

## Hessian geometry of nonequilibrium chemical reaction networks and entropy production decompositions

Tetsuya J. Kobayashi <sup>\*</sup>, 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 17 February 2022; accepted 18 August 2022; published 15 September 2022)

We derive the Hessian geometric structure of nonequilibrium chemical reaction networks on the flux and force spaces induced by the Legendre duality of convex dissipation functions and characterize their dynamics as a generalized flow. With this geometric structure, we can extend theories of nonequilibrium systems with quadratic dissipation functions to more general cases with nonquadratic ones, which are pivotal for studying chemical reaction networks. By applying generalized notions of orthogonality in Hessian geometry to chemical reaction networks, two generalized decompositions of the entropy production rate are obtained, each of which captures gradient-flow and minimum-dissipation aspects in nonequilibrium dynamics.

DOI: [10.1103/PhysRevResearch.4.033208](https://doi.org/10.1103/PhysRevResearch.4.033208)

### I. INTRODUCTION

Thermodynamics aims at establishing the general description of macroscopic systems. Although such a description was obtained for equilibrium situations [1,2], its extension to nonequilibrium ones has been limited to specific systems and models. For near-equilibrium situations, Onsager and Machlup evaluated the entropy production rate using the linear approximation known as the force-flux relation [3–6], which corresponds to a quadratic dissipation function. With the recent development of macroscopic fluctuation theory and stochastic thermodynamics [7–9], this result was extended to far-from equilibrium situations in fluid dynamics and diffusion processes in a continuous space, but the dissipation functions are still quadratic even though they are generalized to be state-dependent. Those systems are characterized geometrically via the inner product structure induced by the quadratic functions and associated formal Riemannian geometry. However, the knowledge obtained from such systems and models is not directly applicable to other systems with a discrete state space or with nonlinearities in the governing equations, where the natural dissipation functions may not be quadratic.<sup>1</sup> In such cases, the inner product structure is no longer an adequate mathematical basis. Although Wasserstein geometry has recently been introduced into the thermodynamics of diffusion processes as a new geometric approach

[12,13], it also employs the formal Riemannian geometric structure associated with the Wasserstein distance via the Benamou-Brenier formula and thereby is not directly applicable to systems characterized by nonquadratic dissipation functions [14–16].

In this paper, we clarify that this problem can be resolved by using the Hessian (information) geometric structure of the flux and force spaces. Hessian geometry or information geometry [17–19] enables us to relate the force and flux by a nonlinear Legendre transformation induced by the convex dissipation functions.<sup>2</sup> Because the Riemannian metric is regarded as the Legendre transformation with quadratic dissipation functions, Hessian geometry can work as a nonlinear extension of the inner product structure in Riemannian geometry. Even with being a nonlinear extension, Hessian geometry still preserves many important aspects of the inner product structure in the form of generalized orthogonalities and Pythagorean theorem, among others [18,19]. Therefore, we can naturally and consistently extend various previous results for systems with quadratic dissipation functions to those with nonquadratic ones.

We derive and demonstrate the structure by focusing mainly on chemical reaction networks (CRN) because they are representative thermodynamic systems with both discrete state space and nonlinearity in the governing rate equations [20–23]. CRN also include Markov jump processes (MJP) on a graph, which is an important class of systems in stochastic thermodynamics with a discrete state space but with a linear governing equation.<sup>3</sup> In addition, CRN are also important in light of their biological and engineering applications [24–26]. The relevant dissipation functions of these systems

<sup>\*</sup>tetsuya@mail.crmind.net; 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.*

<sup>1</sup>Mathematically, for a given governing equation, there can exist multiple dissipation functions, all of which can induce the same governing equation. Thus, we need to find or choose dissipation functions that are physically natural and relevant to the target system and related problems [10,11].

<sup>2</sup>When the dissipation functions are quadratic, the force and the corresponding flux are related by a linear transformation induced by the associated Riemannian metric.

<sup>3</sup>It should be noted that a linear governing equation does not necessarily mean that its dissipation functions are quadratic.

are not necessarily quadratic, but can be more general convex functions.<sup>4</sup>

The nonquadratic property makes it difficult to dissect the equilibrium-like aspects from nonequilibrium dynamics by the notion of orthogonality from the usual inner product structure, as was done for overdamped diffusion processes [28]. As a result, some entropy decomposition formulas have not yet been generalized to MJP and CRN. By employing the generalized notions of orthogonality in Hessian geometry [17–19], we demonstrate that different aspects of nonequilibrium CRN can be dissected as generalized decompositions of the entropy production rate (EPR). In particular, we derive a CRN version of the Maes-Netočný relation by generalizing the Helmholtz decomposition. If the dissipation functions are quadratic, these results are automatically and formally reduced to the original results obtained for diffusion processes. Thus, Hessian geometry provides a natural generalization for resolving the problem. Finally, we also discuss how Hessian geometry can potentially fill the gap between quadratic and nonquadratic cases by geometrically capturing the nonlinear dual relation between force and flux, unify the description of those systems, and thereby extend the applicability of nonequilibrium thermodynamics. As such an example, we mention a relation of our results with the thermodynamic uncertainty relation, and describe how our formulation can contribute to network thermodynamics and variational characterizations of nonequilibrium systems.

This paper is organized as follows. In Sec. II, we define MJP and CRN. In Sec. III, we introduce the Legendre duality between flux and force, and the associated notion of generalized flow. In Sec. IV, we clarify the Hessian geometric structure in the flux-force space and generalized orthogonalities, which are one of the main contributions of this work. In Sec. V, we demonstrate how the generalized notions of orthogonality lead to different types of EPR decompositions and their geometric meaning. In Sec. VI, we verify the obtained decompositions using CRN that can have equilibrium, complex-balanced, and noncomplex balanced steady states. In Sec. VII, we summarize our results and provide possible applications and contributions of the Hessian structure in the flux-force space to other thermodynamic problems.

## II. MODELS

In this section, we define MJP and CRN and show how MJP can be regarded as a special case of CRN.

### A. Markov jump processes

A reversible Markov jump diffusion process describes random jumps of noninteracting particles on a graph  $\mathbb{G}$  consisting of  $N_v$  vertices,  $\{\mathbf{v}_i\}_{i \in [1, N_v]}$ , and  $N_e$  oriented edges  $\{\mathbf{e}_e\}_{e \in [1, N_e]}$ .  $k_e^+ \geq 0$  is the forward jump rate from the head of the oriented

edge  $\mathbf{e}_e$  to its tail.  $k_e^- \geq 0$  is the reverse jump rate from the tail to the head of  $\mathbf{e}_e$ . For infinitely many such particles, we consider  $p_i(t) \in [0, 1]$ , the fraction of particles on vertex  $\mathbf{v}_i$  at time  $t$ . Then, the forward and reverse one-way fluxes on the  $e$ th edge are

$$j_e^+(\mathbf{p}) = k_e^+ p_{\mathbf{v}^+(\mathbf{e}_e)}, \quad j_e^-(\mathbf{p}) = k_e^- p_{\mathbf{v}^-(\mathbf{e}_e)}, \quad (1)$$

where  $\mathbf{v}^+(\mathbf{e}_e)$  and  $\mathbf{v}^-(\mathbf{e}_e)$  are the head and tail vertices of edge  $\mathbf{e}_e$ .<sup>5</sup> The total flux function is the difference of one-way flux functions as  $\mathbf{j}(\mathbf{p}) = \mathbf{j}^+(\mathbf{p}) - \mathbf{j}^-(\mathbf{p}) \in \mathbb{R}^{N_e}$  where  $\mathbf{j}^\pm(\mathbf{p}) = (j_1^\pm(\mathbf{p}), \dots, j_{N_e}^\pm(\mathbf{p}))^T$ . Then, the dynamics of the density vector  $\mathbf{p}(t) := (p_1(t), \dots, p_{N_v}(t))^T \in \mathbb{R}_{\geq 0}^{N_v}$  is represented by the master equation

$$\dot{\mathbf{p}} = -B\mathbf{j}(\mathbf{p}) = -\text{div}_B \mathbf{j}(\mathbf{p}), \quad (2)$$

where  $B \in \{0, \pm 1\}^{N_v \times N_e}$  is the incidence matrix of graph  $\mathbb{G}$  and  $\text{div}_B := B$ . More specifically, for  $B = (b_{i,e})$ ,

$$b_{i,e} := +1 \text{ if } \mathbf{v}_i \text{ is the head of edge } \mathbf{e}_e,$$

$$b_{i,e} := -1 \text{ if } \mathbf{v}_i \text{ is the tail of edge } \mathbf{e}_e,$$

$$b_{i,e} := 0 \text{ otherwise.}$$

Equation (2) is the continuity equation for diffusion on a graph<sup>6</sup> and  $B$  can be regarded as a discrete divergence operator on a graph [29].<sup>7</sup> We also define the head and tail incidence matrices, respectively, as  $B^+ := \max[B, 0]$  and  $B^- := \max[-B, 0]$ . Thus,  $B = B^+ - B^-$ .

Then, the flux functions are compactly described in a vector form as

$$\mathbf{j}^\pm(\mathbf{p}) = \mathbf{k}^\pm \circ (B^\pm)^T \mathbf{p}, \quad (3)$$

where  $\circ$  is the component-wise product of two vectors. In this paper, we assume that all edges describe reversible transitions, i.e.,  $k_e^\pm > 0$  for all  $e$ . While the representation of the master equation by Eq. (2) is different from that used conventionally in stochastic thermodynamics, they are equivalent and this representation suits our purpose of unveiling the relationship between MJP and CRN as well as the underlying geometric structure.

### B. Chemical reaction networks

Deterministic chemical reaction networks (CRN) are an important class of macroscopic models in light of their historical role played in thermodynamics since Gibbs [30] and of their wide range of applications in engineering and biology [24–26,31]. A reversible CRN is composed of a collection of

<sup>5</sup>Here, we have abused the notation  $\mathbf{v}^+(\mathbf{e}_e)$  to indicate the index of the vertex  $\mathbf{v}^+(\mathbf{e}_e)$ .

<sup>6</sup>Mathematically, the same equation describes the time-evolution of the occupation probability of states between which a system transit stochastically.

<sup>7</sup>The usual divergence in vector calculus includes the information of the metric associated with the manifold on which the divergence is defined [29]. In contrast,  $\text{div}_B$  is defined only by the topological information of the underlying graph and does not include metric information. More precisely,  $\text{div}_B$  is the adjoint of exterior derivative on the graph [29].

<sup>4</sup>At least, two types of dissipation functions have been proposed for deterministic CRNs [10,27]. One is quadratic, and the other is nonquadratic. We focus on the latter, which is more natural and relevant in the sense that it was derived from the large deviation rate function of the underlying stochastic CRN models.

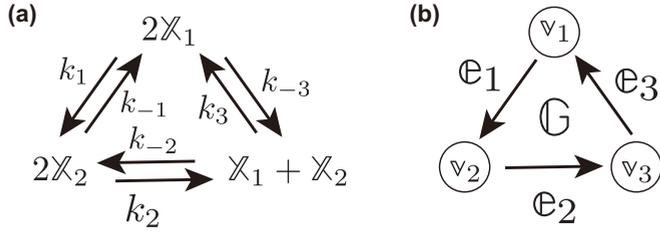
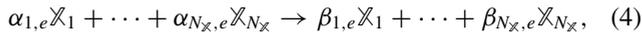


FIG. 1. Diagrammatic illustration of CRN (a) Diagrammatic representation of chemical reaction equations with reaction rate constants  $\{k_{\pm i}\}_{i=1,2,3}$ . (b) The graph-theoretic structure of the CRN in (a).  $v_1$ ,  $v_2$ , and  $v_3$  correspond to complex  $2X_1$ ,  $2X_2$ , and  $X_1 + X_2$ , respectively. Each directed edge represents a pair of forward and reverse reactions, and the direction of the edge indicates the direction of the forward reaction.

forward and reverse reaction pairs [Fig. 1(a)], the  $e$ th forward reaction of which is described by the following chemical reaction equation [24]:



where  $X_i$  is the  $i$ th molecular species,  $N_x$  is the number of different kinds of molecular species, and  $\alpha_{i,e} \in \mathbb{N}_{\geq 0}$  and  $\beta_{i,e} \in \mathbb{N}_{\geq 0}$  are the numbers of molecule  $X_i$  involved as the reactants and products of the  $e$ th reaction, respectively. The stoichiometric vector of the  $e$ th forward reaction is defined as  $s_e := (\beta_{1,e} - \alpha_{1,e}, \dots, \beta_{N_x,e} - \alpha_{N_x,e})^T$ . The stoichiometric matrix is  $S := (s_1, \dots, s_{N_e})$ . The stoichiometric vector of the  $e$ th reverse reaction is obtained by just changing the sign of the forward one:  $-s_e$ . Thus, the stoichiometric matrix  $S$  defined only for the forward reactions is sufficient to characterize a reversible CRN. Let  $j_e^+(\mathbf{x})$  and  $j_e^-(\mathbf{x})$  be the one-way fluxes of the  $e$ th forward and reverse reactions in which  $\mathbf{x} = (x_1, \dots, x_{N_x})^T \in \mathcal{X} := \mathbb{R}_{>0}^{N_x}$  represents the concentration of molecules. The total flux is  $\mathbf{j}(\mathbf{x}) = \mathbf{j}^+(\mathbf{x}) - \mathbf{j}^-(\mathbf{x})$  where  $\mathbf{j}^{\pm}(\mathbf{x}) = (j_1^{\pm}(\mathbf{x}), \dots, j_{N_e}^{\pm}(\mathbf{x}))^T$ . Then, we have the deterministic chemical rate equation (CRE) [23,24,31] as

$$\dot{\mathbf{x}} = S\mathbf{j}(\mathbf{x}) = -\text{div}_S \mathbf{j}(\mathbf{x}). \quad (5)$$

Equation (5) is the continuity equation for the CRN with the divergence operator  $\text{div}_S = -S$ .

The intrinsic graph structure of CRN can be manifested by considering the sets of reactants and products as vertices of the graph  $\mathbb{G}$  connected by the reaction edges [Fig. 1(b)]. Specifically, the set of reactants  $v_e^+ := (\alpha_{1,e}X_1 + \cdots + \alpha_{N_x,e}X_{N_x})$  and the set of products  $v_e^- := (\beta_{1,e}X_1 + \cdots + \beta_{N_x,e}X_{N_x})$  are regarded as the head and tail vertices of the  $e$ th reaction edge, respectively. Such sets (vertices) are called complexes in the CRN theory [31]. Because each complex is a set of molecular species, a CRN is a kind of hypergraph in which multiple molecular species are connected by oriented edges (reactions) [32]. Reflecting this hypergraph nature,  $S$  can be decomposed as

$$S = -\Gamma B = -\Gamma(B^+ - B^-), \quad (6)$$

where  $N_v$  is the number of different vertices (complexes),  $B$  is the incidence matrix of the complex graph  $\mathbb{G}$ , and  $\Gamma = (\gamma_1, \dots, \gamma_{N_v}) \in \mathbb{Z}^{N_x \times N_v}$  where  $\gamma_i$  specifies the molec-

ular species involved in the  $i$ th vertex (complex) as  $v_i := (\gamma_{1,i}X_1 + \cdots + \gamma_{N_x,i}X_{N_x})$ .

If we adopt the law of mass action kinetics, the  $e$ th forward and reverse reaction fluxes can be represented as

$$j_e^{\pm}(\mathbf{x}) = k_e^{\pm} \sum_{i=1}^{N_v} b_{i,e}^{\pm} \prod_{j=1}^{N_x} x_j^{\gamma_{j,i}}, \quad (7)$$

where  $b_{i,e}^{\pm}$  is the  $(i, e)$  component of  $B^{\pm}$  as  $B^{\pm} = (b_{i,e}^{\pm})$  and  $k_e^+ \in \mathbb{R}_{>0}$  and  $k_e^- \in \mathbb{R}_{\geq 0}$  are the reaction rate constants of the  $e$ th forward and reverse reactions, respectively. In vector form, we can compactly represent it as

$$\mathbf{j}^{\pm}(\mathbf{x}) = \mathbf{k}^{\pm} \circ (B^{\pm})^T \mathbf{x}^{\Gamma^T}, \quad (8)$$

where  $\mathbf{x}^{\gamma} := \prod_{j=1}^{N_x} x_j^{\gamma_j} \in \mathbb{R}_{>0}$  and  $\mathbf{x}^{\Gamma^T} := (x^{\gamma_1}, \dots, x^{\gamma_{N_v}})^T$ . We should note an important relation,  $(B^{\pm})^T \mathbf{x}^{\Gamma^T} = \mathbf{x}^{(\Gamma B^{\pm})^T}$ , which holds because every column vector of  $B^{\pm}$  contains only one +1 and the others are 0 (see also Appendix A for the notation).

By comparing Eq. (2) and Eq. (3) with Eq. (5) and Eq. (8), we can see that CRN contain MJP as a special case. Specifically, if we identify  $\mathbf{x}$  with  $\mathbf{p}$  and set  $\Gamma = I$ , where  $I$  is the identity matrix, Eq. (5) and Eq. (8) are reduced to Eq. (2) and Eq. (3), respectively. In other words, a MJP is a CRN, each complex of which contains only one molecular species. Such a CRN is called a monomolecular reaction network and its thermodynamic nature has been investigated in the context of equilibrium and nonequilibrium chemical thermodynamics, especially by Hill [20]. Because CRN include MJP, we work only on CRN in the following sections.<sup>8</sup>

### III. LEGENDRE DUALITY OF FLUX AND FORCE

Next, we introduce the Legendre duality between flux and force for CRN,<sup>9</sup> and summarize their relation to entropy production. The specific type of convex function introduced here was recently derived in the large deviation theory [10,11,27,33–38].

For a given pair of one-way fluxes  $\mathbf{j}^{\pm}$ , the total flux and force are defined as

$$\mathbf{j} := \mathbf{j}^+ - \mathbf{j}^- \in \mathcal{J}, \quad \mathbf{f} := \frac{1}{2} \ln \frac{\mathbf{j}^+}{\mathbf{j}^-} \in \mathcal{F}, \quad (9)$$

where  $\mathcal{J} = \mathbb{V}$  is an  $N_e$ -dimensional vector space and  $\mathcal{F} = \mathbb{V}^*$  is its linear-algebraic dual space.<sup>10</sup> The force of this form comes from the local detailed balance condition [39], and is

<sup>8</sup>The state of MJP is the probability vector  $\mathbf{p}(t)$ . Thus, the conservation of probability  $\sum_i p_i(t) = 1$  should be satisfied. To regard MJP as a CRN, we here do not impose such a conservation in advance because the conservation of probability over time, i.e.,  $\sum_i p_i(t) = \text{const.}$  is automatically satisfied by the property of the incidence matrix  $B$  such that  $\mathbf{1}^T B = \mathbf{0}$ . Thus, by restricting the dynamics of the system with the initial condition satisfying  $\sum_i p_i(0) = 1$ , the constraint  $\sum_i p_i(t) = 1$  is automatically obtained.

