HOKKAIDO UNIVERSITY

Title	Information geometric bound on general chemical reaction networks
Author(s)	Mizohata, Tsuyoshi; Kobayashi, Tetsuya J.; Bouchard, Louis-S. et al.
Citation	Physical Review E, 109(4), 044308 https://doi.org/10.1103/PhysRevE.109.044308
Issue Date	2024-04-11
Doc URL	https://hdl.handle.net/2115/92617
Rights	Copyright (2024) by The American Physical Society.
Type	journal article
File Information	PhysRevE.109.044308.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

Information geometric bound on general chemical reaction networks

Tsuyoshi Mizohata,[1,*] Tetsuya J. Kobayashi,[2,†] Louis-S. Bouchard,[3,4,5,‡] and Hideyuki Miyahara[1,§]
[1]*Graduate School of Information Science and Technology, Hokkaido University, Sapporo, Hokkaido 060-0814, Japan*
[2]*Institute of Industrial Science, The University of Tokyo, 4-6-1, Komaba, Meguro-ku, Tokyo 153-8505 Japan*
[3]*Center for Quantum Science and Engineering, University of California, Los Angeles, California 90095, USA*
[4]*Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90095, USA*
[5]*California NanoSystems Institute, University of California, Los Angeles, California 90095, USA*

(Received 19 September 2023; revised 24 January 2024; accepted 21 March 2024; published 11 April 2024)

We investigate the convergence of chemical reaction networks (CRNs), aiming to establish an upper bound on their reaction rates. The nonlinear characteristics and discrete composition of CRNs pose significant challenges in this endeavor. To circumvent these complexities, we adopt an information geometric perspective, utilizing the natural gradient to formulate a nonlinear system. This system effectively determines an upper bound for the dynamics of CRNs. We corroborate our methodology through numerical simulations, which reveal that our constructed system converges more rapidly than CRNs within a particular class of reactions. This class is defined by the count of chemicals, the highest stoichiometric coefficients in the reactions, and the total number of reactions involved. Further, we juxtapose our approach with traditional methods, illustrating that the latter falls short in providing an upper bound for CRN reaction rates. Although our investigation centers on CRNs, the widespread presence of hypergraphs across various disciplines, ranging from natural sciences to engineering, indicates potential wider applications of our method, including in the realm of information science.

DOI: 10.1103/PhysRevE.109.044308

I. INTRODUCTION

Over the past three decades, extensive research has been dedicated to understanding stochastic and information thermodynamics, specifically focusing on bounds related to entropy production and various physical quantities [1–4]. This trajectory persists, with newer studies shedding light on thermodynamic uncertainty relations [5–9] and establishing thermodynamic bounds on cross correlations [10–12]. Parallel to the work on physical systems, researchers have also explored bounds in nonphysical realms such as biological systems. For example, limits concerning population growth have been studied [13–18].

Recent studies have unveiled the geometric structure of chemical reaction networks (CRNs) and have also extended these concepts to the domain of general hypergraphs [19–21]. Concurrently, topological analyses on CRNs and hypergraphs have been performed [22–25]. Despite these advancements, the intrinsic nonlinearity in CRNs presents a significant challenge for elucidating specific properties, leaving gaps in our understanding.

Information geometry offers a framework that applies differential geometry to probability distributions, facilitating the exploration of their geometric structures [26,27]. Among its significant contributions is the concept of the natural gradient (NG) [28], which has demonstrated effectiveness in optimization problems, particularly in the realm of machine learning. Additional studies have ventured into the acceleration of information gradient flows [29] and have investigated the biological significance of gradient flows [30]. Research has also extended to constraints involving rates of statistical divergences and mutual information [31]. Furthermore, there are some information-geometric studies on CRNs [32–35]. These diverse applications underline the versatility of information geometry, which we leverage in this paper.

In the present paper, we explore the upper bound on reaction rates in general CRNs using NG. Initially, we present a geometrical description of CRN dynamics [20,21]. Subsequently, we categorize CRNs according to the number of chemicals involved, the maximum coefficients in the reactions, and the total number of reactions. Utilizing this classification, we formulate a nonlinear system that provides an upper bound on reaction rates for a given class of CRNs. Through numerical simulations, we find that the constructed system exhibits a steeper gradient, facilitating faster convergence. Importantly, this fast convergence minimizes the Kullback-Leibler (KL) divergence to zero. In contrast, conventional CRNs often maintain a nonzero KL divergence due to nontrivial equilibrium points. We also note that conventional methods are insufficient for achieving these results, underscoring the uniqueness of our approach.

The remainder of this paper is structured as follows. In Sec. II, we furnish an overview of CRNs. Section III elucidates the challenges of establishing an upper bound on CRNs using Newton's method. Section IV is dedicated to

*mizohata.tsuyoshi.t8@elms.hokudai.ac.jp
†tetsuya@sat.t.u-tokyo.ac.jp
‡louis.bouchard@gmail.com
§Corresponding author: miyahara@ist.hokudai.ac.jp, hmiyahara512@gmail.com

 ©2024 American Physical Society

explaining NG. In Sec. V, we introduce a dynamical system that serves as an upper bound for CRNs in a specified class. Numerical simulations are presented in Sec. VI. The paper concludes with Sec. VII.

II. CHEMICAL REACTION NETWORKS

In this section, the primary aim is to formulate the geometric representation of the dynamical equations governing CRNs, as delineated in Refs. [20,21]. We commence by presenting the standard notation for hypergraphs and CRNs. Subsequently, we elucidate the dynamics intrinsic to CRNs, as well as concepts of Legendre duality and detailed balance. These elements are then combined to construct the geometric expression of CRN dynamics.

A. Definition of CRNs

We begin with a hypergraph $(\mathbb{V}, \mathbb{E})$, where $\mathbb{V} \coloneqq \{\mathbb{v}_i\}_{i=1}^{N_\mathbb{v}}$ and $\mathbb{E} \coloneqq \{\mathbb{e}_e\}_{e=1}^{N_\mathbb{e}}$ as a hypergraph provides a mathematical framework to describe a chemical reaction. Suppose that a CRN of interest involves $N_\mathbb{X}$ chemicals, denoted as $\mathbb{X}_1, \mathbb{X}_2, \ldots, \mathbb{X}_{N_\mathbb{X}}$. In the case of a CRN, each hypervertex $\mathbb{v}_i$ is composed of a combination of chemicals $\mathbb{X}_1, \mathbb{X}_2, \ldots, \mathbb{X}_{N_\mathbb{X}}$ and given by

$$\mathbb{v}_i \coloneqq \gamma_{1,i}\mathbb{X}_1 + \gamma_{2,i}\mathbb{X}_2 + \cdots + \gamma_{N_\mathbb{X},i}\mathbb{X}_{N_\mathbb{X}}. \tag{2.1}$$

Each hyperedge $\mathbb{e}_e$ corresponds to a chemical reaction and is defined by a directed pair of two hypervertices $\mathbb{e}_e \coloneqq (\mathbb{v}_e^+, \mathbb{v}_e^-)$, which can be expressed as

$$\begin{aligned}&\alpha_{1,e}\mathbb{X}_1 + \alpha_{2,e}\mathbb{X}_2 + \cdots + \alpha_{N_\mathbb{X},e}\mathbb{X}_{N_\mathbb{X}}\\ &\quad\xrightarrow{\mathbb{e}_e} \beta_{1,e}\mathbb{X}_1 + \beta_{2,e}\mathbb{X}_2 + \cdots + \beta_{N_\mathbb{X},e}\mathbb{X}_{N_\mathbb{X}}.\end{aligned} \tag{2.2}$$

Here, $\mathbb{v}_e^\pm$ are chosen from $\{\mathbb{v}_i\}_{i=1}^{N_\mathbb{v}}$ and, in Eq. (2.2), $\mathbb{v}_e^+ = \alpha_{1,e}\mathbb{X}_1 + \alpha_{2,e}\mathbb{X}_2 + \cdots + \alpha_{N_\mathbb{X},e}\mathbb{X}_{N_\mathbb{X}}$ and $\mathbb{v}_e^- = \beta_{1,e}\mathbb{X}_1 + \beta_{2,e}\mathbb{X}_2 + \cdots + \beta_{N_\mathbb{X},e}\mathbb{X}_{N_\mathbb{X}}$. We also define the order of reaction as follows:

$$m \coloneqq \max_{i,e}\{\alpha_{i,e}, \beta_{i,e}\}. \tag{2.3}$$

To characterize CRNs, m in Eq. (2.3) will play an important role.

When a CRN involves multiple chemical reactions, the description provided above may be inadequate. To describe a complex CRN, the stoichiometric matrix plays a crucial role. The stoichiometric matrix S is defined as an $N_\mathbb{X} \times N_\mathbb{e}$ matrix and is given by

$$S \coloneqq \left[\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_{N_\mathbb{e}}\right], \tag{2.4}$$

where, for $e = 1, 2, \ldots, N_\mathbb{e}$,

$$\boldsymbol{s}_e \coloneqq \begin{bmatrix} \beta_{1,e} - \alpha_{1,e} \\ \beta_{2,e} - \alpha_{2,e} \\ \vdots \\ \beta_{N_\mathbb{X},e} - \alpha_{N_\mathbb{X},e} \end{bmatrix}. \tag{2.5}$$

That is, the (j, e)th element of S is given by $s_{j,e} = \beta_{j,e} - \alpha_{j,e}$ for $j = 1, 2, \ldots, N_\mathbb{X}$ and $e = 1, 2, \ldots, N_\mathbb{e}$. In general, when a CRN involves multiple chemical reactions, the stoichiometric matrix provides a concise representation of the relationships between the reactants and products.

The stoichiometric matrix S is also expressed as $S = -\Gamma B$. Here, $B \in \{1, 0, -1\}^{N_\mathbb{v} \times N_\mathbb{e}}$ is the incidence matrix whose (i, e)th element is given for $i = 1, 2, \ldots, N_\mathbb{v}$ and $e = 1, 2, \ldots, N_\mathbb{e}$ by

$$b_{i,e} \coloneqq \begin{cases} 1 & (\mathbb{v}_i \text{ is the head of hyperedge } \mathbb{e}_e\colon \mathbb{v}_i = \mathbb{v}_e^+) \\ -1 & (\mathbb{v}_i \text{ is the tail of hyperedge } \mathbb{e}_e\colon \mathbb{v}_i = \mathbb{v}_e^-) \\ 0 & (\text{otherwise}), \end{cases} \tag{2.6}$$

and $\Gamma \in \mathbb{Z}_{\geqslant 0}^{N_\mathbb{X} \times N_\mathbb{v}}$ is given by

$$\Gamma \coloneqq \left[\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \ldots, \boldsymbol{\gamma}_{N_\mathbb{v}}\right], \tag{2.7}$$

where, using $\gamma_{1,i}, \gamma_{2,i}, \ldots, \gamma_{N_\mathbb{X},i}$ in Eq. (2.1), $\boldsymbol{\gamma}_i$ is defined as

$$\boldsymbol{\gamma}_i \coloneqq \left[\gamma_{1,i}, \gamma_{2,i}, \ldots, \gamma_{N_\mathbb{X},i}\right]^\mathsf{T}, \tag{2.8}$$

for $i = 1, 2, \ldots, N_\mathbb{v}$. Having defined the necessary variables to describe CRNs, we will now derive the equation that characterizes the dynamics of CRNs in the remainder of this section.

B. Dynamics of CRNs

To analyze the dynamics of a CRN, we introduce fluxes associated with each hyperedge. Let $j_e^+(\boldsymbol{x})$ and $j_e^-(\boldsymbol{x})$ denote the currents from the head to the tail and from the tail to the head of hyperedge $\mathbb{e}_e$, respectively, where $\boldsymbol{x}$ is the chemical concentration vector. We define $\boldsymbol{j}^+(\boldsymbol{x}) \coloneqq [j_1^+(\boldsymbol{x}), j_2^+(\boldsymbol{x}), \ldots, j_{N_\mathbb{e}}^+(\boldsymbol{x})]^\mathsf{T}$ and $\boldsymbol{j}^-(\boldsymbol{x}) \coloneqq [j_1^-(\boldsymbol{x}), j_2^-(\boldsymbol{x}), \ldots, j_{N_\mathbb{e}}^-(\boldsymbol{x})]^\mathsf{T}$.

