



Kinetic derivation of the Hessian geometric structure in chemical reaction networks

Tetsuya J. Kobayashi ^{*}, Dimitri Loutchko, Atsushi Kamimura , and Yuki Sughiyama

Institute of Industrial Science, The University of Tokyo, 4-6-1, Komaba, Meguro-ku, Tokyo 153-8505 Japan



(Received 24 January 2022; accepted 8 June 2022; published 21 July 2022)

The theory of chemical kinetics forms the basis to describe the dynamics of chemical reaction networks. Owing to physical and thermodynamic constraints, the networks possess various structures, which can be utilized to characterize important properties of the networks. In this work, we reveal the Hessian geometry which underlies chemical reaction networks and demonstrate how it originates from the interplay of stoichiometric and thermodynamic constraints. Our derivation is based on kinetics, we assume the law of mass action and characterize the equilibrium states by the detailed balance condition. The obtained geometric structure is then related to thermodynamics via the Hessian geometry appearing in a pure thermodynamic derivation. We demonstrate, based on the fact that both equilibrium and complex balanced states form toric varieties, how the Hessian geometric framework can be extended to nonequilibrium complex balanced steady states. We conclude that Hessian geometry provides a natural framework to capture the thermodynamic aspects of chemical reaction networks.

DOI: [10.1103/PhysRevResearch.4.033066](https://doi.org/10.1103/PhysRevResearch.4.033066)

I. INTRODUCTION

Chemical kinetics constitutes the basis to describe various and complex behaviors of chemically implemented systems such as metabolic networks and intracellular signaling systems [1,2]. Since the formulation and establishment of the law of mass action by C.M. Guldberg and P. Waage, the theories of chemical kinetics and chemical dynamics have been studied and developed using techniques from various different disciplines [1,3–6].

However, the theories of chemical dynamics are not always consistent with thermodynamics. Yet, the consideration of thermodynamics is essential for constructing physically and thermodynamically sound theories for chemical reaction networks. For example, the combination of the law of mass action and the detailed balance condition leads to a kinetic characterization of the equilibrium state in a way that is consistent with chemical thermodynamics, as shown already in 1901 by Wegscheider [7]. Since a chemical reaction network does not necessarily obey the detailed balance condition, it leads to the investigation not only of equilibrium systems but also of a wide range of nonequilibrium reaction systems. Sparked by the pioneering work of Hill and Schnakenberg [8–10], a thermodynamic foundation has been in development for chemical reaction networks out of equilibrium within the

last decades [6,11–16], by employing the knowledge from stochastic thermodynamics [17].

In addition, the consideration of constraints of physical or thermodynamic origin also introduces additional intriguing structures into chemical kinetics. Motivated by the work of Horn and Jackson [18], who extended equilibrium states to complex balanced states, an algebro-geometric structure of chemical reaction networks was discovered and employed to study mass action systems in applied mathematics [19–21]. Other structures in chemical kinetics are also unveiled by using tools from graph theory, homological algebra, and others [3,22–28].

Thus, clarification of the interrelation of chemical kinetics and thermodynamics can be a fruitful source of new physics and mathematics for chemical reaction networks [11,12,14,29–33].

Recently, we found that Hessian geometry provides a natural framework for thermodynamics of chemical reaction systems and used the geometric structure to show that several important results, which were thus far derived only from kinetics, are of pure thermodynamic origin [34]. Nonetheless, it is important to clarify how the Hessian geometric structure is linked to chemical kinetics because the majority of results for chemical reaction networks are based on mass action kinetics rather than thermodynamics and also because kinetic information is essential for working on nonequilibrium situations. This is achieved in this paper.

To this end, we find a close connection between the results of equilibrium systems obtained in algebraic geometry [19] and the thermodynamics which is encoded in the Hessian structure [34]. This combination enables us to grasp the geometric structure of the whole state space, which clarifies the dualistic relation between stoichiometric and thermodynamic constraints manifested as the flatness of the respective dual spaces. Thereby, we generalize and extend the information

^{*}Also at Universal Biology Institute, The University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, 113-8654, Japan.

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

geometric framework for chemical reaction networks limited to a single stoichiometric compatibility class [35].

We derive the Hessian geometric structure [36], which purifies and generalizes some aspects of the information geometric one [37], by starting from the kinetic characterization of the chemical equilibrium state via the law of mass action and the detailed balance condition. In the derivation, we show that the equilibrium states of a chemical reaction network are described by a toric variety [19]. The analytification of the toric variety plays a fundamental role throughout the paper as it constitutes a generalization of the exponential family well-known in statistics [37–41]. Building on the theory of exponential families, we extend the dually flat structure from information geometry [41–43] to chemical reaction networks.

The toric parameter representation of the equilibrium variety naturally leads to a dual space \mathcal{Y} , which is conjugate to the state space \mathcal{X} of molecular concentrations. The Hessian geometric structure and associated convex potential functions $\varphi(\mathbf{x})$ and $\varphi^*(\mathbf{y})$ appear on the state space \mathcal{X} and its dual \mathcal{Y} : in the former, stoichiometric constraints form a linear coordinate system, whereas, in the latter, the equilibrium variety leads to the definition of a dual linear coordinate system. Since the equilibrium variety is defined by the parameters of the reaction system, which are specified thermodynamically by the environmental variables, the dual space mathematically captures the role of thermal reservoirs attached to the system. The spaces \mathcal{X} and \mathcal{Y} are thermodynamically related to chemical densities (concentrations) and chemical potentials verified by comparison with their purely thermodynamic derivation in the accompanying paper [34]. The convex functions and associated Bregman divergences are mapped to the thermodynamic free energy of the system and the difference of total entropy.

In the class of equilibrium systems, our thermodynamic results are special instances of the general theory derived from a purely thermodynamic argument. In this aspect, this paper is supplementary to our accompanying paper [34]. However, owing to the mass action assumption, we find that the equilibrium manifold has the structure of an algebraic variety, which is not necessarily true in nonmass action cases. Thereby, we establish a link to the extensive work carried out in real algebraic geometry [19]. Owing to this, we show how our results extend to nonequilibrium complex balanced steady-states, because they share the algebraic structure with the equilibrium states. We expect this link between Hessian geometry and algebraic geometry to be even more fruitful in the future for thermodynamics and mathematics.

This paper is organized as follows: In Sec. II, we introduce a linear coordinate system in the concentration space \mathcal{X} based on the stoichiometric constraints. In Sec. III, we derive that the set of equilibrium states has the structure of a toric variety and present its parametrization. The equilibrium variety is used to define a linear coordinate system in the dual space \mathcal{Y} , which is also a nonlinear coordinate system of \mathcal{X} space and yields a dual foliation with the stoichiometric constraints. In Sec. IV, we clarify the Hessian geometric structure associated with the dual spaces, and introduce its constituents originating from the first derivatives of the potential functions. In Sec. V, we introduce additional properties originating from the second derivatives. In Sec. VI, we relate the derived geometric

structure to thermodynamics by employing the results in Ref. [34]. In Sec. VII, we show how the framework can be passed on to the nonequilibrium complex balanced state. In the supplementary Sec. VIII, we demonstrate how our work is relevant to the Markov jump processes on a graph, which is often used in stochastic thermodynamics.

To make the theory and results more accessible to researchers in chemical reaction network theory who are not necessarily familiar with information or Hessian geometry, we clarify several implicit and confusing identifications of different objects in conventional textbooks of information geometry [37].

II. CHEMICAL REACTION KINETICS AND STOICHIOMETRIC SUBSPACE

In this work, we consider reversible chemical reaction networks with the stoichiometric matrix given by $S \in \mathbb{Z}^{N \times M}$, where N and M are the number of chemical species and that of pairs of forward and reverse reactions, respectively. Let $\mathbf{x} := (x_1, \dots, x_N)^T \in \mathbb{R}_{\geq 0}^N$ be the concentrations of molecular species involved in the network in a constant volume Ω , e.g., the case when the volume Ω is dominated by solvent.

Let $\mathbf{j}^\pm(\mathbf{x}; \boldsymbol{\theta}) \in \mathbb{R}_{\geq 0}^M$ be the rate functions of the forward and reverse reactions. The vector $\boldsymbol{\theta}$ represents parameters of \mathbf{j}^\pm such as reaction rate constants, the detail of which is specified later and represented abstractly at this point. The reaction rate equation of the network is given by

$$\frac{d\mathbf{x}}{dt} = S\mathbf{j}(\mathbf{x}; \boldsymbol{\theta}), \quad (1)$$

where $\mathbf{j}(\mathbf{x}; \boldsymbol{\theta}) := \mathbf{j}^+(\mathbf{x}; \boldsymbol{\theta}) - \mathbf{j}^-(\mathbf{x}; \boldsymbol{\theta})$ is the total flux [3,6]. In this work, we focus only on the case that $\mathbf{x}(t) > \mathbf{0}$ holds for $t \in [0, \infty]$. In other words, we deal with persistent chemical reaction networks.¹ Thus, we define the state space \mathcal{X} for \mathbf{x} as $\mathcal{X} := \mathbb{R}_{>0}^N$.

A. Conserved quantities and stoichiometric subspace

From the rate equation, we can see that, for any column vector $\mathbf{u}^* \in \text{Ker}[S^T]$, the quantity $(\mathbf{u}^*)^T \mathbf{x}(t)$ is conserved [3]. Let $\{\mathbf{u}_i^*\}_{i \in [1, \dots, \ell]}$ be a complete basis of $\text{Ker}[S^T]$ and ℓ be the dimension of $\text{Ker}[S^T]$. Note that $\{\mathbf{u}_i^*\}$ is generally a nonorthogonal (oblique) basis. We define $U^* := (\mathbf{u}_1^*, \dots, \mathbf{u}_\ell^*)$. From this definition, $S^T U^* = 0$ and $(U^*)^T S = 0$ hold. Then, a vector $\boldsymbol{\eta} = (U^*)^T \mathbf{x}_0$ specifies the values of all stoichiometrically conserved quantities for the initial state \mathbf{x}_0 .² The trajectory $\mathbf{x}(t)$ of Eq. (1) starting from $\mathbf{x}_0 \in \mathcal{X}$ at $t = 0$ satisfies $(U^*)^T \mathbf{x}(t) = (U^*)^T \mathbf{x}_0$. Thus, we define the stoichiometric affine subspace³ (stoichiometric compatibility class) as

¹The verification of the persistence of a given chemical reaction network is a hard problem in general. For networks with mass action kinetics, the persistence property was proven recently for the complex balanced systems, which also includes the equilibrium systems [30].

²The system may have additional (nonlinear) conserved quantities that are determined by the particular structure of $\mathbf{j}(\mathbf{x})$.

³Because $\mathbf{x} > 0$ is required, $\mathcal{P}^{\mathcal{X}}(\boldsymbol{\eta})$ is an affine subspace with boundary. In the theory of algebraic geometry of chemical reaction

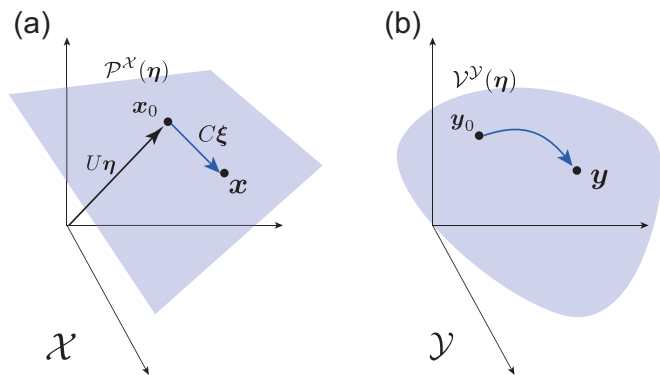


FIG. 1. (a) A linear coordinate system of \mathcal{X} induced by the stoichiometric matrix S . The blue plane represents the stoichiometric subspace $\mathcal{P}^X(\eta)$. (b) The stoichiometric manifold $\mathcal{V}^Y(\eta)$ (the blue curved surface) obtained by mapping $\mathcal{P}^X(\eta)$ into \mathcal{Y} by the Legendre transformation $\partial\phi$.

[Fig. 1(a)]

$$\mathcal{P}^X(\eta) := \{x | (U^*)^T x = (U^*)^T x_0 = \eta\}. \quad (2)$$

To ensure that $x > 0$, we determine the domain of η appropriately, cf. Eq. (7).

B. Extent of chemical reaction

The state of the system $x(t)$ is restricted to the stoichiometric subspace: $x(t) \in \mathcal{P}^X(\eta)$. By using the extent of chemical reaction $\mathbf{J} := \int_0^t \mathbf{j}(x(t'); \boldsymbol{\theta}) dt' \in \mathbb{R}^M$, we can specify $x(t)$ starting from $x_0 \in \mathcal{P}^X(\eta)$ at $t = 0$ as

$$x(t) = x_0 + S\mathbf{J} \in \mathcal{P}^X(\eta). \quad (3)$$

More generally, because $x - x_0 \in \text{Im}[S]$, any point on $\mathcal{P}^X(\eta)$ can be specified as

$$x = x_0 + C\xi \in \mathcal{P}^X(\eta), \quad (4)$$

where $\xi \in \mathbb{R}^{N-\ell}$ and C is defined by using $N - \ell$ independent vectors $\{c_1, \dots, c_{N-\ell}\}$ that form an oblique basis of $\text{Im}[S]$. Note that $N - \ell = \dim \text{Im}[S]$. Because $(u_i^*)^T c_j = 0$ for all i and j , the equality $(U^*)^T C = 0$ holds because of $(U^*)^T S = 0$ [Fig. 1(a)].

We define the dual bases U and C^* of U^* and C such that the orthogonality relations $U^T U^* = I$, $C^T C^* = I$, and $U^T C^* = 0$ are satisfied. Then, these bases span the following linear subspaces: $\langle U^* \rangle = \text{Ker}[S^T]$, $\langle C^* \rangle = \text{Ker}[S^T]^\perp$, $\langle C \rangle = \text{Im}[S]$, and $\langle U \rangle = \text{Im}[S]^\perp$, where $\langle A \rangle$ denotes the linear subspace spanned by column vectors in A and $^\perp$ denotes the complement [44].

Equation (4) is not a canonical representation of a given x . In other words, for any $x \in \mathcal{X}$, the coordinates ξ in Eq. (4) are not uniquely given because they depend on the choice of $x_0 \in \mathcal{P}^X(\eta)$. To make the representation unique, among

networks, $\mathcal{P}^X(\eta)$ is often called a (convex) polyhedron or a polytope if it is bounded. In this work, we use the term affine subspace or simply a subspace to designate $\mathcal{P}^X(\eta)$ or similar geometric objects for notational simplicity.

all $x_0 \in \mathcal{P}^X(\eta)$, we choose one that satisfies $x_0(\eta) = U\eta$. Because $U^T U^* = I$ holds, $\eta = (U^*)^T x_0(\eta)$ is satisfied. This gives a unique linear parametrization of x as

$$x(\eta, \xi) = U\eta + C\xi, \quad (5)$$

where η specifies the position of the origin of stoichiometric subspace and ξ are coordinates on the subspace [Fig. 1(a)].⁴

With this parametrization, the stoichiometric subspace is represented as

$$\mathcal{P}^X(\eta) := \{x | x = U\eta + C\xi, (\eta, \xi) \in \mathcal{E}\}, \quad (6)$$

where we define

$$\mathcal{E} := \{(\eta, \xi) | U\eta + C\xi > \mathbf{0}\}, \quad (7)$$

to ensure $x > 0$. In the following, we consider only $(\eta, \xi) \in \mathcal{E}$.

III. EQUILIBRIUM VARIETY AND DUAL COORDINATE

The conserved quantities η and the extent of chemical reaction \mathbf{J} , or its variant ξ , are commonly used as variables to specify states of chemical systems. As verified in Eq. (5), (η, ξ) is a linear coordinate system of \mathcal{X} . Next, we show how dual coordinates (η^*, ξ^*) can be naturally defined if the system is an equilibrium system.