<sup>9</sup>We abbreviate the thermodynamic force as force hereafter.

<sup>10</sup>The definition of the force here includes 1/2, which does not appear in the conventional definition of thermodynamic force. This is because we adopt the derivation of this form of force from the

also consistent with macroscopic chemical thermodynamics with the mass action kinetics [2,21,23]. If  $\mathbf{j}$  and  $\mathbf{f}$  are defined as in Eq. (9), the entropy production rate (EPR) is obtained as

$$\dot{\Sigma} := 2\langle \mathbf{j}, \mathbf{f} \rangle = (\mathbf{j}^+ - \mathbf{j}^-)^T \ln \frac{\mathbf{j}^+}{\mathbf{j}^-} \geq 0, \quad (10)$$

where  $\langle \mathbf{j}, \mathbf{f} \rangle := \sum_{e=1}^{N_e} j_e f_e$  is the usual bilinear pairing on  $\mathbb{V} \times \mathbb{V}^*$ .

As Eq. (9) implies, the pair of flux and force has a nonlinear relationship. To show that the relation is a Legendre duality, we introduce the frenetic activity [40],<sup>11</sup>

$$\boldsymbol{\omega} := 2\sqrt{\mathbf{j}^+ \circ \mathbf{j}^-} \in \mathbb{R}_{>0}^{N_e}. \quad (11)$$

Then, we have  $\mathbf{j} = \frac{\boldsymbol{\omega}}{2} \circ [e^{\mathbf{f}} - e^{-\mathbf{f}}]$ . Thus, if  $\boldsymbol{\omega}$  is given, the force  $\mathbf{f}$  can be converted to the corresponding flux  $\mathbf{j}$  that satisfies Eq. (9). This relation between the pair  $(\mathbf{j}, \mathbf{f})$  is a one-to-one Legendre duality induced by the following strictly convex smooth functions:

$$\begin{aligned} \Psi_{\boldsymbol{\omega}}^*(\mathbf{f}) &:= \boldsymbol{\omega}^T [\cosh(\mathbf{f}) - \mathbf{1}], \\ \Psi_{\boldsymbol{\omega}}(\mathbf{j}) &:= \mathbf{j}^T \sinh^{-1} \left( \frac{\mathbf{j}}{\boldsymbol{\omega}} \right) - \boldsymbol{\omega}^T \left[ \sqrt{\mathbf{1} + \left( \frac{\mathbf{j}}{\boldsymbol{\omega}} \right)^2} - \mathbf{1} \right], \end{aligned} \quad (12)$$

which lead to the Legendre transformations

$$\mathbf{j} = \partial_{\mathbf{f}} \Psi_{\boldsymbol{\omega}}^*(\mathbf{f}) = \boldsymbol{\omega} \circ \sinh(\mathbf{f}), \quad (13)$$

$$\mathbf{f} = \partial_{\mathbf{j}} \Psi_{\boldsymbol{\omega}}(\mathbf{j}) = \sinh^{-1} \left( \frac{\mathbf{j}}{\boldsymbol{\omega}} \right). \quad (14)$$

Here,  $\cosh(\mathbf{f})$  and  $\sinh(\mathbf{j})$  are the hyperbolic cosine and sin functions applied in a componentwise manner as  $[\cosh(\mathbf{f})]_e = \cosh(f_e) = (e^{f_e} + e^{-f_e})/2$  and  $[\sinh(\mathbf{j})]_e = \sinh(j_e) = (e^{j_e} - e^{-j_e})/2$ . We can easily verify that Eq. (13) and Eq. (14) are equivalent to Eq. (9) by direct computation (see also Appendix B). In the following, we abbreviate  $\partial_{\mathbf{f}} \Psi_{\boldsymbol{\omega}}^*$  and  $\partial_{\mathbf{j}} \Psi_{\boldsymbol{\omega}}$  with  $\partial \Psi_{\boldsymbol{\omega}}^*$  and  $\partial \Psi_{\boldsymbol{\omega}}$  because we do not use differentiation of  $\Psi_{\boldsymbol{\omega}}^*$  and  $\Psi_{\boldsymbol{\omega}}$  with respect to other variables in this work. Moreover, the pair  $(\mathbf{j}, \mathbf{f})$  satisfies the Legendre identity

$$\Psi_{\boldsymbol{\omega}}^*(\mathbf{f}) + \Psi_{\boldsymbol{\omega}}(\mathbf{j}) - \langle \mathbf{j}, \mathbf{f} \rangle = 0. \quad (15)$$

The convex functions  $\Psi_{\boldsymbol{\omega}}^*(\mathbf{f})$  and  $\Psi_{\boldsymbol{\omega}}(\mathbf{j})$  of the form in Eq. (12) were recently derived via large deviation functions of the corresponding microscopic models [10,11,33–38], where they are called dissipation functions. Both dissipation functions are nonnegative and symmetric,

$$\Psi_{\boldsymbol{\omega}}(\mathbf{j}) = \Psi_{\boldsymbol{\omega}}(-\mathbf{j}) \geq 0, \quad \Psi_{\boldsymbol{\omega}}^*(\mathbf{f}) = \Psi_{\boldsymbol{\omega}}^*(-\mathbf{f}) \geq 0, \quad (16)$$

and satisfy  $\min_{\mathbf{f}} \Psi_{\boldsymbol{\omega}}^*(\mathbf{f}) = \Psi_{\boldsymbol{\omega}}^*(\mathbf{0}) = 0$  and  $\min_{\mathbf{j}} \Psi_{\boldsymbol{\omega}}(\mathbf{j}) = \Psi_{\boldsymbol{\omega}}(\mathbf{0}) = 0$ . From these properties, we can verify that  $\mathbf{f} = \mathbf{0}$

large deviation theory [38]. We can remove it by including 1/2 in the definition of the dissipation functions.

<sup>11</sup>The definition of the activity here includes 2, which does not appear in the definition of the activity in [40]. This is because we adopt the derivation of this form of activity from the large deviation theory [38]. We can remove it by including 2 in the definition of the dissipation functions.

and  $\mathbf{j} = \mathbf{0}$  are Legendre dual:  $\mathbf{0} = \partial \Psi_{\boldsymbol{\omega}}^*(\mathbf{0})$  and  $\mathbf{0} = \partial \Psi_{\boldsymbol{\omega}}(\mathbf{0})$ . This is consistent with the thermodynamic requirement that if the force is zero, the corresponding flux becomes zero, and vice versa.

In addition, from the Legendre identity [Eq. (15)], the nonnegativity of the EPR can also be attributed to the nonnegativity of the dissipation functions,

$$\dot{\Sigma}/2 = \langle \mathbf{j}, \mathbf{f} \rangle = \Psi_{\boldsymbol{\omega}}^*(\mathbf{f}) + \Psi_{\boldsymbol{\omega}}(\mathbf{j}) \geq 0. \quad (17)$$

It should be noted that the dissipation functions naturally provide a decomposition of the EPR into nonnegative terms.

In the following, a pair of flux and force with the same decoration, e.g.,  $(\mathbf{j}, \mathbf{f})$ ,  $(\mathbf{j}', \mathbf{f}')$ , or  $(\mathbf{j}(\mathbf{x}), \mathbf{f}(\mathbf{x}))$ , represents a Legendre dual pair linked by Eq. (13) and Eq. (14). Because of the one-to-one Legendre duality, the CRE [Eq. (5)] can be represented as

$$\dot{\mathbf{x}} = S\mathbf{j}(\mathbf{x}) = S\partial \Psi_{\boldsymbol{\omega}(\mathbf{x})}^*[\mathbf{f}(\mathbf{x})] = -\text{div}_S \partial \Psi_{\boldsymbol{\omega}(\mathbf{x})}[\mathbf{f}(\mathbf{x})]. \quad (18)$$

The equation of this form is a generalized flow driven by the force  $\mathbf{f}(\mathbf{x})$  [16,35,38]. The representation of the dynamics in this form is not specific to CRN or MJP. Thus, it can cover other systems such as overdamped diffusion by appropriately defining  $\text{div}$ ,  $\Psi_{\boldsymbol{\omega}}(\mathbf{x})$ ,  $\Psi_{\boldsymbol{\omega}}^*(\mathbf{x})$ , and  $\mathbf{f}(\mathbf{x})$ .

It should be noted that the derivations of Legendre duality and associated quantities and relations are not dependent on the specific functional form of the one-way fluxes, i.e., Eq. (8) or Eq. (3), the former of which is from the kinetic law of mass action assumption. Thus, the result here might be applied to a wider class of kinetic laws.

#### IV. HESSIAN GEOMETRY AND GENERALIZED ORTHOGONALITIES

Because of the nonlinearity of Legendre duality with the nonquadratic dissipation functions, we can no longer employ the inner product structure between flux and force and the associated formal Riemannian geometric notions (Fig. 2). We clarify that this problem is resolved by employing Hessian geometry and its geometric notions.<sup>12</sup> The elucidation of the fundamental roles played by Hessian geometry in nonequilibrium dynamics of MJP and CRN is one of the main contributions of this paper.

##### A. Hessian geometry

Hessian geometry is the geometry induced by convex potential functions  $\Psi(\mathbf{j})$  and  $\Psi^*(\mathbf{f})$  to a pair of dual spaces  $\mathcal{J}$  and  $\mathcal{F}$  [17,18].<sup>13</sup> In this section,  $\Psi(\mathbf{j})$  and  $\Psi^*(\mathbf{f})$  are not restricted to the specific form of Eq. (12) to obtain general

<sup>12</sup>It should be noted that, in differential geometry, Hessian geometry is a class of Riemannian geometry with a metric given by a Hessian matrix, which additionally induces the Legendre dual structure [17].

<sup>13</sup>Hessian geometry is generally defined on affine manifolds. However, because we relate the Hessian geometric structure with inner product one, we here restrict  $\mathcal{J}$  and  $\mathcal{F}$  to be the dual vector spaces for simplicity of presentation.

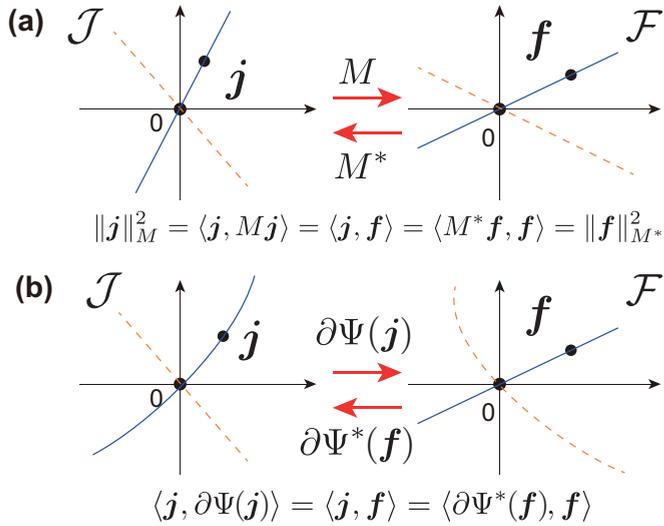


FIG. 2. (a) Linear-algebraic dual spaces ( $\mathcal{J}, \mathcal{F}$ ) with the inner product structure defined by a metric via  $M$  and  $M^* = M^{-1}$ . Linear subspaces in  $\mathcal{J}$ , e.g., blue-solid and red-dashed lines, are mapped to linear subspaces in  $\mathcal{F}$ . (b) Linear-algebraic dual spaces ( $\mathcal{J}, \mathcal{F}$ ) with the Hessian geometric structure defined by Legendre transformations  $\partial\Psi$  and  $\partial\Psi^*$ . A linear subspace in  $\mathcal{J}$ , e.g., red-dashed line, is mapped to a curved subspace in  $\mathcal{F}$  whereas a linear subspace in  $\mathcal{F}$ , e.g., blue-solid line, is mapped to a curved subspace in  $\mathcal{J}$ .

results. We consider the case for each fixed  $\omega$ .<sup>14</sup> Thus, the dependence of  $\Psi(j)$  and  $\Psi^*(f)$  on  $\omega$  is omitted for the sake of notational simplicity. They are assumed more generally to be just smooth and strictly convex functions satisfying the Legendre duality

$$\Psi^*(f) = \max_j [\langle j, f \rangle - \Psi(j)], \quad (19)$$

$$\Psi(j) = \max_f [\langle j, f \rangle - \Psi^*(f)]. \quad (20)$$

Instead of the association of  $j$  and  $f$  by a linear transformation  $j = M^* f$  on the inner product space [Fig. 2(a)],  $j$  and  $f$  are associated by the Legendre transformation [Fig. 2(b)]

$$j = \partial_f \Psi^*(f), \quad f = \partial_j \Psi(j). \quad (21)$$

Hessian geometry is fundamental for capturing nonlinear geometry induced by the convex functions. As important applications, it has played the essential roles in describing the geometry of statistical models in information geometry [18,19] and that of equilibrium thermodynamics [2,41,42].

To regard the Hessian geometric structure as a nonlinear generalization of the inner product structure,<sup>15</sup> we additionally assume that  $\Psi(j)$  satisfies the symmetry condition  $\Psi(j) = \Psi(-j)$ . From this symmetry condition for  $\Psi(j)$ , we also obtain the symmetry of  $\Psi^*(f) = \Psi^*(-f)$  via the

<sup>14</sup>This also means fixed  $x$  if  $\omega$  is a function of  $x$ . This is similar to the inner product structure of the tangent and cotangent spaces at each point  $x$  on a Riemannian manifold.

<sup>15</sup>It should be noted that the Hessian geometric structure is not necessarily restricted to this interpretation as an extension of the inner product structure.

Legendre duality. From the symmetries, the minimums of  $\Psi(j)$  and  $\Psi^*(f)$  are attained at  $\mathbf{0}$ :  $\arg \min_j \Psi(j) = \mathbf{0}$  and  $\arg \min_f \Psi^*(f) = \mathbf{0}$ . Because the Legendre transformation Eq. (21) is independent of a constant in  $\Psi(j)$  and  $\Psi^*(f)$ , without losing generality, we can assume that  $\min_j \Psi(j) = \Psi(\mathbf{0}) = 0$  and  $\min_f \Psi^*(f) = \Psi^*(\mathbf{0}) = 0$ . Thus

$$\Psi(j) = \Psi(-j) \geq 0, \quad \Psi^*(f) = \Psi^*(-f) \geq 0. \quad (22)$$

In the following, we call  $\Psi(j)$  and  $\Psi^*(f)$  with these properties of dissipation functions.

To demonstrate that the Hessian geometric structure includes the inner product structure, suppose that  $\hat{\Psi}(j)$  is a quadratic function as

$$\hat{\Psi}(j) = \frac{1}{2} \langle j, Mj \rangle := \frac{1}{2} \|j\|_M^2, \quad (23)$$

where  $M \in \mathbb{R}^{N_e \times N_e}$  is a positive definite matrix. Then, its Legendre dual becomes

$$\hat{\Psi}^*(f) = \frac{1}{2} \langle M^{-1} f, f \rangle = \frac{1}{2} \langle M^* f, f \rangle := \frac{1}{2} \|f\|_{M^*}^2, \quad (24)$$

where  $M^* := M^{-1}$ . Thus,  $\hat{\Psi}(j)$  and  $\hat{\Psi}^*(f)$  are reduced to the squared norms associated with the metric  $M$  and  $M^*$ . The Legendre transformations become

$$j = Mf = (M^*)^{-1} f, \quad f = M^* j = M^{-1} j. \quad (25)$$

These are the linear pairing of  $j$  and  $f$  via the metric  $M$ . Furthermore, for the paired  $j$  and  $f$  by  $M$ ,  $\hat{\Psi}(j)$  and  $\hat{\Psi}^*(f)$  are essentially identical to the EPR as

$$\frac{1}{2} \langle j, f \rangle = \hat{\Psi}(j) = \hat{\Psi}^*(f) = \dot{\Sigma}/4. \quad (26)$$

However, this identity for  $\hat{\Psi}(j)$  and  $\hat{\Psi}^*(f)$  no longer holds for a nonquadratic  $\Psi(j)$  and  $\Psi^*(f)$  pair. As a result, various notions of geometry in the inner product space are generalized in Hessian geometry.

## B. Generalized distance and Bregman divergence

Owing to the nonlinearity, the notion of distance is generalized into at least two versions.

The first version is to define a generalized distance via the dissipation functions as

$$\mathcal{D}_H(j, j') := \Psi(j - j'), \quad (27)$$

$$\mathcal{D}_H^*(f, f') := \Psi^*(f - f'). \quad (28)$$

From Eq. (22),  $\mathcal{D}_H(j, j')$  is nonnegative, symmetric  $\mathcal{D}_H(j, j') = \mathcal{D}_H(j', j)$  and satisfies  $\mathcal{D}_H(j, j') = 0$  if and only if  $j = j'$ . The same is true for  $\mathcal{D}_H^*(f, f')$ . In addition, they are reduced to the usual distance induced by the squared norm for quadratic cases, as shown in Eq. (23) and Eq. (24).

The second version is the Bregman divergence, which is defined as

$$\mathcal{D}[j||j'] := \Psi(j) - \Psi(j') - \langle j - j', \partial_j \Psi(j') \rangle$$

$$\mathcal{D}^*[f'||f] := \Psi^*(f) - \Psi^*(f') - \langle \partial_f \Psi^*(f'), f - f' \rangle.$$

The Bregman divergence  $\mathcal{D}[j||j']$  is nonnegative and strictly convex for  $j$ <sup>16</sup> and also attains the minimum 0 if and only

<sup>16</sup>Not necessarily for  $j'$ .

if  $j = j'$ .<sup>17</sup> However, it is generally asymmetric  $\mathcal{D}[j; j'] \neq \mathcal{D}[j'; j]$ . If  $\hat{\Psi}(j)$  is quadratic [Eq. (23)],  $\mathcal{D}[j||j']$  is reduced to  $\mathcal{D}[j||j'] = \frac{1}{2}\|j - j'\|_M^2$  and  $\mathcal{D}^*[f||f']$  is reduced to  $\mathcal{D}^*[f||f'] = \frac{1}{2}\|f - f'\|_{M^*}^2$ . It should be noted that for any flux-force pairs  $(j, f)$  and  $(j', f')$ , the divergences  $\mathcal{D}[j||j']$  and  $\mathcal{D}^*[f' || f]$  are the same,

$$\mathcal{D}[j||j'] = \mathcal{D}^*[f' || f]. \quad (29)$$

Thus,  $\mathcal{D}[j||j']$  and  $\mathcal{D}^*[f' || f]$  are different representations of the same quantity in the flux space  $\mathcal{J}$  and the force space  $\mathcal{F}$ , respectively. As a result,  $\mathcal{D}[j||j']$  and  $\mathcal{D}^*[f' || f]$  are also the same as the following mixed representation:

$$\mathcal{D}[j; f'] := \Psi(j) + \Psi^*(f') - \langle j, f' \rangle, \quad (30)$$

i.e.,  $\mathcal{D}[j||j'] = \mathcal{D}[j; f'] = \mathcal{D}^*[f' || f]$  (see also Appendix C).