The law of mass action is widely observed to hold for CRNs and is considered one of the fundamental characteristics that differentiate CRNs from nonchemical hypergraphs. Based on this, we make the assumption of mass action kinetics for the forward and reverse reaction fluxes on hyperedge $\mathbb{e}_e$ in Eq. (2.2),

$$j_e^\pm(\boldsymbol{x}) = k_e^\pm \sum_{i=1}^{N_\mathbb{v}} b_{i,e}^\pm \prod_{j=1}^{N_\mathbb{X}} x_j^{\gamma_{j,i}}, \tag{2.9}$$

where, for $i = 1, 2, \ldots, N_\mathbb{X}$ and $e = 1, 2, \ldots, N_\mathbb{e}$,

$$b_{i,e}^+ \coloneqq \max(b_{i,e}, 0), \tag{2.10}$$

$$b_{i,e}^- \coloneqq -\min(b_{i,e}, 0), \tag{2.11}$$

and $k_e^\pm$ are the reaction rate coefficients for the forward and backward currents on $\mathbb{e}_e$. Expressed in vector notation, Eq. (2.9) can be written as

$$\boldsymbol{j}^\pm(\boldsymbol{x}) = \boldsymbol{k}^\pm \circ (B^\pm)^\mathsf{T}\boldsymbol{x}^{\Gamma^\mathsf{T}} \tag{2.12}$$

$$= \boldsymbol{k}^\pm \circ \boldsymbol{x}^{(\Gamma B^\pm)^\mathsf{T}}, \tag{2.13}$$

where

$$B^+ \coloneqq \max(B, \mathbb{0}), \tag{2.14}$$

$$B^- \coloneqq -\min(B, \mathbb{0}), \tag{2.15}$$

$$\boldsymbol{x}^{\Gamma^\mathsf{T}} \coloneqq [\boldsymbol{x}^{\boldsymbol{\gamma}_1}, \boldsymbol{x}^{\boldsymbol{\gamma}_2}, \ldots, \boldsymbol{x}^{\boldsymbol{\gamma}_{N_\mathbb{v}}}]^\mathsf{T}, \tag{2.16}$$

$$\boldsymbol{x}^{\boldsymbol{\gamma}_i} \coloneqq \prod_{j=1}^{N_{\mathbb{x}}} x_j^{\gamma_{j,i}}, \tag{2.17}$$

$$\boldsymbol{k}^{\pm} \coloneqq \left[k_1^{\pm}, k_2^{\pm}, \ldots, k_{N_{\mathbb{e}}}^{\pm}\right]^{\mathsf{T}}. \tag{2.18}$$

Here, $\mathbb{0}$ represents the zero matrix, which has the same size as matrix B. The functions $\max(\cdot,\cdot)$ and $\min(\cdot,\cdot)$ are applied elementwise, meaning that for each element $[A]_{i,j}$ and $[B]_{i,j}$, we have $[\max(A,B)]_{i,j} = \max([A]_{i,j}, [B]_{i,j})$ and $[\min(A,B)]_{i,j} = \min([A]_{i,j}, [B]_{i,j})$, respectively. The notation $[\cdot]_{i,j}$ represents the element located at the ith row and jth column. Moreover, the symbol $\circ$ denotes the elementwise product, which is defined as follows:

$$\boldsymbol{x} \circ \boldsymbol{y} \coloneqq \begin{bmatrix} x_1 y_1 \\ x_2 y_2 \\ \vdots \\ x_{N_{\mathbb{x}}} y_{N_{\mathbb{x}}} \end{bmatrix}, \tag{2.19}$$

where $\boldsymbol{x} \coloneqq [x_1, x_2, \ldots, x_{N_{\mathbb{x}}}]^{\mathsf{T}}, \boldsymbol{y} \coloneqq [y_1, y_2, \ldots, y_{N_{\mathbb{x}}}]^{\mathsf{T}}$.

The chemical concentration vector $\boldsymbol{x}_t$ at time t satisfies the chemical rate equation (CRE) given by [36–38]

$$\dot{\boldsymbol{x}}_t = S\boldsymbol{j}(\boldsymbol{x}_t), \tag{2.20}$$

where $\boldsymbol{j}(\boldsymbol{x}) \coloneqq \boldsymbol{j}^+(\boldsymbol{x}) - \boldsymbol{j}^-(\boldsymbol{x})$.

C. Legendre duality of fluxes and forces

In the realm of physics, the relationship between fluxes and forces is commonly expressed through Legendre duality, a concept that describes how forces and fluxes are dual aspects of the same system. Their product results in entropy production, denoted as $2\langle \boldsymbol{j}, \boldsymbol{f} \rangle$. In the context of chemical thermodynamics, we define the force on a hyperedge $\mathbb{e}_e$ in a manner consistent with entropy production:

$$f_e(\boldsymbol{x}) \coloneqq \frac{1}{2} \ln \frac{j_e^+(\boldsymbol{x})}{j_e^-(\boldsymbol{x})}, \tag{2.21}$$

for $e = 1, 2, \ldots, N_{\mathbb{e}}$. The corresponding vector form of Eq. (2.21), $\boldsymbol{f}(\boldsymbol{x}) \coloneqq [f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \ldots, f_{N_{\mathbb{e}}}(\boldsymbol{x})]^{\mathsf{T}}$, can be expressed as

$$\boldsymbol{f}(\boldsymbol{x}) = \frac{1}{2} \ln \frac{\boldsymbol{j}^+(\boldsymbol{x})}{\boldsymbol{j}^-(\boldsymbol{x})}, \tag{2.22}$$

where the division and the logarithmic function are computed elementwise.

We introduce a quantity called frenetic activity, particularly on hyperedge $\mathbb{e}_e$, to describe the rate of change in the state of the system $\mathbb{e}_e$ [20,21] as

$$\omega_e(\boldsymbol{x}) \coloneqq 2\sqrt{j_e^+(\boldsymbol{x}) j_e^-(\boldsymbol{x})} \tag{2.23}$$

for $e = 1, 2, \ldots, N_{\mathbb{e}}$. The vector form of Eq. (2.23), denoted as $\boldsymbol{\omega}(\boldsymbol{x}) \coloneqq [\omega_1(\boldsymbol{x}), \omega_2(\boldsymbol{x}), \ldots, \omega_{N_{\mathbb{e}}}(\boldsymbol{x})]^{\mathsf{T}}$, can be expressed as

$$\boldsymbol{\omega}(\boldsymbol{x}) = 2\sqrt{\boldsymbol{j}^+(\boldsymbol{x}) \circ \boldsymbol{j}^-(\boldsymbol{x})}. \tag{2.24}$$

Then, the following strictly convex smooth function $\Psi^*_{\boldsymbol{\omega}(\boldsymbol{x})}(\boldsymbol{f}(\boldsymbol{x}))$, which is called the dissipation function, establishes the Legendre duality between force $\boldsymbol{f}(\boldsymbol{x})$, Eq. (2.22), and flux $\boldsymbol{j}(\boldsymbol{x})$, Eq. (2.24):

$$\Psi^*_{\boldsymbol{\omega}(\boldsymbol{x})}(\boldsymbol{f}(\boldsymbol{x})) \coloneqq \boldsymbol{\omega}(\boldsymbol{x})^{\mathsf{T}}[\cosh(\boldsymbol{f}(\boldsymbol{x})) - \boldsymbol{1}], \tag{2.25}$$

where

$$\cosh(\boldsymbol{f}(\boldsymbol{x})) \coloneqq \begin{bmatrix} \cosh(f_1(\boldsymbol{x})) \\ \cosh(f_2(\boldsymbol{x})) \\ \vdots \\ \cosh(f_{N_{\mathbb{e}}}(\boldsymbol{x})) \end{bmatrix}, \tag{2.26}$$

$$\boldsymbol{f}(\boldsymbol{x}) \coloneqq \left[f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \ldots, f_{N_{\mathbb{e}}}(\boldsymbol{x})\right]^{\mathsf{T}}, \tag{2.27}$$

$$\boldsymbol{1} \coloneqq [\underbrace{1, 1, \ldots, 1}_{N_{\mathbb{e}}}]^{\mathsf{T}}. \tag{2.28}$$

As a result, we have[1]

$$\boldsymbol{j}(\boldsymbol{x}) = \partial_f \Psi^*_{\boldsymbol{\omega}(\boldsymbol{x})}(\boldsymbol{f}(\boldsymbol{x})). \tag{2.29}$$

Note that

$$\partial_f \Psi^*_{\boldsymbol{\omega}(\boldsymbol{x})}(\boldsymbol{f}(\boldsymbol{x})) = \boldsymbol{\omega}(\boldsymbol{x}) \circ \sinh(\boldsymbol{f}(\boldsymbol{x})) \tag{2.30}$$

$$= \begin{bmatrix} \omega_1(\boldsymbol{x}) \sinh(f_1(\boldsymbol{x})) \\ \omega_2(\boldsymbol{x}) \sinh(f_2(\boldsymbol{x})) \\ \vdots \\ \omega_{N_{\mathbb{e}}}(\boldsymbol{x}) \sinh(f_{N_{\mathbb{e}}}(\boldsymbol{x})) \end{bmatrix}. \tag{2.31}$$

Combining Eqs. (2.20) and (2.29), we get

$$\dot{\boldsymbol{x}}_t = S\partial_f \Psi^*_{\boldsymbol{\omega}(\boldsymbol{x}_t)}(\boldsymbol{f}(\boldsymbol{x}_t)). \tag{2.32}$$

While Eq. (2.32) is a well-defined differential equation, it lacks an explicit functional form for $\boldsymbol{f}(\boldsymbol{x})$, thus limiting its predictive capability. The functional form of $\boldsymbol{f}(\boldsymbol{x})$ based on thermodynamics and kinetics will be elaborated in the subsequent subsection.

D. Chemical reaction dynamics

Until this point, the discussion has centered on the general description of dynamics on hypergraphs. Going forward, the focus will be exclusively on CRNs. In the realm of chemical thermodynamics, it is a common assumption to employ mass action kinetics to describe reaction rates. Within this framework, a specific definition of force is accepted and widely used [20,21,36,37]:

$$\boldsymbol{f}(\boldsymbol{x}) = -\frac{1}{2}\left(S^{\mathsf{T}} \ln \boldsymbol{x} - \ln \frac{\boldsymbol{k}^+}{\boldsymbol{k}^-}\right). \tag{2.33}$$

To clarify the geometric meaning of Eq. (2.33), we introduce the Bregman divergence $\mathcal{D}_\phi(\boldsymbol{x}\|\boldsymbol{y})$ associated with potential $\phi(\cdot)$:[2]

$$\mathcal{D}_\phi(\boldsymbol{x}\|\boldsymbol{y}) \coloneqq \phi(\boldsymbol{x}) - \phi(\boldsymbol{y}) - \langle \boldsymbol{x} - \boldsymbol{y}, \partial_{\boldsymbol{x}}\phi(\boldsymbol{y}) \rangle. \tag{2.34}$$

The derivative of Eq. (2.34) is given by

$$\partial_{\boldsymbol{x}} \mathcal{D}_\phi(\boldsymbol{x}\|\boldsymbol{y}) = \partial_{\boldsymbol{x}}\phi(\boldsymbol{x}) - \partial_{\boldsymbol{x}}\phi(\boldsymbol{y}). \tag{2.35}$$

[1]We have used the following notation: $\partial_f \Psi^*_{\boldsymbol{\omega}(\boldsymbol{x})}(\boldsymbol{f}(\boldsymbol{x})) = \partial_f \Psi^*_{\boldsymbol{\omega}(\boldsymbol{x})}(\boldsymbol{f})|_{\boldsymbol{f}=\boldsymbol{f}(\boldsymbol{x})}$.

[2]We have used the notation $\partial_{\boldsymbol{x}}\phi(\boldsymbol{y}) = \partial_{\boldsymbol{x}}\phi(\boldsymbol{x})|_{\boldsymbol{x}=\boldsymbol{y}}$.