We point out that, to obtain the result, we will use the detailed balance condition together with the kinetic law of mass action to characterize the equilibrium state and the system. In our accompanying paper, by contrast, we derive the same result only from a thermodynamic argument without these kinetic assumptions [34]. We use the kinetic assumptions here because they prevail in chemical reaction network theory and stochastic thermodynamics [12,16,17], and also because we want to link these disciplines to the results in Ref. [34].

A. Equilibrium variety

The positive equilibrium states of the system [Eq. (1)] are defined here as the states that satisfy the detailed balance condition:

$$\mathcal{V}_{\text{eq}}^X(\boldsymbol{\theta}) := \{x > 0 | \mathbf{j}(x; \boldsymbol{\theta}) = \mathbf{0}\}. \quad (8)$$

Note that $\mathcal{V}_{\text{eq}}^X(\boldsymbol{\theta})$ can be empty if a specified $\boldsymbol{\theta}$ admits no equilibrium state. We define the set of parameters Θ_{eq} such that $\mathcal{V}_{\text{eq}}^X(\boldsymbol{\theta})$ is not empty if $\boldsymbol{\theta} \in \Theta_{\text{eq}}$.

We additionally assume that $\mathbf{j}^\pm(x; \boldsymbol{\theta})$ satisfy the law of mass action:

$$j_m^+(x; \boldsymbol{\theta}) = k_m^+ x^{\boldsymbol{\gamma}_m^+}, \quad j_m^-(x; \boldsymbol{\theta}) = k_m^- x^{\boldsymbol{\gamma}_m^-}, \quad (9)$$

where $k_m^\pm \in \mathbb{R}_{>0}$ are the rate constants of the m th forward and reverse reactions, respectively. The integer vectors $\boldsymbol{\gamma}_m^+, \boldsymbol{\gamma}_m^- \in \mathbb{Z}^N$ specify reactants and products of the m th forward reaction. Thus, $\boldsymbol{\gamma}_m^- - \boldsymbol{\gamma}_m^+ = \mathbf{s}_m$ where \mathbf{s}_m is the m th column vector of the stoichiometric matrix S . For a pair of vectors $x \in \mathcal{X}$ and $\boldsymbol{\alpha} \in \mathbb{Z}^N$, the exponential $x^{\boldsymbol{\alpha}}$ represents the

⁴Note that Eq. (5) is obtained without any kinetic information.

monomial⁵

$$\mathbf{x}^{\alpha^T} := \prod_{i=1}^N x_i^{\alpha_i}. \quad (10)$$

We write Eq. (9) in a vector form as

$$\mathbf{j}^+(\mathbf{x}; \boldsymbol{\theta}) = \mathbf{k}^+ \circ \mathbf{x}^{(\Gamma^+)^T}, \quad \mathbf{j}^-(\mathbf{x}; \boldsymbol{\theta}) = \mathbf{k}^- \circ \mathbf{x}^{(\Gamma^-)^T}, \quad (11)$$

where $\mathbf{x}^{(\Gamma^\pm)^T} := (\mathbf{x}^{(\nu_1^\pm)^T}, \dots, \mathbf{x}^{(\nu_M^\pm)^T})^T$ and \circ is the component-wise product of vectors. Then $\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\boldsymbol{\theta})$ is given by

$$\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\boldsymbol{\theta}) := \{\mathbf{x} > 0 \mid \mathbf{k}^+ \circ \mathbf{x}^{(\Gamma^+)^T} = \mathbf{k}^- \circ \mathbf{x}^{(\Gamma^-)^T}\}. \quad (12)$$

Now, $\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\boldsymbol{\theta})$ is an algebraic variety, i.e., the manifold defined as the zeros of algebraic equations.⁶ Thus, $\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\boldsymbol{\theta})$ is called an equilibrium variety or equilibrium manifold. In addition, we note that \mathbf{k}^\pm are the parameters of \mathbf{j}^\pm and thus $\boldsymbol{\theta} = (\mathbf{k}^\pm)$.

B. Parameter conditions for equilibrium

Next, we derive the necessary and sufficient condition that the set of parameters Θ_{eq} must satisfy to have nonempty $\mathcal{V}_{\text{eq}}(\boldsymbol{\theta})$ for $\boldsymbol{\theta} \in \Theta_{\text{eq}}$. In other words, we characterize the parameter set that admits equilibrium states. By rearranging $\mathbf{k}^+ \circ \mathbf{x}^{(\Gamma^+)^T} = \mathbf{k}^- \circ \mathbf{x}^{(\Gamma^-)^T}$, we obtain

$$\ln \mathbf{K} := \ln \frac{\mathbf{k}^+}{\mathbf{k}^-} = S^T \ln \mathbf{x}, \quad (13)$$

where $\Gamma^- - \Gamma^+ = S$ was used. From the Fredholm alternative, Eq. (13) has a solution if and only if $\ln \mathbf{K} \in \text{Im}[S^T]$. The condition $\ln \mathbf{K} \in \text{Im}[S^T]$ is an abstract representation of the Wegscheider condition [7,12]. Then, we can represent Θ_{eq} as

$$\Theta_{\text{eq}} = \left\{ \boldsymbol{\theta} = (\mathbf{k}^\pm) \mid \mathbf{K} = \frac{\mathbf{k}^+}{\mathbf{k}^-}, \ln \mathbf{K} \in \text{Im}[S^T] \right\}. \quad (14)$$

Hereby, \mathbf{K} is the vector of equilibrium constants. Thus, this representation means that, among all parameters, only the equilibrium constants \mathbf{K} are relevant for the existence of equilibrium states. This is natural because the equilibrium state of a system should be characterized statically without specifying any kinetic information of the system.

C. Toric parametrization of the equilibrium variety

For a given $\mathbf{K} \in \Theta_{\text{eq}}$, there exists a particular solution $\tilde{\mathbf{x}}_{\text{eq}}$ of Eq. (13). Then the equilibrium variety $\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\mathbf{K})$, i.e., the set of \mathbf{x} that satisfies Eq. (13), can be represented as

$$\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\mathbf{K}) = \{\mathbf{x} \mid \ln \mathbf{x} = \ln \tilde{\mathbf{x}}_{\text{eq}} + U^* \boldsymbol{\eta}^*, \boldsymbol{\eta}^* \in \mathbb{R}^\ell\}, \quad (15)$$

⁵The transpose in \mathbf{x}^{α^T} may look confusing. But this transpose is for the notational consistency with the application of logarithm to \mathbf{x}^{α^T} as $\ln \mathbf{x}^{\alpha^T} = \boldsymbol{\alpha}^T \ln \mathbf{x}$. See also Appendix B for a more detailed explanation.

⁶It should be noted that, hereafter, we use the analytification of the variety to work on differential geometric aspects of $\mathcal{V}_{\text{eq}}^{\mathcal{X}}$. But we abuse the word variety to emphasize the fact that $\mathcal{V}_{\text{eq}}^{\mathcal{X}}$ is given by algebraic equations (which are derived from the detailed balance condition).

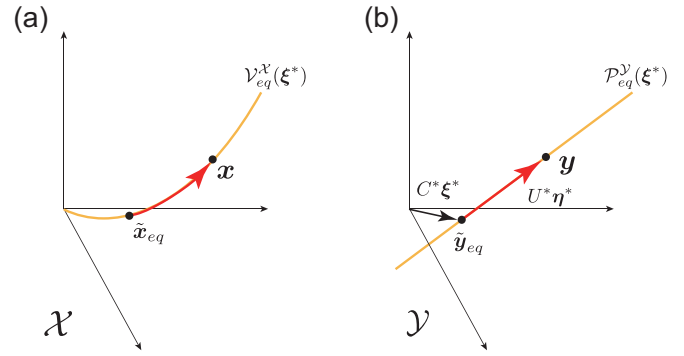


FIG. 2. (a) A curved manifold in \mathcal{X} defined by the equilibrium variety $\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\xi^*)$ (orange curve). $\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\xi^*)$ is generally high-dimensional, but in this figure, it is one-dimensional because of the limitation of three-dimensional space for visualization. (b) The equilibrium variety shown in \mathcal{Y} space. In \mathcal{Y} space, it is a flat affine subspace. The linear coordinate system induced by the equilibrium variety is also shown.

because U^* is a basis of $\text{Ker}[S^T]$. This representation is known as the affine toric parametrization of the variety $\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\mathbf{K})$. In algebraic statistics, U^* is also called the design matrix of the toric variety. A toric variety is characterized as being generated by a toric ideal, i.e., a prime binomial ideal in the coordinate ring of the ambient space $\mathbb{R}^N \supset \mathcal{X}$ [45–47]. Because the detailed balance condition is nothing but a set of binomial equations, a toric variety is a natural representation of the equilibrium states.

From Eq. (15), we see that $\boldsymbol{\eta}^*$ works as a coordinate of the variety $\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\mathbf{K})$. However, similarly to the case of \mathbf{x}_0 , the actual value of $\boldsymbol{\eta}^*$ depends on the choice of $\tilde{\mathbf{x}}_{\text{eq}}$, which is not uniquely specified because it is just a particular solution. Among all $\tilde{\mathbf{x}}_{\text{eq}}$ satisfying $\ln \mathbf{K} = S^T \ln \tilde{\mathbf{x}}_{\text{eq}}$, we choose one such that $(\ln \tilde{\mathbf{x}}_{\text{eq}} + \hat{\mathbf{y}}) \in \text{Ker}[S^T]^\perp$ where $\hat{\mathbf{y}}$ determines a reference point and works as a free parameter. We associate it with the standard chemical potential in the later section. Because $\langle C^* \rangle = \text{Ker}[S^T]^\perp$, we can write $\ln \tilde{\mathbf{x}}_{\text{eq}} = C^* \xi^* - \hat{\mathbf{y}}$. Since the equation

$$\ln \mathbf{K} = S^T (C^* \xi^* - \hat{\mathbf{y}}) \quad (16)$$

uniquely determines ξ^* if $\hat{\mathbf{y}}$ is fixed and $\ln \mathbf{K} \in \text{Im} S^T$, we can use ξ^* instead of \mathbf{K} to specify the equilibrium variety, which we denote by $\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\xi^*)$ from now on [Fig. 2(a)]:

$$\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\xi^*) = \{\mathbf{x} \mid \ln \mathbf{x} = C^* \xi^* + U^* \boldsymbol{\eta}^* - \hat{\mathbf{y}}, \boldsymbol{\eta}^* \in \mathbb{R}^\ell\}. \quad (17)$$

Now we introduce the space \mathcal{Y} obtained by the nonlinear transformation

$$\mathbf{y} = \ln \mathbf{x} + \hat{\mathbf{y}}, \quad (18)$$

i.e., $\mathcal{Y} := \hat{\mathbf{y}} + \ln \mathcal{X} = \mathbb{R}^N$. Any point on $\mathbf{y} \in \mathcal{Y}$ can be linearly parametrized as

$$\mathbf{y}(\boldsymbol{\eta}^*, \xi^*) = U^* \boldsymbol{\eta}^* + C^* \xi^*. \quad (19)$$

Thus, $(\boldsymbol{\eta}^*, \xi^*)$ is a linear coordinate system of \mathcal{Y} [Fig. 2(b)]. By the reverse transformation of Eq. (18), any $\mathbf{x} \in \mathcal{X}$ can also

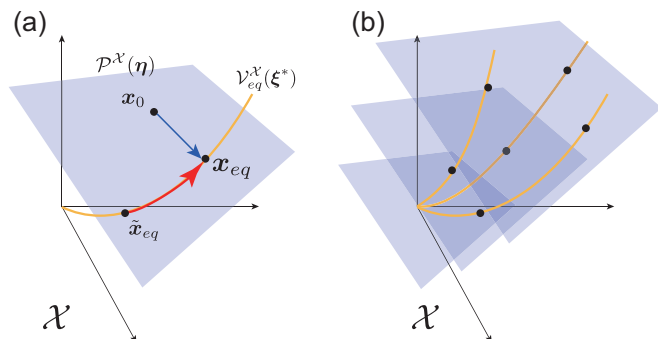


FIG. 3. (a) Intersection of the stoichiometric subspace $\mathcal{P}^{\mathcal{X}}(\eta)$ (blue hyperplane) and the equilibrium variety $\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\xi^*)$ (orange curve) at \mathbf{x}_{eq} . (b) Foliation (a nonlinear coordinate system) formed by the stoichiometric subspaces (blue hyperplanes) and the equilibrium varieties (orange curves).

be parametrized nonlinearly as

$$\mathbf{x}(\boldsymbol{\eta}^*, \boldsymbol{\xi}^*) = \exp[U^* \boldsymbol{\eta}^* + C^* \boldsymbol{\xi}^* - \hat{\mathbf{y}}]. \quad (20)$$

As the form of the transformation in Eq. (18) implies, \mathcal{Y} is the space of chemical potentials, which is thermodynamically conjugate to the concentration (density) space \mathcal{X} [48]. Moreover, $\mathbf{x}(\boldsymbol{\eta}, \boldsymbol{\xi})$ and $\mathbf{y}(\boldsymbol{\eta}^*, \boldsymbol{\xi}^*)$ are Legendre dual as we will show in Sec. IV.

D. Foliation and mixed coordinate

Finally, we show that $(\boldsymbol{\eta}, \boldsymbol{\xi}^*)$, a mixture of the two coordinate systems $(\boldsymbol{\eta}, \boldsymbol{\xi})$ and $(\boldsymbol{\eta}^*, \boldsymbol{\xi}^*)$, can also work as a nonlinear coordinate system for \mathcal{X} .

For a specific value of kinetic parameters satisfying $\mathbf{K} \in \Theta_{\text{eq}}$, the variety $\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\mathbf{K}) = \mathcal{V}_{\text{eq}}^{\mathcal{X}}(\boldsymbol{\xi}^*)$ specifies the set of equilibrium states for the given parameter value. Also, for any initial state \mathbf{x}_0 of the system, its time evolution $\mathbf{x}(t)$ is constrained to the stoichiometric subspace $\mathcal{P}^{\mathcal{X}}(\mathbf{x}_0) = \mathcal{P}^{\mathcal{X}}(\boldsymbol{\eta})$. Thus, the reachable equilibrium point must lie in their intersection [see Fig. 3(a)]

$$\mathbf{x}_{\text{eq}} \in \mathcal{P}^{\mathcal{X}}(\boldsymbol{\eta}) \cap \mathcal{V}_{\text{eq}}^{\mathcal{X}}(\boldsymbol{\xi}^*). \quad (21)$$

Because $\mathcal{P}^{\mathcal{X}}(\boldsymbol{\eta})$ and $\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\boldsymbol{\xi}^*)$ are characterized by the same structural matrix U^* , cf. Eqs. (2) and (17), their intersection is assured to be unique and transversal by Birch's theorem for the exponential family in statistics [49].

The uniqueness enables us to specify \mathbf{x}_{eq} by $\boldsymbol{\eta}$ and $\boldsymbol{\xi}^*$ as $\mathbf{x}_{\text{eq}}(\boldsymbol{\eta}, \boldsymbol{\xi}^*)$. Because both $\mathcal{P}^{\mathcal{X}}(\boldsymbol{\eta})$ and $\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\boldsymbol{\xi}^*)$ can cover the whole state space \mathcal{X} by changing $\boldsymbol{\eta}$ and $\boldsymbol{\xi}^*$, respectively, they form a foliation of \mathcal{X} [Fig. 3(b)]. In other words, $(\boldsymbol{\eta}, \boldsymbol{\xi}^*)$ works as a nonlinear coordinate system of \mathcal{X} . Physically, this means that any equilibrium state \mathbf{x}_{eq} can be characterized by the stoichiometric subspace $\mathcal{P}^{\mathcal{X}}(\boldsymbol{\eta})$ and the variety $\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\boldsymbol{\xi}^*)$, each of which has the corresponding value of $\boldsymbol{\eta}$ and $\boldsymbol{\xi}^*$ explicitly given by

$$\boldsymbol{\eta} = (U^*)^T \mathbf{x}_{\text{eq}}, \quad \boldsymbol{\xi}^* = C^T (\ln \mathbf{x}_{\text{eq}} + \hat{\mathbf{y}}). \quad (22)$$

This mixed coordinate system is intensively used in information geometry [37] and the existence of the analogous coordinate system for chemical reaction networks emphasizes

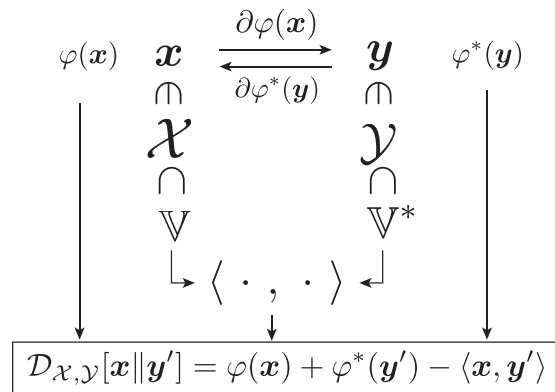


FIG. 4. Relations between \mathcal{X} and \mathcal{Y} , \mathbb{V} and \mathbb{V}^* , $\varphi(\mathbf{x})$ and $\varphi(\mathbf{x})^*$, and $\partial\varphi(\mathbf{x})$ and $\partial\varphi(\mathbf{y})^*$. The Bregman divergence \mathcal{D} is induced by using all the constituents.

their information geometric properties. Moreover, the real algebraic geometry of toric varieties has been employed in computational and algebraic statistics to handle exponential families with linear constraints [49]. Thus, information geometry and algebraic geometry capture the same entity from different aspects.