### C. Generalized Hilbert orthogonality

The two generalized distances naturally lead to two generalized notions of orthogonality.

For any two force vectors  $f_S, f_A \in \mathcal{F}$ , the first orthogonality is defined by

$$\Psi^*(f_S + f_A) = \Psi^*(f_S - f_A). \quad (31)$$

This definition comes from the fact that, if  $f_S$  and  $f_A$  are orthogonal in an inner product space, then  $\|f_S + f_A\|_{M^*} = \|f_S - f_A\|_{M^*}$  holds. This orthogonality is called a generalized Hilbert orthogonality or simply Hilbert orthogonality in Ref. [43]. Then an orthogonal decomposition of a given force  $f$  into  $f = f_S + f_A$  is obtained by finding an isodissipation force  $f_{\text{iso}}$  satisfying  $\Psi^*(f) = \Psi^*(f_{\text{iso}})$  and by computing  $f_S$  and  $f_A$  as

$$f_S = \frac{f + f_{\text{iso}}}{2}, \quad f_A = \frac{f - f_{\text{iso}}}{2}. \quad (32)$$

Note that  $f_{\text{iso}} = f_S - f_A$  and also that we can have different decompositions of  $f$  by choosing different isodissipation forces  $f'_{\text{iso}}$ . In addition, from the symmetry of  $\Psi^*$ , we have  $\Psi^*(f) = \Psi^*(f_{\text{iso}}) = \Psi^*(-f_{\text{iso}}) = \Psi^*(-f)$ . As a consequence of the orthogonality, we obtain

$$2\langle j, f_A \rangle = \mathcal{D}[j||j_{\text{iso}}] \geq 0, \quad (33)$$

where  $j_{\text{iso}}$  is the Legendre transform of  $f_{\text{iso}}$  (see Appendix D 1 for a proof). Similarly, by considering the symmetry of  $\Psi^*(f)$ , we also have

$$2\langle j, f_S \rangle = \mathcal{D}[j||-j_{\text{iso}}] \geq 0. \quad (34)$$

From these relations, we have a decomposition of  $\langle j, f \rangle$  into two nonnegative terms,

$$\begin{aligned} \langle j, f \rangle &= \langle j, f_S \rangle + \langle j, f_A \rangle \\ &= \frac{\mathcal{D}[j||-j_{\text{iso}}]}{2} + \frac{\mathcal{D}[j||j_{\text{iso}}]}{2}. \end{aligned} \quad (35)$$

Because  $\langle j, f \rangle$  is proportional to the EPR [Eq. (10)], this decomposition provides a way to obtain nonnegative decompositions of the EPR. It should be noted that we have

infinitely many decompositions of this type. In addition, by choosing  $j_{\text{iso}} = j$ , the decomposition reduces to the fluctuation relation  $\Sigma = \mathcal{D}[j||-j]$ , where we used  $j = j_{\text{iso}}$ . This orthogonality has been proposed [34,36–38] to characterize the quasipotential and the gradient flow aspects of equilibrium and nonequilibrium systems as we see in the next section. The isodissipation hypersurface is known as a level surface of the convex function in Hessian geometry [17] whose Legendre transform is a centro-affine surface in affine differential geometry [44]. In both cases, it works as a central geometric object. Moreover, this centro-affine surface plays an important role in isobaric thermodynamics, where volume can change in conjunction with reactions [42]. The Hilbert orthogonality is also defined in the flux space using  $\Psi(j)$  as  $\Psi(j_A + j_S) = \Psi(j_A - j_S)$  while we do not use it in this paper. It should be noted that the orthogonality of  $f_A$  and  $f_S$  does not mean the orthogonality of the corresponding  $j_A$  and  $j_S$  in general.

### D. Information geometric orthogonality

The second orthogonality comes from the information geometry [17–19]. For three flux vectors  $j, j', j'' \in \mathcal{J}$  satisfying  $j = j' + j''$ , the Bregman divergence  $\mathcal{D}[j||\mathbf{0}]$  can be decomposed as

$$\mathcal{D}[j||\mathbf{0}] = \mathcal{D}[j||j''] + \mathcal{D}[j''||\mathbf{0}] + \langle j', f'' \rangle, \quad (36)$$

where  $f''$  is the Legendre dual of  $j''$ . The second orthogonality of  $j'$  and  $j''$  is defined by  $\langle j', f'' \rangle = 0$ . Then the following generalized Pythagorean theorem (GPT) holds:

$$\mathcal{D}[j||\mathbf{0}] = \mathcal{D}[j||j''] + \mathcal{D}[j''||\mathbf{0}]. \quad (37)$$

For quadratic dissipation functions, the GPT is reduced to the conventional Pythagorean theorem

$$\|j\|_M^2 = \|j'\|_M^2 + \|j''\|_M^2. \quad (38)$$

Thus, Eq. (37) is a generalization of the usual orthogonality. We call this orthogonality the information geometric orthogonality. Similarly, for  $f = f^\dagger + f^\ddagger$ , the dual orthogonality between  $f^\dagger, f^\ddagger \in \mathcal{F}$  is also defined by  $\langle j^\dagger, f^\ddagger \rangle = 0$ , leading to the dual version of GPT,

$$\mathcal{D}^*[f||\mathbf{0}] = \mathcal{D}^*[f||f^\dagger] + \mathcal{D}^*[f^\dagger||\mathbf{0}]. \quad (39)$$

The Bregman divergence and the information geometric orthogonality are the central geometric tools in information geometry for analyzing manifolds of statistical models and for conducting statistically meaningful projections onto submanifolds [18,19]. In addition, they have recently been employed in thermodynamics of diffusion processes, MJP, and CRN [2,41,42,45–47]. However, these works investigate the information and Hessian geometric structures of state space, i.e., the space of the probability vector  $p$  or the concentration vector  $x$ , in which the relevant thermodynamic potential function such as the Gibbs free energy works as the convex potential function inducing the Hessian structure. As a result, the Hessian structure of the state space captures the energetic and equilibrium aspects of the systems. In contrast, the Hessian structure of the flux-force space in this work captures the kinetic and nonequilibrium aspects of the systems.

It should be noted that the Hilbert and information geometric orthogonalities are reduced to usual orthogonality

<sup>17</sup>It should be noted that the symmetry property of the dissipation function Eq. (22) is not required to derive these properties.

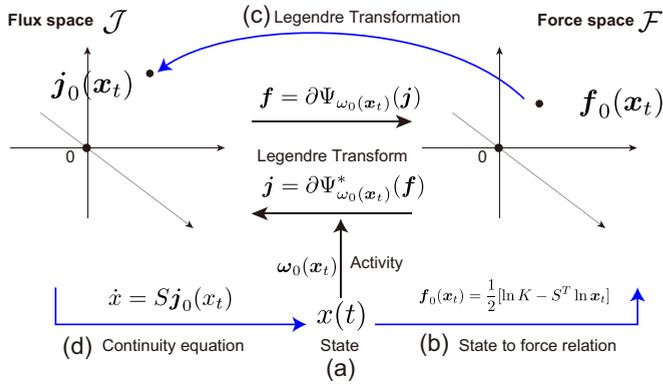


FIG. 3. Schematic representation of the flux-force relationship for CRN with mass action kinetics and induced CRE [Eq. (18)]. Depending on the current state  $\mathbf{x}(t)$ , (a), the force  $\mathbf{f}_0(\mathbf{x}_t)$  is induced (b). The force is mapped to the corresponding flux  $\mathbf{j}_0(\mathbf{x}_t)$  by the Legendre transformation  $\partial \Psi_{\omega_0(\mathbf{x}_t)}^*[\mathbf{f}_0(\mathbf{x}_t)]$  (c). The flux induces the change in  $\mathbf{x}_t$  via the continuity equation (d).

when the dissipation functions are quadratic (see also Appendix C 1).

## V. CHEMICAL REACTION DYNAMICS AND ENTROPY PRODUCTION DECOMPOSITIONS

In this section, we demonstrate how the two generalizations of orthogonality enable us to extend the different entropy production decompositions to CRN (and MJP). From the explicit functional forms of the one-way fluxes under mass action kinetics [Eq. (8)], the total flux, force, and activity defined in Eq. (9) and Eq. (11) are expressed as

$$\begin{aligned} \mathbf{j}_0(\mathbf{x}; \mathbf{k}^\pm) &= (\mathbf{k}^+ \circ \mathbf{B}_+^T - \mathbf{k}^- \circ \mathbf{B}_-^T) \mathbf{x}^{\Gamma^T}, \\ \mathbf{f}_0(\mathbf{x}; \mathbf{k}^\pm) &= \frac{1}{2} \left[ \ln \frac{\mathbf{k}^+}{\mathbf{k}^-} - S^T \ln \mathbf{x} \right], \\ \omega_0(\mathbf{x}; \mathbf{k}^\pm) &= 2\sqrt{\mathbf{k}^+ \circ \mathbf{k}^-} \circ \mathbf{x}^{[\Gamma(B^+ + B^-)]^T/2}. \end{aligned} \quad (40)$$

We use the subscript  $(\cdot)_0$  to designate the flux, force, and activity given by these particular forms. With these specific forms of force and activity, the CRN dynamics is described as a generalized flow [Eq. (18) and Fig. 3].

We further transform the kinetic parameters  $\mathbf{k}^\pm$  into force part ( $\mathbf{K}$ ) and activity part ( $\boldsymbol{\kappa}$ ) as

$$\mathbf{k}^+ = \boldsymbol{\kappa} \circ \mathbf{K}^{1/2}, \quad \mathbf{k}^- = \boldsymbol{\kappa} \circ \mathbf{K}^{-1/2}, \quad (41)$$

where  $\boldsymbol{\kappa} := \sqrt{\mathbf{k}^+ \circ \mathbf{k}^-}$  and  $\mathbf{K} := \mathbf{k}^+ / \mathbf{k}^-$ . Thus,  $(\boldsymbol{\kappa}, \mathbf{K})$  has the same information as  $\mathbf{k}^\pm$ . Moreover, we can verify that the force and activity are dependent only on  $\mathbf{K}$  and  $\boldsymbol{\kappa}$ , respectively, i.e.,  $\mathbf{f}_0(\mathbf{x}; \mathbf{k}^\pm) = \mathbf{f}_0(\mathbf{x}; \mathbf{K})$  and  $\omega_0(\mathbf{x}; \mathbf{k}^\pm) = \omega_0(\mathbf{x}; \boldsymbol{\kappa})$ . This clear and physically relevant separation of parameters is a good property of the nonquadratic dissipation function.

### A. Equilibrium dynamics

First, we consider the case in which the dynamics is equilibrium. For CRN with mass action kinetics, the equilibrium states are defined as the set  $\mathcal{M}_{\text{eq}}(\mathbf{K}, \boldsymbol{\kappa})$  satisfying the detailed

balance condition (DBC)

$$\mathcal{M}_{\text{eq}}(\mathbf{K}, \boldsymbol{\kappa}) := \{\mathbf{x} | \mathbf{j}_0(\mathbf{x}; \mathbf{K}, \boldsymbol{\kappa}) = 0\}. \quad (42)$$

For a parameter set  $(\mathbf{K}, \boldsymbol{\kappa})$  that admits the existence of equilibrium states, i.e.,  $\mathcal{M}_{\text{eq}}(\mathbf{K}, \boldsymbol{\kappa}) \neq \emptyset$ , the CRN becomes an equilibrium CRN. The condition  $\mathcal{M}_{\text{eq}}(\mathbf{K}, \boldsymbol{\kappa}) \neq \emptyset$  is satisfied if and only if the parameter  $\mathbf{K}$  satisfies the Wegscheider equilibrium (EQ) condition [2,23,41],

$$\ln \mathbf{K} \in \text{Im} S^T, \quad (43)$$

where  $\text{Im} S^T$  means the image of matrix  $S^T$ . We denote such  $\mathbf{K}$  as  $\mathbf{K}_{\text{eq}}$ . Note that the EQ condition is independent of the activity parameter  $\boldsymbol{\kappa}$ . Then, for each initial condition  $\mathbf{x}_0 \in \mathcal{X}$ , a unique equilibrium state  $\mathbf{x}_{\text{eq}}$  exists, which is determined by the intersection of  $\mathcal{M}_{\text{eq}}(\mathbf{K}_{\text{eq}})$  and the stoichiometric compatibility class [41,48]

$$\mathcal{P}(\mathbf{x}_0) := \{\mathbf{x} | (\mathbf{x} - \mathbf{x}_0) \in \text{Im} S\} \quad (44)$$

as  $\mathbf{x}_{\text{eq}} = \mathcal{M}_{\text{eq}}(\mathbf{K}_{\text{eq}}) \cap \mathcal{P}(\mathbf{x}_0)$  [41,48]. Moreover, from the DBC, we have  $\ln \mathbf{K}_{\text{eq}} = S^T \ln \mathbf{x}_{\text{eq}}$ . Then, the force  $\mathbf{f}_0(\mathbf{x})$  of the equilibrium dynamics is represented as

$$\mathbf{f}_0(\mathbf{x}) = -\frac{1}{2} S^T \ln \left( \frac{\mathbf{x}}{\mathbf{x}_{\text{eq}}} \right) = -\frac{1}{2} S^T \partial_{\mathbf{x}} \varphi_{\text{eq}}(\mathbf{x}), \quad (45)$$

where  $\varphi_{\text{eq}}(\mathbf{x}) := \mathcal{D}_{KL}[\mathbf{x} || \mathbf{x}_{\text{eq}}] := \mathbf{x}^T \ln \frac{\mathbf{x}}{\mathbf{x}_{\text{eq}}} - \mathbf{1}^T (\mathbf{x} - \mathbf{x}_{\text{eq}})$  is the Gibbs free energy of the equilibrium CRN. This representation indicates that  $\mathbf{f}_0(\mathbf{x}) \in \text{Im} S^T$  for the equilibrium dynamics. The matrix  $S^T$  can be considered as the discrete version of the gradient  $\text{grad}_S := -S^T$  for the chemical hypergraph because  $-S^T$  is the adjoint operator of  $\text{div}_S = -S$ . Thus, the equilibrium force is a gradient of  $\varphi_{\text{eq}}(\mathbf{x})$ :  $\mathbf{f}_0(\mathbf{x}) = \frac{1}{2} \text{grad}_S [\partial_{\mathbf{x}} \varphi_{\text{eq}}(\mathbf{x})]$ . Additionally, the Wegscheider condition [Eq. (43)] is interpreted as  $\ln \mathbf{K}_{\text{eq}} \in \text{Im} [\text{grad}_S]$ . Then the dynamics [Eq. (18)] under the EQ parameter condition becomes a generalized gradient flow (GF) of  $\varphi_{\text{eq}}(\mathbf{x})$  [10,27],

$$\dot{\mathbf{x}} = S \mathbf{j}_0(\mathbf{x}) = S \partial \Psi_{\mathbf{x}}^* \left[ -\frac{1}{2} S^T \partial_{\mathbf{x}} \varphi_{\text{eq}}(\mathbf{x}) \right] \quad (46)$$

$$= -\text{div}_S \partial \Psi_{\mathbf{x}}^* \left[ \frac{1}{2} \text{grad}_S [\partial_{\mathbf{x}} \varphi_{\text{eq}}(\mathbf{x})] \right], \quad (47)$$

where  $\Psi_{\mathbf{x}}^* := \Psi_{\omega_0(\mathbf{x})}^*$ . Along  $\mathbf{x}_t$ ,  $\varphi_{\text{eq}}(\mathbf{x}_t)$  is decreasing because  $d\varphi_{\text{eq}}(\mathbf{x}_t)/dt$  is always nonpositive as

$$\begin{aligned} -\frac{d\varphi_{\text{eq}}(\mathbf{x}_t)}{dt} &= \langle \dot{\mathbf{x}}, -\partial_{\mathbf{x}} \varphi_{\text{eq}}(\mathbf{x}_t) \rangle = \langle S \mathbf{j}_0(\mathbf{x}_t), -\partial_{\mathbf{x}} \varphi_{\text{eq}}(\mathbf{x}_t) \rangle \\ &= \langle \mathbf{j}_0(\mathbf{x}_t), -S^T \partial_{\mathbf{x}} \varphi_{\text{eq}}(\mathbf{x}_t) \rangle = 2 \langle \mathbf{j}_0(\mathbf{x}_t), \mathbf{f}_0(\mathbf{x}_t) \rangle \\ &= 2 [\Psi_{\mathbf{x}_t}(\mathbf{j}_0(\mathbf{x}_t)) + \Psi_{\mathbf{x}_t}^*(\mathbf{f}_0(\mathbf{x}_t))] = \dot{\Sigma}_t \geq 0, \end{aligned} \quad (48)$$

where we use Eq. (15) and Eq. (10). By integration, we have the relation between the change of free energy and the entropy production,

$$\varphi_{\text{eq}}(\mathbf{x}_0) - \varphi_{\text{eq}}(\mathbf{x}_t) = \int_{t'=0}^t \dot{\Sigma}_{t'} dt'. \quad (49)$$

<sup>18</sup>The uniqueness of the intersection is a consequence of Hessian geometric structure on the concentration-chemical potential space.

### B. Complex balanced dynamics, Hilbert orthogonality, and Hatano-Sasa decomposition

Next, we consider the nonequilibrium complex balanced (CB) steady state, and the EPR decomposition induced by the Hilbert orthogonality. The result here is a reinterpretation of the result in Refs. [35,36,38] from the viewpoint of Hessian geometry. Nevertheless, we include it here because the decomposition is contrasted with the information geometric decomposition in the next subsection.