The KL divergence is Eq. (2.34) with the following potential:[3]

$$\phi_{\mathrm{KL}}(\boldsymbol{x}) \coloneqq \sum_{i=1}^{N_{\mathscr{X}}} x_i \ln x_i. \tag{2.36}$$

Then, the KL divergence is defined by $\mathcal{D}_{\phi_{\mathrm{KL}}}(\cdot\|\cdot) \coloneqq \mathcal{D}_{\mathrm{KL}}(\cdot\|\cdot)$ and reads

$$\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}\|\boldsymbol{y}) = \sum_{i=1}^{N_{\mathscr{X}}} x_i \ln \frac{x_i}{y_i} - \sum_{i=1}^{N_{\mathscr{X}}} x_i + \sum_{i=1}^{N_{\mathscr{X}}} y_i, \tag{2.37}$$

and its derivative takes the following form:

$$\partial_{\boldsymbol{x}} \mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}\|\boldsymbol{y}) = \begin{bmatrix} \ln x_1 - \ln y_1 \\ \ln x_2 - \ln y_2 \\ \vdots \\ \ln x_{N_{\mathscr{X}}} - \ln y_{N_{\mathscr{X}}} \end{bmatrix}. \tag{2.38}$$

Then, Eq. (2.33) is rewritten as

$$\boldsymbol{f}(\boldsymbol{x}) = -\tfrac{1}{2} S^{\mathsf{T}} \partial_{\boldsymbol{x}} \mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}\|\hat{\boldsymbol{x}}) + \boldsymbol{f}_{\mathrm{ne}}. \tag{2.39}$$

The definition of $\hat{\boldsymbol{x}}$ will be given in the following subsection, and $\boldsymbol{f}_{\mathrm{ne}} \notin \mathrm{Im}[S^{\mathsf{T}}]$ represents the nonequilibrium force incurred to the system [19].

Mass action kinetics also offers the following definitions of the flux and activity [20,21,36,37]:

$$\boldsymbol{j}(\boldsymbol{x}) = (\boldsymbol{k}^+ \circ (B^+)^{\mathsf{T}} - \boldsymbol{k}^- \circ (B^-)^{\mathsf{T}}) \boldsymbol{x}^{\Gamma^{\mathsf{T}}}. \tag{2.40}$$

Substituting Eq. (2.40) into Eq. (2.24), we also get the activity for CRNs:

$$\boldsymbol{\omega}(\boldsymbol{x}) = 2\sqrt{\boldsymbol{k}^+ \circ \boldsymbol{k}^-} \circ \boldsymbol{x}^{R^{\mathsf{T}}/2}, \tag{2.41}$$

where

$$R \coloneqq \Gamma(B^+ + B^-). \tag{2.42}$$

In the remaining part of this section, we will present the geometric expression of the equation for CRNs.

E. Geometric expression of an equilibrium CRE

Up to this point, the discussion has centered on the geometric relationships that exist among the chemical concentration, potential, force, and flux in a CRN. Subsequently, the CRE specified in Eq. (2.20) can be reformulated into a geometric expression [36–38]. To accomplish this, the detailed balance condition (DBC) must be taken into account. The DBC, a criterion for the dynamic stability of a system at equilibrium, is described in the following section [20,21]:

$$\ln \frac{\boldsymbol{k}^+}{\boldsymbol{k}^-} = S^{\mathsf{T}} \ln \boldsymbol{x}_{\mathrm{eq}}. \tag{2.43}$$

Here, $\boldsymbol{x}_{\mathrm{eq}}$ represents the equilibrium chemical concentration vector, which is dependent on both the initial concentration vector $\boldsymbol{x}_{\mathrm{ini}}$ and the specific CRE under consideration. Additionally, if Eq. (2.43) is met, then $\boldsymbol{f}_{\mathrm{ne}} = \boldsymbol{0}$. Generally, at equilibrium, net fluxes cease ($\boldsymbol{j} = \boldsymbol{0}$), allowing us to define a set of equilibrium chemical concentration vectors as follows:

$$V_{\mathrm{eq}} \coloneqq \{\boldsymbol{x} > \boldsymbol{0} | \boldsymbol{j}(\boldsymbol{x}) = \boldsymbol{0}\}. \tag{2.44}$$

From Eq. (2.43), Eq. (2.44) is transformed into

$$V_{\mathrm{eq}} = \{\boldsymbol{x} > \boldsymbol{0} | \exists \boldsymbol{\eta} \in \mathbb{R}^{|\ker(S^{\mathsf{T}})|}, \ln \boldsymbol{x} = \ln \boldsymbol{x}_{\mathrm{eq}} + U\boldsymbol{\eta}\}, \tag{2.45}$$

where $U \coloneqq [\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{|\ker(S^{\mathsf{T}})|}]$ and $\{\boldsymbol{u}_i\}_{i=1}^{|\ker(S^{\mathsf{T}})|}$ are the bases of $\ker(S^{\mathsf{T}})$. We have introduced $\hat{\boldsymbol{x}}$ in Eq. (2.39). We here impose the following relation to $\hat{\boldsymbol{x}}$:

$$\hat{\boldsymbol{x}} \in V_{\mathrm{eq}}. \tag{2.46}$$

Then Eq. (2.39) describes dynamics of gradient flow to V_{eq}. Equation (2.46) is equivalently written as

$$\ln \frac{\boldsymbol{k}^+}{\boldsymbol{k}^-} = S^{\mathsf{T}} \ln \hat{\boldsymbol{x}}. \tag{2.47}$$

Note that using $\hat{\boldsymbol{x}}$ instead of $\boldsymbol{x}_{\mathrm{eq}}$ provides us with a generalized expression of the dynamical system.

Finally, we have arrived at the geometric expression of a CRE, namely, combining Eqs. (2.32), (2.39), (2.41), and (2.43), we get[4]

$$\dot{\boldsymbol{x}}_t = S \partial_f \Psi^*_{\boldsymbol{\omega}(\boldsymbol{x}_t)}\big(-\tfrac{1}{2} S^{\mathsf{T}} \partial_{\boldsymbol{x}} \mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t\|\hat{\boldsymbol{x}})\big), \tag{2.48}$$

where $\hat{\boldsymbol{x}} \in V_{\mathrm{eq}}$. Note that in Eq. (2.48), replacing $\hat{\boldsymbol{x}}$ with $\boldsymbol{x}_{\mathrm{eq}}$ does not affect the dynamics of CRNs because $S^{\mathsf{T}} U \eta = \boldsymbol{0}$. In Ref. [38], it is also mentioned that the KL divergence is a Lyapunov function for the dynamics of CRNs.

III. DIFFICULTY OF CONSTRUCTING AN UPPER BOUND ON THE REACTION RATES OF CRNS

In this section, we briefly revisit Newton's method and present a counterexample illustrating its limitations in establishing an upper bound on the reaction rates of CRNs.

A. Newton's method

As stated in Sec. I, the objective of this paper is to determine an upper bound on the reaction rates of CRNs. One might assume that straightforward optimization methods could achieve this. However, before discussing NG, we elucidate the challenges of using Newton's method [39] as an optimization technique for this purpose. While the gradient method is another elementary optimization technique, its indeterminate step size precludes its consideration in this paper. We now turn to a specific optimization problem:

$$\min_{\boldsymbol{x}} f(\boldsymbol{x}). \tag{3.1}$$

Letting $\boldsymbol{x}_t$ be the state at the tth iteration for $t \in \mathbb{Z}_{\geqslant 0}$, Newton's method for Eq. (3.1) is given by

$$\boldsymbol{x}_{t+1} = \boldsymbol{x}_t - \big[\partial_{\boldsymbol{x}}^2 f(\boldsymbol{x}_t)\big]^{-1} \partial_{\boldsymbol{x}} f(\boldsymbol{x}_t). \tag{3.2}$$

In the case of CRNs, we have $f(\boldsymbol{x}) = \mathcal{D}_\phi(\boldsymbol{x}\|\hat{\boldsymbol{x}})$; then Eq. (3.2) reads

$$\boldsymbol{x}_{t+1} = \boldsymbol{x}_t - G_\phi^{-1}(\boldsymbol{x}_t) \partial_{\boldsymbol{x}} \mathcal{D}_\phi(\boldsymbol{x}_t\|\hat{\boldsymbol{x}}), \tag{3.3}$$

where G_ϕ is the Hessian of $\phi(\cdot)$.

[3] See Appendix A for details.

[4] We have used the following notation: $\partial_{\boldsymbol{x}} \mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t\|\hat{\boldsymbol{x}}) = \partial_{\boldsymbol{x}} \mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}\|\hat{\boldsymbol{x}})|_{\boldsymbol{x}=\boldsymbol{x}_t}$.

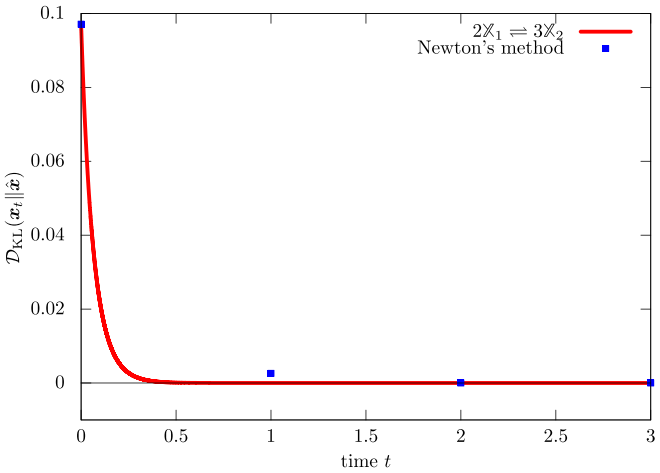


FIG. 1. Dependence of $\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t\|\hat{\boldsymbol{x}})$ on t for the CRN in Eq. (3.4) and their upper bound in the case of $\boldsymbol{x}_{\mathrm{eq}} = \hat{\boldsymbol{x}}$.

B. Counterexample

We will demonstrate a counterexample to show that Eq. (3.3) does not yield an upper bound for a CRN. We consider the following CRN with $N_{\mathbb{X}} = 2$, $m = 3$, and $N_{\mathrm{e}} = 1$:

$$2\mathbb{X}_1 \rightleftharpoons 3\mathbb{X}_2. \tag{3.4}$$

For the simulations of Eq. (2.48), we set $k_e^{\pm} = 1$, $\Delta t = 1.0 \times 10^{-4}$, $\boldsymbol{x}_{\mathrm{ini}} = [3/4, 11/8]^{\mathsf{T}}$, and $\hat{\boldsymbol{x}} = [1.0, 1.0]^{\mathsf{T}}$. In Fig. 1, we plot the dynamics of Eq. (3.4) as well as the dynamics obtained using Newton's method. At $t = 1$, the divergence of Newton's method is greater than that of the CRN, indicating that Newton's method fails to bound the dynamics. This observation is illustrated in the figure. The reason for this discrepancy lies in the nonlinearity of Eq. (3.4).

IV. NATURAL GRADIENT

In this section, we explore the NG method and its applicability to the problem of constraining reaction rates in CRNs. As our proposed methodology hinges on NG, understanding its theoretical underpinnings and its distinction from Newton's method is crucial.