IV. HESSIAN GEOMETRIC STRUCTURE OF EQUILIBRIUM CHEMICAL REACTION NETWORKS

In the previous section, we have introduced the dual coordinates and their mixture for equilibrium chemical reaction networks starting from a conventional chemical kinetics formulation. Here, we deductively clarify their Hessian geometric structure [34,36,37].

A. Dually flat state space

We regard the state space \mathcal{X} as a subspace embedded in the N -dimensional vector space $\mathbb{V} = \mathbb{R}^N$. Let $\mathcal{Y} = \mathbb{V}^*$ be the dual vector space of \mathbb{V} and $\langle \cdot, \cdot \rangle$ be the bilinear form defined on $\mathbb{V} \times \mathbb{V}^*$ (Fig. 4).⁷ On \mathcal{X} and \mathcal{Y} , we define the following two strictly convex smooth potential functions, cf. Fig. 4:

$$\varphi(\mathbf{x}) := [\ln \mathbf{x} - \hat{\mathbf{y}} - \mathbf{1}]^T \mathbf{x}, \quad (23)$$

$$\varphi^*(\mathbf{y}) := \mathbf{1}^T e^{\mathbf{y} + \hat{\mathbf{y}}}. \quad (24)$$

This yields the one-to-one Legendre duality between $\mathbf{x} \in \mathcal{X}$ and $\mathbf{y} \in \mathcal{Y}$ as

$$\mathbf{y} = \partial_{\mathbf{x}} \varphi(\mathbf{x}) := \left\{ \frac{\partial \varphi(\mathbf{x})}{\partial \mathbf{x}} \right\} = \ln \mathbf{x} - \hat{\mathbf{y}}, \quad (25)$$

$$\mathbf{x} = \partial_{\mathbf{y}} \varphi^*(\mathbf{y}) := \left\{ \frac{\partial \varphi^*(\mathbf{y})}{\partial \mathbf{y}} \right\} = e^{\mathbf{y} + \hat{\mathbf{y}}}, \quad (26)$$

such that $\varphi(\mathbf{x})$ and $\varphi^*(\mathbf{y})$ are also dual satisfying Legendre identity:

$$\varphi(\mathbf{x}) + \varphi^*(\mathbf{y}) - \langle \mathbf{x}, \mathbf{y} \rangle = 0. \quad (27)$$

⁷Note that the result in this section is obtained without assuming any inner product structure.

Thus, the pair $(\mathcal{X}, \mathcal{Y})$ is equipped with two dualities: one is linear algebraic duality and the other is nonlinear Legendre duality induced by the convex functions, $\varphi(\mathbf{x})$ and $\varphi^*(\mathbf{y})$. This is a Hessian structure [36], which is a mathematical basis underlying information geometry [37].

We use the following notation: a pair (\mathbf{x}, \mathbf{y}) is always treated as the Legendre dual pair defined by Eqs. (25) and (26) (Fig. 4). Pairs with different decorations, e.g., $(\mathbf{x}', \mathbf{y}')$ and $(\mathbf{x}'', \mathbf{y}'')$, will be treated as distinct Legendre dual pairs. We also abbreviate $\partial_{\mathbf{x}}\varphi(\mathbf{x})$ with $\partial\varphi(\mathbf{x})$ for the sake of notational simplicity.

B. Duality in subspaces

This duality is inherited by the linear coordinate systems, which were derived by a chemical kinetic argument [Figs. 1(a) and 2(b)]:

$$\mathbf{x}(\boldsymbol{\eta}, \boldsymbol{\xi}) = U\boldsymbol{\eta} + C\boldsymbol{\xi}, \quad (28)$$

$$\mathbf{y}(\boldsymbol{\eta}^*, \boldsymbol{\xi}^*) = U^*\boldsymbol{\eta}^* + C^*\boldsymbol{\xi}^*. \quad (29)$$

In particular, we have the following partial Legendre duality between $\boldsymbol{\xi}$ and $\boldsymbol{\xi}^*$ and $\boldsymbol{\eta}$ and $\boldsymbol{\eta}^*$:

$$\partial_{\boldsymbol{\eta}}\varphi[\mathbf{x}(\boldsymbol{\eta}, \boldsymbol{\xi})] = U^T\mathbf{y} = \boldsymbol{\eta}^*, \quad (30)$$

$$\partial_{\boldsymbol{\xi}}\varphi[\mathbf{x}(\boldsymbol{\eta}, \boldsymbol{\xi})] = C^T\mathbf{y} = \boldsymbol{\xi}^*, \quad (31)$$

$$\partial_{\boldsymbol{\eta}^*}\varphi^*[\mathbf{y}(\boldsymbol{\eta}^*, \boldsymbol{\xi}^*)] = (U^*)^T\mathbf{x} = \boldsymbol{\eta}, \quad (32)$$

$$\partial_{\boldsymbol{\xi}^*}\varphi^*[\mathbf{y}(\boldsymbol{\eta}^*, \boldsymbol{\xi}^*)] = (C^*)^T\mathbf{x} = \boldsymbol{\xi}, \quad (33)$$

where $\partial_{\boldsymbol{\eta}}$ is Legendre transform with respect to $\boldsymbol{\eta}$. We also have

$$\varphi(\boldsymbol{\eta}, \boldsymbol{\xi}) + \varphi^*(\boldsymbol{\eta}^*, \boldsymbol{\xi}^*) = \langle \boldsymbol{\eta}, \boldsymbol{\eta}^* \rangle + \langle \boldsymbol{\xi}, \boldsymbol{\xi}^* \rangle. \quad (34)$$

This means that the linear coordinate systems defined on \mathcal{X} and \mathcal{Y} are preserved under the additional structure of Legendre duality. Such \mathcal{X} and \mathcal{Y} are called dually flat spaces in information geometry [37] and Hessian geometry [36].

C. Bregman divergence

Using the Legendre dual potential functions, the Bregman divergence between two points \mathbf{x} and \mathbf{x}' is defined as [37,50]

$$\mathcal{D}_{\mathcal{X}}[\mathbf{x}||\mathbf{x}'] := \varphi(\mathbf{x}) - \varphi(\mathbf{x}') - \langle \mathbf{x} - \mathbf{x}', \partial\varphi(\mathbf{x}') \rangle. \quad (35)$$

Due to the convexity of $\varphi(\mathbf{x})$, the function $\mathcal{D}[\mathbf{x}||\mathbf{x}']$ is nonnegative and measures the extent of convexity as the deviation of $\varphi(\mathbf{x})$ from its linear extrapolation $\varphi(\mathbf{x}') + \langle \mathbf{x} - \mathbf{x}', \partial\varphi(\mathbf{x}') \rangle$, evaluated at \mathbf{x}' . The minimum of $\mathcal{D}_{\mathcal{X}}[\mathbf{x}||\mathbf{x}']$ is 0, which is achieved if and only if $\mathbf{x} = \mathbf{x}'$. Moreover, $\mathcal{D}_{\mathcal{X}}[\mathbf{x}||\mathbf{x}']$ is convex with respect to \mathbf{x} (not necessarily with respect to \mathbf{x}'). A direct computation, using Eqs. (23) and (24), gives

$$\mathcal{D}_{\mathcal{X}}[\mathbf{x}||\mathbf{x}'] = \left(\ln \frac{\mathbf{x}}{\mathbf{x}'} \right)^T \mathbf{x} - \mathbf{1}^T (\mathbf{x} - \mathbf{x}'). \quad (36)$$

This indicates that $\mathcal{D}[\mathbf{x}||\mathbf{x}']$ is reduced to the generalized Kullback-Leibler divergence for positive measures on a discrete space.

Similarly, for \mathbf{y} and \mathbf{y}' , the dual Bregman divergence is given by

$$\mathcal{D}_{\mathcal{Y}}[\mathbf{y}||\mathbf{y}'] := \varphi^*(\mathbf{y}) - \varphi^*(\mathbf{y}') - \langle \partial\varphi^*(\mathbf{y}'), \mathbf{y} - \mathbf{y}' \rangle. \quad (37)$$

If \mathbf{y} and \mathbf{y}' are Legendre dual to \mathbf{x} and \mathbf{x}' , respectively, then the Bregman divergences satisfy the symmetry property

$$\mathcal{D}_{\mathcal{X}}[\mathbf{x}||\mathbf{x}'] = \varphi(\mathbf{x}) + \varphi^*(\mathbf{y}') - \langle \mathbf{x}, \mathbf{y}' \rangle = \mathcal{D}_{\mathcal{Y}}[\mathbf{y}'||\mathbf{y}], \quad (38)$$

where the Legendre identity $\varphi(\mathbf{x}') + \varphi^*(\mathbf{y}') = \langle \mathbf{x}', \mathbf{y}' \rangle = \langle \mathbf{x}', \partial\varphi(\mathbf{x}') \rangle$ was used. Thus, we are led to define $\mathcal{D}_{\mathcal{X},\mathcal{Y}}[\mathbf{x}||\mathbf{y}']$ as (Fig. 4)

$$\mathcal{D}_{\mathcal{X},\mathcal{Y}}[\mathbf{x}||\mathbf{y}'] := \varphi(\mathbf{x}) + \varphi^*(\mathbf{y}') - \langle \mathbf{x}, \mathbf{y}' \rangle. \quad (39)$$

Note that $\mathcal{D}_{\mathcal{X}}[\mathbf{x}||\mathbf{x}']$, $\mathcal{D}_{\mathcal{Y}}[\mathbf{y}'||\mathbf{y}]$, and $\mathcal{D}_{\mathcal{X},\mathcal{Y}}[\mathbf{x}||\mathbf{y}']$ are just different representations of the same geometric quantity because \mathbf{x} and \mathbf{y} as well as \mathbf{x}' and \mathbf{y}' are in one-to-one correspondence by the Legendre transformation. In the following, we switch among the three equivalent notations depending on the purpose. We emphasize that

$$\mathbf{x}' = \arg \min_{\mathbf{x}} \mathcal{D}_{\mathcal{X}}[\mathbf{x}||\mathbf{x}'] = \arg \min_{\mathbf{x}} \mathcal{D}_{\mathcal{X},\mathcal{Y}}[\mathbf{x}||\mathbf{y}']. \quad (40)$$

As we will show later, the relation between the potential function φ and the Bregman divergence is the mathematical reason why the Kullback-Leibler divergence appears as the difference of the total entropy in equilibrium systems.

D. Dual orthogonality

We demonstrate that the Bregman divergence and the mixed coordinate representation play a central role when determining the equilibrium state of the system.

Consider a chemical reaction network [Eq. (1)] with an equilibrium parameter $\mathbf{K} \in \Theta_{\text{eq}}$ and an initial state \mathbf{x}_0 . Let \mathbf{x}_0 be in stoichiometric subspace $\mathcal{P}^{\mathcal{X}}(\boldsymbol{\eta})$ and let the equilibrium variety corresponding to \mathbf{K} be $\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\boldsymbol{\xi}^*)$. The correspondence between \mathbf{K} and $\boldsymbol{\xi}^*$ is given explicitly in Eqs. (22) and (16). Then, the equilibrium state that the system should converge to is determined by $\mathbf{x}_{\text{eq}}(\boldsymbol{\eta}, \boldsymbol{\xi}^*) \in \mathcal{P}^{\mathcal{X}}(\boldsymbol{\eta}) \cap \mathcal{V}_{\text{eq}}^{\mathcal{X}}(\boldsymbol{\xi}^*)$. Any point on $\mathcal{P}^{\mathcal{X}}(\boldsymbol{\eta})$, including the initial state \mathbf{x}_0 , can be represented uniquely as $\mathbf{x}_p(\boldsymbol{\eta}, \boldsymbol{\xi}_p) \in \mathcal{P}^{\mathcal{X}}(\boldsymbol{\eta})$. Similarly, any point on $\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\boldsymbol{\xi}^*)$ can be written as $\mathbf{x}_q^{\mathcal{X}}(\boldsymbol{\eta}_q^*, \boldsymbol{\xi}^*) \in \mathcal{V}_{\text{eq}}^{\mathcal{X}}(\boldsymbol{\xi}^*)$. Then, from the definition of $\mathcal{P}^{\mathcal{X}}(\boldsymbol{\eta})$ and $\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\boldsymbol{\xi}^*)$, the relations

$$\mathbf{x}_p(\boldsymbol{\eta}, \boldsymbol{\xi}_p) - \mathbf{x}_{\text{eq}}(\boldsymbol{\eta}, \boldsymbol{\xi}) = C(\boldsymbol{\xi}_p - \boldsymbol{\xi}), \quad (41)$$

$$\mathbf{y}_q(\boldsymbol{\eta}_q^*, \boldsymbol{\xi}^*) - \mathbf{y}_{\text{eq}}(\boldsymbol{\eta}^*, \boldsymbol{\xi}^*) = U^*(\boldsymbol{\eta}_q^* - \boldsymbol{\eta}^*), \quad (42)$$

hold. Here, \mathbf{y}_q and \mathbf{y}_{eq} are the Legendre duals of \mathbf{x}_q and \mathbf{x}_{eq} , respectively. This yields the orthogonality

$$\langle \mathbf{x}_p(\boldsymbol{\eta}, \boldsymbol{\xi}_p) - \mathbf{x}_{\text{eq}}(\boldsymbol{\eta}, \boldsymbol{\xi}), \mathbf{y}_q(\boldsymbol{\eta}_q^*, \boldsymbol{\xi}^*) - \mathbf{y}_{\text{eq}}(\boldsymbol{\eta}^*, \boldsymbol{\xi}^*) \rangle = 0, \quad (43)$$

which follows from $C^T U^* = 0$. Without using the coordinate representation, this relation means that

$$\langle \mathbf{x}_p - \mathbf{x}_{\text{eq}}, \mathbf{y}_q - \mathbf{y}_{\text{eq}} \rangle = 0, \quad (44)$$

if $\mathbf{x}_p \in \mathcal{P}^{\mathcal{X}}(\boldsymbol{\eta})$, $\mathbf{x}_q \in \mathcal{V}_{\text{eq}}^{\mathcal{X}}(\boldsymbol{\xi}^*)$, and $\mathbf{x}_{\text{eq}} \in \mathcal{P}^{\mathcal{X}}(\boldsymbol{\eta}) \cap \mathcal{V}_{\text{eq}}^{\mathcal{X}}(\boldsymbol{\xi}^*)$. For any \mathbf{x}, \mathbf{x}' and \mathbf{x}'' , the divergence $\mathcal{D}_{\mathcal{X}}[\mathbf{x}||\mathbf{x}']$ satisfies

$$\begin{aligned} \mathcal{D}_{\mathcal{X}}[\mathbf{x}||\mathbf{x}'] + \mathcal{D}_{\mathcal{X}}[\mathbf{x}'||\mathbf{x}'] \\ = \mathcal{D}_{\mathcal{X}}[\mathbf{x}||\mathbf{x}'] + \langle (\mathbf{x} - \mathbf{x}'), (\mathbf{y}' - \mathbf{y}'') \rangle. \end{aligned} \quad (45)$$

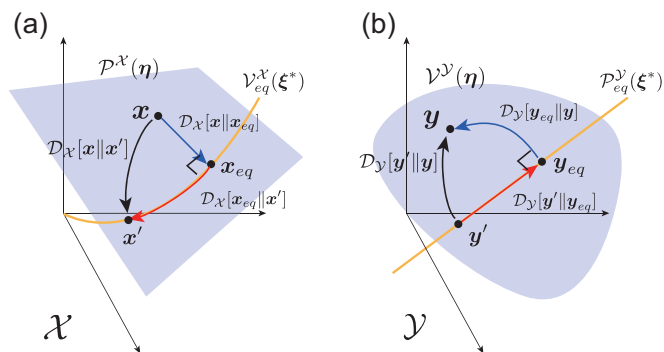


FIG. 5. (a) Graphic representation of the generalized Pythagorean theorem and orthogonality relation on \mathcal{X} space [Eq. (46)]. The equilibrium point \mathbf{x}_{eq} is characterized by either projection of \mathbf{x} onto $\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\xi^*)$ along the flat subspace $\mathcal{P}^{\mathcal{X}}(\eta)$ or that of \mathbf{x}' on the $\mathcal{P}^{\mathcal{X}}(\eta)$ along the curved manifold $\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\xi^*)$ [Eq. (48)]. (b) The same Pythagorean theorem and orthogonality relation in \mathcal{Y} space [Eq. (52)]. The equilibrium point \mathbf{y}_{eq} is characterized by either projection of \mathbf{y} onto $\mathcal{P}_{\text{eq}}^{\mathcal{Y}}(\xi^*)$ along the curved manifold $\mathcal{V}^{\mathcal{Y}}(\eta)$ or that of \mathbf{y}' on the $\mathcal{V}^{\mathcal{Y}}(\eta)$ along the flat manifold $\mathcal{P}_{\text{eq}}^{\mathcal{Y}}(\xi^*)$ [Eq. (54)].