The CB states of CRN are defined as the set satisfying

$$\mathcal{M}_{\text{cb}}(\mathbf{K}, \boldsymbol{\kappa}) := \{\mathbf{x} | B\mathbf{j}_0(\mathbf{x}; \mathbf{K}, \boldsymbol{\kappa}) = \mathbf{0}\}. \quad (50)$$

Similarly to the equilibrium case, a parameter set  $(\mathbf{K}_{\text{cb}}, \boldsymbol{\kappa}_{\text{cb}})$  induces CB dynamics if the CB condition  $\mathcal{M}_{\text{cb}}(\mathbf{K}_{\text{cb}}, \boldsymbol{\kappa}_{\text{cb}}) \neq \emptyset$  holds.<sup>19</sup> Then, for each initial condition  $\mathbf{x}_0$ , the CB steady state  $\mathbf{x}_{\text{cb}}$  is uniquely determined by the intersection of  $\mathcal{M}_{\text{cb}}(\mathbf{K}_{\text{cb}}, \boldsymbol{\kappa}_{\text{cb}})$  and the stoichiometric compatibility class  $\mathcal{P}(\mathbf{x}_0)$  as  $\mathbf{x}_{\text{cb}} = \mathcal{M}_{\text{cb}}(\mathbf{K}_{\text{cb}}, \boldsymbol{\kappa}_{\text{cb}}) \cap \mathcal{P}(\mathbf{x}_0)$  [41,48].

The CB state inherits important properties of the equilibrium state such as the uniqueness, global stability, and the gradient-flow-like aspect [22,23,41,48]. By definition, an EQ state is a CB state, but not vice versa. Additionally, the steady states of MJP always satisfy the CB condition because  $B\mathbf{j}_0(\mathbf{x}; \mathbf{K}, \boldsymbol{\kappa}) = \mathbf{0}$  is nothing but the steady state condition for MJP [Eq. (2)]. Thus, MJP are unconditionally complex balanced. In contrast, CRN are not always complex balanced depending on the parameter value.

Using  $\mathbf{x}_{\text{cb}}$ , the force  $\mathbf{f}_0(\mathbf{x})$  is represented as

$$\mathbf{f}_0(\mathbf{x}) = -\frac{1}{2}S^T \partial_{\mathbf{x}} \varphi_{\text{cb}}(\mathbf{x}) + \frac{1}{2} \ln \frac{\mathbf{K}_{\text{cb}}}{\bar{\mathbf{K}}_{\text{eq}}}, \quad (51)$$

where  $\bar{\mathbf{K}}_{\text{eq}} := S^T \ln \mathbf{x}_{\text{cb}}$  and  $\varphi_{\text{cb}}(\mathbf{x}) := \mathcal{D}_{KL}[\mathbf{x} || \mathbf{x}_{\text{cb}}]$ . Let  $\mathbf{f}_S(\mathbf{x}) := -\frac{1}{2}S^T \partial_{\mathbf{x}} \varphi_{\text{cb}}(\mathbf{x})$  and  $\mathbf{f}_A := \frac{1}{2} \ln \frac{\mathbf{K}_{\text{cb}}}{\bar{\mathbf{K}}_{\text{eq}}}$ . Then, for all  $\mathbf{x} \in \mathcal{X}$ , the generalized Hilbert orthogonality holds (Fig. 4):

$$\Psi_{\mathbf{x}}^*(\mathbf{f}_S(\mathbf{x}) + \mathbf{f}_A) = \Psi_{\mathbf{x}}^*(\mathbf{f}_S(\mathbf{x}) - \mathbf{f}_A), \quad (52)$$

because of the CB condition  $B\mathbf{j}_0(\mathbf{x}_{\text{cb}}) = \mathbf{0}$  [34,36–38] (see Appendix D 2 for a proof). Thus, Eq. (35) admits a nonnegative decomposition of the EPR,

$$\dot{\Sigma}(\mathbf{x}) = \dot{\Sigma}_{\text{ex}}^{\text{GF}}(\mathbf{x}) + \dot{\Sigma}_{\text{hk}}^{\text{GF}}(\mathbf{x}), \quad (53)$$

where  $\dot{\Sigma}_{\text{ex}}^{\text{GF}}(\mathbf{x}) := 2\langle \mathbf{j}_0(\mathbf{x}), \mathbf{f}_S(\mathbf{x}) \rangle$  and  $\dot{\Sigma}_{\text{hk}}^{\text{GF}}(\mathbf{x}) := 2\langle \mathbf{j}_0(\mathbf{x}), \mathbf{f}_A \rangle$ . Owing to this nonnegative decomposition,  $\varphi_{\text{cb}}(\mathbf{x})$  works as a Lyapunov function (quasipotential),

$$-\frac{d\varphi_{\text{cb}}(\mathbf{x}_t)}{dt} = \dot{\Sigma}_{\text{ex}}^{\text{GF}}(\mathbf{x}_t) = \mathcal{D}[\mathbf{j}_0(\mathbf{x}_t) || -\mathbf{f}_{\text{iso}}(\mathbf{x}_t)] \geq 0,$$

and  $\dot{\Sigma}_{\text{ex}}^{\text{GF}}(\mathbf{x})$  behaves like an excess EPR, i.e.,  $\lim_{t \rightarrow \infty} \dot{\Sigma}_{\text{ex}}^{\text{GF}}(\mathbf{x}_t) = 0$ . Thus,  $\varphi_{\text{cb}}(\mathbf{x})$  as a Lyapunov function guarantees the stability of the CB state  $\mathbf{x}_{\text{cb}}$  and the gradient-flow-like property of the CB dynamics.

This decomposition is equivalent to those proposed for CRN in [22,23] as a generalization of the Hatano-Sasa decomposition [49]. Therefore, this result is just a reinterpretation

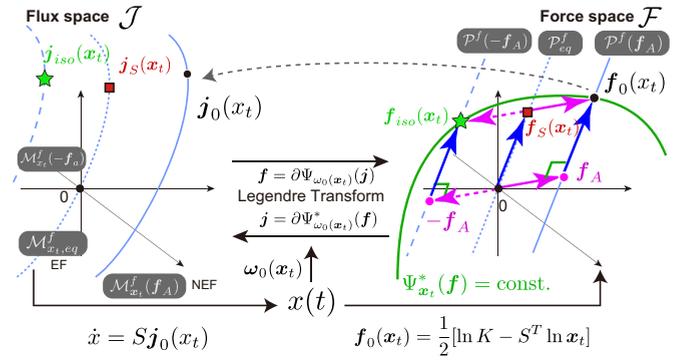


FIG. 4. Schematic illustration of the Hilbert orthogonality between  $\mathbf{f}_S(\mathbf{x}_t)$  (blue arrows) and  $\mathbf{f}_A$  (magenta arrows). The green-solid curve represents the isodissipation hypersurface:  $\{\mathbf{f} | \Psi_{\mathbf{x}_t}^*(\mathbf{f}) = \Psi_{\mathbf{x}_t}^*(\mathbf{f}_0(\mathbf{x}_t)) = \text{const.}\}$ . The black circle, red square, and green star in  $\mathcal{F}$  are  $\mathbf{f}_0(\mathbf{x}_t)$ ,  $\mathbf{f}_S(\mathbf{x}_t)$ , and  $\mathbf{f}_{\text{iso}}(\mathbf{x}_t)$ , respectively. The black circle, red square, and green star in  $\mathcal{J}$  are their Legendre transform, i.e.,  $\mathbf{j}_0(\mathbf{x}_t)$ ,  $\mathbf{j}_S(\mathbf{x}_t)$ , and  $\mathbf{j}_{\text{iso}}(\mathbf{x}_t)$ , respectively. The light blue lines in  $\mathcal{F}$  are the NEF subspaces,  $\mathcal{P}^f(\mathbf{f}_A)$  (solid line) and  $\mathcal{P}^f(-\mathbf{f}_A)$  (dashed line), and the EF subspace  $\mathcal{P}_{\text{eq}}^f$  (dotted line). In  $\mathcal{J}$ , the corresponding NEF manifolds,  $\mathcal{M}_{\mathbf{x}_t}^f(\mathbf{f}_A)$  (the light-solid blue curve) and  $\mathcal{M}_{\mathbf{x}_t}^f(-\mathbf{f}_A)$  (the light-dashed-blue curve), and the EF manifold  $\mathcal{M}_{\mathbf{x}_t, \text{eq}}^f$  (the light-dotted-blue curve) are also depicted.

of the previous ones. However, owing to the geometric perspective endowed by Hessian geometry, the nonnegativity of  $\dot{\Sigma}_{\text{ex}}^{\text{GF}}(\mathbf{x})$  and  $\dot{\Sigma}_{\text{hk}}^{\text{GF}}(\mathbf{x})$  becomes transparent, the proof of which in previous papers required more technical and less transparent computations [22,23].

### C. Information geometric orthogonality and Maes-Netočný decomposition

While the gradient flow aspect and associated EPR decomposition by the quasipotential have been investigated in CRN theory, large deviation theory, and applied mathematics [22,23,36,48], the decomposition may not capture other important thermodynamic aspects such as minimum dissipation or minimum entropy production principles. In addition, the Hilbert orthogonality for  $\varphi_{\text{cb}}(\mathbf{x})$  [Eq. (52)] holds only when the CB condition is satisfied. In CRN, we can have noncomplex-balanced steady states depending on the parameter values.

As another type of EPR decomposition, we have the Maes-Netočný decomposition, which generalizes the Komatsu-Nakagawa-Sasa-Tasaki approach [50,51]. Moreover, its geometric interpretation for overdamped diffusion processes was recently provided in terms of the formal Riemannian geometry of the flux-force space [28]. Its extension to MJP and CRN has yet to be achieved because of their nonquadratic dissipation functions. We show that the information geometric orthogonality is central to this extension.

In the Maes-Netočný decomposition, for a given  $\mathbf{x} = \mathbf{x}_t$ ,<sup>20</sup> the nonequilibrium flux  $\mathbf{j}_0(\mathbf{x})$  is decomposed into the quasisteady flux  $\mathbf{j}_{\text{st}}(\mathbf{x})$  and the remaining part as  $\mathbf{j}_0(\mathbf{x}) = \mathbf{j}_{\text{st}}(\mathbf{x}) +$

<sup>19</sup>In contrast to the equilibrium case, the CB condition  $\mathcal{M}_{\text{cb}}(\mathbf{K}, \boldsymbol{\kappa}) \neq \emptyset$  depends on both  $\mathbf{K}$  and  $\boldsymbol{\kappa}$ .

<sup>20</sup> $\mathbf{x}_t$  is not necessarily a steady state here.

$(\mathbf{j}_0(\mathbf{x}) - \mathbf{j}_{st}(\mathbf{x}))$ . The quasisteady flux  $\mathbf{j}_{st}(\mathbf{x})$  is the flux that makes a given state  $\mathbf{x}$  steady obtained by modulating the conserved force, i.e., being generated by the gradient of potential force in the case of overdamped diffusion. This means that  $\mathbf{j}_{st}(\mathbf{x}) \in \text{Ker}[\text{div}]$  and  $\mathbf{f}_0(\mathbf{x}) - \mathbf{f}_{st}(\mathbf{x}) \in \text{Im}[\text{grad}]$  hold where  $\text{Ker}[\text{div}]$  and  $\text{Im}[\text{grad}]$  formally mean the kernel and image of the divergence and gradient operators. Furthermore,  $\mathbf{j}_{st}(\mathbf{x})$  is also characterized as the minimum EPR flux under the constraint of  $\mathbf{f}_0(\mathbf{x}) - \mathbf{f}_{st}(\mathbf{x}) \in \text{Im}[\text{grad}]$  [50]. From the viewpoint of vector calculus, the decomposition is a special case of the Helmholtz-Hodge decomposition [52]. With the Riemannian inner product structure between  $\mathcal{J}$  and  $\mathcal{F}$  for overdamped diffusion, we can also regard  $\mathbf{j}_{st}(\mathbf{x})$  as the orthogonal projection of  $\mathbf{j}_0(\mathbf{x})$  onto  $\text{Ker}[\text{div}]$  along  $\text{Im}[\text{grad}]$ . Because the Maes-Netočný decomposition is tightly linked to the orthogonality in the inner product space, it is not trivial how to extend the decomposition to other systems such as CRN and MJP, if they are characterized by nonquadratic dissipation functions. We show how information-geometric orthogonality can resolve this problem.

First, we define a linear affine subspace  $\mathcal{P}^v(\mathbf{j})$  as

$$\mathcal{P}^v(\mathbf{j}) = \{\mathbf{j}' | \mathbf{j}' - \mathbf{j} \in \text{Ker}S\} \subset \mathcal{J}. \quad (54)$$

Because  $\mathbf{j}' \in \mathcal{P}^v(\mathbf{j})$  satisfies  $S\mathbf{j}' = S\mathbf{j}$ ,  $\mathcal{P}^v(\mathbf{j})$  is the set of the fluxes that induces the same instantaneous velocity  $\dot{\mathbf{x}} = S\mathbf{j}$  as  $\mathbf{j}$  does. Thus, we call  $\mathcal{P}^v(\mathbf{j})$  an isovelocity subspace in  $\mathcal{J}$ . Because  $\dot{\mathbf{x}} = 0$  for the steady state flux, we also denote  $\mathcal{P}_{st}^v := \mathcal{P}^v(\mathbf{0})$  as the steady state (zero-velocity) subspace.

As the complementary subspace of  $\mathcal{P}^v(\mathbf{j})$ , we define a linear affine subspace  $\mathcal{P}^f(\mathbf{f}) \subset \mathcal{F}$  as

$$\mathcal{P}^f(\mathbf{f}) := \{\mathbf{f}' | \mathbf{f}' - \mathbf{f} \in \text{Im}S^T\}. \quad (55)$$

Because  $S^T$  is the gradient of a CRN, i.e.,  $\text{grad} = -S^T$ ,  $\mathcal{P}^f(\mathbf{f})$  is the subspace obtained by shifting  $\text{Im}[\text{grad}] = \text{Im}[S^T]$  by  $\mathbf{f}$ . We call  $\mathcal{P}^f(\mathbf{f})$  a nonequilibrium force (NEF) subspace in  $\mathcal{F}$  if  $\mathbf{f} \notin \text{Im}S^T$  and the equilibrium force (EF) subspace if  $\mathbf{f} \in \text{Im}S^T$  because, if the equilibrium condition [Eq. (42)] is satisfied,  $\mathbf{f}_0(\mathbf{x})$  is always in  $\text{Im}[S^T]$  as in Eq. (45). We denote the EF subspace by  $\mathcal{P}_{eq}^f := \mathcal{P}^f(\mathbf{0})$ .

To obtain the generalized version of  $\mathbf{j}_{st}(\mathbf{x})$ , we transform  $\mathcal{P}^f(\mathbf{f})$  and  $\mathcal{P}_{eq}^f$  to  $\mathcal{J}$  via the Legendre transformation

$$\mathcal{M}_x^f(\mathbf{f}) := \partial\Psi_x^*[\mathcal{P}^f(\mathbf{f})] \subset \mathcal{J}, \quad (56)$$

$$\mathcal{M}_{x,eq}^f := \partial\Psi_x^*[\mathcal{P}_{eq}^f] \subset \mathcal{J}. \quad (57)$$

It should be noted that  $\mathcal{M}_x^f(\mathbf{f})$  depends on the current state  $\mathbf{x}$  of the system even though  $\mathcal{P}^f(\mathbf{f})$  does not because the Legendre transformation is dependent on  $\mathbf{x}$  via the activity  $\omega_0(\mathbf{x})$ . From the definition,  $\mathcal{M}_x^f(\mathbf{f})$  and  $\mathcal{M}_{x,eq}^f$  are the set of nonequilibrium and equilibrium fluxes obtained by modulating the gradient force.

Then the generalized quasisteady flux  $\mathbf{j}_{st}(\mathbf{x})$  is obtained by the intersection of  $\mathcal{P}_{st}^v$  and  $\mathcal{M}_x^f(\mathbf{f}_0(\mathbf{x}))$ ,

$$\mathbf{j}_{st}(\mathbf{x}) := \mathcal{P}_{st}^v \cap \mathcal{M}_x^f(\mathbf{f}_0(\mathbf{x})). \quad (58)$$

For the case of overdamped diffusion processes, the Legendre transformation of the corresponding quadratic dissipation function is linear and, therefore,  $\mathcal{M}_x^f(\mathbf{f}_0(\mathbf{x}))$  becomes a flat

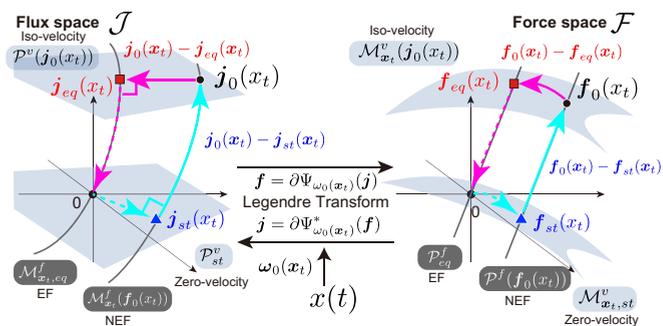


FIG. 5. Schematic illustration of the information geometric orthogonality. (Left) The orthogonality between  $\mathbf{j}_{st}(x_t)$  and  $\mathbf{j}_0(x_t) - \mathbf{j}_{st}(x_t)$  (blue arrows) and the dual orthogonality between  $\mathbf{j}_{eq}(x_t)$  and  $\mathbf{j}_0(x_t) - \mathbf{j}_{eq}(x_t)$  (red arrows) in  $\mathcal{J}$ . The blue planes are the isovelocity subspace  $\mathcal{P}^v(\mathbf{j}_0(x_t))$  (the upper plane) and the zero-velocity subspace  $\mathcal{P}_{st}^v$  (the lower plane). The grey curves are the NEF manifold  $\mathcal{M}_x^f(\mathbf{f}_0(x_t))$  and the EF manifold  $\mathcal{M}_{x,eq}^f$ . (Right) The same orthogonality shown in  $\mathcal{F}$  space. The orthogonality between  $\mathbf{f}_{st}(x_t)$  and  $\mathbf{f}_0(x_t) - \mathbf{f}_{st}(x_t)$  (blue arrows) and the dual orthogonality between  $\mathbf{f}_{eq}(x_t)$  and  $\mathbf{f}_0(x_t) - \mathbf{f}_{eq}(x_t)$  (red arrows) in  $\mathcal{F}$ . The blue surfaces are the isovelocity manifold  $\mathcal{M}_x^v(\mathbf{j}_0(x_t))$  (the upper surface) and the zero-velocity manifold  $\mathcal{M}_{x,eq}^v$  (the lower surface). The grey lines are the NEF subspaces  $\mathcal{P}^f(\mathbf{f}_0(x_t))$  and the EF subspace  $\mathcal{P}_{x,eq}^f$ .

subspace. Then, the existence and uniqueness of  $\mathbf{j}_{st}(\mathbf{x})$  follow from the linear algebra. However, for the generalized  $\mathbf{j}_{st}(\mathbf{x})$ , the existence and uniqueness of  $\mathbf{j}_{st}(\mathbf{x})$  are not guaranteed *a priori*, because  $\mathcal{M}_x^f(\mathbf{f}_0(\mathbf{x}))$  is a curved manifold as shown in Fig. 5. Nonetheless,  $\mathbf{j}_{st}(\mathbf{x})$  in Eq. (58) exists uniquely. More generally, for any  $\mathbf{j}'$  and  $\mathbf{f}''$ , the following intersection is unique and transversal (see Appendix D 3 for proof):

$$\mathbf{j}_{int}(\mathbf{j}', \mathbf{f}'') = \mathcal{P}^v(\mathbf{j}') \cap \mathcal{M}_x^f(\mathbf{f}''). \quad (59)$$

This result generalizes the uniqueness of the intersection of complementary subspaces under an inner product structure. This is one of the notable properties of Hessian geometry. In information geometry and Hessian geometry,  $\mathcal{P}^v(\mathbf{j})$  and  $\mathcal{M}_x^f(\mathbf{f})$  are called dually flat subspaces [17–19].