A. Derivation of NG

In this section, we outline the derivation of the NG method, which is grounded in information geometry. Specifically, we will elucidate how the dynamics of a given vector $\boldsymbol{x}_t$ at time t are updated within the framework of NG:

$$\boldsymbol{x}_{t+\Delta t} = \boldsymbol{x}_t + \Delta\boldsymbol{x}_t(\epsilon), \tag{4.1}$$

where $\Delta\boldsymbol{x}_t(\epsilon)$ is defined as.[5]

$$\Delta\boldsymbol{x}_t(\epsilon) = \underset{\Delta\boldsymbol{x}:\mathcal{D}_{\phi'}(\boldsymbol{x}_t+\Delta\boldsymbol{x}\|\boldsymbol{x}_t)\leqslant\epsilon}{\arg\min} [f(\boldsymbol{x}_t+\Delta\boldsymbol{x}) - f(\boldsymbol{x}_t)] \tag{4.2}$$

$$\approx \underset{\Delta\boldsymbol{x}:\frac{1}{2}\Delta\boldsymbol{x}^{\mathsf{T}}G_{\phi'}(\boldsymbol{x}_t)\Delta\boldsymbol{x}\leqslant\epsilon}{\arg\min} \partial_{\boldsymbol{x}} f(\boldsymbol{x}_t)^{\mathsf{T}}\Delta\boldsymbol{x}. \tag{4.3}$$

Note that we have used the lowest-order approximation for $f(\boldsymbol{x}_t+\Delta\boldsymbol{x}) - f(\boldsymbol{x}_t)$ and $\mathcal{D}_{\phi'}(\boldsymbol{x}_t+\Delta\boldsymbol{x}\|\boldsymbol{x}_t)$. Here, $G_{\phi'}(\boldsymbol{x}_t)$ is the Hessian given by

$$[G_{\phi'}(\boldsymbol{x}_t)]_{i,j} := \frac{\partial^2}{\partial x_i \partial x_j}\phi'(\boldsymbol{x}_t), \tag{4.4}$$

where $[\cdot]_{i,j}$ is the (i,j)th element. In the case of Eq. (2.36), Eq. (4.4) reads

$$[G_{\phi'}(\boldsymbol{x}_t)]_{i,j} = \delta_{i,j}\frac{1}{[\boldsymbol{x}_t]_i}, \tag{4.5}$$

where $\delta_{i,j}$ is the Kronecker delta function and $[\cdot]_i$ is the ith element. To derive Eq. (4.3), we have used the following expansion of the Bregman divergence:

$$\mathcal{D}_{\phi'}(\boldsymbol{x}_t+\Delta\boldsymbol{x}\|\boldsymbol{x}_t) = \phi'(\boldsymbol{x}_t+\Delta\boldsymbol{x}) - \phi'(\boldsymbol{x}_t) - \langle(\boldsymbol{x}_t+\Delta\boldsymbol{x}) - \boldsymbol{x}_t, \partial_{\boldsymbol{x}}\phi'(\boldsymbol{x}_t)\rangle \tag{4.6}$$

$$\approx \phi'(\boldsymbol{x}_t) + \partial_{\boldsymbol{x}}\phi(\boldsymbol{x}_t)^{\mathsf{T}}\Delta\boldsymbol{x} + \tfrac{1}{2}\Delta\boldsymbol{x}^{\mathsf{T}}G_{\phi'}(\boldsymbol{x}_t)\Delta\boldsymbol{x} - \phi'(\boldsymbol{x}_t) - \langle(\boldsymbol{x}_t+\Delta\boldsymbol{x}) - \boldsymbol{x}_t, \partial_{\boldsymbol{x}}\phi'(\boldsymbol{x}_t)\rangle \tag{4.7}$$

$$= \tfrac{1}{2}\Delta\boldsymbol{x}^{\mathsf{T}}G_{\phi'}(\boldsymbol{x}_t)\Delta\boldsymbol{x}. \tag{4.8}$$

Note that Δt in Eq. (4.1) is set to unity in the conventional formulation of NG; in the following section, we will impose a specific relationship between Δt and ϵ in Eq. (4.1) to connect NG and CRNs.

To find the solution of Eq. (4.3), we employ the method of Lagrange multipliers where the Lagrange function reads

$$L(\Delta\boldsymbol{x},\lambda) := \partial_{\boldsymbol{x}} f(\boldsymbol{x}_t)^{\mathsf{T}}\Delta\boldsymbol{x} - \lambda\left(\tfrac{1}{2}\Delta\boldsymbol{x}^{\mathsf{T}}G_{\phi'}(\boldsymbol{x}_t)\Delta\boldsymbol{x} - \epsilon\right). \tag{4.9}$$

The derivative of Eq. (4.9) with respect to $\Delta\boldsymbol{x}$ takes the following form:

$$\frac{\partial}{\partial\Delta\boldsymbol{x}}L(\Delta\boldsymbol{x},\lambda) = \partial_{\boldsymbol{x}} f(\boldsymbol{x}_t) - \lambda G_{\phi'}(\boldsymbol{x}_t)\Delta\boldsymbol{x}. \tag{4.10}$$

Then, the solution of Eq. (4.10) is given by

$$\Delta\boldsymbol{x} = \frac{1}{\lambda}G_{\phi'}^{-1}(\boldsymbol{x}_t)\partial_{\boldsymbol{x}} f(\boldsymbol{x}_t). \tag{4.11}$$

The derivative of Eq. (4.9) with respect to λ has the following form:

$$\frac{\partial}{\partial\lambda}L(\Delta\boldsymbol{x},\lambda) = -\left(\frac{1}{2}\Delta\boldsymbol{x}^{\mathsf{T}}G_{\phi'}(\boldsymbol{x}_t)\Delta\boldsymbol{x} - \epsilon\right). \tag{4.12}$$

Taking Eq. (4.11) into account, the solution of Eq. (4.12) is written as

$$\lambda^2 = \frac{\partial_{\boldsymbol{x}} f(\boldsymbol{x}_t)^{\mathsf{T}}G_{\phi'}^{-1}(\boldsymbol{x}_t)\partial_{\boldsymbol{x}} f(\boldsymbol{x}_t)}{2\epsilon}. \tag{4.13}$$

Combining Eqs. (4.11) and (4.13) and taking account of the nature of the minimization problems, the solution of Eq. (4.3) takes the following form:

$$\Delta\boldsymbol{x}_t(\epsilon) = -\sqrt{\frac{2\epsilon}{\partial_{\boldsymbol{x}} f(\boldsymbol{x}_t)^{\mathsf{T}}G_{\phi'}^{-1}(\boldsymbol{x}_t)\partial_{\boldsymbol{x}} f(\boldsymbol{x}_t)}}\,G_{\phi'}^{-1}(\boldsymbol{x}_t)\partial_{\boldsymbol{x}} f(\boldsymbol{x}_t). \tag{4.14}$$

Here, we have added a negative sign to Eq. (4.14) to decrease the value of $f(\boldsymbol{x})$ by Eq. (4.1). Note that $\phi'(\cdot)$ in Eq. (4.14) may be different from $\phi(\cdot)$ appearing in Sec. II. In the case of CRNs, $f(\boldsymbol{x}_t)$ in Eq. (4.14) represents $\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t\|\hat{\boldsymbol{x}})$. As shown in Eq. (4.14), ϵ is a key parameter in NG. From the perspective of applying NG to CRNs, the relationship between ϵ in NG and

[5]We have used the following notation: $\partial_{\boldsymbol{x}} f(\boldsymbol{x}_t) = \partial_{\boldsymbol{x}} f(\boldsymbol{x})|_{\boldsymbol{x}=\boldsymbol{x}_t}$

Δt in CRNs, when discretized, is still missing. Therefore, NG cannot be directly applied to CRNs. In the following section, we will explain how to address this challenge and develop a general upper bound on the dynamics of CRNs.

B. Comparison with Newton's method

In this section, we compare NG with Newton's method. Newton's method is a special case of NG when Eq. (4.14) is adjusted according to certain conditions. Specifically, the conditions are $\phi(\cdot) = \phi'(\cdot)$ and $\epsilon = \partial_{\boldsymbol{x}} f(\boldsymbol{x}_t)^{\mathsf{T}} G_{\phi'}^{-1}(\boldsymbol{x}_t)\partial_{\boldsymbol{x}} f(\boldsymbol{x}_t)$. The equation thus becomes equivalent to Eq. (3.3). This equivalency leads us to introduce a systematic NG-based method to determine the direction and step size for a gradient system that bounds CRNs of a specific class.

V. UPPER BOUND ON REACTION RATES

In this section, we construct a nonlinear system that gives an upper bound on reaction rates of CRNs in a given class. The class is characterized by several topological numbers of CRNs: $N_{\mathbb{v}}$, $N_{\mathbb{e}}$, and m.

A. Upper bound system

Comparing discretized CRE dynamics with NG dynamics, represented by Eq. (4.1), presents a challenge. The difficulty arises from the absence of an established relationship between ϵ, the constraint parameter in NG, and Δt, the time step in the discretized CRE. To address this issue, we propose the following relationship between ϵ and Δt:

$$\epsilon = \mathcal{D}_{\phi'}(\boldsymbol{x}_t + \|\dot{\boldsymbol{x}}_t\|_{\mathrm{F}} \boldsymbol{e}_t \Delta t \| \boldsymbol{x}_t), \tag{5.1}$$

where $\|\cdot\|_{\mathrm{F}}$ is the Frobenius norm and $\boldsymbol{e}_t$ is a vector that satisfies $\|\boldsymbol{e}_t\|_{\mathrm{F}} = 1$. In NG, $\Delta \boldsymbol{x}_t(\epsilon)$ in Eq. (4.1) becomes larger when ϵ is increased. Therefore, for the rest of this subsection, we describe procedures for computing $\|\dot{\boldsymbol{x}}_t\|_{\mathrm{F}}$ and $\boldsymbol{e}_t$ that maximize Eq. (5.1).

Then, we try to compute the maximum value of ϵ in Eq. (5.1). Note that $S : \mathbb{R}^{N_{\mathbb{e}}} \to \mathbb{R}^{N_{\mathbb{x}}}$ and S is a $N_{\mathbb{x}} \times N_{\mathbb{e}}$ matrix. From Eq. (2.48), we get

$$\begin{aligned}
\|\dot{\boldsymbol{x}}_t\|_{\mathrm{F}} &= \left\| S \partial_f \Psi^*_{\boldsymbol{\omega}(\boldsymbol{x}_t)}\left(-\tfrac{1}{2} S^{\mathsf{T}} \partial_{\boldsymbol{x}} \mathcal{D}_\phi(\boldsymbol{x}_t \| \hat{\boldsymbol{x}})\right) \right\|_{\mathrm{F}} && (5.2)\\
&\leqslant \|S\|_{\mathrm{F}} \left\| \partial_f \Psi^*_{\boldsymbol{\omega}(\boldsymbol{x}_t)}\left(-\tfrac{1}{2} S^{\mathsf{T}} \partial_{\boldsymbol{x}} \mathcal{D}_\phi(\boldsymbol{x}_t \| \hat{\boldsymbol{x}})\right) \right\|_{\mathrm{F}} && (5.3)\\
&\leqslant \|S\|_{\mathrm{F}} \left\| \partial_f \Psi^*_{\boldsymbol{\omega}(\boldsymbol{x}_t)}\left(\left\| -\tfrac{1}{2} S^{\mathsf{T}} \partial_{\boldsymbol{x}} \mathcal{D}_\phi(\boldsymbol{x}_t \| \hat{\boldsymbol{x}}) \right\|_{\mathrm{abs}}\right) \right\|_{\mathrm{F}} && (5.4)\\
&\leqslant \|S\|_{\mathrm{F}} \left\| \partial_f \Psi^*_{\boldsymbol{\omega}(\boldsymbol{x}_t)}\left(\tfrac{1}{2} \|S^{\mathsf{T}}\|_{\mathrm{F}} \|\partial_{\boldsymbol{x}} \mathcal{D}_\phi(\boldsymbol{x}_t \| \hat{\boldsymbol{x}})\|_{\mathrm{F}}^{N_{\mathbb{x}} \to N_{\mathbb{e}}}\right) \right\|_{\mathrm{F}}. && (5.5)
\end{aligned}$$

Here, $\|\cdot\|_{\mathrm{abs}}$ and $\|\cdot\|_{\mathrm{F}}^{N_{\mathbb{x}} \to N_{\mathbb{e}}}$ are defined as, respectively,

$$\|\boldsymbol{v}\|_{\mathrm{abs}} := [|v_1|, |v_2|, \ldots, |v_{N_{\mathbb{x}}}|]^{\mathsf{T}}, \tag{5.6}$$

$$\|\boldsymbol{v}\|_{\mathrm{F}}^{N_{\mathbb{x}} \to N_{\mathbb{e}}} := [\underbrace{\|\boldsymbol{v}\|_{\mathrm{F}}, \|\boldsymbol{v}\|_{\mathrm{F}}, \ldots, \|\boldsymbol{v}\|_{\mathrm{F}}}_{N_{\mathbb{e}}}]^{\mathsf{T}} \tag{5.7}$$

for $\boldsymbol{v} := [v_1, v_2, \ldots, v_{N_{\mathbb{x}}}]^{\mathsf{T}}$. From Eq. (2.31), we have

$$\partial_f \Psi^*_{\boldsymbol{\omega}(\boldsymbol{x})}(\|\boldsymbol{f}(\boldsymbol{x})\|_{\mathrm{abs}}) = \boldsymbol{\omega}(\boldsymbol{x}) \circ \sinh(\|\boldsymbol{f}(\boldsymbol{x})\|_{\mathrm{abs}}), \tag{5.8}$$

$$\partial_f \Psi^*_{\boldsymbol{\omega}(\boldsymbol{x})}\left(\|\boldsymbol{f}(\boldsymbol{x})\|_{\mathrm{F}}^{N_{\mathbb{x}} \to N_{\mathbb{e}}}\right) = \boldsymbol{\omega}(\boldsymbol{x}) \circ \sinh\left(\|\boldsymbol{f}(\boldsymbol{x})\|_{\mathrm{F}}^{N_{\mathbb{x}} \to N_{\mathbb{e}}}\right). \tag{5.9}$$