Thus, for $\mathbf{x}_p, \mathbf{x}_q$, and \mathbf{x}_{eq} satisfying $\mathbf{x}_p \in \mathcal{P}^{\mathcal{X}}(\eta)$, $\mathbf{x}_q \in \mathcal{V}_{\text{eq}}^{\mathcal{X}}(\xi^*)$, and $\mathbf{x}_{\text{eq}} \in \mathcal{P}^{\mathcal{X}}(\eta) \cap \mathcal{V}_{\text{eq}}^{\mathcal{X}}(\xi^*)$, the generalized Pythagorean theorem on \mathcal{X} space holds [Fig. 5(a)]:

$$\mathcal{D}_{\mathcal{X}}[\mathbf{x}_p||\mathbf{x}_{\text{eq}}] + \mathcal{D}_{\mathcal{X}}[\mathbf{x}_{\text{eq}}||\mathbf{x}_q] = \mathcal{D}_{\mathcal{X}}[\mathbf{x}_p||\mathbf{x}_q]. \quad (46)$$

This relation is geometric in the sense that it is independent of the choice of η and ξ^* or the choice of coordinate systems on $\mathcal{P}^{\mathcal{X}}$ and \mathcal{V}_{eq} .

From the Pythagorean theorem, we obtain two variational characterizations of the equilibrium state [Fig. 5(a)]:

$$\begin{aligned} \mathbf{x}_{\text{eq}}(\eta, \xi^*) &= \arg \min_{\mathbf{x} \in \mathcal{P}^{\mathcal{X}}(\eta)} \mathcal{D}_{\mathcal{X}}[\mathbf{x}||\mathbf{x}'] \\ &\times \text{for any fixed } \mathbf{x}' \in \mathcal{V}_{\text{eq}}^{\mathcal{X}}(\xi^*) \quad (47) \\ &= \arg \min_{\mathbf{x}' \in \mathcal{V}_{\text{eq}}^{\mathcal{X}}(\xi^*)} \mathcal{D}_{\mathcal{X}}[\mathbf{x}||\mathbf{x}'] \\ &\times \text{for any fixed } \mathbf{x} \in \mathcal{P}^{\mathcal{X}}(\eta). \quad (48) \end{aligned}$$

The former means that the equilibrium point can be obtained as the point in the stoichiometric subspace $\mathcal{P}^{\mathcal{X}}$ at which $\mathcal{D}_{\mathcal{X}}$ is minimized. This is related to the relaxation process of $\mathbf{x}(t)$ toward the equilibrium point in the stoichiometric compatibility class determined by the initial state \mathbf{x}_0 .

In contrast, the latter means that the equilibrium point, to which a given initial state \mathbf{x}_0 converges, can be obtained as the point in the equilibrium variety $\mathcal{V}_{\text{eq}}^{\mathcal{X}}$ at which $\mathcal{D}_{\mathcal{X}}$ is minimized. The role of this equation will be clarified when we consider its thermodynamic meaning in Sec. VI.

Note that the convergence of $\mathbf{x}(t)$ to \mathbf{x}_{eq} is not ensured by this geometric argument alone. The geometry just ensures the uniqueness of \mathbf{x}_{eq} . Thermodynamically, the convergence is required from the second law of thermodynamics. Kinetically, if the mass action kinetic is assumed, one can directly prove the convergence by showing that the Bregman divergence

$\mathcal{D}_{\mathcal{X}}[\mathbf{x}(t)||\mathbf{x}_{\text{eq}}]$ is a Lyapunov function [12,51,52]⁸:

$$\frac{d\mathcal{D}_{\mathcal{X}}[\mathbf{x}(t)||\mathbf{x}_{\text{eq}}]}{dt} = -\mathbf{j}(\mathbf{x}(t))^T \ln \frac{\mathbf{j}^+(\mathbf{x}(t))}{\mathbf{j}^-(\mathbf{x}(t))} \leq 0. \quad (49)$$

Thus, the mass action kinetics is consistent with thermodynamics.

E. Variety, subspace, and orthogonality in the dual space

Because \mathcal{X} and \mathcal{Y} are in bijection to each other, all the results obtained previously on \mathcal{X} space can be carried over to \mathcal{Y} space. On \mathcal{Y} , the geometric nature of the stoichiometric subspace and the equilibrium variety is swapped due to the logarithmic Legendre transformation. After transforming $\mathcal{P}^{\mathcal{X}}(\eta)$ and $\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\xi^*)$ by $\partial\varphi(\mathbf{x})$, we obtain the stoichiometric manifold and the equilibrium subspace in \mathcal{Y} space as [Figs. 1(b) and 2(b)]⁹

$$\begin{aligned} \mathcal{V}^{\mathcal{Y}}(\eta) &:= \partial\varphi(\mathcal{P}^{\mathcal{X}}(\eta)) \quad (50) \\ &= \{\mathbf{y}|\mathbf{y} = \ln[U\eta + C\xi] - \hat{\mathbf{y}}, (\xi, \eta) \in \mathcal{E}\} \\ \mathcal{P}_{\text{eq}}^{\mathcal{Y}}(\xi^*) &:= \partial\varphi(\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\xi^*)) \\ &= \{\mathbf{y}|\mathbf{y} = C^*\xi^* + U^*\eta^*, \eta^* \in \mathbb{R}^{\ell}\}. \quad (51) \end{aligned}$$

Now, the curved variety $\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\xi^*)$ in \mathcal{X} space is a flat subspace $\mathcal{P}_{\text{eq}}^{\mathcal{Y}}(\xi^*)$ in \mathcal{Y} space. Conversely, the flat subspace $\mathcal{P}^{\mathcal{X}}(\eta)$ in \mathcal{X} is a curved manifold $\mathcal{V}^{\mathcal{Y}}(\eta)$ in \mathcal{Y} .¹⁰ This is the essence of the dually flat structure. In the conventional formulation of information geometry by Amari [37], the theory is typically formulated by using only \mathcal{X} space without explicitly introducing \mathcal{X} and \mathcal{Y} as a pair, which obscures the relation between \mathcal{X} and \mathcal{Y} . Moreover, as we will see in Sec. VI, the relation becomes fundamental when we consider thermodynamics.

In \mathcal{Y} space, for $\mathbf{y}_p, \mathbf{y}_q$, and \mathbf{y}_{eq} satisfying $\mathbf{y}_p \in \mathcal{V}^{\mathcal{Y}}(\eta)$, $\mathbf{y}_q \in \mathcal{P}_{\text{eq}}^{\mathcal{Y}}(\xi^*)$, and $\mathbf{y}_{\text{eq}} \in \mathcal{V}^{\mathcal{Y}}(\eta) \cap \mathcal{P}_{\text{eq}}^{\mathcal{Y}}(\xi^*)$, the Pythagorean relation and the corresponding variational relations [Fig. 5(b)] are given by

$$\mathcal{D}_{\mathcal{Y}}[\mathbf{y}_q||\mathbf{y}_{\text{eq}}] + \mathcal{D}_{\mathcal{Y}}[\mathbf{y}_{\text{eq}}||\mathbf{y}_p] = \mathcal{D}_{\mathcal{Y}}[\mathbf{y}_q||\mathbf{y}_p]. \quad (52)$$

Analogously to Eqs. (47) and (48), we have two variational characterizations of the equilibrium state in \mathcal{Y} space [Fig. 5(b)] as

$$\begin{aligned} \mathbf{y}_{\text{eq}}(\xi^*, \eta) &= \arg \min_{\mathbf{y} \in \mathcal{V}^{\mathcal{Y}}(\eta)} \mathcal{D}_{\mathcal{Y}}[\mathbf{y}'||\mathbf{y}] \\ &\times \text{for any } \mathbf{y}' \in \mathcal{P}_{\text{eq}}^{\mathcal{Y}}(\xi^*), \quad (53) \\ &= \arg \min_{\mathbf{y}' \in \mathcal{P}_{\text{eq}}^{\mathcal{Y}}(\xi^*)} \mathcal{D}_{\mathcal{Y}}[\mathbf{y}'||\mathbf{y}] \\ &\times \text{for any } \mathbf{y} \in \mathcal{V}^{\mathcal{Y}}(\eta). \quad (54) \end{aligned}$$

⁸We omit the proof because this result is commonly known.

⁹It should be noted that the symbols used for $\mathcal{V}^{\mathcal{Y}}(\eta)$ and $\mathcal{P}_{\text{eq}}^{\mathcal{Y}}$ are different from those in our accompanying paper [34]. In Ref. [34], $\mathcal{V}^{\mathcal{Y}}$ and $\mathcal{P}_{\text{eq}}^{\mathcal{Y}}$ are denoted as $\mathcal{P}^{\mathcal{Y}}$ and $\mathcal{V}_{\text{EQ}}^{\mathcal{Y}}$, respectively. In this work, we use \mathcal{P} for flat subspaces and \mathcal{V} for curved manifolds.

¹⁰The space $\mathcal{V}^{\mathcal{Y}}(\eta)$ is not necessarily an algebraic variety because it is not defined by algebraic equations. Thus, we use the word manifold here.

The implication of these equation, especially that of the latter, will be discussed in Sec. VI after clarifying their connection to thermodynamics.

V. TANGENT AND COTANGENT SPACES AND THEIR RIEMANNIAN METRIC

In the previous section, we have obtained the geometric structure by exploiting only the information of the first derivative of $\varphi(\mathbf{x})$. In this section, we explore the information contained in the second derivative of $\varphi(\mathbf{x})$, i.e., its Hessian.

A. Hessian and Fisher information

Because $\varphi(\mathbf{x})$ is strictly convex and smooth, its Hessian matrix (second derivative matrix)

$$G_{\mathcal{X}}(\mathbf{x}) := \left(\frac{\partial^2 \varphi(\mathbf{x})}{\partial x_i \partial x_j} \right)_{i,j} = \text{diag} \left(\frac{1}{\mathbf{x}} \right), \quad (55)$$

is positive definite, which endows the tangent space of \mathcal{X} with a natural inner product structure:

$$\langle \Delta \mathbf{x}, \Delta \mathbf{x}' \rangle_{\mathcal{X}} := \langle \Delta \mathbf{x}, G_{\mathcal{X}}(\mathbf{x}) \Delta \mathbf{x}' \rangle, \quad (56)$$

where $\Delta \mathbf{x}$ and $\Delta \mathbf{x}'$ are elements of the tangent space $\mathcal{T}_{\mathbf{x}}\mathcal{X}$ at $\mathbf{x} \in \mathcal{X}$. Moreover, $\langle \cdot, \cdot \rangle$ is the bilinear form defined on $\mathcal{T}_{\mathbf{x}}\mathcal{X} \times \mathcal{T}_{\mathbf{x}}^*\mathcal{X}$, where $\mathcal{T}_{\mathbf{x}}^*\mathcal{X}$ is the corresponding cotangent space. The induced norm is $|\Delta \mathbf{x}|_{\mathcal{X}} := \sqrt{\langle \Delta \mathbf{x}, \Delta \mathbf{x} \rangle_{\mathcal{X}}}$. Also, $G_{\mathcal{X}}(\mathbf{x})$ induces a mapping from the tangent space $\mathcal{T}_{\mathbf{x}}\mathcal{X}$ to the cotangent space $\mathcal{T}_{\mathbf{x}}^*\mathcal{X}$ via $\Delta \mathbf{y} = G_{\mathcal{X}}(\mathbf{x}) \Delta \mathbf{x} \in \mathcal{T}_{\mathbf{x}}^*\mathcal{X}$.

The Hessian of $\varphi^*(\mathbf{y})$ is computed as

$$G_{\mathcal{Y}}(\mathbf{y}) := \left(\frac{\partial^2 \varphi^*(\mathbf{y})}{\partial y_i \partial y_j} \right)_{i,j} = \text{diag } e^{y+\hat{y}} \quad (57)$$

and satisfies $G_{\mathcal{X}}(\mathbf{x})G_{\mathcal{Y}}(\mathbf{y}) = I$ because $\mathbf{x} = e^{y+\hat{y}}$. Thus, $G_{\mathcal{X}}^{-1}(\mathbf{x}) = G_{\mathcal{Y}}(\mathbf{y})$ and $G_{\mathcal{Y}}^{-1}(\mathbf{y}) = G_{\mathcal{X}}(\mathbf{x})$. The inner product induced by $G_{\mathcal{Y}}(\mathbf{y})$ on the dual tangent space $\mathcal{T}_{\mathbf{y}}\mathcal{Y}$ is

$$\langle \Delta \mathbf{y}, \Delta \mathbf{y}' \rangle_{\mathcal{Y}} := \langle G_{\mathcal{Y}}(\mathbf{y}) \Delta \mathbf{y}, \Delta \mathbf{y}' \rangle. \quad (58)$$

Because of the linear duality between \mathcal{X} and \mathcal{Y} , we have relations: $\mathcal{T}_{\mathbf{x}} \cong \mathcal{T}_{\mathbf{y}}^*$ and $\mathcal{T}_{\mathbf{x}}^* \cong \mathcal{T}_{\mathbf{y}}$.