By virtue of the uniqueness, we obtain a unique decomposition of the flux  $\mathbf{j}_0(\mathbf{x})$ ,

$$\mathbf{j}_0(\mathbf{x}) = \mathbf{j}_{st}(\mathbf{x}) + (\mathbf{j}_0(\mathbf{x}) - \mathbf{j}_{st}(\mathbf{x})). \quad (60)$$

Because  $\mathbf{j}_{st}(\mathbf{x}) \in \text{Ker}S$  and  $\mathbf{j}_{st}(\mathbf{x}) \in \mathcal{M}_x^f(\mathbf{f}_0(\mathbf{x}))$  where  $\mathcal{M}_x^f(\mathbf{f}_0(\mathbf{x})) = \partial_j\Psi_x^*[\mathbf{f}_0(\mathbf{x}) + \text{Im}[S^T]]$ ,  $\mathbf{j}_{st}(\mathbf{x})$  is the steady flux that is obtained by modulating the nonequilibrium force  $\mathbf{f}_0(\mathbf{x})$  by adding the gradient equilibrium force. We also have  $\mathbf{f}_0(\mathbf{x}) - \mathbf{f}_{st}(\mathbf{x}) \in \text{Im}S^T$ . Thus, Eq. (60) is a generalization of the Maes-Netočný decomposition and also of the Helmholtz decomposition. Moreover, from  $\langle \mathbf{j}_{st}(\mathbf{x}), \mathbf{f}_0(\mathbf{x}) - \mathbf{f}_{st}(\mathbf{x}) \rangle = 0$ , we have the generalized Pythagorean relation

$$\mathcal{D}_x^*[\mathbf{f}_0(\mathbf{x}) \|\mathbf{0}\rangle] = \mathcal{D}_x^*[\mathbf{f}_0(\mathbf{x}) \|\mathbf{f}_{st}(\mathbf{x})\rangle] + \mathcal{D}_x^*[\mathbf{f}_{st}(\mathbf{x}) \|\mathbf{0}\rangle], \quad (61)$$

which is further reduced to a decomposition of the dual dissipation function

$$\Psi_x^*(\mathbf{f}_0(\mathbf{x})) = \mathcal{D}_x^*[\mathbf{f}_0(\mathbf{x}) \|\mathbf{f}_{st}(\mathbf{x})\rangle] + \Psi_x^*(\mathbf{f}_{st}(\mathbf{x})). \quad (62)$$

Then,  $\mathbf{f}_{\text{st}}$  is characterized variationally as (see Appendix D 4)

$$\mathbf{f}_{\text{st}}(\mathbf{x}) = \arg \min_{\mathbf{f} \in \mathcal{P}^f(\mathbf{j}_0(\mathbf{x}))} \Psi_{\mathbf{x}}^*(\mathbf{f}). \quad (63)$$

This variational formula indicates that  $\mathbf{f}_{\text{st}}(\mathbf{x})$  is the minimum dissipation force that can be obtained by tilting the nonequilibrium force  $\mathbf{j}_0(\mathbf{x})$  with the equilibrium gradient force. Thus,  $\mathbf{j}_{\text{st}}(\mathbf{x})$  is the minimum dissipation (MD) steady state flux. For quadratic cases where the equivalence of the dissipation functions and the EPR holds [Eq. (26)], the result formally reduces to the minimum entropy production principle of Maes and Netočný.

However, for nonquadratic cases, Eq. (62) alone does not provide a relevant EPR decomposition because EPR is not determined only by  $\Psi_{\mathbf{x}}^*(\mathbf{f}_0(\mathbf{x}))$ . To resolve the problem, we consider the dual decomposition of  $\mathbf{j}_0(\mathbf{x})$  as shown in Fig. 5,

$$\mathbf{j}_0(\mathbf{x}) = \mathbf{j}_{\text{eq}}(\mathbf{x}) + (\mathbf{j}_0(\mathbf{x}) - \mathbf{j}_{\text{eq}}(\mathbf{x})), \quad (64)$$

where we define  $\mathbf{j}_{\text{eq}}(\mathbf{x})$  as

$$\mathbf{j}_{\text{eq}}(\mathbf{x}) := \mathcal{P}^v(\mathbf{j}_0(\mathbf{x})) \cap \mathcal{M}_{\mathbf{x}, \text{eq}}^f. \quad (65)$$

Because  $\mathbf{j}_{\text{eq}}(\mathbf{x}) \in \mathcal{P}^v(\mathbf{j}_0(\mathbf{x}))$ , the flux  $\mathbf{j}_{\text{eq}}(\mathbf{x})$  induces the same instantaneous velocity  $\dot{\mathbf{x}}$  as  $\mathbf{j}_0(\mathbf{x})$  does. Moreover, because  $\mathbf{j}_{\text{eq}}(\mathbf{x}) \in \mathcal{M}_{\mathbf{x}, \text{eq}}^f$ , it is the flux induced by an equilibrium force, i.e., a pure gradient force. Thus,  $\mathbf{j}_{\text{eq}}(\mathbf{x})$  is the equilibrium flux that induces the same dynamics as  $\mathbf{j}_0(\mathbf{x})$ . It should be noted that  $\mathbf{j}_0(\mathbf{x})$  is generally induced by a nonequilibrium force. In addition, we have  $\mathbf{f}_{\text{eq}}(\mathbf{x}) \in \text{Im}S^T$  and  $\mathbf{j}(\mathbf{x}) - \mathbf{j}_{\text{eq}}(\mathbf{x}) \in \text{Ker}S$  from  $\mathbf{j}_{\text{eq}}(\mathbf{x}) \in \mathcal{M}_{\mathbf{x}, \text{eq}}^f$  and  $\mathbf{j}_{\text{eq}}(\mathbf{x}) \in \mathcal{P}^v(\mathbf{j}_0(\mathbf{x}))$ , respectively. Thus, Eq. (64) is another generalization of Helmholtz-Hodge decomposition of the flux  $\mathbf{j}_0(\mathbf{x})$ . Moreover, owing to  $(\mathbf{j}_0(\mathbf{x}) - \mathbf{j}_{\text{eq}}(\mathbf{x}), \mathbf{f}_{\text{eq}}(\mathbf{x})) = 0$ , we have the Pythagorean relation

$$\mathcal{D}_{\mathbf{x}}[\mathbf{j}_0(\mathbf{x})\|\mathbf{0}] = \mathcal{D}_{\mathbf{x}}[\mathbf{j}_0(\mathbf{x})\|\mathbf{j}_{\text{eq}}(\mathbf{x})] + \mathcal{D}_{\mathbf{x}}[\mathbf{j}_{\text{eq}}(\mathbf{x})\|\mathbf{0}], \quad (66)$$

which is reduced to a decomposition of the primal dissipation function

$$\Psi_{\mathbf{x}}(\mathbf{j}_0(\mathbf{x})) = \mathcal{D}_{\mathbf{x}}[\mathbf{j}_0(\mathbf{x})\|\mathbf{j}_{\text{eq}}(\mathbf{x})] + \Psi_{\mathbf{x}}(\mathbf{j}_{\text{eq}}(\mathbf{x})). \quad (67)$$

From this,  $\mathbf{j}_{\text{eq}}$  can be characterized variationally (see Appendix D 4 for proof)

$$\mathbf{j}_{\text{eq}}(\mathbf{x}) = \arg \min_{\mathbf{j} \in \mathcal{P}^v(\mathbf{j}_0(\mathbf{x}))} \Psi_{\mathbf{x}}(\mathbf{j}). \quad (68)$$

This variational formula means that, among all fluxes that induce the same velocity as  $\mathbf{j}_0(\mathbf{x})$ , the equilibrium flux  $\mathbf{j}_{\text{eq}}(\mathbf{x})$  is the flux that minimizes the primal dissipation function. Thus,  $\mathbf{j}_{\text{eq}}(\mathbf{x})$  is the MD flux.

By combining Eq. (67) and Eq. (62), we have an EPR decomposition (see Appendix D 5 for proof)

$$\dot{\Sigma}(\mathbf{x}) = \dot{\Sigma}_{\text{hk}}^{\text{MD}}(\mathbf{x}) + \dot{\Sigma}_{\text{ex}}^{\text{MD}}(\mathbf{x}) \quad (69)$$

where we define housekeeping and excess EPR as

$$\begin{aligned} \dot{\Sigma}_{\text{hk}}^{\text{MD}}(\mathbf{x}) &:= 2[\mathcal{D}_{\mathbf{x}}[\mathbf{j}_0(\mathbf{x})\|\mathbf{j}_{\text{eq}}(\mathbf{x})] + \Psi_{\mathbf{x}}^*(\mathbf{f}_{\text{st}}(\mathbf{x}))] \geq 0, \\ \dot{\Sigma}_{\text{ex}}^{\text{MD}}(\mathbf{x}) &:= 2[\mathcal{D}_{\mathbf{x}}^*[\mathbf{f}_0(\mathbf{x})\|\mathbf{f}_{\text{st}}(\mathbf{x})] + \Psi_{\mathbf{x}}(\mathbf{j}_{\text{eq}}(\mathbf{x}))] \geq 0. \end{aligned}$$

TABLE I. The parameter values used for simulations.

Type	$k_1^+$	$k_1^-$	$k_2^+$	$k_2^-$	$k_3^+$	$k_3^-$
Equilibrium	4	1	$\frac{3}{\sqrt{2}}$	$3\sqrt{2}$	$2\sqrt{2}$	$4\sqrt{2}$
Complex balanced	$\frac{1}{2}$	2	4	$\frac{47}{4}$	$\sqrt{2}$	$\frac{15}{2} + 2\sqrt{2}$
Noncomplex balanced	$\frac{1}{2}$	2	$\frac{1}{17}$	$\frac{85}{8}$	$\frac{273}{68}$	$\frac{137}{68}$

If the state  $\mathbf{x}_t$  converges to a steady state  $\mathbf{x}_{\text{st}}$ ,<sup>21</sup> then

$$\lim_{t \rightarrow \infty} \dot{\Sigma}_{\text{hk}}(\mathbf{x}_t) = \dot{\Sigma}_{\text{hk}}(\mathbf{x}_{\text{st}}) = 2(\mathbf{j}_0(\mathbf{x}_{\text{st}}), \mathbf{f}_0(\mathbf{x}_{\text{st}})), \quad (70)$$

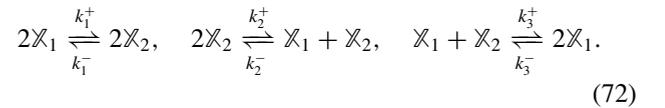
$$\lim_{t \rightarrow \infty} \dot{\Sigma}_{\text{ex}}(\mathbf{x}_t) = \dot{\Sigma}_{\text{ex}}(\mathbf{x}_{\text{st}}) = 0, \quad (71)$$

where we used  $\mathbf{j}_0(\mathbf{x}_{\text{st}}) = \mathbf{j}_{\text{st}}(\mathbf{x}_{\text{st}})$  and  $\mathbf{j}_{\text{eq}}(\mathbf{x}_{\text{st}}) = \mathbf{0}$ . Thus, this EPR decomposition geometrically generalizes the Maes-Netočný one to MJP and CRN and clarifies its dualistic minimum dissipation principles. If the state  $\mathbf{x}_t$  converges to a CB state  $\mathbf{x}_{\text{cb}}$ , then  $\lim_{t \rightarrow \infty} \dot{\Sigma}_{\text{hk}}(\mathbf{x}_t) = 2(\mathbf{j}_A, \mathbf{f}_A)$  holds.

## VI. NUMERICAL DEMONSTRATION

We numerically demonstrate the two decompositions in Eq. (53) and Eq. (69) and the geometric relations among  $\mathbf{j}_0(\mathbf{x})$ ,  $\mathbf{j}_{\text{iso}}(\mathbf{x})$ ,  $\mathbf{j}_S(\mathbf{x})$ , and  $\mathbf{j}_A$  and those among  $\mathbf{j}_0(\mathbf{x})$ ,  $\mathbf{j}_{\text{st}}(\mathbf{x})$ , and  $\mathbf{j}_{\text{eq}}(\mathbf{x})$  by using the CRN in Fig. 1(a) [53,54].

The CRN depicted in Fig. 1(a) is defined by the following set of chemical reaction equations:



The corresponding structural quantities are

$$S = \begin{pmatrix} -2 & +1 & +1 \\ +2 & -1 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} +1 & 0 & -1 \\ -1 & +1 & 0 \\ 0 & -1 & +1 \end{pmatrix}, \quad (73)$$

$$\Gamma = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \end{pmatrix}. \quad (74)$$

One can show that this simple nonlinear CRN has any of the equilibrium, CB, and non-CB steady states depending on the kinetic parameter values [53,54]. In addition, the flux and force spaces are three dimensional, and, thereby, the relevant geometric objects and quantities can be visualized computationally. We used the parameter values in Table I for the simulations in Fig. 6 and in Fig. 7. Because the state space is two-dimensional and the stoichiometric compatibility class is one-dimensional, i.e.,  $\mathcal{P}(\mathbf{x}_0) := \{\mathbf{x}|\mathbf{x} - \mathbf{x}_0 \in \text{Im}[S]\} = \{\mathbf{x}|\mathbf{x} = \mathbf{x}_0 + \xi(1, -1)^T, \xi \in \mathbb{R}\}$ , the trajectory is restricted in the one-dimensional  $\mathcal{P}(\mathbf{x}_0)$  independently of the parameter values. We also selected the parameter values in Table I so that all the sets of equilibrium, CB and non-CB steady states become identical to the set,  $\{\mathbf{x}|\mathbf{x}_2 = 2\mathbf{x}_1\}$ . Thus, the topological properties of the state space are the same for the three sets of parameters.

Figure 6 shows the values of the EPR and the associated dissipation functions in the state space. For the equilibrium

<sup>21</sup> $\mathbf{x}_{\text{st}}$  is not necessarily a CB state but can be a general steady state.

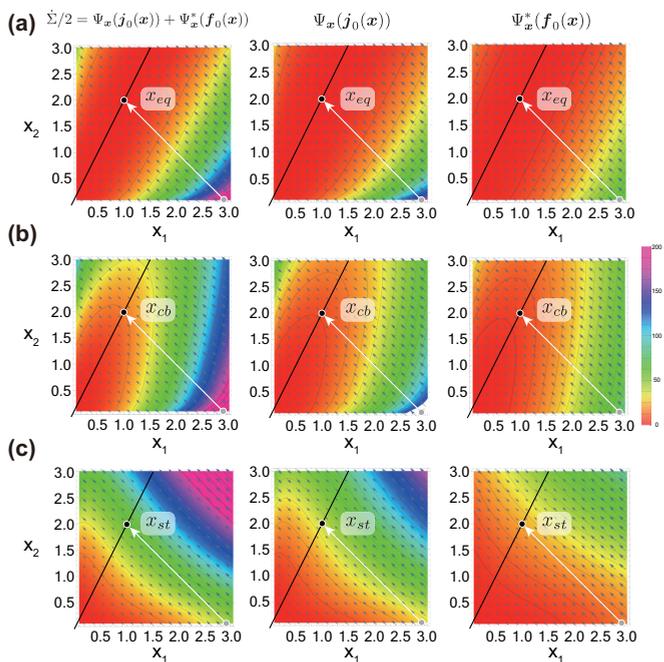


FIG. 6. Heatmap plots of EPR  $\dot{\Sigma}$  (left panels) and the two dissipation functions,  $\Psi_x(j_0(x))$  (center panels) and  $\Psi_x^*(f_0(x))$  (right panels) as functions of  $x$  for the equilibrium parameter set (a), the complex balanced parameter set (b), and the noncomplex balanced parameter set (c) shown in Table I. In each panel, gray and black points are the initial and the steady states, respectively. The white arrow is the trajectory of  $x_t$ . The black line in each panel is the set of equilibrium states (a), complex-balanced states (b), and noncomplex balanced steady states (c), respectively.

parameter [Fig. 6(a)], the EPR and dissipation functions attain their global minimum value of zero on the equilibrium states (black line). For the CB parameter [Fig. 6(b)], the global minimum is not necessarily attained, which results in the nonzero EPR at the CB steady state  $x_{cb}$  [Fig. 7(a)]. For the non-CB parameter [Fig. 6(c)], the landscapes of the EPR and dissipation functions become more complicated, and their values along the trajectory  $x_t$  are not monotonous as shown in Fig. 7(c).

For the CB and non-CB cases, the two decompositions, Eq. (53) and Eq. (69), are computed [Figs. 7(b) and 7(d)]. For the CB case, we verify that both decompositions provide the expected behaviors: both housekeeping and excess components stay nonnegative; the excess ones converge to 0; the housekeeping ones do to a finite nonnegative value [Fig. 7(b)].