Given $S : \mathbb{R}^{N_{\mathbb{e}}} \to \mathbb{R}^{N_{\mathbb{x}}}$ and $\boldsymbol{v} \in \mathbb{R}^{N_{\mathbb{x}}}$, we have the following inequality for $e = 1, 2, \ldots, N_{\mathbb{e}}$:

$$\begin{aligned}
[\|S^{\mathsf{T}} \boldsymbol{v}\|_{\mathrm{abs}}]_e &\leqslant \|S^{\mathsf{T}}\|_{\mathrm{F}} \|\boldsymbol{v}\|_{\mathrm{F}} && (5.10)\\
&= \left[\|S^{\mathsf{T}}\|_{\mathrm{F}} \|\boldsymbol{v}\|_{\mathrm{F}}^{N_{\mathbb{x}} \to N_{\mathbb{e}}}\right]_e, && (5.11)
\end{aligned}$$

where $[\cdot]_e$ is the eth element. Then, we have finished computing the bound on $\|\dot{\boldsymbol{x}}_t\|_{\mathrm{F}}$ within a given class of CRNs.

Next, we compute $\boldsymbol{e}_t$ as follows:

$$\begin{aligned}
\boldsymbol{e}_t &= \operatorname*{arg\,max}_{\boldsymbol{e}:\|\boldsymbol{e}\|_{\mathrm{F}}=1} \mathcal{D}_{\phi'}(\boldsymbol{x}_t + \|\dot{\boldsymbol{x}}_t\|_{\mathrm{F}} \boldsymbol{e} \Delta t \| \boldsymbol{x}_t) && (5.12)\\
&\approx \operatorname*{arg\,max}_{\boldsymbol{e}:\|\boldsymbol{e}\|_{\mathrm{F}}=1} \left(\tfrac{1}{2} \|\dot{\boldsymbol{x}}_t\|_{\mathrm{F}}^2 (\Delta t)^2 \boldsymbol{e}^{\mathsf{T}} G_{\phi'}(\boldsymbol{x}_t) \boldsymbol{e}\right) && (5.13)\\
&= \operatorname*{arg\,max}_{\boldsymbol{e}:\|\boldsymbol{e}\|_{\mathrm{F}}=1} \boldsymbol{e}^{\mathsf{T}} G_{\phi'}(\boldsymbol{x}_t) \boldsymbol{e}. && (5.14)
\end{aligned}$$

Thus, $\boldsymbol{e}_t$ is the eigenvector associated with the maximum eigenvalue of $G_{\phi'}(\boldsymbol{x}_t)$. Substituting Eq. (5.5) and the solution of Eq. (5.14) into Eq. (5.1), we can calculate the maximum value of ϵ within a given class of CRNs.

B. S and R of an upper bound system

To identify the upper bound described by Eq. (5.5) for CRNs under certain constraints, both S in Eq. (2.4) and R in Eq. (2.42) must be carefully designed. We introduce a method for determining S_{ub} and R_{ub} specific to a class of CRNs characterized by $N_{\mathbb{x}}$ as the number of chemicals, m as the highest coefficient in chemical reactions, and $N_{\mathbb{e}}$ as the number of reactions. The S_{ub} and R_{ub} matrices are of dimensions $N_{\mathbb{x}} \times N_{\mathbb{e}}$, and their elements at the (i, e)-th position are defined as follows:

$$[S_{\mathrm{ub}}]_{i,e} := m, \tag{5.15}$$

$$[R_{\mathrm{ub}}]_{i,e} := \mathbb{1}[x_i \leqslant 1] \min_i([R]_{i,e}) + \mathbb{1}[x_i > 1] \max_i([R]_{i,e}). \tag{5.16}$$

Here, $\mathbb{1}[\cdot]$ denotes the indicator function, and $[\cdot]_{i,e}$ represents the (i, e)-th element. The reader may think that $\mathbb{1}[\cdot]$ is not necessary. This reflects the fact that $x^n \geqslant x^m$ for $x \in [1, \infty)$ and $n \geqslant m$ but $x^n \leqslant x^m$ for $x \in (0, 1]$ and $n \geqslant m$. By solving Eq. (4.14) with Eqs. (5.1), (5.5), (5.14), (5.15), and (5.16), we can compute the upper bound for a given class. In other words, we use the following inequality to construct an upper bound system:

$$\|\dot{\boldsymbol{x}}_t\|_{\mathrm{F}} \leqslant \|S_{\mathrm{ub}}\|_{\mathrm{F}} \left\| \partial_f \Psi^*_{\boldsymbol{\omega}_{\mathrm{ub}}(\boldsymbol{x}_t)}\left(\tfrac{1}{2} \|S_{\mathrm{ub}}^{\mathsf{T}}\|_{\mathrm{F}} \|\partial_{\boldsymbol{x}} \mathcal{D}_\phi(\boldsymbol{x}_t \| \hat{\boldsymbol{x}})\|_{\mathrm{F}}^{N_{\mathbb{x}} \to N_{\mathbb{e}}}\right) \right\|_{\mathrm{F}}, \tag{5.17}$$

where

$$\boldsymbol{\omega}_{\mathrm{ub}}(\boldsymbol{x}) := 2\sqrt{\boldsymbol{k}^+ \circ \boldsymbol{k}^-} \circ \boldsymbol{x}^{R_{\mathrm{ub}}^{\mathsf{T}}/2}. \tag{5.18}$$

C. Upper bound system with the KL constraint

We utilize Eq. (2.36), represented as $\phi'(\cdot) = \phi_{\mathrm{KL}}(\cdot)$, as the potential function for the Bregman divergence

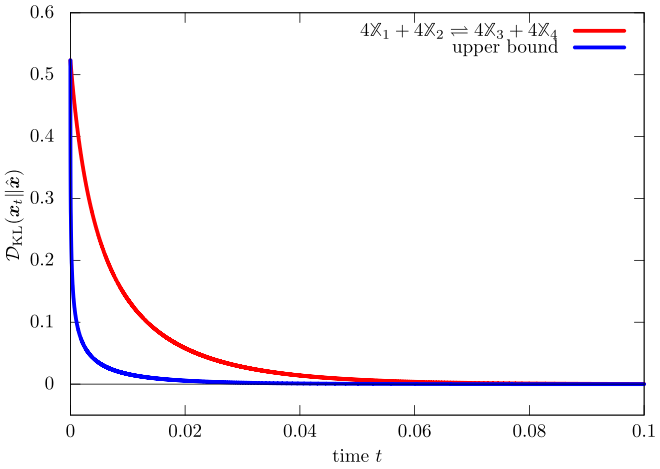


FIG. 2. Dependence of $\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t\|\hat{\boldsymbol{x}})$ on time t for the CRN in Eq. (6.1) and its upper bound in the case of $\boldsymbol{x}_{\mathrm{eq}} = \hat{\boldsymbol{x}}$.

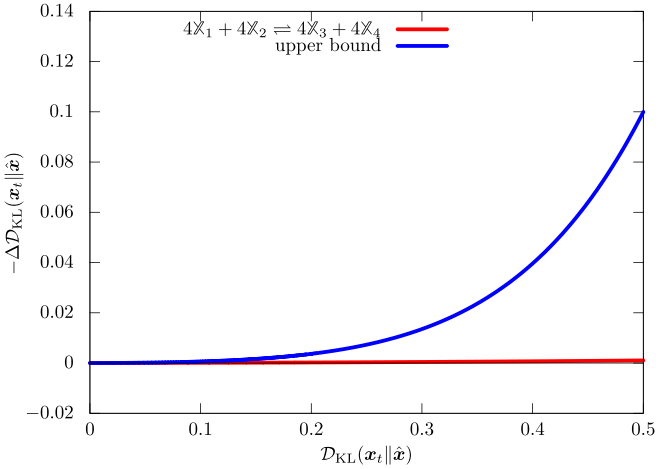


FIG. 3. Relationship between $\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t\|\hat{\boldsymbol{x}})$ and $-\Delta\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t\|\hat{\boldsymbol{x}})$ for the CRN in Eq. (6.1) and its upper bound in the case of $\boldsymbol{x}_{\mathrm{eq}} = \hat{\boldsymbol{x}}$. We have used $\boldsymbol{x}_t$ on the solution of Eq. (6.1).

in the constraint of NG.[6]. Subsequently, by substituting $\|\partial_{\boldsymbol{x}}\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t\|\hat{\boldsymbol{x}})\|_{\mathrm{F}}^{N_{\mathbb{X}}\to N_{\mathrm{e}}}$ into Eq. (5.5), we can determine the maximum value of $\|\dot{\boldsymbol{x}}_t\|_{\mathrm{F}}$ as stated in Eq. (5.5).

VI. NUMERICAL SIMULATIONS

In this section, numerical simulations are conducted to elucidate the upper-bound dynamics for a specified class of CRNs. We consider CRNs with various values of $N_{\mathbb{X}}$, N_{e}, and m. The initial condition is chosen as $\hat{\boldsymbol{x}} = [1.0, 1.0, 1.0, 1.0]^{\intercal}$ and the time step as $\Delta t = 1.0 \times 10^{-5}$. The rate constants $k_e^{\pm}$ are fixed at 1 for all e ranging from 1 to N_{e}. Simulations are executed for a total of 3.0×10^4 steps. The initial chemical concentration vector at time $t = 0$ is denoted as $\boldsymbol{x}_{\mathrm{ini}}$.

A. Case where CRNs that convergence to $\hat{\boldsymbol{x}}$

First, we consider the following CRN and set $\boldsymbol{x}_{\mathrm{ini}} = [1/2, 1/2, 3/2, 3/2]^{\intercal}$, $\hat{\boldsymbol{x}} = [1.0, 1.0, 1.0, 1.0]^{\intercal}$, and $\Delta t = 1.0 \times 10^{-5}$:

$$4\mathbb{X}_1 + 4\mathbb{X}_2 \rightleftharpoons 4\mathbb{X}_3 + 4\mathbb{X}_4. \tag{6.1}$$

In this case, we have $\boldsymbol{x}_{\mathrm{eq}} = \hat{\boldsymbol{x}}$. In Fig. 2, we plot the dynamics of Eq. (6.1) and that of the system constructed in Sec. V. The system constructed in Sec. V provides a tighter bound. In Fig. 3, we show the time difference of the KL divergence $-\Delta\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t\|\hat{\boldsymbol{x}})$ per Δt. We have used $\boldsymbol{x}_t$ on the solution of Eq. (6.1) with $\boldsymbol{x}_{\mathrm{ini}} = [1/2, 1/2, 3/2, 3/2]^{\intercal}$; that is, $-\Delta\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t\|\hat{\boldsymbol{x}})$ of the CRN in Eq. (6.1) and the system constructed in Sec. V on the orbit of the CRN in Eq. (6.1). As shown in Fig. 3, the system constructed in Sec. V shows faster convergence at each $\boldsymbol{x}_t$. Note that in the above example, $\lim_{t\to\infty}\boldsymbol{x}_t = \hat{\boldsymbol{x}}$; as a result, $\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t\|\hat{\boldsymbol{x}})$ goes to zero for large t. In general, this does not hold. We provide illustrative examples in the following subsection.