In information geometry, $G_{\mathcal{X}}(\mathbf{x})$ is known as Fisher information matrix. It is related to the infinitesimal change of the Bregman divergence as

$$ds^2 = 2\mathcal{D}_{\mathcal{X}}[\mathbf{x}||\mathbf{x} + \Delta \mathbf{x}] = \langle \Delta \mathbf{x}, G_{\mathcal{X}}(\mathbf{x}) \Delta \mathbf{x} \rangle \quad (59)$$

$$= 2\mathcal{D}_{\mathcal{Y}}[\mathbf{y} + \Delta \mathbf{y}||\mathbf{y}] = \langle \Delta \mathbf{y}, G_{\mathcal{Y}}(\mathbf{y}) \Delta \mathbf{y} \rangle. \quad (60)$$

B. Fisher information for the dual and mixed coordinates

By inserting $\Delta \mathbf{x} = U \Delta \boldsymbol{\eta} + C \Delta \boldsymbol{\xi}$ or $\Delta \mathbf{y} = U^* \Delta \boldsymbol{\eta}^* + C^* \Delta \boldsymbol{\xi}^*$, we obtain Fisher information matrices for $(\boldsymbol{\eta}, \boldsymbol{\xi})$ and $(\boldsymbol{\eta}^*, \boldsymbol{\xi}^*)$:

$$G_{\boldsymbol{\eta}, \boldsymbol{\xi}}(\mathbf{x}) := \begin{pmatrix} U^T G_{\mathcal{X}}(\mathbf{x}) U & U^T G_{\mathcal{X}}(\mathbf{x}) C \\ C^T G_{\mathcal{X}}(\mathbf{x}) U & C^T G_{\mathcal{X}}(\mathbf{x}) C \end{pmatrix}, \quad (61)$$

$$G_{\boldsymbol{\eta}^*, \boldsymbol{\xi}^*}(\mathbf{y}) := \begin{pmatrix} (U^*)^T G_{\mathcal{Y}}(\mathbf{y}) U^* & (U^*)^T G_{\mathcal{Y}}(\mathbf{y}) C^* \\ (C^*)^T G_{\mathcal{Y}}(\mathbf{y}) U^* & (C^*)^T G_{\mathcal{Y}}(\mathbf{y}) C^* \end{pmatrix}. \quad (62)$$

We can verify that $G_{\boldsymbol{\eta}^*, \boldsymbol{\xi}^*}(\mathbf{y})$ is the inverse matrix of $G_{\boldsymbol{\eta}, \boldsymbol{\xi}}(\mathbf{x})$ by directly computing $G_{\boldsymbol{\eta}^*, \boldsymbol{\xi}^*}(\mathbf{y})G_{\boldsymbol{\eta}, \boldsymbol{\xi}}(\mathbf{x}) = I$ where we use the fact that

$$P_U := U(U^*)^T, \quad P_C := C(C^*)^T, \quad (63)$$

$$P_U^* := U^*U^T, \quad P_C^* := C^*C^T \quad (64)$$

are orthogonal projection matrices and satisfy $P_U + P_C = P_U^* + P_C^* = I$. The effective metric matrix can be further simplified by using the mixed coordinates. With the mixed coordinates, $\Delta \mathbf{x}$ is represented by

$$\Delta \mathbf{x} = U \Delta \boldsymbol{\eta} + G_{\mathcal{Y}}(\mathbf{y}) C^* \Delta \boldsymbol{\xi}^* \quad (65)$$

$$= G_{\mathcal{Y}}(\mathbf{x}) U^* \Delta \boldsymbol{\eta}^* + C \Delta \boldsymbol{\xi}. \quad (66)$$

The cross terms in ds^2 disappear due to the dual orthogonality:

$$ds^2 \sim \langle \Delta \boldsymbol{\eta}, G_{\boldsymbol{\eta}}(\mathbf{x}) \Delta \boldsymbol{\eta} \rangle + \langle \Delta \boldsymbol{\xi}^*, G_{\boldsymbol{\xi}^*}(\mathbf{y}) \Delta \boldsymbol{\xi}^* \rangle, \quad (67)$$

$$\sim \langle \Delta \boldsymbol{\eta}^*, G_{\boldsymbol{\eta}^*}(\mathbf{y}) \Delta \boldsymbol{\eta}^* \rangle + \langle \Delta \boldsymbol{\xi}, G_{\boldsymbol{\xi}}(\mathbf{x}) \Delta \boldsymbol{\xi} \rangle, \quad (68)$$

where $G_{\boldsymbol{\eta}}(\mathbf{x}) := U^T G_{\mathcal{X}}(\mathbf{x}) U$ and $G_{\boldsymbol{\xi}}(\mathbf{x}) := C^T G_{\mathcal{X}}(\mathbf{x}) C$ are the diagonal blocks of $G_{\boldsymbol{\eta}, \boldsymbol{\xi}}(\mathbf{x})$ [Eq. (61)], whereas $G_{\boldsymbol{\eta}^*}(\mathbf{y}) := (U^*)^T G_{\mathcal{Y}}(\mathbf{y}) U^*$ and $G_{\boldsymbol{\xi}^*}(\mathbf{y}) := (C^*)^T G_{\mathcal{Y}}(\mathbf{y}) C^*$ are those of $G_{\boldsymbol{\eta}^*, \boldsymbol{\xi}^*}(\mathbf{y})$ [Eq. (62)]. This is a tangent and cotangent space version of the generalized Pythagorean relation and the orthogonality between the stoichiometric subspaces and equilibrium varieties.¹¹

VI. LINK TO CHEMICAL THERMODYNAMICS

In this section, we clarify how the Hessian geometric structure and its constituents can be related to equilibrium chemical thermodynamics. To this end, because the results were derived kinetically from the law of mass action and detailed balancing, we have to rederive the same results from thermodynamics, if at all possible, without assuming any kinetics. This is achieved in our accompanying paper [34]. Here, we make the correspondence precise. Refer to Ref. [34] for the more general results derived from purely thermodynamic arguments.

A. Gibbs free energy and chemical potential

For an ideal dilute solution, the functional form of the Gibbs free energy density is known as

$$\mathcal{G}(\mathbf{x}) = RT[\ln \mathbf{x} + \hat{\mathbf{v}}^0/RT - \mathbf{1}]^T \mathbf{x} + \mathcal{G}_0. \quad (69)$$

Here, \mathcal{G}_0 is constant, which represents solvent contribution to the free energy [12]. By comparing this equation with Eq. (23), the free energy relates to the convex potential function $\varphi(\mathbf{x})$ as

$$\mathcal{G}(\mathbf{x}) = RT\varphi(\mathbf{x}) + \mathcal{G}_0, \quad (70)$$

where we identify $\hat{\mathbf{y}} = -\hat{\mathbf{v}}^0/RT$.

The chemical potential is calculated from the free energy as

$$\boldsymbol{\mu}(\mathbf{x}) = \frac{\partial \mathcal{G}(\mathbf{x})}{\partial \mathbf{x}} = RT \partial_{\mathbf{x}} \varphi(\mathbf{x}) = RT \mathbf{y}, \quad (71)$$

¹¹Note that neither $G_{\boldsymbol{\eta}}(\mathbf{x})G_{\boldsymbol{\eta}^*}(\mathbf{y}) = I$ nor $G_{\boldsymbol{\xi}}(\mathbf{x})G_{\boldsymbol{\xi}^*}(\mathbf{y}) = I$ generally holds.

where we use Eq. (25). Thus, \mathcal{Y} is the space of chemical potentials, which is thermodynamically conjugate to the concentration space of molecular species \mathcal{X} .

B. Total entropy and Bregman divergence

To clarify the physical setting of our chemical reaction network, we distinguish confined and open chemicals. Here, the concentrations of the confined chemicals are denoted by \mathbf{x} . They evolve with the stoichiometric matrix S by the reaction rate equation, Eq. (1). Similarly, the concentrations of open chemicals change with a stoichiometric matrix O by the chemical reactions. However, we assume that the exchange of the open chemicals with the reservoir is much faster than the time scale of the reactions. Therefore, the concentrations of the open chemicals can be regarded as a constant specified by the chemical potentials of the reservoir $\tilde{\boldsymbol{\mu}}$.

The entropy function of the total system with the reservoir is related to $\varphi(\mathbf{x})$ as

$$\Sigma^{\text{tot}}(\mathbf{x}; \tilde{\mathbf{y}}_{\text{res}}) = R\Omega[\tilde{\mathbf{y}}_{\text{res}}(\mathbf{x}) - \varphi(\mathbf{x})] + \text{const}, \quad (72)$$

where $\tilde{\mathbf{y}}_{\text{res}}$ is specified by the state of the reservoir through the simultaneous equations

$$\frac{1}{RT}O^T\tilde{\boldsymbol{\mu}} = -S^T\tilde{\mathbf{y}}_{\text{res}}, \quad (73)$$

see Eq. (41) in Ref. [34].

The existence of the solution $\tilde{\mathbf{y}}_{\text{res}}$ of Eq. (73) is guaranteed by the equilibrium parameter condition, Eq. (14) in this paper, and Eq. (78) of Ref. [34] (see Appendix A for details).

To assure the uniqueness of $\tilde{\mathbf{y}}_{\text{res}}$, we assume $\text{Ker}[S^T] = \{0\}$ in this subsection. The case $\text{Ker}[S^T] \neq \{0\}$ is discussed the next subsection.

By using the relation between the potential $\varphi(\mathbf{x})$ and the Bregman divergence $\mathcal{D}_{\mathcal{X},\mathcal{Y}}[\mathbf{x}||\mathbf{y}']$, the difference of total entropy between \mathbf{x} and \mathbf{x}' can be associated with the difference of Bregman divergences:

$$\begin{aligned} \Sigma^{\text{tot}}(\mathbf{x}; \tilde{\mathbf{y}}_{\text{res}}) - \Sigma^{\text{tot}}(\mathbf{x}'; \tilde{\mathbf{y}}_{\text{res}}) \\ = -R\Omega[\mathcal{D}_{\mathcal{X},\mathcal{Y}}[\mathbf{x}||\tilde{\mathbf{y}}_{\text{res}}] - \mathcal{D}_{\mathcal{X},\mathcal{Y}}[\mathbf{x}'||\tilde{\mathbf{y}}_{\text{res}}]]. \end{aligned} \quad (74)$$

Since the second law of thermodynamics states that the total entropy function $\Sigma^{\text{tot}}(\mathbf{x}; \tilde{\mathbf{y}}_{\text{res}})$ with respect to \mathbf{x} must be maximized at the equilibrium state \mathbf{x}_{eq} , we obtain it, without stoichiometric constraints, i.e., $\text{Ker}[S^T] = \{0\}$, as

$$\mathbf{x}_{\text{eq}} = \arg \max_{\mathbf{x}} \Sigma^{\text{tot}}(\mathbf{x}; \tilde{\mathbf{y}}_{\text{res}}). \quad (75)$$

By inserting $\mathbf{x}' = \tilde{\mathbf{x}}_{\text{res}}$ into Eq. (74), where $\tilde{\mathbf{x}}_{\text{res}}$ is the Legendre transform of $\tilde{\mathbf{y}}_{\text{res}}$, we obtain

$$\Sigma^{\text{tot}}(\mathbf{x}; \tilde{\mathbf{y}}_{\text{res}}) = \Sigma^{\text{tot}}(\tilde{\mathbf{x}}_{\text{res}}; \tilde{\mathbf{y}}_{\text{res}}) - R\Omega\mathcal{D}_{\mathcal{X},\mathcal{Y}}[\mathbf{x}||\tilde{\mathbf{y}}_{\text{res}}]. \quad (76)$$

Thus,

$$\mathbf{x}_{\text{eq}} = \arg \min_{\mathbf{x}} \mathcal{D}_{\mathcal{X},\mathcal{Y}}[\mathbf{x}||\tilde{\mathbf{y}}_{\text{res}}] = \tilde{\mathbf{x}}_{\text{res}}, \quad (77)$$

and the total entropy production is

$$\begin{aligned} \Delta \Sigma^{\text{tot}}(\mathbf{x} \rightarrow \mathbf{x}_{\text{eq}}; \tilde{\mathbf{y}}_{\text{res}}) &:= \Sigma^{\text{tot}}(\tilde{\mathbf{x}}_{\text{res}}; \tilde{\mathbf{y}}_{\text{res}}) - \Sigma^{\text{tot}}(\mathbf{x}; \tilde{\mathbf{y}}_{\text{res}}) \\ &= R\Omega\mathcal{D}_{\mathcal{X},\mathcal{Y}}[\mathbf{x}||\tilde{\mathbf{y}}_{\text{res}}]. \end{aligned} \quad (78)$$

Equation (77) indicates that specifying a certain reservoir parameter $\tilde{\mathbf{y}}_{\text{res}} \in \mathcal{Y}$ by modulating the reservoir chemical potential $\tilde{\boldsymbol{\mu}}$ is equivalent to specifying the equilibrium state $\mathbf{x}_{\text{eq}} \in \mathcal{X}$ under the reservoir parameter $\tilde{\mathbf{y}}_{\text{res}}$ because of the one-to-one correspondence between $\tilde{\mathbf{x}}_{\text{res}} = \mathbf{x}_{\text{eq}}$ and $\tilde{\mathbf{y}}_{\text{res}}$.

Furthermore, if $\mathbf{x} = \tilde{\mathbf{x}}_{\text{res}} = \mathbf{x}_{\text{eq}}$, then we have

$$\Sigma^{\text{tot}}(\tilde{\mathbf{x}}_{\text{res}}; \tilde{\mathbf{y}}_{\text{res}}) = R\Omega\varphi^*(\tilde{\mathbf{y}}_{\text{res}}) + \text{const}, \quad (79)$$

where we use Eqs. (27) and (72). It indicates that $\varphi^*(\tilde{\mathbf{y}}_{\text{res}})$ is the total entropy at the equilibrium point \mathbf{x}_{eq} specified by the parameter $\tilde{\mathbf{y}}_{\text{res}}$.

C. Entropy production under stoichiometric constraints

If stoichiometric constraints exist, i.e., $\text{Ker}[S^T] \neq \{0\}$, then $\tilde{\mathbf{y}}_{\text{res}}$ cannot be determined uniquely from the reservoir parameter by Eq. (73). Nevertheless, the theoretical formulation in the previous subsection can be applied and the equilibrium state \mathbf{x}_{eq} is specified by the reservoir condition.

Let $\tilde{\mathbf{y}}$ be a particular solution of Eq. (73). A system $\mathbf{x}(t)$ with initial condition \mathbf{x}_0 is restricted to the stoichiometric subspace $\mathcal{P}^{\mathcal{X}}(\boldsymbol{\eta}_0)$ which contains \mathbf{x}_0 . Then, the equilibrium state is characterized as

$$\mathbf{x}_{\text{eq}} = \arg \max_{\mathbf{x} \in \mathcal{P}^{\mathcal{X}}(\boldsymbol{\eta}_0)} \Sigma^{\text{tot}}(\mathbf{x}; \tilde{\mathbf{y}}) \quad (80)$$

$$= \arg \min_{\mathbf{x} \in \mathcal{P}^{\mathcal{X}}(\boldsymbol{\eta}_0)} \mathcal{D}_{\mathcal{X},\mathcal{Y}}[\mathbf{x}||\tilde{\mathbf{y}}]. \quad (81)$$

Let $\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\tilde{\boldsymbol{\xi}}^*)$ be the equilibrium variety to which $\tilde{\mathbf{y}}$ belongs, i.e., $\tilde{\mathbf{x}}$, the Legendre transform of $\tilde{\mathbf{y}}$, satisfies $\tilde{\mathbf{x}} \in \mathcal{V}_{\text{eq}}^{\mathcal{X}}(\tilde{\boldsymbol{\xi}}^*)$. Moreover, let \mathbf{x}_{int} be the intersection point of the subspace and the variety, i.e., $\mathbf{x}_{\text{int}} \in \mathcal{P}^{\mathcal{X}}(\boldsymbol{\eta}_0) \cap \mathcal{V}_{\text{eq}}^{\mathcal{X}}(\tilde{\boldsymbol{\xi}}^*)$. As shown before, this intersection is unique and can be specified as $\mathbf{x}_{\text{int}} = \mathbf{x}_{\text{int}}(\boldsymbol{\eta}_0, \tilde{\boldsymbol{\xi}}^*)$. For any $\mathbf{x} \in \mathcal{P}^{\mathcal{X}}(\boldsymbol{\eta}_0)$, the generalized Pythagorean relation holds:

$$\mathcal{D}_{\mathcal{X},\mathcal{Y}}[\mathbf{x}||\tilde{\mathbf{y}}] = \mathcal{D}_{\mathcal{X},\mathcal{Y}}[\mathbf{x}||\mathbf{y}_{\text{int}}] + \mathcal{D}_{\mathcal{X},\mathcal{Y}}[\mathbf{x}_{\text{int}}||\tilde{\mathbf{y}}]. \quad (82)$$

Thus,

$$\begin{aligned} \mathbf{x}_{\text{eq}} &= \arg \min_{\mathbf{x} \in \mathcal{P}^{\mathcal{X}}(\boldsymbol{\eta}_0)} [\mathcal{D}_{\mathcal{X},\mathcal{Y}}[\mathbf{x}||\mathbf{y}_{\text{int}}] + \mathcal{D}_{\mathcal{X},\mathcal{Y}}[\mathbf{x}_{\text{int}}||\tilde{\mathbf{y}}]] \\ &= \arg \min_{\mathbf{x} \in \mathcal{P}^{\mathcal{X}}(\boldsymbol{\eta}_0)} \mathcal{D}_{\mathcal{X},\mathcal{Y}}[\mathbf{x}||\mathbf{y}_{\text{int}}] = \mathbf{x}_{\text{int}}. \end{aligned} \quad (83)$$

This verifies that the equilibrium state \mathbf{x}_{eq} , which maximizes the total entropy, is characterized by the intersection point of the subspace and variety. Also, the entropy production becomes

$$\begin{aligned} \Delta \Sigma^{\text{tot}}(\mathbf{x} \rightarrow \mathbf{x}_{\text{eq}}; \tilde{\mathbf{y}}) &:= \Sigma^{\text{tot}}(\mathbf{x}_{\text{eq}}; \tilde{\mathbf{y}}) - \Sigma^{\text{tot}}(\mathbf{x}; \tilde{\mathbf{y}}) \\ &= R\Omega\mathcal{D}_{\mathcal{X},\mathcal{Y}}[\mathbf{x}||\mathbf{y}_{\text{eq}}], \end{aligned} \quad (84)$$

where the relevant part is independent of $\tilde{\mathbf{y}}$ owing to the Pythagorean relation. The convergence to \mathbf{x}_{eq} is then attributed to the second law, which was also proven kinetically for the mass action system [see Eq. (49)].