For the non-CB case, the expected behaviors are also produced [Fig. 7(d)] for the information-geometric decomposition, Eq. (69) (a generalization of Maes-Netočný decomposition). On the contrary, the housekeeping component of the Hilbert decomposition, Eq. (53), becomes negative within a certain time window because the Hilbert orthogonality, Eq. (52), does not hold in this case. It should be noted that this result does not necessarily mean that the Hilbert decomposition is not applicable to the non-CB case. The result only indicates that the force decomposition by  $f_S(x) := -\frac{1}{2}S^T \partial_x \varphi_{cb}(x)$  and  $f_A := \frac{1}{2} \ln \frac{K_{cb}}{K_{eq}}$  with the specific form of  $\varphi_{cb}(x) := \mathcal{D}_{KL}[x||x_{cb}]$  no longer satisfies the Hilbert orthogonality. We may be able to recover the orthogonality and the

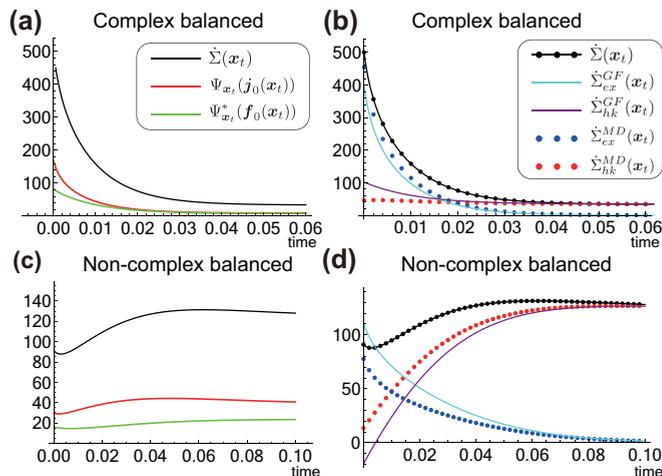


FIG. 7. The EPR and its decompositions along the trajectory  $x_t$  for complex balanced [(a),(b)] and noncomplex balanced [(c),(d)] parameter sets. The decomposition of EPR into two dissipation functions [(a),(c)]. The decompositions of EPR by Hilbert ( $\dot{\Sigma}_{hk}^{GF}(x)$ ,  $\dot{\Sigma}_{ex}^{GF}(x)$ ) and information geometric ( $\dot{\Sigma}_{ex}^{MD}(x)$ ,  $\dot{\Sigma}_{hk}^{MD}(x)$ ) orthogonalities [(b),(d)].

associated decomposition [38] by choosing other functional form of  $\varphi_{cb}(x)$ . Finding such  $\varphi_{cb}(x)$  is linked to the computation of quasipotential, but this problem is still challenging for generic non-CB cases.

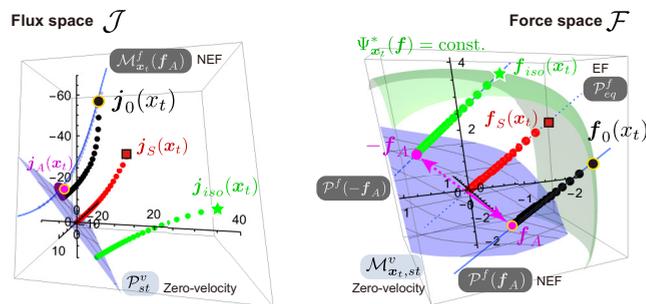


FIG. 8. Computational visualization of the Hilbert orthogonality between  $f_S(x_t)$  and  $f_A$  for the CRN in Fig 1(a). The green surface in  $\mathcal{F}$  represents the isodissipation hypersurface:  $\Psi_{x_t}^*(f) = \Psi_{x_t}^*(f_0(x_t)) = \text{const.}$ . The black circle with yellow border, red square with black border, green star, and magenta circle with yellow border in  $\mathcal{F}$  are respectively  $f_0(x_t)$ ,  $f_S(x_t)$ ,  $f_{iso}(x_t)$ , and  $f_A$  evaluated at  $x_t = (5/2, 1/2)^T$ . The black circle, red square, green star, and magenta circle with border in  $\mathcal{J}$  are their Legendre transform, i.e.,  $j_0(x_t)$ ,  $j_S(x_t)$ ,  $j_{iso}(x_t)$ , and  $j_A(x_t)$ , respectively. The black, red, and green circles without border in  $\mathcal{F}$  are the trajectories of  $\{f_0(x_{t'})\}$ ,  $\{f_S(x_{t'})\}$ , and  $\{f_{iso}(x_{t'})\}$ . The black, red, green, and magenta circles without border in  $\mathcal{J}$  are the trajectories of  $\{j_0(x_{t'})\}$ ,  $\{j_S(x_{t'})\}$ ,  $\{j_{iso}(x_{t'})\}$ , and  $\{j_A(x_{t'})\}$ , respectively. The light blue lines in  $\mathcal{F}$  are the NEF subspaces,  $\mathcal{P}^f(f_A)$  (solid line) and  $\mathcal{P}^f(-f_A)$  (dashed line), and the EF subspace  $\mathcal{P}_{eq}^f$  (dotted line). In  $\mathcal{J}$ , the corresponding NEF manifolds,  $\mathcal{M}_x^f(f_A)$  (the light blue curve with dots), is also depicted. The blue plane in  $\mathcal{J}$  is the zero-velocity subspace  $\mathcal{P}_{st}^v$ , and the blue surface in  $\mathcal{F}$  is its Legendre transformation:  $\mathcal{M}_{x_t, st}^v$ .

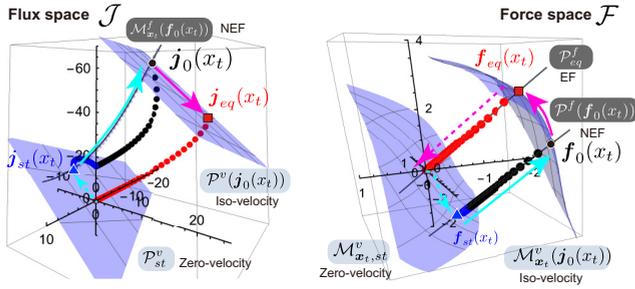


FIG. 9. Computational visualization of the information geometric orthogonality for the CRN in Fig 1(a). (Left) The orthogonality between  $\mathbf{j}_{st}(\mathbf{x}_t)$  (cyan-dashed arrow) and  $\mathbf{j}_0(\mathbf{x}_t) - \mathbf{j}_{st}(\mathbf{x}_t)$  (cyan-solid arrow) and the dual orthogonality between  $\mathbf{j}_{eq}(\mathbf{x}_t)$  and  $\mathbf{j}_0(\mathbf{x}_t) - \mathbf{j}_{eq}(\mathbf{x}_t)$  (purple-solid arrow) in  $\mathcal{J}$  evaluated at  $\mathbf{x}_t = (5/2, 1/2)^T$ . The blue planes are the isovelocity subspace  $\mathcal{P}^v(\mathbf{j}_0(\mathbf{x}_t))$  (the upper plane) and the zero-velocity subspace  $\mathcal{P}^v_{st}$  (the lower plane). The grey curve with dots is the NEF manifold  $\mathcal{M}_{x_t}^f(\mathbf{f}_0(\mathbf{x}_t))$ . Black, red, and blue circles are the trajectories of  $\{\mathbf{j}_0(\mathbf{x}_{t'})\}$ ,  $\{\mathbf{j}_{eq}(\mathbf{x}_{t'})\}$ , and  $\{\mathbf{j}_{st}(\mathbf{x}_{t'})\}$ , respectively. (Right) The same orthogonalities shown in  $\mathcal{F}$  space. The orthogonality between  $\mathbf{f}_{st}(\mathbf{x}_t)$  (cyan-curved arrow) and  $\mathbf{f}_0(\mathbf{x}_t) - \mathbf{f}_{st}(\mathbf{x}_t)$  (cyan-solid arrow) and the dual orthogonality between  $\mathbf{f}_{eq}(\mathbf{x}_t)$  (purple-dashed arrow) and  $\mathbf{f}_0(\mathbf{x}_t) - \mathbf{f}_{eq}(\mathbf{x}_t)$  (purple-solid arrow) in  $\mathcal{F}$ . Black, red circles are the trajectories of  $\{\mathbf{f}_0(\mathbf{x}_{t'})\}$ ,  $\{\mathbf{f}_{eq}(\mathbf{x}_{t'})\}$ , respectively. The blue surfaces are the isovelocity manifold  $\mathcal{M}_{x_t}^v(\mathbf{j}_0(\mathbf{x}_t))$  (the upper surface) and the zero-velocity manifold  $\mathcal{M}_{x_t, st}^v$  (the lower surface). The grey lines are the NEF subspaces  $\mathcal{P}^f(\mathbf{f}_0(\mathbf{x}_t))$  and the EF subspace  $\mathcal{P}^f_{eq}$ .

Finally, we computationally visualized the various geometric objects introduced in Figs. 8 and 9 for the CB parameter set in Table I.

First, we show the behaviors of the forces,  $\mathbf{f}_0(\mathbf{x}_t)$ ,  $\mathbf{f}_{iso}(\mathbf{x}_t)$ ,  $\mathbf{f}_S(\mathbf{x}_t)$ , and  $\mathbf{f}_A$ , which satisfy the Hilbert orthogonality, and the corresponding fluxes. In the force space  $\mathcal{F}$  (the right panel of Fig. 8), we observe that the trajectories of  $\mathbf{f}_0(\mathbf{x}_t)$ ,  $\mathbf{f}_{iso}(\mathbf{x}_t)$ , and  $\mathbf{f}_S(\mathbf{x}_t)$  are actually restricted on one-dimensional subspaces (lines),  $\mathcal{P}^f(\mathbf{f}_A)$ ,  $\mathcal{P}^f(-\mathbf{f}_A)$ , and  $\mathcal{P}^f_{eq} = \mathcal{P}^f(\mathbf{0})$ , respectively.  $\mathbf{f}_0(\mathbf{x}_t)$  and  $\mathbf{f}_{iso}(\mathbf{x}_t)$  are also on the isodissipation hypersurface (the green surface) satisfying  $\Psi_{x_t}^*(\mathbf{f}_0(\mathbf{x}_t)) = \Psi_{x_t}^*(\mathbf{f}_{iso}(\mathbf{x}_t))$ . In the flux space  $\mathcal{J}$  (the left panel of Fig. 8),  $\mathcal{P}^f(\mathbf{f}_A)$  is transformed to the one-dimensional curve,  $\mathcal{M}_{x_t}^f(\mathbf{f}_A)$ . The trajectories of  $\mathbf{j}_0(\mathbf{x}_t)$ ,  $\mathbf{j}_{iso}(\mathbf{x}_t)$ , and  $\mathbf{j}_S(\mathbf{x}_t)$  are also curved in  $\mathcal{J}$ . All the trajectories converge onto the zero-velocity subspace  $\mathcal{P}^v_{st}$ .

Next, we verify the information-geometric orthogonality in Fig. 9. In the flux space  $\mathcal{J}$  (the left panel in Fig. 9), we observe that  $\mathbf{j}_0(\mathbf{x}_t)$  and the corresponding equilibrium flux  $\mathbf{j}_{eq}(\mathbf{x}_t)$  are on the isovelocity subspace  $\mathcal{P}^v(\mathbf{j}_0(\mathbf{x}_t))$ . The steady state flux  $\mathbf{j}_{st}(\mathbf{x}_t)$  are on the intersection of the zero-velocity subspace  $\mathcal{P}^v_{st}$  and the NEF manifold  $\mathcal{M}_{x_t}^f(\mathbf{f}_0(\mathbf{x}_t))$ . By mapping these objects into  $\mathcal{F}$  by the Legendre transform, we observe that  $\mathbf{f}_0(\mathbf{x}_t)$  is on the intersection of the NEF subspace  $\mathcal{P}^f(\mathbf{f}_0(\mathbf{x}_t))$  and the isovelocity manifold  $\mathcal{M}_{x_t}^v(\mathbf{j}_0(\mathbf{x}_t))$ . Similarly,  $\mathbf{f}_{eq}(\mathbf{x}_t)$  is on the intersection of the EF subspace  $\mathcal{P}^f_{eq}$  and the isovelocity manifold  $\mathcal{M}_{x_t}^v(\mathbf{j}_0(\mathbf{x}_t))$ . By the Legendre transformation, the isovelocity manifold  $\mathcal{M}_{x_t}^v(\mathbf{j}_0(\mathbf{x}_t))$ , and the zero-velocity manifold  $\mathcal{M}_{x_t, st}^v$  are curved in  $\mathcal{F}$ . These graph-

ical representations demonstrate how the nonlinear Legendre transform relates the subspaces and manifolds in  $\mathcal{J}$  and  $\mathcal{F}$  and also how the generalized notions of orthogonality of Hessian geometry are realized in these spaces.

## VII. SUMMARY AND DISCUSSION

We have clarified that Hessian geometry is a natural geometric structure of nonequilibrium and nonlinear CRN as well as MJP. More generally, it can capture the geometry induced by nonquadratic convex dissipation functions. By employing generalized notions of orthogonality, various aspects of nonequilibrium dynamics are dissected as decompositions of EPR, which generalize the well-established ones [28,50,51] to CRN and MJP.

### A. Thermodynamic uncertainty relation and Fisher information

The nonquadratic property of the dissipation functions of CRN and MJP also appears implicitly in different problems of thermodynamics. One such example is the thermodynamic uncertainty relation (TUR) and its extension to MJP. TUR is a relation that bounds the fluctuation of a generalized current in a nonequilibrium steady state by the entropy production [55],

$$\frac{\mathbb{E}[\mathcal{J}_d^{ss}]^2}{\mathbb{V}\text{ar}[\mathcal{J}_d^{ss}]} \leq \frac{1}{2} \int_0^\tau \dot{\Sigma}(t') dt' \quad (75)$$

where  $\mathcal{J}_d^{ss}$  is a time-integrated generalized current at the steady state, and  $\mathbb{E}[\mathcal{J}_d^{ss}]$  and  $\mathbb{V}\text{ar}[\mathcal{J}_d^{ss}]$  are the mean and variance of  $\mathcal{J}_d^{ss}$ . Since the proposal of the TUR conjecture and its first proof [56,57], the TUR has been extended to various models and situations, including MJP [55] and CRN [58]. However, even though Eq. (75) is tight for overdamped diffusion processes, it is not tight for MJP. Instead, for MJP, the fluctuation of the current is bounded tightly by the pseudoentropy production rate (pEPR) [58],

$$\frac{\mathbb{E}[\mathcal{J}_d^{ss}]^2}{\mathbb{V}\text{ar}[\mathcal{J}_d^{ss}]} \leq \frac{1}{2} \int_0^\tau \dot{\Pi}(t') dt' \quad (76)$$

where the pEPR is defined as

$$\dot{\Pi}(t) := 2 \sum_e \frac{(j_e^+(\mathbf{x}_t) - j_e^-(\mathbf{x}_t))^2}{j_e^+(\mathbf{x}_t) + j_e^-(\mathbf{x}_t)}. \quad (77)$$

and  $\dot{\Pi}(t) \leq \dot{\Sigma}(t)$  holds [55]. We here mention that this gap between the EPR and pEPR is linked to the Fisher information of the dissipation function. The Fisher information matrix (or metric) for a strictly convex function  $\phi^*(\mathbf{f})$  is defined by its Hessian as

$$G_{e,e'}^*(\mathbf{f}) := \frac{\partial^2 \phi^*(\mathbf{f})}{\partial f_e \partial f_{e'}}. \quad (78)$$

For  $\phi^*(\mathbf{f}) = \Psi_\omega^*(\mathbf{f})$  defined by Eq. (12), we have

$$G^*(\omega, \mathbf{f}) = \text{diag}[\omega \circ \cosh(\mathbf{f})] = \text{diag}[\mathbf{j}^+ + \mathbf{j}^-], \quad (79)$$

where we used Eq. (9) and Eq. (11). The diagonal part,  $\mathbf{j}^+ + \mathbf{j}^-$ , is often called dynamical activity. Similarly, for  $\phi(\mathbf{j}) = \Psi_\omega(\mathbf{j})$ , we have the Fisher information matrix

$$G(\omega, \mathbf{j}) = G^*(\omega, \mathbf{f})^{-1} = \text{diag}\left[\frac{1}{\mathbf{j}^+ + \mathbf{j}^-}\right]. \quad (80)$$

Then, the pEPR is represented as

$$\dot{\Pi}(j) = 2\langle j, G(\omega, j)j \rangle = 2\langle j, f^p \rangle, \quad (81)$$

which allows us to regard  $\dot{\Pi}(t)$  as an approximation of the EPR by replacing the actual force  $f$  associated with  $j$  with a pseudoforce  $f^p := G(\omega, j)j$ . If the dissipation function is quadratic, the actual force  $f$  and the pseudoforce  $f^p$  coincide. Thus, the nontightness of the TUR for MJP is a manifestation of the nonquadratic nature of the dissipation functions.

### B. Network thermodynamics

The information geometric decomposition that we introduced is also related to the network thermodynamics [20,21,59–61]. The network thermodynamics is an attempt to extend the methodology of network theory developed mainly for linear electric circuits to other physical systems [62,63]. While the linear network theory successfully works for linear electric circuits, which have a linear force and flux relation (Ohm’s law) and resulting quadratic dissipation functions, its application to MJP and CRN has encountered difficulties due to the nonlinear relation between force and flux, especially when we evaluate the EPR decomposition for a transient and far from equilibrium state. The Hessian geometric structure and the generalized decompositions in this work are extensions of those investigated in network thermodynamics. Specifically, the information geometric decomposition is a nonlinear generalization of the cycle-cocycle decomposition [64]. However, network thermodynamics and network theory accommodate a wide variety of methods that are not exploited in this work [64,65]. For example, the algebraic structure underlying the network is explored using the integral basis defined by cycles, cords, and spanning trees in the network [21,66]. As an application, the steady EPR is related to cycle fluxes and affinities in the network. We may incorporate and exploit the algebraic structure of MJP and CRN more explicitly into our framework.

### C. Relevance of nonquadratic dissipation functions

As mentioned in Introduction, CRN and MJP can also be characterized by other quadratic dissipation functions, i.e., the force-flux relation [Eq. (9)] for CRN and MJP can be obtained by a quadratic function. Specifically, we can choose  $\hat{\Psi}_x^*(f) := \frac{1}{2}\langle M_x^* f, f \rangle$  where  $M_x^* := \text{diag}[\frac{j^+(x)-j^-(x)}{\ln j^+(x)-\ln j^-(x)}]$ . This type of dissipation function was proposed in [67–70] ahead of the nonquadratic ones we used. In addition, this quadratic dissipation function as well as the associated formal Riemannian geometry have been used in stochastic thermodynamics [71–73]. From the different choice of the dissipation function, different results, e.g., different EPR decompositions, are derived mathematically. It would be the next primary mission to resolve the issue about how to choose physically relevant dissipation functions.

