, Negative differential response in chemical reactions, New J. Phys. **21**, 073005 (2019).

[55] F. Avanzini, G. Falasco, and M. Esposito, Thermodynamics of chemical waves, J. Chem. Phys. **151**, 234103 (2019).

[56] E. Penocchio, R. Rao, and M. Esposito, Thermodynamic efficiency in dissipative chemistry, Nat. Commun. **10**, 3865 (2019).

[57] Y. Peng, H. Qian, D. A. Beard, and H. Ge, Universal relation between thermodynamic driving force and one-way fluxes in a nonequilibrium chemical reaction with complex mechanism, Phys. Rev. Res. **2**, 033089 (2020).

[58] G. Falasco, R. Rao, and M. Esposito, Information Thermodynamics of Turing Patterns, Phys. Rev. Lett. **121**, 108301 (2018).

[59] I. Prigogine and R. Lefever, Symmetry breaking instabilities in dissipative systems. II, J. Chem. Phys. **48**, 1695 (1968).

[60] E. W. Lund, Guldberg and Waage and the law of mass action, J. Chem. Educ. **42**, 548 (1965).

[61] T. De Donder and P. van Rysselberghe, *Thermodynamic Theory of Affinity* (Stanford University Press, Stanford, 1936).

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

[63] I. Csiszar, Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems, Ann. Stat. **19**, 2032 (1991).

[64] N. N. Čencov, *Statistical Decision Rules and Optimal Decisions* (Nauka, Moscow, 1972).

[65] R. Lefever, G. Nicolis, and P. Borckmans, The Brusselator: It does oscillate all the same, J. Chem. Soc., Faraday Trans. 1, **84**, 1013 (1988).