The independence of the choice of $\tilde{\mathbf{y}}$ can be understood more clearly in \mathcal{Y} space from the dual variational equation:

$$\mathbf{y}_{\text{eq}} = \arg \min_{\tilde{\mathbf{y}} \in \mathcal{P}_{\text{eq}}^{\mathcal{Y}}(\tilde{\boldsymbol{\xi}}^*)} \mathcal{D}_{\mathcal{X},\mathcal{Y}}[\mathbf{x}_0||\tilde{\mathbf{y}}]$$

$$\begin{aligned}
 &= \arg \min_{\tilde{\mathbf{y}} \in \mathcal{P}_{\text{eq}}^{\mathcal{Y}}(\tilde{\boldsymbol{\xi}}^*)} [\mathcal{D}_{\mathcal{X},\mathcal{Y}}[\mathbf{x}_0 \parallel \mathbf{y}_{\text{int}}] + \mathcal{D}_{\mathcal{X},\mathcal{Y}}[\mathbf{x}_{\text{int}} \parallel \tilde{\mathbf{y}}]] \\
 &= \arg \min_{\tilde{\mathbf{y}} \in \mathcal{P}_{\text{eq}}^{\mathcal{Y}}(\tilde{\boldsymbol{\xi}}^*)} \mathcal{D}_{\mathcal{X},\mathcal{Y}}[\mathbf{x}_{\text{int}} \parallel \tilde{\mathbf{y}}] = \mathbf{y}_{\text{int}}, \tag{85}
 \end{aligned}$$

which corresponds to Eq. (54). Note that the equilibrium subspace $\mathcal{P}_{\text{eq}}^{\mathcal{Y}}(\tilde{\boldsymbol{\xi}}^*)$ defined by Eq. (51) is composed of the solutions of Eq. (73). The coordinate $\tilde{\boldsymbol{\xi}}^*$ is uniquely determined by $\tilde{\boldsymbol{\xi}}^* = C^T \tilde{\mathbf{y}}$ for an arbitrary particular solution $\tilde{\mathbf{y}}$. This fact represents that, if there exist stoichiometric constraints, the relevant quantity specified by the reservoir chemical potentials $\tilde{\boldsymbol{\mu}}$ is no longer a point or value but geometric objects, i.e., the equilibrium variety $\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\tilde{\boldsymbol{\xi}}^*)$ and subspace $\mathcal{P}_{\text{eq}}^{\mathcal{Y}}(\tilde{\boldsymbol{\xi}}^*)$. Furthermore, the coordinate $\tilde{\boldsymbol{\xi}}^*$ specifies the equilibrium state with the initial condition \mathbf{x}_0 .

D. Linear response of total entropy

Finally, we investigate responses of the total entropy to infinitesimal changes of either \mathbf{x} or \mathbf{y}_{res} . First, suppose that state \mathbf{x} is perturbed to $\mathbf{x}' = \mathbf{x} + \Delta\mathbf{x}$. For a general perturbation, which is not restricted by stoichiometric constraints, we obtain

$$\begin{aligned}
 \frac{\Delta \Sigma^{\text{tot}}(\mathbf{x} \rightarrow \mathbf{x}'; \tilde{\mathbf{y}}_{\text{res}})}{R\Omega} &= \mathcal{D}_{\mathcal{X}}[\mathbf{x} + \Delta\mathbf{x} \parallel \mathbf{x}_{\text{eq}}] - \mathcal{D}_{\mathcal{X}}[\mathbf{x} \parallel \mathbf{x}_{\text{eq}}] \\
 &\approx \langle \Delta\mathbf{x}, \mathbf{y} - \mathbf{y}_{\text{eq}} \rangle. \tag{86}
 \end{aligned}$$

Here, $\Delta\mathbf{x}$ and $\mathbf{y} - \mathbf{y}_{\text{eq}}$ are treated as elements of \mathcal{X} and \mathcal{Y} , respectively. Moreover, we use an implicit identification of the base spaces with their tangent and cotangent spaces. Because \mathcal{X} and \mathcal{Y} inherit the affine structure of the dual vector spaces \mathbb{V} and \mathbb{V}^* , and because the tangent and cotangent spaces are isomorphic to these vector spaces, we can identify these spaces (noncanonically) as

$$\mathcal{T}_{\mathbf{y}}^* \mathcal{Y} \cong \mathcal{T}_{\mathbf{x}} \mathcal{X} \cong \mathbb{V} \supset \mathcal{X}, \quad \mathcal{T}_{\mathbf{x}}^* \mathcal{X} \cong \mathcal{T}_{\mathbf{y}} \mathcal{Y} \cong \mathbb{V}^* = \mathcal{Y}.$$

Because of this isomorphism, we can also regard $\Delta\mathbf{x}$ and $\mathbf{y} - \mathbf{y}_{\text{eq}}$ as elements of tangent and cotangent spaces. By combining this fact with the Cauchy-Schwartz inequality, we obtain

$$\left| \frac{\Delta \Sigma^{\text{tot}}(\mathbf{x} \rightarrow \mathbf{x}'; \tilde{\mathbf{y}}_{\text{res}})}{R\Omega} \right| = \left| \langle \Delta\mathbf{x}, \mathbf{y} - \mathbf{y}_{\text{eq}} \rangle \right| \tag{87}$$

$$\leq |\mathbf{y} - \mathbf{y}_{\text{eq}}|_{\mathcal{Y}} |\Delta\mathbf{x}|_{\mathcal{X}}, \tag{88}$$

where $|\cdot|_{\mathcal{X}}$ and $|\cdot|_{\mathcal{Y}}$ are the metrics on the tangent and cotangent spaces induced by the Fisher information [Eqs. (56) and (58)]. From this, we see that $|\mathbf{y} - \mathbf{y}_{\text{eq}}|_{\mathcal{Y}}$ is an upper bound of the sensitivity of entropy production proposed in Ref. [35] and that entropy production is maximized if $\Delta\mathbf{x}$ is parallel to $\mathbf{y} - \mathbf{y}_{\text{eq}}$ in the sense of linear duality.

VII. COMPLEX-BALANCED SYSTEMS

The Hessian geometric structure can also be extended beyond equilibrium chemical reaction networks to complex balanced ones.

Complex balanced networks, introduced by Horn and Jackson [18], are a class of nonequilibrium reaction networks, which retains several properties of equilibrium chemical

networks.¹² A complex balanced network has a unique steady state in each stoichiometric compatibility class called complex balanced state, which is also globally stable [29,30]. Moreover, the generalized Kullback-Leibler divergence works as the Lyapunov function of the network [18]. This similarity is partially attributed to the shared geometric structure between equilibrium and complex balanced networks.

To define the complex balanced state, we note that the stoichiometric matrix S can be decomposed as $S = -\Gamma B$ by exploiting the underlying hypergraph-theoretic structure of chemical reaction networks. To derive Γ and B , let

$$\gamma_{1,m}^+ X_1 + \dots + \gamma_{N,m}^+ X_N \rightarrow \gamma_{1,m}^- X_1 + \dots + \gamma_{N,m}^- X_N \tag{89}$$

be the chemical reaction equation of m th reaction. Because reactants and products of a reaction generally include multiple molecular species, a reaction is a directed hyperedge linking multiple species (vertices). Thus, a chemical reaction network is an oriented hypergraph. The sets of molecules composing the reactants and products of a reaction are called complexes. In the case of Eq. (89), $\gamma_{1,m}^+ X_1 + \dots + \gamma_{N,m}^+ X_N$ and $\gamma_{1,m}^- X_1 + \dots + \gamma_{N,m}^- X_N$ are complexes. By regarding complexes as vertices, a chemical reaction network is interpreted as an oriented graph in which each reaction links two complexes. Suppose that a chemical reaction network composed of H complexes and $\gamma_{1,h} X_1 + \dots + \gamma_{N,h} X_N$ be the h th complex. Then, the matrix $\Gamma = (\gamma_{i,h}) \in \mathbb{Z}_{\geq 0}^{N \times H}$ defines all the complexes. Next, suppose that m th reaction has h_r th complex as the reactant and h_p th complex as the product. This information is encoded by a vector $\mathbf{b}_m \in \{\pm 1\}^{H \times 1}$, whose components $b_{h,m}$ are +1 if $h = h_r$, -1 if $h = h_p$, and 0 otherwise. The matrix $B = (b_{h,m}) \in \{\pm 1\}^{H \times M}$ is the incidence matrix of the oriented graph for the chemical reaction network. It is straightforward to verify that $S = -\Gamma B$ and $\Gamma^\pm = \Gamma B^\pm$ hold where $B^\pm = \max[\pm B, 0]$.

With the assumption that the reaction flux $\mathbf{j}^\pm(\mathbf{x}_{cb}; \boldsymbol{\theta})$ satisfies the law of mass action, the set of complex balanced states is defined by

$$\mathcal{V}_{cb}^{\mathcal{X}}(\boldsymbol{\theta}) := \{\mathbf{x}_{cb} > 0 \mid B\mathbf{j}(\mathbf{x}_{cb}; \boldsymbol{\theta}) = \mathbf{0}\}. \tag{90}$$

Because $\mathbf{j}^\pm(\mathbf{x}_{cb}; \boldsymbol{\theta})$ consists of monomials, $\mathcal{V}_{cb}^{\mathcal{X}}(\boldsymbol{\theta})$ is an algebraic variety. As before, we define the parameter sets Θ_{cb} in which $\mathcal{V}_{cb}^{\mathcal{X}}(\boldsymbol{\theta})$ is nonempty, i.e.,

$$\Theta_{cb} := \{\boldsymbol{\theta} \mid \mathcal{V}_{cb}^{\mathcal{X}}(\boldsymbol{\theta}) \neq \emptyset\}. \tag{91}$$

Obviously, $\Theta_{\text{eq}} \subset \Theta_{cb}$ and thus an equilibrium variety is a special class of complex balanced varieties. Compared with the definition of an equilibrium variety, $B\mathbf{j}(\mathbf{x}_{cb}; \boldsymbol{\theta}) = \mathbf{0}$ is in general not given by binomial equations. Nevertheless, $\mathcal{V}_{cb}^{\mathcal{X}}(\boldsymbol{\theta})$ is a toric variety, meaning that $B\mathbf{j}(\mathbf{x}_{cb}; \boldsymbol{\theta}) = \mathbf{0}$ can be converted to binomial equations by appropriate algebraic manipulation similar to the Gaussian elimination for linear equations [19]. In other words, a complex balanced state is defined by hidden detailed balance (binomial) equations (see also Appendix B for the derivation). Thus, as a toric variety, $\mathcal{V}_{cb}^{\mathcal{X}}(\boldsymbol{\theta})$ has an explicit parametrization in the same way as the equilibrium

¹²Whether a given network is complex balanced one or not depends on the value of kinetic parameters. Depending on the parameter value, the same network can be noncomplex balanced one.

variety [Eq. (15)]. Moreover, the complex balanced variety has the same design matrix and thus is given by

$$\mathcal{V}_{cb}^{\mathcal{X}}(\boldsymbol{\theta}) = \{\mathbf{x} | \ln \mathbf{x} = \ln \tilde{\mathbf{x}}_{cb} + U^* \boldsymbol{\eta}^*\}, \quad (92)$$

where $\tilde{\mathbf{x}}_{cb}$ is determined by the actual values of kinetic parameters [19]. It should be noted that multiple parameter values $\boldsymbol{\theta}, \boldsymbol{\theta}' \in \Theta_{cb}$ can generate the same variety $\mathcal{V}_{cb}^{\mathcal{X}}(\boldsymbol{\theta}) = \mathcal{V}_{cb}^{\mathcal{X}}(\boldsymbol{\theta}')$.

As geometrical objects, the complex balanced variety and equilibrium variety are indistinguishable and therefore the embedding of each complex balanced variety into \mathcal{X} is parametrized analogously by $\boldsymbol{\xi}^*$. From now on, we write $\mathcal{V}_{cb}^{\mathcal{X}}(\boldsymbol{\xi}^*)$ for the complex balanced variety embedded in \mathcal{X} . As a result, the Hessian geometric structure of the equilibrium variety described in Sec. IV is preserved in the more general case of complex balancing.

For example, for a given initial state $\mathbf{x}_0 \in \mathcal{P}^{\mathcal{X}}(\boldsymbol{\eta})$, the complex balanced steady state \mathbf{x}_{cb} is the intersection of the subspace $\mathcal{P}^{\mathcal{X}}(\boldsymbol{\eta})$, determined by the initial state, and the complex balanced variety $\mathcal{V}_{cb}^{\mathcal{X}}(\boldsymbol{\xi}^*)$:

$$\mathbf{x}_{cb} \in \mathcal{P}^{\mathcal{X}}(\boldsymbol{\eta}) \cap \mathcal{V}_{cb}^{\mathcal{X}}(\boldsymbol{\xi}^*). \quad (93)$$

For the same reason as the equilibrium state, \mathbf{x}_{cb} is unique and the intersection is transversal [19]. By using the same potential function $\varphi(\mathbf{x})$ as in Eq. (23), define the dual space \mathcal{Y} , the Legendre transformed function $\varphi^*(\mathbf{x})$, the Bregman divergence $\mathcal{D}_{\mathcal{X}}$, the tangent and cotangent spaces $\mathcal{T}\mathcal{X}$ and $\mathcal{T}\mathcal{Y}$, and also the Fisher information given by $G_{\mathcal{X}}(\mathbf{x})$ and $G_{\mathcal{Y}}(\mathbf{y})$. In other words, we can naturally embed a $\mathcal{V}_{cb}^{\mathcal{X}}(\boldsymbol{\theta})$ into the Hessian geometric structure constructed for the equilibrium varieties. As a result, the same Pythagorean relation as Eq. (46) holds for the complex balanced case:

$$\mathcal{D}_{\mathcal{X}}[\mathbf{x}_p || \mathbf{x}_{cb}] + \mathcal{D}_{\mathcal{X}}[\mathbf{x}_{cb} || \mathbf{x}_q] = \mathcal{D}_{\mathcal{X}}[\mathbf{x}_p || \mathbf{x}_q], \quad (94)$$

where $\mathbf{x}_p \in \mathcal{P}^{\mathcal{X}}(\boldsymbol{\eta})$, $\mathbf{x}_q \in \mathcal{V}_{cb}^{\mathcal{X}}(\boldsymbol{\xi}^*)$, and $\mathbf{x}_{cb} \in \mathcal{P}^{\mathcal{X}}(\boldsymbol{\eta}) \cap \mathcal{V}_{cb}^{\mathcal{X}}(\boldsymbol{\xi}^*)$. Thus, the complex balanced state admits the variational characterizations

$$\begin{aligned} \mathbf{x}_{cb}(\boldsymbol{\eta}, \boldsymbol{\xi}^*) &= \arg \min_{\mathbf{x} \in \mathcal{P}^{\mathcal{X}}(\boldsymbol{\eta})} \mathcal{D}_{\mathcal{X}}[\mathbf{x} || \mathbf{x}'] \\ &\times \text{for any fixed } \mathbf{x}' \in \mathcal{V}_{cb}^{\mathcal{X}}(\boldsymbol{\xi}^*) \end{aligned} \quad (95)$$

$$\begin{aligned} &= \arg \min_{\mathbf{x}' \in \mathcal{V}_{cb}^{\mathcal{X}}(\boldsymbol{\xi}^*)} \mathcal{D}_{\mathcal{X}}[\mathbf{x} || \mathbf{x}'] \\ &\times \text{for any fixed } \mathbf{x} \in \mathcal{P}^{\mathcal{X}}(\boldsymbol{\eta}). \end{aligned} \quad (96)$$

The analogous variational characterization also holds in the dual space.