### D. Duality and variational characterizations in thermodynamics

The notion of duality is the core of thermodynamics. In equilibrium thermodynamics, the duality of extensive and intensive variables induced by thermodynamic potential func-

tions characterizes the energetic aspect of physical systems. The roles and differences of dual pairs of thermodynamic potential functions are well recognized. In addition, Hessian geometry is central to comprehend the geometric properties of equilibrium thermodynamics [2,41]. For the duality between the force and flux explored in this paper, the dual dissipation functions have been recognized for a while in macroscopic fluctuation theory [7] and metric gradient flow theory [16]. Nevertheless, nonquadratic dissipation functions have been investigated only very recently even in large deviation theory [10,11,27,33–38] and also in other physics communities [74]. As in the case of equilibrium thermodynamics, grasping the roles played by the dual functions should be essential for understanding nonequilibrium and kinetic aspects of thermodynamics. Hence, Hessian geometry works as an indispensable tool for investigating the geometry induced by the duality and also the variational aspects of nonequilibrium phenomena. In thermodynamics, there exist continued attempts to characterize nonequilibrium states and relations variationally, which go by the names of minimum entropy production principle [75–77], maximum entropy production principle [78,79], the least dissipation principle [5,80], and others. However, all of these principles still have limitations in their applicability for nonlinear and far from equilibrium situations [81]. Hessian geometry and also information geometry, which have an ability to handle nonlinearity induced by convex functions, may contribute to resolving a part of problems in such principles. Moreover, they can serve as a natural language to integrate the equilibrium (energetic) and nonequilibrium (kinetic) descriptions, and thereby provide us with a more universal understanding of thermodynamic systems and extend the applicability of nonequilibrium thermodynamics.

### ACKNOWLEDGMENTS

This research is supported by JST (Grants No. JP-MJCR2011, No. JPMJCR1927) and JSPS (Grant No. 19H05799). We thank Shuhei Horiguchi for valuable suggestions.

### APPENDIX A: NOTATION

We clarify the notation used in the main text. The bold letters represent vectors, and we define operations for and between them as follows:

$$\langle j, f \rangle := \sum_{e=1}^{N_e} j_e f_e, \quad (A1)$$

$$j \circ j' := (j_1 j'_1, \dots, j_{N_e} j'_{N_e})^T, \quad (A2)$$

$$\frac{j}{j'} := \left( \frac{j_1}{j'_1}, \dots, \frac{j_{N_e}}{j'_{N_e}} \right)^T, \quad (A3)$$

$$e^f := (e^{f_1}, \dots, e^{f_{N_e}})^T, \quad (A4)$$

$$\sqrt{j} := (\sqrt{j_1}, \dots, \sqrt{j_{N_e}})^T, \quad (A5)$$

$$x^\alpha := \prod_{i=1}^{N_x} x_i^{\alpha_i} \in \mathbb{R}, \quad (A6)$$

$$x^{A^T} := (x^{\alpha_1}, \dots, x^{\alpha_{N_v}})^T, \quad (A7)$$

where  $A = (\alpha_1, \dots, \alpha_{N_c})$  is a matrix.  $N_x$ ,  $N_e$ , and  $N_v$  are the numbers of molecular species, reactions, and complexes. The last notation is consistent with the application of the component-wise logarithmic function as  $\ln(x^{A^T}) = A^T \ln x$ .

## APPENDIX B: DUALITY IN THE FLUX-FORCE DUAL SPACE

The relations among  $j^\pm$ ,  $j$ ,  $f$ , and  $\omega$  in Eq. (9) and Eq. (11) are immediately derived by noting that  $j^\pm = \frac{\omega}{2} \circ e^{\pm f}$  holds. The Legendre identity in Eq. (15) is verified as

$$\begin{aligned} \Psi_\omega(j) + \Psi_\omega^*(f) &= j^T \sinh^{-1}\left(\frac{j}{\omega}\right) - \omega^T \left[ \sqrt{1 + \left(\frac{j}{\omega}\right)^2} - 1 \right] \\ &\quad + \omega^T [\cosh(f) - 1], \end{aligned} \quad (\text{B1})$$

$$\begin{aligned} &= \langle j, f \rangle - \omega^T \left[ \cosh\left(\sinh^{-1}\left(\frac{j}{\omega}\right)\right) - 1 \right] \\ &\quad + \omega^T [\cosh(f) - 1] = \langle j, f \rangle, \end{aligned} \quad (\text{B2})$$

where we used  $f = \partial_j \Psi_\omega(j) = \sinh^{-1}\left(\frac{j}{\omega}\right)$  and  $\cosh(\sinh^{-1}(x)) = \sqrt{1 + x^2}$ . These relations were obtained mainly via large deviation rate functions [10,11,27,33–38]:

## APPENDIX C: BREGMAN DIVERGENCE

We show that Eq. (29) and Eq. (30) hold [18,19]. From the definition,

$$\begin{aligned} \mathcal{D}[j||j'] &:= \Psi(j) - \Psi(j') - \langle j - j', \partial \Psi[j'] \rangle \\ &= \Psi(j) - \Psi(j') - \langle j - j', f' \rangle \\ &= \Psi(j) - (\Psi(j') - \langle j', f' \rangle) - \langle j, f' \rangle \\ &= \Psi(j) + \Psi^*(f') - \langle j, f' \rangle = \mathcal{D}[j; f'], \end{aligned} \quad (\text{C1})$$

holds where we used Eq. (14) and Eq. (15), namely,

$$f' = \partial \Psi[j'], \quad \Psi(j') + \Psi^*(f') = \langle j', f' \rangle. \quad (\text{C2})$$

Similarly,

$$\begin{aligned} \mathcal{D}^*[f'||f] &:= \Psi^*(f') - \Psi^*(f) - \langle \partial_f \Psi^*(f), f' - f \rangle \\ &= \Psi^*(f') - \Psi^*(f) - \langle j, f' - f \rangle \\ &= \Psi^*(f') - (\Psi^*(f) - \langle j, f \rangle) - \langle j, f' \rangle \\ &= \Psi(j) + \Psi^*(f') - \langle j, f' \rangle = \mathcal{D}[j; f']. \end{aligned} \quad (\text{C3})$$

### The two orthogonalities for a quadratic dissipation function

If the dissipation functions are quadratic, the Hilbert orthogonality condition in Eq. (31) can be represented as

$$\begin{aligned} \hat{\Psi}^*(f_S + f_A) &= \hat{\Psi}^*(f_S - f_A) \\ &\iff \langle M^{-1}(f_S + f_A), (f_S + f_A) \rangle \\ &= \langle M^{-1}(f_S - f_A), (f_S - f_A) \rangle \\ &\iff \langle M^{-1}f_S, f_A \rangle + \langle M^{-1}f_A, f_S \rangle = 0 \\ &\iff \langle j_S, f_A \rangle + \langle j_A, f_S \rangle = 0. \end{aligned} \quad (\text{C4})$$

Thus,  $\langle j_S, f_A \rangle = \langle j_A, f_S \rangle = 0$ .

The information geometric orthogonality does not require modification because it is already defined by the bilinear form  $\langle j', f'' \rangle = 0$ . If the dissipation functions are quadratic, the Bregman divergence leads to

$$\begin{aligned} \mathcal{D}[j; f'] &= \frac{1}{2} \langle j, Mj \rangle + \frac{1}{2} \langle M^{-1}f', f' \rangle - \langle j, f' \rangle \\ &= \frac{1}{2} \|j\|_M^2 + \frac{1}{2} \|j'\|_M^2 - \langle j, f' \rangle = \frac{1}{2} \|j - j'\|_M^2. \end{aligned} \quad (\text{C5})$$

Taking into account the fact that  $\mathcal{D}[j||j'] = \mathcal{D}[j; f']$ , the above result illustrates that  $\mathcal{D}[j||j']$  is the square norm between  $j$  and  $j'$  if  $\Psi$  is quadratic. For  $j = j' + j''$ , the GPT reduces to the usual form of Pythagorean theorem,

$$\begin{aligned} \mathcal{D}[j||\mathbf{0}] &= \mathcal{D}[j||j''] + \mathcal{D}[j''||\mathbf{0}]. \\ \iff \|j - \mathbf{0}\|_M^2 &= \|j - j''\|_M^2 + \|j'' - \mathbf{0}\|_M^2. \end{aligned} \quad (\text{C6})$$

These computations illustrate that the generalized orthogonalities are actually generalized versions of the conventional ones.

## APPENDIX D: PROOF OF MAIN RESULTS

We provide the proof of the main results.

### 1. Proof of positivity of decomposition by Hilbert orthogonality

We prove Eq. (33) for  $f_S$  and  $f_A$  satisfying the Hilbert orthogonality:  $\Psi^*(f_S + f_A) = \Psi^*(f_S - f_A)$ . For  $f = f_S + f_A$ ,  $f_{\text{iso}} = j_S - j_A$ , and their Legendre transforms  $j$  and  $j_{\text{iso}}$ , we have

$$\begin{aligned} \mathcal{D}[j||j_{\text{iso}}] &= \mathcal{D}[j; f_{\text{iso}}] = \Psi(j) + \Psi^*(f_{\text{iso}}) - \langle j, f_{\text{iso}} \rangle \\ &= \Psi(j) + \Psi^*(f) - \langle j, f_{\text{iso}} \rangle \\ &= \langle j, f \rangle - \langle j, f_{\text{iso}} \rangle \\ &= \langle j, f - f_{\text{iso}} \rangle = 2\langle j, f_A \rangle, \end{aligned} \quad (\text{D1})$$

where we used  $\Psi^*(f_{\text{iso}}) = \Psi^*(f)$  and the Legendre identity [Eq. (15)]. Similarly,

$$\begin{aligned} \mathcal{D}[j||-j_{\text{iso}}] &= \mathcal{D}[j; -f_{\text{iso}}] \\ &= \Psi(j) + \Psi^*(-f_{\text{iso}}) - \langle j, -f_{\text{iso}} \rangle \\ &= \Psi(j) + \Psi^*(f) + \langle j, f_{\text{iso}} \rangle \\ &= \langle j, f \rangle + \langle j, f_{\text{iso}} \rangle \\ &= \langle j, f + f_{\text{iso}} \rangle = 2\langle j, f_S \rangle, \end{aligned} \quad (\text{D2})$$

where we used  $\Psi^*(-f_{\text{iso}}) = \Psi^*(f_{\text{iso}}) = \Psi^*(f)$ .

### 2. Proof of Hilbert orthogonality for complex-balanced state

For the CB state, we prove the orthogonality:  $\Psi_x^*(f_S(x) + f_A) = \Psi_x^*(f_S(x) - f_A)$  for all  $x \in \mathcal{X}$ . It should be noted that this orthogonality was proven to hold for the CB condition in the context of large deviation theory [36]. We describe its details with our notation here for clarity.

The orthogonality condition is equivalent to

$$\Psi_x^*(f_0(x)) = \Psi_x^*(f_0(x) - 2f_S(x)) \quad (\text{D3})$$

where we used  $f_0(\mathbf{x}) = f_S(\mathbf{x}) + f_A$  and the symmetry  $\Psi_x^*(\mathbf{f}) = \Psi_x^*(-\mathbf{f})$ . This can be converted to

$$\langle \mathbf{j}_0^+(\mathbf{x}), e^{-2f_S(\mathbf{x})} - \mathbf{1} \rangle + \langle \mathbf{j}_0^-(\mathbf{x}), e^{2f_S(\mathbf{x})} - \mathbf{1} \rangle = 0. \quad (\text{D4})$$

Inserting  $2f_S(\mathbf{x}) = -S^T \ln \frac{\mathbf{x}}{\mathbf{x}_{cb}}$ , we have

$$\left\langle \mathbf{j}^+(\mathbf{x}), \left( \frac{\mathbf{x}}{\mathbf{x}_{cb}} \right)^{S^T} - \mathbf{1} \right\rangle + \left\langle \mathbf{j}^-(\mathbf{x}), \left( \frac{\mathbf{x}}{\mathbf{x}_{cb}} \right)^{-S^T} - \mathbf{1} \right\rangle = 0 \quad (\text{D5})$$

By computing the left-hand side in a componentwise manner, we have

$$\begin{aligned} & \sum_{e=1}^{N_e} j_e^+(\mathbf{x}) \left[ \left( \frac{\mathbf{x}}{\mathbf{x}_{cb}} \right)^{-(\Gamma b_e)^T} - 1 \right] \\ &= \sum_{e=1}^{N_e} k_e^+ \mathbf{x}_{cb}^{\gamma_e^+} \left[ \left( \frac{\mathbf{x}}{\mathbf{x}_{cb}} \right)^{\gamma_e^-} - \left( \frac{\mathbf{x}}{\mathbf{x}_{cb}} \right)^{\gamma_e^+} \right], \end{aligned} \quad (\text{D6})$$

$$\begin{aligned} & \sum_{e=1}^{N_e} j_e^-(\mathbf{x}) \left[ \left( \frac{\mathbf{x}}{\mathbf{x}_{cb}} \right)^{(\Gamma b_e)^T} - 1 \right] \\ &= \sum_{e=1}^{N_e} k_e^- \mathbf{x}_{cb}^{\gamma_e^-} \left[ \left( \frac{\mathbf{x}}{\mathbf{x}_{cb}} \right)^{\gamma_e^+} - \left( \frac{\mathbf{x}}{\mathbf{x}_{cb}} \right)^{\gamma_e^-} \right]. \end{aligned} \quad (\text{D7})$$

Using these expressions, we can evaluate the left-hand side as

$$\begin{aligned} & \left\langle \mathbf{j}^+(\mathbf{x}), \left( \frac{\mathbf{x}}{\mathbf{x}_{cb}} \right)^{S^T} - \mathbf{1} \right\rangle + \left\langle \mathbf{j}^-(\mathbf{x}), \left( \frac{\mathbf{x}}{\mathbf{x}_{cb}} \right)^{-S^T} - \mathbf{1} \right\rangle \\ &= \sum_{e=1}^{N_e} \left( k_e^+ \mathbf{x}_{cb}^{\gamma_e^+} - k_e^- \mathbf{x}_{cb}^{\gamma_e^-} \right) \left[ \left( \frac{\mathbf{x}}{\mathbf{x}_{cb}} \right)^{\gamma_e^-} - \left( \frac{\mathbf{x}}{\mathbf{x}_{cb}} \right)^{\gamma_e^+} \right] \\ &= \left\langle (\mathbf{j}^+(\mathbf{x}_{cb}) - \mathbf{j}^-(\mathbf{x}_{cb})), \left[ \left( \frac{\mathbf{x}}{\mathbf{x}_{cb}} \right)^{(\Gamma B^-)^T} - \left( \frac{\mathbf{x}}{\mathbf{x}_{cb}} \right)^{(\Gamma B^+)^T} \right] \right\rangle \\ &= \left\langle \mathbf{j}(\mathbf{x}_{cb}), B_-^T \left( \frac{\mathbf{x}}{\mathbf{x}_{cb}} \right)^{\Gamma^T} - B_+^T \left( \frac{\mathbf{x}}{\mathbf{x}_{cb}} \right)^{\Gamma^T} \right\rangle \\ &= \left\langle B\mathbf{j}(\mathbf{x}_{cb}), \left( \frac{\mathbf{x}}{\mathbf{x}_{cb}} \right)^{\Gamma^T} \right\rangle. \end{aligned} \quad (\text{D8})$$

Thus, if  $\mathbf{x}_{cb}$  satisfies the CB condition  $B\mathbf{j}(\mathbf{x}_{cb}) = \mathbf{0}$ , then  $\Psi_x^*(f_S(\mathbf{x}) + f_A) = \Psi_x^*(f_S(\mathbf{x}) - f_A)$  holds for all  $\mathbf{x}$ .

### 3. Uniqueness of intersections

Here we outline the proof of the uniqueness of  $\mathbf{j}_{\text{int}}$  in Eq. (59) as the intersection of  $\mathcal{P}^v(\mathbf{j}')$  and  $\mathcal{M}_x^f(\mathbf{f}'')$ . To this end, we first define  $\mathbf{j}_{\text{int}}$  variationally, and then show that  $\mathbf{j}_{\text{int}}$  defined in this way is exactly the intersection.

Bregman divergence  $\mathcal{D}_x[\mathbf{j}||\mathbf{j}'']$  as a function of  $\mathbf{j}$  is strictly convex and lower bounded by the minimum 0. We define  $\mathbf{j}_{\text{int}}$  by the following variational problem with a linear constraint:

$$\mathbf{j}_{\text{int}} := \arg \min_{\mathbf{j} \in \mathcal{P}^v(\mathbf{j}')} \mathcal{D}_x[\mathbf{j}||\mathbf{j}'']. \quad (\text{D9})$$

The minimizer  $\mathbf{j}_{\text{int}}$  is unique from the property of convex optimization. By definition,  $\mathbf{j}_{\text{int}} \in \mathcal{P}^v(\mathbf{j}')$ . Because  $\mathbf{j}_{\text{int}}$  minimizes  $\mathcal{D}_x[\mathbf{j}||\mathbf{j}'']$  under the constraint  $\mathcal{P}^v(\mathbf{j}') = \mathbf{j}' + \text{Ker}S$ , the normal vector of the level hypersurface of the divergence  $\mathcal{S}_{\text{int}} := \{\mathbf{j} | \mathcal{D}_x[\mathbf{j}_{\text{int}}||\mathbf{j}''] = \mathcal{D}_x[\mathbf{j}||\mathbf{j}'']\}$  evaluated at  $\mathbf{j}_{\text{int}}$  is

$$\begin{aligned} \partial_{\mathbf{j}} \mathcal{D}_x[\mathbf{j}||\mathbf{j}'']|_{\mathbf{j}=\mathbf{j}_{\text{int}}} &= \partial_{\mathbf{j}} [\Psi_x(\mathbf{j}) + \Psi_x^*(\mathbf{f}'') - \langle \mathbf{j}, \mathbf{f}'' \rangle]|_{\mathbf{j}=\mathbf{j}_{\text{int}}} \\ &= \mathbf{f}_{\text{int}} - \mathbf{f}''. \end{aligned} \quad (\text{D10})$$

The normal vector should be orthogonal to  $\text{Ker}S$ . This means that  $\mathbf{f}_{\text{int}} - \mathbf{f}'' \in \text{Im}S^T = \mathcal{P}^f(\mathbf{0})$ . Therefore,  $\mathbf{f}_{\text{int}} \in \mathcal{P}^f(\mathbf{f}'')$ . By the Legendre transform, we have  $\mathbf{j}_{\text{int}} \in \mathcal{M}_x^f(\mathbf{f}'')$ . Now, we obtain  $\mathbf{j}_{\text{int}} \in \mathcal{P}^v(\mathbf{j}') \cap \mathcal{M}_x^f(\mathbf{f}'')$ .