[6]While there are many different candidates for $\phi'(\cdot)$, the L^2 constraint is often used. Then, we explain the case of the L^2 constraint in Appendix B

B. Case where CRNs that do not convergence to $\hat{\boldsymbol{x}}$

We initially consider CRNs that are bounded by the upper bound dynamics with $N_{\mathbb{X}} = 3$, $m \leqslant 2$, and $N_{\mathrm{e}} = 1$ and compare them from the perspective of reaction rates. Here, we contemplate six distinct reactions with identical topological characteristics ($N_{\mathbb{X}} = 3$, $m \leqslant 2$, and $N_{\mathrm{e}} = 1$):

$$\mathbb{X}_1 + \mathbb{X}_2 \rightleftharpoons \mathbb{X}_3, \tag{6.2a}$$

$$\mathbb{X}_1 + \mathbb{X}_2 \rightleftharpoons 2\mathbb{X}_3, \tag{6.2b}$$

$$\mathbb{X}_1 + 2\mathbb{X}_2 \rightleftharpoons \mathbb{X}_3, \tag{6.2c}$$

$$2\mathbb{X}_1 + 2\mathbb{X}_2 \rightleftharpoons \mathbb{X}_3. \tag{6.2d}$$

In Fig. 4, we present numerical simulations of CRNs for the system (6.2). We set $\boldsymbol{x}_{\mathrm{ini}} = [15/16, 15/16, 9/8]$, $\hat{\boldsymbol{x}} = [1.0, 1.0, 1.0]^{\intercal}$, and $\Delta t = 10^{-4}$. Figure 4 demonstrates that the system constructed in Sec. V functions effectively as the upper bound for the given CRNs.

Next, we consider CRNs with higher nonlinearity. Here we consider the CRN in Eq. (6.1) and the following five different

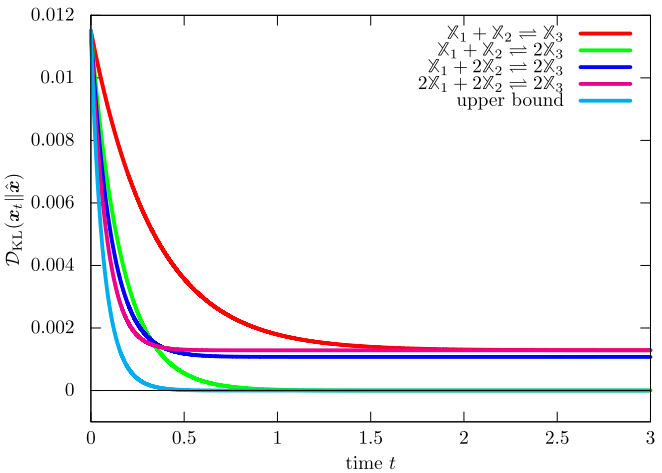


FIG. 4. Dependence of $\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t\|\hat{\boldsymbol{x}})$ on time t for several CRNs in Eq. (6.2) and their upper bound in the case of $\boldsymbol{x}_{\mathrm{eq}} \neq \hat{\boldsymbol{x}}$.

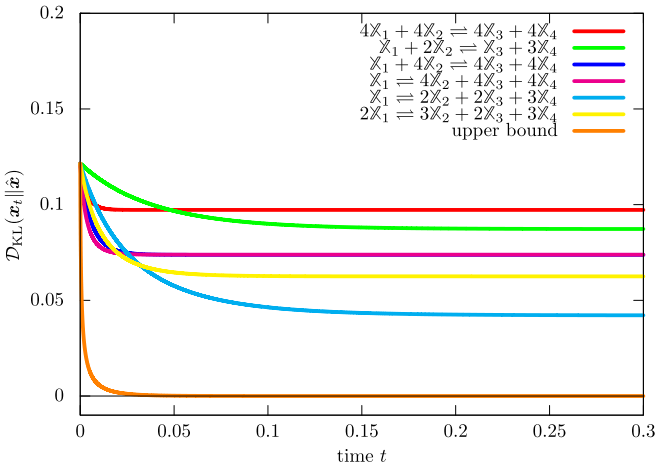


FIG. 5. Dependence of $\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t\|\hat{\boldsymbol{x}})$ on time t for several CRNs in Eqs. (6.3) and their upper bound in the case of $\boldsymbol{x}_{\mathrm{eq}} \neq \hat{\boldsymbol{x}}$.

reactions, which have the same topological quantities ($N_{\mathbb{X}} = 4$, $m \leqslant 4$, and $N_{\mathbb{e}} = 1$) as Eq. (6.1):

$$\mathbb{X}_1 + 4\mathbb{X}_2 \rightleftharpoons 4\mathbb{X}_3 + 4\mathbb{X}_4, \tag{6.3a}$$

$$\mathbb{X}_1 + 2\mathbb{X}_2 \rightleftharpoons \mathbb{X}_3 + 3\mathbb{X}_4, \tag{6.3b}$$

$$4\mathbb{X}_1 \rightleftharpoons 4\mathbb{X}_2 + 4\mathbb{X}_3 + 4\mathbb{X}_4, \tag{6.3c}$$

$$\mathbb{X}_1 \rightleftharpoons 2\mathbb{X}_2 + 2\mathbb{X}_3 + 3\mathbb{X}_4, \tag{6.3d}$$

$$2\mathbb{X}_1 \rightleftharpoons 3\mathbb{X}_2 + 2\mathbb{X}_3 + 3\mathbb{X}_4. \tag{6.3e}$$

Note that $\boldsymbol{x}_{\mathrm{eq}}$ is a function of $\boldsymbol{x}_{\mathrm{ini}}$, and CRNs in Eqs. (6.3) have different $\boldsymbol{x}_{\mathrm{eq}}$.

We set $\boldsymbol{x}_{\mathrm{ini}} = [9/8, 87/80, 27/20, 27/20]^{\intercal}$, $\hat{\boldsymbol{x}} = [1.0, 1.0, 1.0, 1.0]^{\intercal}$, and $\Delta t = 1.0 \times 10^{-5}$. In Fig. 5, we plot the dynamics of Eqs. (6.3) and that of the system constructed in Sec. V. It clearly shows that the system constructed in Sec. V gives an upper bound on CRNs. The CRNs in Eqs. (6.3) have equilibrium states different from $\hat{\boldsymbol{x}}$ because of $\ker(S^{\intercal})$; then the gap in $\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t\|\hat{\boldsymbol{x}})$ remains for $t \gg 0$ and the upper bound is relatively loose. In this scenario, the discrepancy between the upper bound and individual

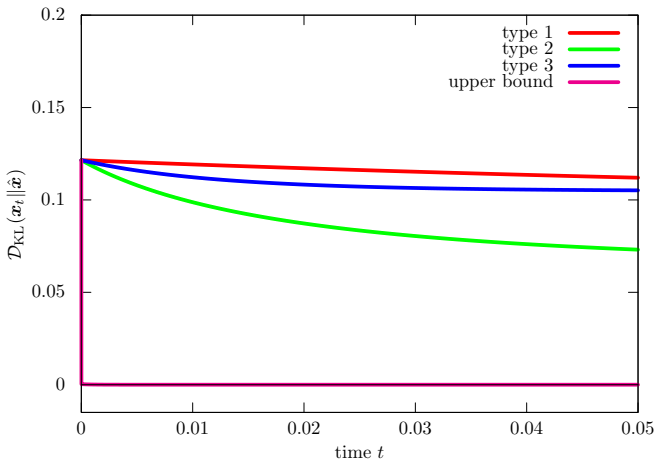


FIG. 6. Dependence of $\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t\|\hat{\boldsymbol{x}})$ on time t for the CRNs in Eq. (6.4) and its upper bound in the case of $\boldsymbol{x}_{\mathrm{eq}} \neq \hat{\boldsymbol{x}}$.

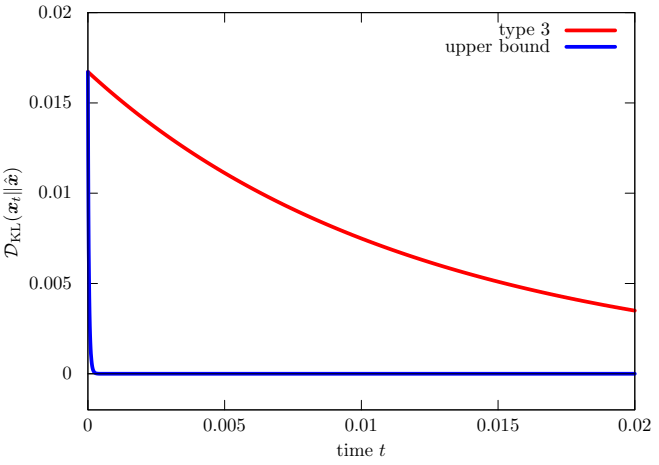


FIG. 7. Dependence of $\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t\|\hat{\boldsymbol{x}})$ on time t for the CRN in Eq. (6.4c) and its upper bound in the case of $\boldsymbol{x}_{\mathrm{eq}} = \hat{\boldsymbol{x}}$.

CRNs is broader compared to previous examples. This can be intuitively understood as follows. With a larger $N_{\mathbb{X}}$, the dimension of $\ker S^{\intercal}$ increases, leading to a greater number of conserved quantities. Conversely, in the case of the upper bound system, the dimension of $\ker S^{\intercal}$ remains unchanged. Therefore, for substantial $N_{\mathbb{X}}$ values, CRN dynamics are subject to more stringent constraints. Consequently, it is logical for the gap between the upper bound and the dynamics of a specific CRN to widen. To conclude, it is noteworthy that in the aforementioned examples, $\lim_{t\to\infty} \boldsymbol{x}_t \neq \hat{\boldsymbol{x}}$. This implies that $\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t\|\hat{\boldsymbol{x}})$ remains positive for large t.

C. Case of $N_{\mathbb{e}} > 1$

We have considered the case of $N_{\mathbb{e}} = 1$. Then, we consider the fully connected CRNs whose hypervertices are given by

$$\mathbb{V}_1 = \{\mathbb{X}_1 + \mathbb{X}_2, \mathbb{X}_2 + \mathbb{X}_3, \mathbb{X}_3 + \mathbb{X}_4, \mathbb{X}_4 + \mathbb{X}_1\}, \tag{6.4a}$$

$$\mathbb{V}_2 = \{\mathbb{X}_1 + 3\mathbb{X}_2 + 4\mathbb{X}_3, \mathbb{X}_2 + 2\mathbb{X}_3, 4\mathbb{X}_1 + \mathbb{X}_3 + \mathbb{X}_4, \mathbb{X}_1 + 3\mathbb{X}_2 + \mathbb{X}_4\}, \tag{6.4b}$$

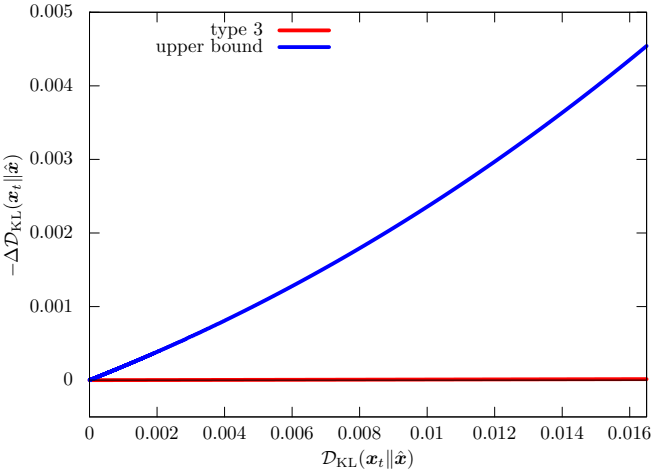


FIG. 8. Dependence of $\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t\|\hat{\boldsymbol{x}})$ on time t for the CRN in Eq. (6.4c) and its upper bound in the case of $\boldsymbol{x}_{\mathrm{eq}} = \hat{\boldsymbol{x}}$.