While the geometric structure is inherited, the thermodynamic aspects are not. The potential function $\varphi(\mathbf{x})$ and the space \mathcal{Y} can admit the same interpretation as the Gibbs free energy function and the chemical potential space. By contrast, the potential function $\varphi(\mathbf{x})$ is no longer associated with total entropy as in Eq. (72) because Eq. (73) may not have a solution in the complex balanced case, Eq. (91). Therefore, the Bregman divergence $\mathcal{D}_{\mathcal{X}}$ does not correspond to the difference of total entropy as in Eq. (74). All of this is rooted in the general lack of a thermodynamically consistent characterization of nonequilibrium states. While there have been continuous attempts to achieve a thermodynamic characterization of complex balanced states [12,16,35], it is still an open problem. We

believe that the Hessian geometric structure can contribute to the resolution of the problem [53].

VIII. STOCHASTIC THERMODYNAMICS ON GRAPH

Finally, we demonstrate how our formalism, as a special case, includes finite-state Markov jump processes.

A reversible Markov jump process on an oriented graph with a finite number of vertices and edges is given by

$$\frac{dp_i}{dt} = \sum_{j=1}^N [W_{i||j} p_j - W_{j||i} p_i]. \quad (97)$$

This process is often employed as a model in stochastic thermodynamics [8,10,54,55]. Here, i is the index of vertices of the graph, and N is the total number of vertices. $W_{i||j}$ is the transition rate from state j to i . The transition from a vertex to itself is usually prohibited, i.e., $W_{i||i} = 0$.

We label the edges by the running variable m , and determine the orientation of each edge arbitrarily. Let B be the incidence matrix of the oriented graph. We define functions $h(m)$ and $t(m)$ that return the index of head and tail vertices of the m th oriented edge, respectively. Then, Eq. (97) can be mapped to a chemical rate equation [Eq. (1)] by defining $x_i = p_i$, $\Gamma = I$, $S = -\Gamma B$, and

$$j_m^+(\mathbf{x}) = W_{t(m)||h(m)} x_{h(m)}, \quad (98)$$

$$j_m^-(\mathbf{x}) = W_{h(m)||t(m)} x_{t(m)}. \quad (99)$$

All fluxes are linear in this case. In stochastic thermodynamics, we often assume the ergodicity of the Markov process,¹³ to assure that Eq. (97) has a uniquely and globally stable steady state $\boldsymbol{\pi}$ by the Perron-Frobenius theorem. Since this assumption guarantees the connectivity of the incidence matrix, B has the unique left null vector $\mathbf{1}$ as $\mathbf{1}^T B = \mathbf{0}$. Thus, the system has only one conserved quantity $\eta = \mathbf{1}^T \mathbf{p}$, which reflects the conservation of total probability, i.e., $\mathbf{1}^T \mathbf{p} = 1$. Thus, $U^* = \mathbf{1}$, and $\mathcal{P}^{\mathcal{X}}(\boldsymbol{\eta}) = \{\mathbf{x} | \eta = (U^*)^T \mathbf{x}\} = \{\mathbf{x} | \eta = \mathbf{1}^T \mathbf{x}\}$. The stoichiometric subspace is nothing but the $N - 1$ dimensional simplex of probability distributions if we fix $\eta = 1$.

The detailed balance condition of Eq. (97) is typically defined as

$$W_{i||j} \pi_j = W_{j||i} \pi_i \quad \text{for all } i \text{ and } j. \quad (100)$$

This is equivalent to the detailed balance condition of the chemical rate equation: $j_m^+(\mathbf{x}_{eq}) = j_m^-(\mathbf{x}_{eq})$ for all m , i.e., $\mathbf{j}(\mathbf{x}_{eq}) = \mathbf{0}$, whereby $\mathbf{x}_{eq} = \boldsymbol{\pi}$. The equilibrium constant characterizing an equilibrium state becomes $K_m = W_{i(m)||h(m)} / W_{h(m)||t(m)}$, which also satisfies

$$\ln K = -B^T \ln \boldsymbol{\pi} \quad (101)$$

as follows from the detailed balance condition. Thus, the equilibrium variety can be defined for the Markov jump process.

¹³To be more precise, the ergodicity here is mathematically defined as follows. Define a matrix \hat{W} whose off-diagonal elements consists of $\hat{W}_{ij} := W_{i||j}$ and diagonal ones are $\hat{W}_{ii} := \sum_{j=1}^N W_{j||i}$. If the matrix \hat{W} is irreducible, then we say that the Markov process, Eq. (97), is ergodic.

Its explicit parametrization [Eq. (15)] is given by

$$\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\mathbf{K}) = \{\mathbf{x} | \mathbf{x} = \tilde{\mathbf{x}}_{\text{eq}} \circ \exp[\mathbf{1}\eta^*]\}. \quad (102)$$

Here, we used that $U = \mathbf{1}^T$ and that $\tilde{\mathbf{x}}_{\text{eq}}$ satisfies $\ln \mathbf{K} = S^T \ln \tilde{\mathbf{x}}_{\text{eq}} = -B^T \ln \tilde{\mathbf{x}}_{\text{eq}}$. From Eq. (101), choosing the identification $\tilde{\mathbf{x}}_{\text{eq}} = \boldsymbol{\pi}$ yields

$$\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\mathbf{K}) = \{\mathbf{x} | \mathbf{x} = \boldsymbol{\pi} \circ \exp[\mathbf{1}\eta^*]\}. \quad (103)$$

The variety $\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\mathbf{K})$ is one-dimensional because the codimension of $\mathcal{P}^{\mathcal{X}}(\eta)$ is one. The intersection of $\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\mathbf{K})$ and $\mathcal{P}^{\mathcal{X}}(\eta)$ is trivially $\boldsymbol{\pi}$ if $\eta = 1$. This means that, in the case of stochastic dynamics on a graph, the Hessian structure is not fully evident because the equilibrium variety is just one-dimensional and the subspace is always fixed at $\eta = 1$, i.e., $\mathbf{1}^T \mathbf{p} = 1$. Under this constraint, the potential function is reduced to the conventional Kullback-Leibler divergence, $\varphi(\mathbf{p}) = \mathbf{p}^T \ln(\mathbf{p}/\hat{\mathbf{p}})$, where we define $\ln \hat{\mathbf{p}} := \hat{y}$. The Legendre dual of \mathbf{p} is $\mathbf{y} = \partial\varphi(\mathbf{p}) = \ln \mathbf{p} - \ln \hat{\mathbf{p}}$. Thus, even though $\ln \mathbf{p}$ looks as if it is just the logarithm of \mathbf{p} , the vectors \mathbf{p} and $\ln \mathbf{p}$ should be discriminated as objects defined on different spaces. The dual of φ is $\varphi^*(\mathbf{y}) = \mathbf{1}^T e^{\mathbf{y}+\hat{y}}$. The Bregman divergence is now reduced to the conventional Kullback-Leibler divergence $\mathcal{D}[\mathbf{p}||\mathbf{p}'] = \mathbf{p}^T \ln(\mathbf{p}/\mathbf{p}')$. Even though $\varphi(\mathbf{p})$ and $\mathcal{D}[\mathbf{p}||\mathbf{p}']$ are different quantities, both have the same expression as a Kullback-Leibler divergence, which can cause confusion. Finally, the Pythagorean relation does not provide nontrivial information because the subspace $\mathcal{P}^{\mathcal{X}}(\eta)$ is fixed as the probability simplex and the variety $\mathcal{V}_{\text{eq}}^{\mathcal{X}}(\mathbf{K})$ is one-dimensional. Thus, the Hessian geometric structure is not fully exhibited under the restricted setting of the Markov jump processes. Nonetheless, the above discussion is beneficial because it allows to discriminate degenerate quantities that have different meanings, such as \mathbf{p} and $\ln \mathbf{p}$ or $\varphi(\mathbf{x})$ and $\mathcal{D}_{\mathcal{X}}$ and also to point out their locations in dual spaces [56].

IX. DISCUSSION

In this work, we derived the Hessian geometric structure of chemical reaction networks which satisfy the law of mass action. When establishing the geometry, a crucial role is played by the fact that the equilibrium and complex balanced states are given by toric varieties. A purely thermodynamic argument in Ref. [34] assures, however, that this geometric structure holds for a broader class of chemical reaction systems than those obeying mass action kinetics. In the last century, there have been continuous attempts to extend the properties of equilibrium and complex balanced systems beyond the law of mass action in physics and applied mathematics [15,20,57]. We expect that our results can provide a new theoretical basis for the development of chemical reaction network theory to such generalized models and situations.

Our current approach relies on the fact that the state space of the system is represented by finite-dimensional vectors: the concentration vector in chemical reaction networks and the probability vector in the Markov jump processes. However, in reaction systems with a spatially inhomogeneous structure or in general stochastic thermodynamic models, the system is described by a positive measure or a probability measure on a continuous space. It is an open question whether the structure

presented here can be extended to such cases. It is naturally expected that information or Hessian geometry on an infinite-dimensional space becomes necessary to address the problem [58,59].

For the theory of information geometry and Hessian geometry [36,37], chemical reaction networks provide a new and fertile field to apply and develop the theory. While major applications of the theory have been restricted to statistics, information science, and other applied mathematical topics, more recently, applications to stochastic thermodynamics have been attempted [55,56,60,61]. Now, we add chemical reaction networks to the list [62]. In addition, from the viewpoint of real algebraic geometry, our results on chemical reaction networks may suggest a way to generalize Birch's theorem to more general kinetic models [20].

The mathematical structure of chemical kinetics and chemical thermodynamics, since their establishment, has been continuously developed within various fields such as physics, applied mathematics, applied chemistry, and systems biology over the past century [1,3,6,12,63]. However, the developments were mostly separated and shared only within the individual fields. In applied mathematics, there is chemical reaction network theory by Feinberg [3], which is based on the work of Aris, Horn, and Jackson [18,63]. In real algebraic geometry, chemical reaction networks and toric geometry are becoming important research topics [29–32]. In systems biology, a new theory emerged, which connects properties of reaction networks with the network topology [23–26,28]. In physics, network thermodynamics by Hill and Schnakenberg [8,10] and stochastic thermodynamics of chemical reaction theory by Qian and Esposito have been studied [6,12,64]. Now, information and Hessian geometry can be added to this variety of applications [62]. Even though the theories have been developed to explain the same physical object, i.e., chemical reaction networks, the interrelationships between them are not clear yet. It will be the next important step to integrate these theories from a unified perspective, which is expected to boost a further development of chemical reaction theory.

ACKNOWLEDGMENTS

This research is supported by JSPS KAKENHI Grants No. 19H05799 and No. 21K21308, and by JST CREST Grants No. JPMJCR2011 and No. JPMJCR1927. We thank the members of our laboratory for fruitful discussions.

APPENDIX A

In thermodynamic mass action kinetics, the rate constants k^{\pm} satisfies

$$\log \frac{k^+}{k^-} = -\frac{1}{RT} \{S^T \hat{y}^o + O^T \tilde{\boldsymbol{\mu}}\}, \quad (A1)$$

which is known as the local detailed balance condition [11,12]. This condition bridges kinetics and thermodynamics. Since $O^T \tilde{\boldsymbol{\mu}} \in \text{Im}[S^T]$ holds from Eq. (14), the simultaneous equations, Eq. (73), with respect to $\tilde{\mathbf{y}}_{\text{res}}$, must be consistent. Accordingly, the existence of the solution is guaranteed.

APPENDIX B

The complex balanced variety defined in Eq. (90) is the toric variety with the same design matrix as the equilibrium variety [Eq. (15)]. This fact was derived in Ref. [19]. We outline the derivation here (refer to Ref. [19] for the more rigorous derivation). By using the decomposition $S = -\Gamma B$, the chemical rate equation with mass action kinetics [Eq. (1)] can be represented as

$$\frac{d\mathbf{x}}{dt} = -\Gamma B[\text{diag} \mathbf{k}^+(B^+)^T - \text{diag} \mathbf{k}^-(B^-)^T] \mathbf{x}^{\Gamma T}, \quad (\text{B1})$$

where $B^+ = \max[B, 0]$ and $B^- = \max[-B, 0]$ are the matrices obtained by using only +1 or -1 components in B . Thus, $B = B^+ - B^-$ holds, and B^+ and B^- encode the complexes of reactants and products, respectively. $\mathbf{x}^{\Gamma T}$ is a H dimensional vector whose h th component is defined as

$$\mathbf{x}^{\Gamma T} := \left(\prod_{i=1}^N x_i^{\gamma_{i,1}}, \dots, \prod_{i=1}^N x_i^{\gamma_{i,H}} \right)^T. \quad (\text{B2})$$

Thus, the h th component of $\mathbf{x}^{\Gamma T}$ represents the monomial corresponding to the h th complex. The transpose on $\mathbf{x}^{\Gamma T}$ is for the notational consistency with the following relation

$$\ln \mathbf{x}^{\Gamma T} = \Gamma^T \ln \mathbf{x},$$

where $\ln \mathbf{x}$ is the component-wise logarithm of \mathbf{x} . The matrices, Γ and B^\pm are related to Γ^\pm in Eq. (11) as $\Gamma^\pm = \Gamma B^\pm$. We define an asymmetric weighted graph-Laplacian L_{k^\pm} as

$$L_{k^\pm} := -B[\text{diag} \mathbf{k}^+(B^+)^T - \text{diag} \mathbf{k}^-(B^-)^T]. \quad (\text{B3})$$

Then, Eq. (B1) can be represented as

$$\frac{d\mathbf{x}}{dt} = \Gamma L_{k^\pm} \mathbf{x}^{\Gamma T}. \quad (\text{B4})$$

With this, the complex balanced variety can be rewritten as

$$\mathcal{V}_{cb}^{\mathcal{X}}(\boldsymbol{\theta}) := \{\mathbf{x}_{cb} > 0 | L_{k^\pm} \mathbf{x}_{cb}^{\Gamma T} = \mathbf{0}\}, \quad (\text{B5})$$

where $\boldsymbol{\theta} = \mathbf{k}^\pm$. Thus, $\mathbf{x}_{cb}^{\Gamma T} \in \text{Ker} L_{k^\pm}$ holds. Because $L_{k^\pm} \mathbf{x}_{cb}^{\Gamma T} = \mathbf{0}$ reduces to $L_{k^\pm} \mathbf{x}_{cb} = \mathbf{0}$ if $\Gamma = I$, L_{k^\pm} is equivalent to the operator of the Markov jump process with the jump rate parameter \mathbf{k}^\pm .

For simplicity, we consider only the case that L_{k^\pm} is irreducible, which means that the directed graph defined by L_{k^\pm} is strictly connected. Because we consider only reversible reactions, this is also equivalent to the condition that the unoriented graph defined by B is connected.¹⁴ The extension to the case with multiple connected components in the graph is straightforward but tiresome.