To show the uniqueness, suppose that there exists another vector  $\mathbf{j}_{\text{int}}^\dagger$  such that  $\mathbf{j}_{\text{int}}^\dagger \in \mathcal{P}^v(\mathbf{j}') \cap \mathcal{M}_x^f(\mathbf{f}'')$ . Using the same argument as  $\mathbf{j}_{\text{int}}$ ,  $\mathbf{j}_{\text{int}}^\dagger$  satisfies the same variational characterization as Eq. (D9). Because of the uniqueness of the solution of the variational problem,  $\mathbf{j}_{\text{int}}^\dagger = \mathbf{j}_{\text{int}}$ .

### 4. Proof of variational characterizations

From Eq. (62) and Eq. (67), the variational characterizations of  $\mathbf{f}_{\text{st}}$  and  $\mathbf{j}_{\text{eq}}$  are obtained as

$$\begin{aligned} & \arg \min_{\mathbf{f} \in \mathcal{P}^f(f_0(\mathbf{x}))} \Psi_x^*(\mathbf{f}) \\ &= \arg \min_{\mathbf{f} \in \mathcal{P}^f(f_0(\mathbf{x}))} [\mathcal{D}_x[\mathbf{j}_{\text{st}}(\mathbf{x}); \mathbf{f}] + \Psi_x^*(\mathbf{f}_{\text{st}}(\mathbf{x}))] \\ &= \arg \min_{\mathbf{f} \in \mathcal{P}^f(f_0(\mathbf{x}))} \mathcal{D}_x[\mathbf{j}_{\text{st}}(\mathbf{x}); \mathbf{f}] = \mathbf{f}_{\text{st}}(\mathbf{x}), \end{aligned} \quad (\text{D11})$$

$$\begin{aligned} & \arg \min_{\mathbf{j} \in \mathcal{P}^v(j_0(\mathbf{x}))} \Psi_x(\mathbf{j}) \\ &= \arg \min_{\mathbf{j} \in \mathcal{P}^v(j_0(\mathbf{x}))} [\mathcal{D}_x[\mathbf{j}; \mathbf{f}_{\text{eq}}(\mathbf{x})] + \Psi_x(\mathbf{j}_{\text{eq}}(\mathbf{x}))] \\ &= \arg \min_{\mathbf{j} \in \mathcal{P}^v(j_0(\mathbf{x}))} \mathcal{D}_x[\mathbf{j}; \mathbf{f}_{\text{eq}}(\mathbf{x})] = \mathbf{j}_{\text{eq}}(\mathbf{x}). \end{aligned} \quad (\text{D12})$$

### 5. Proof of EPR decomposition

From Eq. (62) and Eq. (67), the EPR decomposition is obtained as

$$\begin{aligned} \dot{\Sigma}(\mathbf{x})/2 &= \Psi_x(\mathbf{j}_0(\mathbf{x})) + \Psi_x^*(\mathbf{f}_0(\mathbf{x})) \\ &= [\mathcal{D}_x[\mathbf{j}_0(\mathbf{x}); \mathbf{f}_{\text{eq}}(\mathbf{x})] + \Psi_x(\mathbf{j}_{\text{eq}}(\mathbf{x}))] \\ &\quad + [\mathcal{D}_x[\mathbf{j}_{\text{st}}(\mathbf{x}); \mathbf{f}_0(\mathbf{x})] + \Psi_x^*(\mathbf{f}_{\text{st}}(\mathbf{x}))], \\ &= [\mathcal{D}_x[\mathbf{j}_0(\mathbf{x}); \mathbf{f}_{\text{eq}}(\mathbf{x})] + \Psi_x^*(\mathbf{f}_{\text{st}}(\mathbf{x}))] \\ &\quad + [\mathcal{D}_x[\mathbf{j}_{\text{st}}(\mathbf{x}); \mathbf{f}_0(\mathbf{x})] + \Psi_x(\mathbf{j}_{\text{eq}}(\mathbf{x}))] \\ &= \dot{\Sigma}_{\text{hk}}^{\text{MD}}(\mathbf{x})/2 + \dot{\Sigma}_{\text{ex}}^{\text{MD}}(\mathbf{x})/2, \end{aligned} \quad (\text{D13})$$

where the positions of  $\Psi_x(\mathbf{j}_{\text{eq}}(\mathbf{x}))$  and  $\Psi_x^*(\mathbf{f}_{\text{st}}(\mathbf{x}))$  are swapped from the second to the third line.

- [1] H. B. Callen and H. B. Callen, *Thermodynamics and an Introduction to Thermostatistics* (Wiley, Hoboken, NJ, 1985).  
[2] Y. Sughiyama, D. Loutchko, A. Kamimura, and T. J. Kobayashi, Hessian geometric structure of chemical thermodynamic

- systems with stoichiometric constraints, *Phys. Rev. Res.* **4**, 033065 (2022).  
[3] L. Onsager, Reciprocal relations in irreversible processes. I., *Phys. Rev.* **37**, 405 (1931).

- [4] L. Onsager, Reciprocal relations in irreversible processes. II., *Phys. Rev.* **38**, 2265 (1931).
- [5] L. Onsager and S. Machlup, Fluctuations and irreversible processes, *Phys. Rev.* **91**, 1505 (1953).
- [6] S. Machlup and L. Onsager, Fluctuations and irreversible process. II. Systems with kinetic energy, *Phys. Rev.* **91**, 1512 (1953).
- [7] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim, Macroscopic fluctuation theory, *Rev. Mod. Phys.* **87**, 593 (2015).
- [8] T. Taniguchi and E. G. D. Cohen, Onsager-machlup theory for nonequilibrium steady states and fluctuation theorems, *J Stat Phys* **126**, 1 (2007).
- [9] T. Schmiedl and U. Seifert, Stochastic thermodynamics of chemical reaction networks, *J. Chem. Phys.* **126**, 044101 (2007).
- [10] A. Mielke, M. A. Peletier, and D. R. M. Renger, On the relation between gradient flows and the large-deviation principle, with applications to Markov chains and diffusion, *Potential Anal.* **41**, 1293 (2014).
- [11] M. A. Peletier and A. Schlichting, Cosh gradient systems and tilting [arXiv:2203.05435](https://arxiv.org/abs/2203.05435).
- [12] E. Aurell, C. Mejía-Monasterio, and P. Muratore-Ginanneschi, Optimal Protocols and Optimal Transport in Stochastic Thermodynamics, *Phys. Rev. Lett.* **106**, 250601 (2011).
- [13] M. Nakazato and S. Ito, Geometrical aspects of entropy production in stochastic thermodynamics based on Wasserstein distance, *Phys. Rev. Res.* **3**, 043093 (2021).
- [14] R. Jordan, D. Kinderlehrer, and F. Otto, Free energy and the Fokker-Planck equation, *Physica D* **107**, 265 (1997).
- [15] J.-D. Benamou and Y. Brenier, A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem, *Numer. Math.* **84**, 375 (2000).
- [16] L. Ambrosio, N. Gigli, and G. Savare, *Gradient Flows: In Metric Spaces and in the Space of Probability Measures* (Springer Science & Business Media, New York, 2006).
- [17] H. Shima, *The Geometry of Hessian Structures* (World Scientific, Singapore, 2007).
- [18] S.-i. Amari, *Information Geometry and Its Applications* (Springer, New York, 2016).
- [19] F. Nielsen, An elementary introduction to information geometry, *Entropy* **22**, 1100 (2020).
- [20] T. L. Hill, *Free Energy Transduction and Biochemical Cycle Kinetics* (Courier Corporation, New York, 2005).
- [21] J. Schnakenberg, Network theory of microscopic and macroscopic behavior of master equation systems, *Rev. Mod. Phys.* **48**, 571 (1976).
- [22] 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).
- [23] R. Rao and M. Esposito, Nonequilibrium Thermodynamics of Chemical Reaction Networks: Wisdom from Stochastic Thermodynamics, *Phys. Rev. X* **6**, 041064 (2016).
- [24] D. A. Beard and H. Qian, *Chemical Biophysics: Quantitative Analysis of Cellular Systems*, Cambridge Texts in Biomedical Engineering (Cambridge University Press, Cambridge, 2008).
- [25] A. S. Mikhailov and G. Ertl, *Chemical Complexity: Self-Organization Processes in Molecular Systems* (Springer, New York, 2017).
- [26] U. Alon, *An Introduction to Systems Biology: Design Principles of Biological Circuits* (CRC Press, Boca Raton, FL, 2019).
- [27] A. Mielke, R. I. A. Patterson, M. A. Peletier, and D. R. Michiel Renger, Non-equilibrium thermodynamical principles for chemical reactions with mass-action kinetics, *SIAM J. Appl. Math.* **77**, 1562 (2017).
- [28] A. Dechant, S.-i. Sasa, and S. Ito, Geometric decomposition of entropy production in out-of-equilibrium systems, *Phys. Rev. Res.* **4**, L012034 (2022).
- [29] L. J. Grady and J. R. Polimeni, *Discrete Calculus: Applied Analysis on Graphs for Computational Science* (Springer Science & Business Media, New York, 2010).
- [30] J. W. Gibbs, On the equilibrium of heterogeneous substances, *Am. J. Sci.* **s3-16**, 441 (1878).
- [31] M. Feinberg, *Foundations of Chemical Reaction Network Theory* (Springer, New York, 2019).
- [32] S. Klamt, U.-U. Haus, and F. Theis, Hypergraphs and cellular networks, *PLoS Comput. Biol.* **5**, e1000385 (2009).
- [33] C. Maes, Frenetic Bounds on the Entropy Production, *Phys. Rev. Lett.* **119**, 160601 (2017).
- [34] M. Kaiser, R. L. Jack, and J. Zimmer, Canonical structure and orthogonality of forces and currents in irreversible Markov chains, *J Stat Phys* **170**, 1019 (2018).
- [35] D. R. M. Renger, Gradient and generic systems in the space of fluxes, applied to reacting particle systems, *Entropy* **20**, 596 (2018).
- [36] D. R. M. Renger and J. Zimmer, Orthogonality of fluxes in general nonlinear reaction networks, *Discrete Contin. Dyn. Syst. - S* **14**, 205 (2021).
- [37] M. A. Peletier, R. Rossi, G. Savaré, and O. Tse, Jump processes as generalized gradient flows, *Calc. Var.* **61**, 33 (2022).
- [38] R. I. A. Patterson, D. R. M. Renger, and U. Sharma, Variational structures beyond gradient flows: A macroscopic fluctuation-theory perspective, [arXiv:2103.14384](https://arxiv.org/abs/2103.14384).
- [39] C. Maes, Local detailed balance, *SciPost Phys. Lect. Notes*, **32** (2021).
- [40] C. Maes, *Non-Dissipative Effects in Nonequilibrium Systems* (Springer, New York, 2017).
- [41] T. J. Kobayashi, D. Loutchko, A. Kamimura, and Y. Sughiyama, Kinetic derivation of the Hessian geometric structure in chemical reaction networks, *Phys. Rev. Res.* **4**, 033066 (2022).
- [42] Y. Sughiyama, A. Kamimura, D. Loutchko, and T. J. Kobayashi, Chemical thermodynamics for growing systems, *Phys. Rev. Res.* **4**, 033191 (2022).
- [43] D. R. M. Renger, Anisothermal Chemical Reactions: Onsager-Machlup and macroscopic fluctuation theory, [arXiv:2111.12164](https://arxiv.org/abs/2111.12164).
- [44] K. Nomizu, *Affine Differential Geometry: Geometry of Affine Immersions* (Cambridge University Press, Cambridge, 1994).
- [45] S. Ito, Stochastic Thermodynamic Interpretation of Information Geometry, *Phys. Rev. Lett.* **121**, 030605 (2018).
- [46] A. Kolchinsky and D. H. Wolpert, Work, Entropy Production, and Thermodynamics of Information under Protocol Constraints, *Phys. Rev. X* **11**, 041024 (2021).
- [47] N. Ohga and S. Ito, Information-geometric Legendre duality in stochastic thermodynamics, [arXiv:2112.11008](https://arxiv.org/abs/2112.11008).
- [48] G. Craciun, A. Dickenstein, A. Shiu, and B. Sturmfels, Toric dynamical systems, *J. Symb. Comput.* **44**, 1551 (2009).
- [49] T. Hatano and S.-i. Sasa, Steady-State Thermodynamics of Langevin Systems, *Phys. Rev. Lett.* **86**, 3463 (2001).

- [50] C. Maes and K. Netočný, A nonequilibrium extension of the Clausius heat theorem, *J Stat Phys* **154**, 188 (2014).
- [51] T. S. Komatsu, N. Nakagawa, S.-i. Sasa, and H. Tasaki, Steady-State Thermodynamics for Heat Conduction: Microscopic Derivation, *Phys. Rev. Lett.* **100**, 230602 (2008).
- [52] H. Bhatia, G. Norgard, V. Pascucci, and P.-T. Bremer, The helmholtz-hodge decomposition—A survey, *IEEE Trans. Vis. Comput. Graph.* **19**, 1386 (2013).
- [53] F. Horn and J. S. Rowlinson, Stability and complex balancing in mass-action systems with three short complexes, *Proc. R. Soc. Lond. Math. Phys. Sci.* **334**, 331 (1973).
- [54] M. Pérez Millán, A. Dickenstein, A. Shiu, and C. Conradi, Chemical reaction systems with toric steady states, *Bull. Math. Biol.* **74**, 1027 (2012).
- [55] N. Shiraishi, Optimal thermodynamic uncertainty relation in Markov jump processes, *J. Stat. Phys.* **185**, 19 (2021).
- [56] T. R. Gingrich, J. M. Horowitz, N. Perunov, and J. L. England, Dissipation Bounds All Steady-State Current Fluctuations, *Phys. Rev. Lett.* **116**, 120601 (2016).
- [57] A. C. Barato and U. Seifert, Thermodynamic Uncertainty Relation for Biomolecular Processes, *Phys. Rev. Lett.* **114**, 158101 (2015).
- [58] K. Yoshimura and S. Ito, Thermodynamic Uncertainty Relation and Thermodynamic Speed Limit in Deterministic Chemical Reaction Networks, *Phys. Rev. Lett.* **127**, 160601 (2021).
- [59] G. Oster, A. Perelson, and A. Katchalsky, Network thermodynamics, *Nature (London)* **234**, 393 (1971).
- [60] J. Schnakenberg, *Thermodynamic Network Analysis of Biological Systems* (Springer Science & Business Media, New York, 2012).
- [61] 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).
- [62] V. Belevitch, Summary of the history of circuit theory, *Proc. IRE* **50**, 848 (1962).
- [63] W.-K. Chen, *Applied Graph Theory* (Elsevier, Amsterdam, 2012).
- [64] A. Wachtel, J. Vollmer, and B. Altaner, Fluctuating currents in stochastic thermodynamics. I. Gauge invariance of asymptotic statistics, *Phys. Rev. E* **92**, 042132 (2015).
- [65] B. Altaner, S. Grosskinsky, S. Herminghaus, L. Katthän, M. Timme, and J. Vollmer, Network representations of nonequilibrium steady states: Cycle decompositions, symmetries, and dominant paths, *Phys. Rev. E* **85**, 041133 (2012).
- [66] M. Poletini, System/Environment Duality of Nonequilibrium Network Observables, in *Mathematical Technology of Networks*, Springer Proceedings in Mathematics & Statistics, edited by D. Mugnolo (Springer International Publishing, Cham, 2015) pp. 191–205.
- [67] A. Mielke, A gradient structure for reaction–diffusion systems and for energy-drift-diffusion systems, *Nonlinearity* **24**, 1329 (2011).
- [68] S.-N. Chow, W. Huang, Y. Li, and H. Zhou, Fokker–Planck equations for a free energy functional or Markov process on a graph, *Arch. Ration. Mech. Anal.* **203**, 969 (2012).
- [69] J. Maas, Gradient flows of the entropy for finite Markov chains, *J. Funct. Anal.* **261**, 2250 (2011).
- [70] A. Mielke, Geodesic convexity of the relative entropy in reversible Markov chains, *Calc. Var.* **48**, 1 (2013).
- [71] T. Van Vu and Y. Hasegawa, Geometrical Bounds of the Irreversibility in Markovian Systems, *Phys. Rev. Lett.* **126**, 010601 (2021).
- [72] K. Yoshimura, A. Kolchinsky, A. Dechant, and S. Ito, House-keeping and excess entropy production for general nonlinear dynamics [arXiv:2205.15227](https://arxiv.org/abs/2205.15227).
- [73] T. Van Vu and K. Saito, Thermodynamic unification of optimal transport: Thermodynamic uncertainty relation, minimum dissipation, and thermodynamic speed limits, [arXiv:2206.02684](https://arxiv.org/abs/2206.02684).
- [74] M. Suzuki, Irreversibility and entropy production in transport phenomena, II: Statistical–mechanical theory on steady states including thermal disturbance and energy supply, *Physica A* **391**, 1074 (2012).
- [75] I. Prigogine, *Introduction to Thermodynamics of Irreversible Processes* (Wiley, Hoboken, NJ, 1968).
- [76] M. J. Klein and P. H. E. Meijer, Principle of minimum entropy production, *Phys. Rev.* **96**, 250 (1954).
- [77] M. Poletini, Macroscopic constraints for the minimum entropy production principle, *Phys. Rev. E* **84**, 051117 (2011).
- [78] P. Županović, D. Juretić, and S. Botrić, Kirchhoff’s loop law and the maximum entropy production principle, *Phys. Rev. E* **70**, 056108 (2004).
- [79] L. M. Martyushev, Maximum entropy production principle: History and current status, *Phys.-Usp.* **64**, 558 (2021).
- [80] I. Gyarmati, *Non-Equilibrium Thermodynamics: Field Theory and Variational Principles* (Springer Science & Business Media, New York, 2013).
- [81] M. Poletini, Fact-checking Ziegler’s maximum entropy production principle beyond the linear regime and towards steady states, *Entropy* **15**, 2570 (2013).