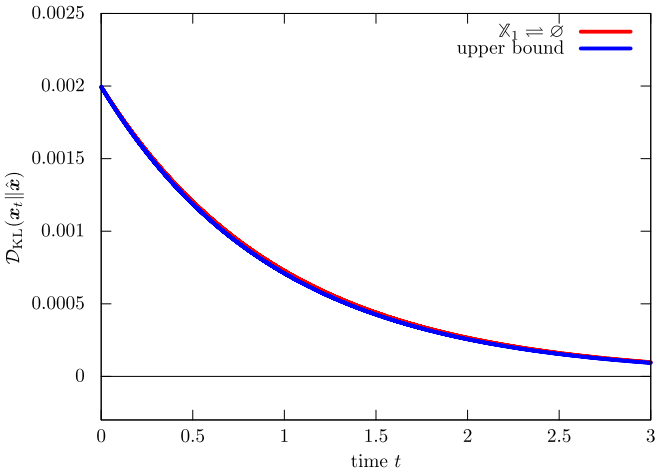


FIG. 9. Dependence of $\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t\|\hat{\boldsymbol{x}})$ on time t for the CRN in Eq. (6.5) and its upper bound.

$$\mathbb{V}_3 = \{4\mathbb{X}_1 + 3\mathbb{X}_2 + 4\mathbb{X}_3, 4\mathbb{X}_2 + 2\mathbb{X}_3 + 4\mathbb{X}_4, 4\mathbb{X}_1 + 4\mathbb{X}_3 + \mathbb{X}_4, 2\mathbb{X}_1 + 3\mathbb{X}_2 + 4\mathbb{X}_4\}. \quad (6.4c)$$

The CRNs in Eqs. (6.4) belong to the class of CRNs labeled by $N_{\mathbb{X}} = 4$, $N_{\mathbb{e}} = 6$, and $m = 4$. We call the CRNs in Eqs. (6.4) types 1, 2, and 3 from above.

We plot the dynamics of the CRNs in Eqs. (6.4) and its upper bound in the case of $\boldsymbol{x}_{\mathrm{eq}} \neq \hat{\boldsymbol{x}}$. In Fig. 6, we set $\boldsymbol{x}_{\mathrm{ini}} = [9/8, 87/80, 27/20, 27/20]^\intercal$, $\hat{\boldsymbol{x}} = [1.0, 1.0, 1.0, 1.0]^\intercal$, $k_e^{\pm} = 1$, and $\Delta t = 1.0 \times 10^{-5}$. Figure 6 clearly demonstrates the upper bound holds for $N_{\mathbb{e}} > 1$.

We show the dependence of $\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t\|\hat{\boldsymbol{x}})$ on time t for the CRN in Eq. (6.4c) and its upper bound in the case of $\boldsymbol{x}_{\mathrm{eq}} = \hat{\boldsymbol{x}}$. In Fig. 7, we set $\hat{\boldsymbol{x}} = [1.2547, 1.1021, 1.1951, 1.3388]^\intercal$. In Fig. 8, we also show the dependence of $\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t\|\hat{\boldsymbol{x}})$ on time t for the CRN in Eq. (6.4c) and its upper bound in the case of $\boldsymbol{x}_{\mathrm{eq}} = \hat{\boldsymbol{x}}$.

D. Examples of tighter bounds

Thus far, the upper bound can be rather loose, depending on numerical conditions, more specifically for large $N_{\mathbb{x}}$, $N_{\mathbb{e}}$,

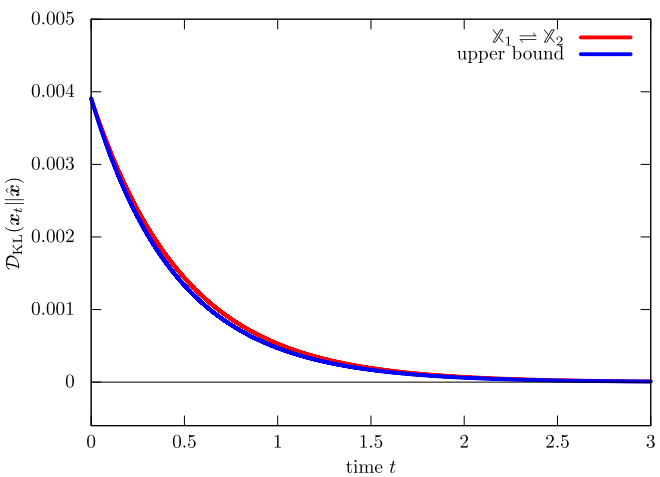


FIG. 10. Dependence of $\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t\|\hat{\boldsymbol{x}})$ on time t for the CRN in Eq. (6.6) and its upper bound.

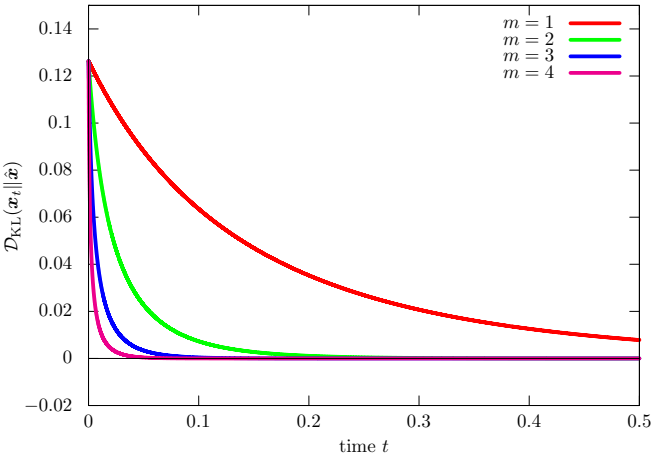


FIG. 11. Dependence of $\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t\|\hat{\boldsymbol{x}})$ for t for $m = 1, 2, 3, 4$. We set $N_{\mathbb{X}} = 4$ and $N_{\mathbb{e}} = 1$.

and m. In this section, we introduce a set of straightforward examples where the upper bound system yields a tight bound. We examine a CRN characterized by $N_{\mathbb{e}} = 1$, $m = 1$, and $N_{\mathbb{X}} = 1$:

$$\mathbb{X}_1 \rightleftharpoons \varnothing. \quad (6.5)$$

In Fig. 9, we plot the dynamics of Eq. (6.5) and its upper bound. We set $\Delta t = 1.0 \times 10^{-4}$, $\boldsymbol{x}_{\mathrm{ini}} = [15/16]$, and $\hat{\boldsymbol{x}} = [1.0]$.

We consider the following CRN with $N_{\mathbb{e}} = 1$, $m = 1$, and $N_{\mathbb{X}} = 2$:

$$\mathbb{X}_1 \rightleftharpoons \mathbb{X}_2. \quad (6.6)$$

In Fig. 10, we plot the dynamics of Eq. (6.6) and its upper bound. We set $\Delta t = 1.0 \times 10^{-4}$, $\boldsymbol{x}_{\mathrm{ini}} = [15/16, 17/16]^\intercal$, and $\hat{\boldsymbol{x}} = [1.0, 1.0]^\intercal$.

E. Comparison of the upper bounds

In this section, we examine the behavior of the upper bound under varying parameters. The parameters are

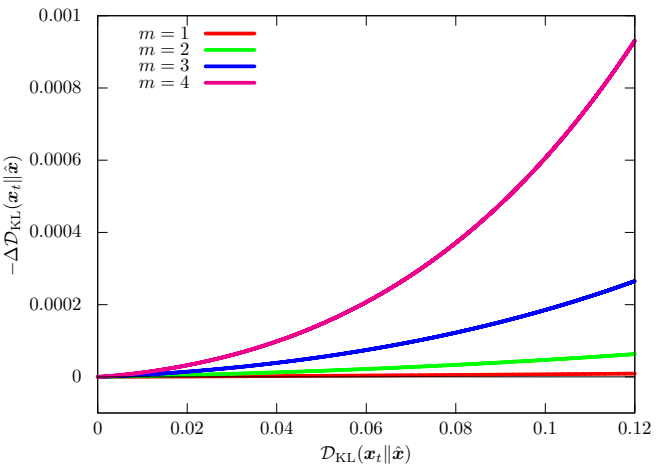


FIG. 12. Relationship between $\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t\|\hat{\boldsymbol{x}})$ and $-\Delta\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t\|\hat{\boldsymbol{x}})$ for $N_{\mathbb{X}} = 4$ and $N_{\mathbb{e}} = 1$.

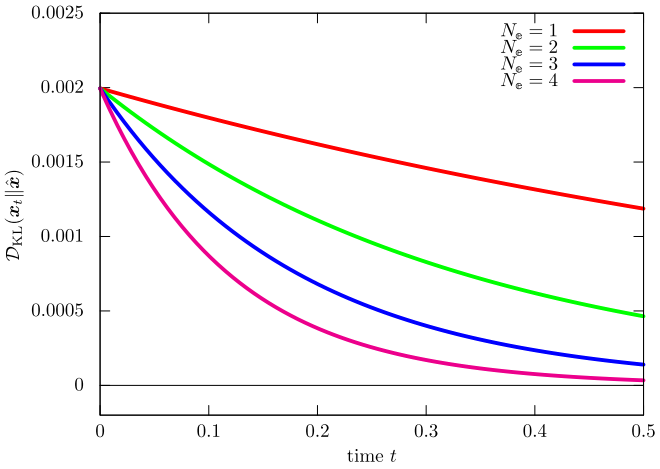


FIG. 13. Dependence of $\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t\|\hat{\boldsymbol{x}})$ for t for $m=1,2,3,4$. We set $N_{\mathbb{X}}=1$ and $m=1$.

$N_{\mathbb{X}}=4$, $N_{\mathrm{e}}=1$, $\boldsymbol{x}_{\mathrm{ini}}=[3/4,3/4,5/4,5/4]^{\intercal}$, and $\boldsymbol{x}_{\mathrm{eq}}=[1.0,1.0,1.0,1.0]^{\intercal}$. Figure 11 depicts the dependence of $\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t\|\hat{\boldsymbol{x}})$ on t for $m=1,2,3,4$. Figure 12 portrays the relationship between $\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t\|\hat{\boldsymbol{x}})$ and $-\Delta\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t\|\hat{\boldsymbol{x}})$ for $N_{\mathbb{X}}=4$ and $N_{\mathrm{e}}=1$. The figures indicate that higher values of m are associated with increased rates of convergence. This behavior is consistent with the expectation that nonlinearity in CRNs tends to influence reaction rates.

We further confirm this trend by varying $N_{\mathbb{X}}$, N_{e}, and m, respectively. We consider the case of $N_{\mathbb{X}}=1$ and $m=1$. In Fig. 13, we plot the dynamics of the upper bound system. We set $\Delta t=1.0\times10^{-4}$, $\boldsymbol{x}_{\mathrm{ini}}=[15/16]$, and $\hat{\boldsymbol{x}}=[1.0]$. We consider the case of $N_{\mathbb{X}}=1$ and $N_{\mathrm{e}}=1$. In Fig. 14, we plot the dynamics of the upper bound system. We set $\Delta t=1.0\times10^{-4}$, $\boldsymbol{x}_{\mathrm{ini}}=[15/16]$, and $\hat{\boldsymbol{x}}=[1.0]$. We consider the case of $N_{\mathrm{e}}=1$ and $m=1$. In Fig. 15, we plot the dynamics of the upper bound system. We set $\Delta t=1.0\times10^{-4}$, $\boldsymbol{x}_{\mathrm{ini}}=[15/16,\underbrace{1.0,\ldots,1.0}_{N_{\mathbb{X}}-1}]^{\intercal}$, and $\hat{\boldsymbol{x}}=[1.0,\underbrace{1.0,\ldots,1.0}_{N_{\mathbb{X}}-1}]^{\intercal}$. As shown in Figs. 13–15, the upper bound system becomes faster with $N_{\mathbb{X}}$, N_{e}, and m increased, respectively.