By the Perron-Frobenius theorem for an irreducible L_{k^\pm} , we have the unique nonnegative eigenvector $\boldsymbol{\Sigma} > \mathbf{0}$ associated with the largest eigenvalue 0, i.e., $L_{k^\pm} \boldsymbol{\Sigma} = \mathbf{0}$. Thus, the vector $\boldsymbol{\Sigma}$ parametrizes the kernel space $\text{Ker} L_{k^\pm}$.¹⁵ We can explicitly obtain $\boldsymbol{\Sigma}$ as a function of the kinetic parameter \mathbf{k}^\pm via the Markov chain tree theorem for weighted digraphs [65]. Using $\boldsymbol{\Sigma}$, the complex balanced variety becomes

$$\mathcal{V}_{cb}^{\mathcal{X}}(\boldsymbol{\theta}) := \{\mathbf{x}_{cb} > 0 | \mathbf{x}_{cb}^{\Gamma T} \in \langle \boldsymbol{\Sigma} \rangle\}, \quad (\text{B6})$$

where $\langle \boldsymbol{\Sigma} \rangle = \{\mathbf{x} | \mathbf{x} = r \boldsymbol{\Sigma}, r \in \mathbb{R}\}$ is a vector space generated by $\boldsymbol{\Sigma}$. If $\mathbf{x}_{cb} > \mathbf{0}$ exists for a fixed parameter value of \mathbf{k}^\pm , then it satisfies $\mathbf{x}_{cb}^{\Gamma T} = \frac{1}{\alpha} \boldsymbol{\Sigma}$ for a certain scalar $\alpha > 0$. By deleting α from this relation, we get, for $h, h' \in [1, H]$,

$$\sum_h \mathbf{x}_{cb}^{\gamma_h^T} = \sum_{h'} \mathbf{x}_{cb}^{\gamma_{h'}^T}, \quad (\text{B7})$$

holds where γ_h is the h th column vector of Γ . Because this is a binomial equation for \mathbf{x} , $\mathcal{V}_{cb}^{\mathcal{X}}(\boldsymbol{\theta})$ is generated by binomial ideal, which leads to the toric property of $\mathcal{V}_{cb}^{\mathcal{X}}(\boldsymbol{\theta})$.

Next, we investigate the parametric condition such that the complex balanced states exist: $\mathcal{V}_{cb}^{\mathcal{X}}(\boldsymbol{\theta}) \neq \emptyset$. By taking the logarithm of $\mathbf{x}_{cb}^{\Gamma T} = \frac{1}{\alpha} \boldsymbol{\Sigma}$, we have

$$\Gamma^T \ln \mathbf{x}_{cb} = \ln \boldsymbol{\Sigma} - \ln \alpha \mathbf{1}. \quad (\text{B8})$$

By defining Cayley matrix Γ_C and the extended vector $\boldsymbol{\chi}_{cb}$ as

$$\Gamma_C := \begin{pmatrix} \Gamma \\ \mathbf{1}^T \end{pmatrix}, \quad \boldsymbol{\chi}_{cb} := \begin{pmatrix} \mathbf{x}_{cb} \\ \alpha \end{pmatrix}, \quad (\text{B9})$$

Eq. (B8) can be represented as

$$\Gamma_C^T \ln \boldsymbol{\chi}_{cb} = \ln \boldsymbol{\Sigma}. \quad (\text{B10})$$

Thus, $\boldsymbol{\chi}_{cb}$ exists if and only if $\ln \boldsymbol{\Sigma} \in \text{Im}[\Gamma_C^T]$, which is also equivalent to $\text{Ker}[\Gamma_C] \perp \ln \boldsymbol{\Sigma}$.

Suppose that the kinetic parameter satisfies $\ln \boldsymbol{\Sigma} \in \text{Im}[\Gamma_C^T]$. Then, we have a particular solution $\tilde{\mathbf{x}}_{cb}$ of Eq. (B8), and the general solution can be described as

$$\ln \mathbf{x}_{cb} - \ln \tilde{\mathbf{x}}_{cb} \in \mathcal{C}, \quad (\text{B11})$$

where $\mathcal{C} := \{\mathbf{y} | \Gamma^T \mathbf{y} \in \langle \mathbf{1} \rangle\}$. Because the incidence matrix B represents a connected graph from our assumption, it has the unique vector \mathbf{v} up to scale, such that $B^T \mathbf{v} = \mathbf{0}$ and $\mathbf{v} = \mathbf{1}$. Therefore, we have

$$\mathcal{C} = \{\mathbf{y} | B^T \Gamma^T \mathbf{y} = \mathbf{0}\} = \{\mathbf{y} | S^T \mathbf{y} = \mathbf{0}\} \quad (\text{B12})$$

$$= \text{Ker}[S^T] = \langle U^* \rangle. \quad (\text{B13})$$

Thus, we obtain the parametrization of \mathbf{x}_{cb} as in Eq. (92).

¹⁵It should be noted that $\boldsymbol{\Sigma}$ is a function of \mathbf{k}^\pm as $\boldsymbol{\Sigma}(\mathbf{k}^\pm)$. We abbreviate this dependency for the sake of notational simplicity.

¹⁴We here ignore the information of edge orientation in Γ .

[1] U. Alon, *An Introduction to Systems Biology: Design Principles of Biological Circuits* (CRC Press, Boca Raton, FL, 2019).

[2] A. S. Mikhailov and G. Ertl, *Chemical Complexity: Self-Organization Processes in Molecular Systems* (Springer, Berlin, 2017).

[3] M. Feinberg, *Foundations of Chemical Reaction Network Theory* (Springer, Berlin, 2019).

[4] J. D. Murray, *Mathematical Biology: I. An Introduction* (Springer Science & Business Media, Berlin, 2011).

- [5] I. R. Epstein and J. A. Pojman, *An Introduction to Nonlinear Chemical Dynamics: Oscillations, Waves, Patterns, and Chaos*, Topics in Physical Chemistry (Oxford University Press, New York, NY, 1998).
- [6] D. A. Beard and H. Qian, *Chemical Biophysics: Quantitative Analysis of Cellular Systems*, Cambridge Texts in Biomedical Engineering (Cambridge University Press, Cambridge, UK, 2008).
- [7] R. Wegscheider, Über simultane Gleichgewichte und die Beziehungen zwischen Thermodynamik und Reaktionskinetik homogener Systeme, *Z. Phys. Chem.* **39U**, 257 (1902).
- [8] T. L. Hill, *Free Energy Transduction and Biochemical Cycle Kinetics* (Courier Corporation, Chelmsford, MA, 2005).
- [9] T. L. Hill, Studies in irreversible thermodynamics IV. Diagrammatic representation of steady state fluxes for unimolecular systems, *J. Theor. Biol.* **10**, 442 (1966).
- [10] J. Schnakenberg, Network theory of microscopic and macroscopic behavior of master equation systems, *Rev. Mod. Phys.* **48**, 571 (1976).
- [11] M. Poletini and M. Esposito, Irreversible thermodynamics of open chemical networks. I. Emergent cycles and broken conservation laws, *J. Chem. Phys.* **141**, 024117 (2014).
- [12] R. Rao and M. Esposito, Nonequilibrium Thermodynamics of Chemical Reaction Networks: Wisdom from Stochastic Thermodynamics, *Phys. Rev. X* **6**, 041064 (2016).
- [13] R. Rao and M. Esposito, Conservation laws and work fluctuation relations in chemical reaction networks, *J. Chem. Phys.* **149**, 245101 (2018).
- [14] R. Rao and M. Esposito, Conservation laws shape dissipation, *New J. Phys.* **20**, 023007 (2018).
- [15] F. Avanzini, E. Penocchio, G. Falasco, and M. Esposito, Nonequilibrium thermodynamics of nonideal chemical reaction networks, *J. Chem. Phys.* **154**, 094114 (2021).
- [16] 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).
- [17] T. Schmiedl and U. Seifert, Stochastic thermodynamics of chemical reaction networks, *J. Chem. Phys.* **126**, 044101 (2007).
- [18] F. Horn and R. Jackson, General mass action kinetics, *Arch. Ration. Mech. Anal.* **47**, 81 (1972).
- [19] G. Craciun, A. Dickenstein, A. Shiu, and B. Sturmfels, Toric dynamical systems, *J. Symbol. Comput.* **44**, 1551 (2009).
- [20] G. Craciun, S. Muller, C. Pantea, P. Y. Yu, G. Craciun, S. Muller, C. Pantea, and P. Y. Yu, A generalization of Birchs theorem and vertex-balanced steady states for generalized mass-action systems, *Math. Biosci. Eng.* **16**, 8243 (2019).
- [21] G. Craciun and M.-S. Sorea, The structure of the moduli spaces of toric dynamical systems, [arXiv:2008.11468](https://arxiv.org/abs/2008.11468) (2020).
- [22] A. van der Schaft, S. Rao, and B. Jayawardhana, On the mathematical structure of balanced chemical reaction networks governed by mass action kinetics, *SIAM J. Appl. Math.* **73**, 953 (2013).
- [23] T. Okada and A. Mochizuki, Law of Localization in Chemical Reaction Networks, *Phys. Rev. Lett.* **117**, 048101 (2016).
- [24] A. Mochizuki and B. Fiedler, Sensitivity of chemical reaction networks: A structural approach. 1. Examples and the carbon metabolic network, *J. Theor. Biol.* **367**, 189 (2015).
- [25] B. Fiedler and A. Mochizuki, Sensitivity of chemical reaction networks: A structural approach. 2. Regular monomolecular systems, *Math. Methods Appl. Sci.* **38**, 3519 (2015).
- [26] G. Shinar and M. Feinberg, Structural sources of robustness in biochemical reaction networks, *Science* **327**, 1389 (2010).
- [27] R. P. Araujo and L. A. Liotta, The topological requirements for robust perfect adaptation in networks of any size, *Nat. Commun.* **9**, 1757 (2018).
- [28] Y. Hirono, T. Okada, H. Miyazaki, and Y. Hidaka, Structural reduction of chemical reaction networks based on topology, *Phys. Rev. Research* **3**, 043123 (2021).
- [29] M. Gopalkrishnan, E. Miller, and A. Shiu, A geometric approach to the global attractor conjecture, *SIAM J. Appl. Dyn. Syst.* **13**, 758 (2014).
- [30] G. Craciun, Toric differential inclusions and a proof of the global attractor conjecture, [arXiv:1501.02860](https://arxiv.org/abs/1501.02860) (2016).
- [31] A. Dickenstein, M. P. Millán, A. Shiu, and X. Tang, Multistationarity in structured reaction networks, *Bull. Math. Biol.* **81**, 1527 (2019).
- [32] B. Joshi and A. Shiu, Which small reaction networks are multistationary? *SIAM J. Appl. Dyn. Syst.* **16**, 802 (2017).
- [33] I. Otero-Muras, P. Yordanov, and J. Stelling, Chemical reaction network theory elucidates sources of multistability in interferon signaling, *PLoS Comput. Biol.* **13**, e1005454 (2017).
- [34] Y. Sughiyama, D. Loutchko, A. Kamimura, and T. J. Kobayashi, Hessian geometric structure of chemical thermodynamic systems with stoichiometric constraints, *Phys. Rev. Research* **4**, 033065 (2022).
- [35] K. Yoshimura and S. Ito, Information geometric inequalities of chemical thermodynamics, *Phys. Rev. Research* **3**, 013175 (2021).
- [36] H. Shima, *The Geometry of Hessian Structures* (World Scientific, Singapore, 2007).
- [37] S.-i. Amari, *Information Geometry and Its Applications* (Springer, Berlin, 2016).
- [38] S.-i. Amari, Differential geometry of curved exponential families—curvatures and information loss, *Ann. Stat.* **10**, 357 (1982).
- [39] S.-i. Amari, Differential geometry of statistical models, in *Differential-Geometrical Methods in Statistics*, Lecture Notes in Statistics, edited by S.-i. Amari (Springer, New York, 1985), pp. 11–65.
- [40] O. E. Barndorff-Nielsen, D. R. Cox, and N. Reid, The role of differential geometry in statistical theory, *Int. Stat. Rev./Revue Internationale de Statistique* **54**, 83 (1986).
- [41] S.-i. Amari and H. Nagaoka, *Methods of Information Geometry* (American Mathematical Society, Providence, RI, 2000).
- [42] S.-I. Amari and T. Han, Statistical inference under multiterminal rate restrictions: A differential geometric approach, *IEEE Trans. Inf. Theory* **35**, 217 (1989).
- [43] I. Okamoto, S.-I. Amari, and K. Takeuchi, Asymptotic theory of sequential estimation: Differential geometrical approach, *Ann. Stat.* **19**, 961 (1991).
- [44] R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, UK, 2013).
- [45] F. Sottile, Toric ideals, real toric varieties, and the algebraic moment map, [arXiv:math/0212044](https://arxiv.org/abs/math/0212044) (2008).
- [46] D. Eisenbud and B. Sturmfels, Binomial ideals, *Duke Math. J.* **84**, 1 (1996).

- [47] D. A. Cox, J. B. Little, and H. K. Schenck, *Toric Varieties* (American Mathematical Society, Providence, RI, 2011).
- [48] H. B. Callen and H. B. Callen, *Thermodynamics and an Introduction to Thermostatistics* (Wiley, New York, NY, 1985).
- [49] L. Pachter and B. Sturmfels, eds., *Algebraic Statistics for Computational Biology* (Cambridge University Press, Cambridge, UK, 2005).
- [50] L. M. Bregman, The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming, *USSR Comput. Math. Math. Phys.* **7**, 200 (1967).
- [51] D. Shear, An analog of the Boltzmann H-theorem (a Liapunov function) for systems of coupled chemical reactions, *J. Theor. Biol.* **16**, 212 (1967).
- [52] J. Higgins, Some remarks on Shear's Liapunov function for systems of chemical reactions, *J. Theor. Biol.* **21**, 293 (1968).
- [53] T. J. Kobayashi, D. Loutchko, A. Kamimura, and Y. Sughiyama, Geometry of nonequilibrium chemical reaction networks and generalized entropy production decompositions (2022), [arXiv:2206.00863](https://arxiv.org/abs/2206.00863).
- [54] U. Seifert, Stochastic thermodynamics, fluctuation theorems and molecular machines, *Rep. Prog. Phys.* **75**, 126001 (2012).
- [55] S. Ito, Stochastic Thermodynamic Interpretation of Information Geometry, *Phys. Rev. Lett.* **121**, 030605 (2018).
- [56] N. Ohga and S. Ito, Information-geometric Legendre duality in stochastic thermodynamics, [arXiv:2112.11008](https://arxiv.org/abs/2112.11008) (2021).
- [57] M. F. Adamer and M. Helmer, Families of toric chemical reaction networks, *J. Math. Chem.* **58**, 2061 (2020).
- [58] N. Ay, J. Jost, H. V. Lê, and L. Schwachhöfer, *Information Geometry* (Springer, Berlin, 2017).
- [59] N. J. Newton, An infinite-dimensional statistical manifold modelled on Hilbert space, *J. Funct. Anal.* **263**, 1661 (2012).
- [60] G. E. Crooks, Measuring Thermodynamic Length, *Phys. Rev. Lett.* **99**, 100602 (2007).
- [61] A. Kolchinsky and D. H. Wolpert, Work, Entropy Production, and Thermodynamics of Information Under Protocol Constraints, *Phys. Rev. X* **11**, 041024 (2021).
- [62] N. Ohga and S. Ito, Information-geometric dual affine coordinate systems for chemical thermodynamics, [arXiv:2112.13813](https://arxiv.org/abs/2112.13813) (2021).
- [63] R. Aris, *Introduction to the Analysis of Chemical Reactors* (Prentice-Hall, Hoboken, NJ, 1965).
- [64] H. Qian and H. Ge, *Stochastic Chemical Reaction Systems in Biology* (Springer International Publishing, Berlin, 2021).
- [65] F. T. Leighton and R. L. Rivest, The Markov chain tree theorem, MIT Lab. Comput. Sci. Tech. Rep. **MIT/LCS/TM-249** (1983).