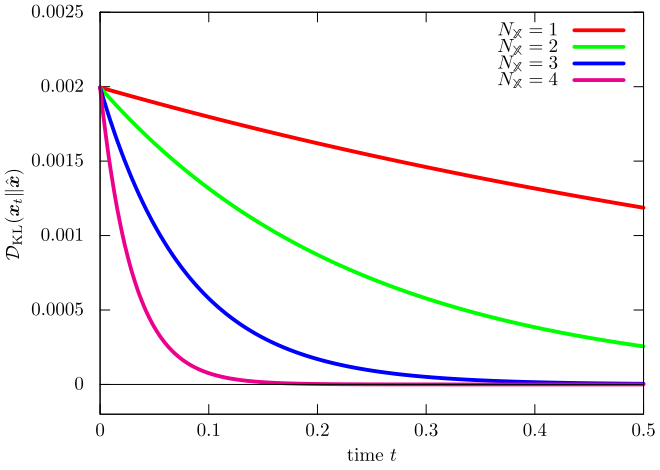


FIG. 14. Dependence of $\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t\|\hat{\boldsymbol{x}})$ for t for $m=1,2,3,4$. We set $N_{\mathbb{X}}=1$ and $N_{\mathrm{e}}=1$.

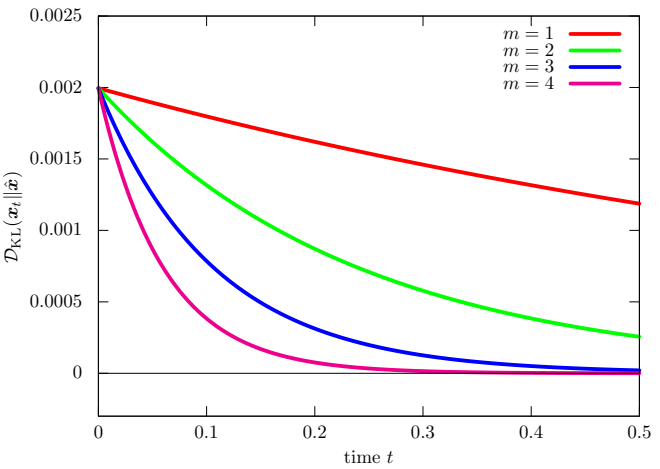


FIG. 15. Dependence of $\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t\|\hat{\boldsymbol{x}})$ for t for $m=1,2,3,4$. We set $N_{\mathrm{e}}=1$ and $m=1$.

VII. CONCLUSIONS

In this paper, we developed a framework based on NG to establish an upper bound on the dynamics of a specific subset of CRNs. The nonlinearity commonly present in CRNs presents a challenge, which is addressed here. While the primary focus has been on CRNs, the methods and discussions are applicable to a wider range of hypergraph dynamics. The paper holds implications for fields beyond chemistry and physics, including information science and machine learning.

ACKNOWLEDGMENTS

H.M. was supported by JSPS KAKENHI Grant No. JP23H04489. T.J.K. was supported by JST (Grants No. JPMJCR2011 and No. JPMJCR1927) and JSPS (Grant No. 19H05799). L.-S.B. was partially funded by NSF Award No. CHE-2002313.

APPENDIX A: DERIVATION OF THE KL DIVERGENCE FROM THE BREGMAN DIVERGENCE

In this Appendix, we show that the Bregman divergence, Eq. (2.34), with Eq. (2.36) is equivalent to the KL divergence, Eq. (2.37). Let us define the following potential for $\alpha\in\mathbb{R}$:

$$\phi^{(\alpha)}_{\mathrm{KL}}(\boldsymbol{x}) := \sum_{i=1}^{N_{\mathbb{X}}} x_i(\ln x_i-\alpha). \tag{A1}$$

The Bregman divergence, Eq. (2.34), with Eq. (A1) is computed as follows:

$$\mathcal{D}_{\phi^{(\alpha)}_{\mathrm{KL}}}(\boldsymbol{x}\|\boldsymbol{y}) = \phi^{(\alpha)}_{\mathrm{KL}}(\boldsymbol{x})-\phi^{(\alpha)}_{\mathrm{KL}}(\boldsymbol{y})-\langle(\boldsymbol{x}-\boldsymbol{y}),\nabla\phi^{(\alpha)}_{\mathrm{KL}}(\boldsymbol{y})\rangle \tag{A2}$$

$$= \sum_{i=1}^{N_{\mathbb{X}}} x_i(\ln x_i-\alpha)-\sum_{i=1}^{N_{\mathbb{X}}} y_i(\ln y_i-\alpha)$$

$$-\sum_{i=1}^{N_{\mathbb{X}}}(x_i-y_i)(\ln y_i-\alpha+1) \tag{A3}$$

$$= \sum_{i=1}^{N_{\mathrm{X}}} x_i \ln x_i - \sum_{i=1}^{N_{\mathrm{X}}} y_i \ln y_i$$
$$- \sum_{i=1}^{N_{\mathrm{X}}} (x_i - y_i) \ln y_i - \sum_{i=1}^{N_{\mathrm{X}}} (x_i - y_i) \tag{A4}$$
$$= \sum_{i=1}^{N_{\mathrm{X}}} x_i \ln \frac{x_i}{y_i} - \sum_{i=1}^{N_{\mathrm{X}}} (x_i - y_i) \tag{A5}$$
$$= \mathcal{D}_{\mathrm{KL}}(\boldsymbol{x} \| \boldsymbol{y}). \tag{A6}$$

Thus, the Bregman divergence, Eq. (2.34), with Eq. (A1) is equivalent to the KL divergence, Eq. (2.37), independently from α. Furthermore, Eq. (2.36) is the special case of Eq. (A1) with $\alpha = 0$.

APPENDIX B: UPPER BOUND SYSTEM WITH THE L^2 CONSTRAINT

In Sec. V, we have considered $\phi_{\mathrm{KL}}(\cdot)$, Eq. (2.36), as the potential of the Bregman divergence in the constraint term since the KL divergence is minimized in CRNs. However, we are not limited to this choice, and it is expected that a different potential in the constraint may give us a different bound. Another simple candidate for the potential of the Bregman divergence in the constraint is the L^2 norm given by

$$\phi_{L^2}(\boldsymbol{x}) \coloneqq \sum_{i=1}^{N_{\mathrm{X}}} |x_i|^2. \tag{B1}$$

In this case, $\mathcal{D}_{\mathrm{KL}}(\boldsymbol{x}_t + \|\dot{\boldsymbol{x}}_t\|_{\mathrm{F}} \boldsymbol{e}_t \Delta t \| \boldsymbol{x}_t)$ does not depend on $\boldsymbol{e}_t$ and the Hessian $G_{\phi_{L^2}}(\boldsymbol{x}_t)$ becomes the identity matrix: $G_{\phi_{L^2}}(\boldsymbol{x}_t) = \mathbb{1}$.

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

[2] N. Shiraishi, *An Introduction to Stochastic Thermodynamics: From Basic to Advanced* (Springer Nature, Shingapore, 2023), Vol. 212.

[3] R. Kawai, J. M. R. Parrondo, and C. Van den Broeck, Dissipation: The phase-space perspective, Phys. Rev. Lett. **98**, 080602 (2007).

[4] H. Miyahara and K. Aihara, Work relations with measurement and feedback control on nonuniform temperature systems, Phys. Rev. E **98**, 042138 (2018).

[5] A. C. Barato and U. Seifert, Thermodynamic uncertainty relation for biomolecular processes, Phys. Rev. Lett. **114**, 158101 (2015).

[6] 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).

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

[8] T. Van Vu and Y. Hasegawa, Uncertainty relations for underdamped Langevin dynamics, Phys. Rev. E **100**, 032130 (2019).

[9] T. Van Vu and K. Saito, Thermodynamic unification of optimal transport: Thermodynamic uncertainty relation, minimum dissipation, and thermodynamic speed limits, Phys. Rev. X **13**, 011013 (2023).

[10] N. Ohga, S. Ito, and A. Kolchinsky, Thermodynamic bound on the asymmetry of cross-correlations, Phys. Rev. Lett. **131**, 077101 (2023).

[11] T. Van Vu, V. Tuan Vo, and K. Saito, Dissipation, quantum coherence, and asymmetry of finite-time cross-correlations, Phys. Rev. Res. **6**, 013273 (2024).

[12] A. Dechant, J. Garnier-Brun, and S.-i. Sasa, Thermodynamic bounds on correlation times, arXiv:2303.13038.

[13] T. J. Kobayashi and Y. Sughiyama, Fluctuation relations of fitness and information in population dynamics, Phys. Rev. Lett. **115**, 238102 (2015).

[14] H. Miyahara, Many-body perturbation theory and fluctuation relations for interacting population dynamics, Phys. Rev. E **99**, 042415 (2019).

[15] Y. Sughiyama and T. J. Kobayashi, Steady-state thermodynamics for population growth in fluctuating environments, Phys. Rev. E **95**, 012131 (2017).

[16] H. Miyahara, Steady-state thermodynamics for population dynamics in fluctuating environments with side information, J. Stat. Mech.: Theory Exp. (2022) 013501.

[17] K. Adachi, R. Iritani, and R. Hamazaki, Universal constraint on nonlinear population dynamics, Commun. Phys. **5**, 129 (2022).

[18] M. Hoshino, R. Nagayama, K. Yoshimura, J. F. Yamagishi, and S. Ito, Geometric speed limit for acceleration by natural selection in evolutionary processes, Phys. Rev. Res. **5**, 023127 (2023).

[19] T. J. Kobayashi, D. Loutchko, A. Kamimura, S. A. Horiguchi, and Y. Sughiyama, Information geometry of dynamics on graphs and hypergraphs, Inf. Geom. (2023), doi:10.1007/s41884-023-00125-w.

[20] 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).

[21] T. J. Kobayashi, D. Loutchko, A. Kamimura, and Y. Sughiyama, Hessian geometry of nonequilibrium chemical reaction networks and entropy production decompositions, Phys. Rev. Res. **4**, 033208 (2022).

[22] T. Okada and A. Mochizuki, Law of localization in chemical reaction networks, Phys. Rev. Lett. **117**, 048101 (2016).

[23] Y. Hirono, T. Okada, H. Miyazaki, and Y. Hidaka, Structural reduction of chemical reaction networks based on topology, Phys. Rev. Res. **3**, 043123 (2021).

[24] Y. Hirono, T. Okada, H. Miyazaki, and Y. Hidaka, Structural reduction of hypergraphs based on topology, JPS Conf. Proc. **36**, 011008 (2021), Proceedings of Blockchain in Kyoto 2021 (BCK21).

[25] Y. Hirono, H. Hong, and J. K. Kim, Robust perfect adaptation of reaction fluxes ensured by network topology, arXiv:2302.01270.

[26] S.-i. Amari and H. Nagaoka, *Methods of Information Geometry* (American Mathematical Society, 2000), Vol. 191.

[27] S.-i. Amari, *Information Geometry and Its Applications*, Applied Mathematical Sciences, Vol. 194 (Springer, Japan, 2016).

[28] S.-I. Amari, Natural gradient works efficiently in learning, Neural Comput. **10**, 251 (1998).
[29] Y. Wang and W. Li, Accelerated information gradient flow, J. Sci. Comput. **90**, 11 (2022).
[30] S. A. Horiguchi and T. J. Kobayashi, Cellular gradient flow structure linking single-cell-level rules and population-level dynamics, Phys. Rev. Res. **5**, L022052 (2023).
[31] J. Karbowski, Bounds on the rates of statistical divergences and mutual information, arXiv:2308.05597.
[32] T. Schmiedl and U. Seifert, Stochastic thermodynamics of chemical reaction networks, J. Chem. Phys. **126**, 044101 (2007).
[33] K. Yoshimura and S. Ito, Thermodynamic uncertainty relation and thermodynamic speed limit in deterministic chemical reaction networks, Phys. Rev. Lett. **127**, 160601 (2021).
[34] K. Yoshimura and S. Ito, Information geometric inequalities of chemical thermodynamics, Phys. Rev. Res. **3**, 013175 (2021).
[35] T. Van Vu and K. Saito, Topological speed limit, Phys. Rev. Lett. **130**, 010402 (2023).
[36] D. A. Beard and H. Qian, *Chemical Biophysics: Quantitative Analysis of Cellular Systems* (Cambridge University Press, Cambridge, 2008), Vol. 126.
[37] M. Feinberg, *Foundations of Chemical Reaction Network Theory* (Springer International Publishing, Germany, 2019).
[38] R. Rao and M. Esposito, Nonequilibrium thermodynamics of chemical reaction networks: Wisdom from stochastic thermodynamics, Phys. Rev. X **6**, 041064 (2016).
[39] R. Hamming, *Numerical Methods for Scientists and Engineers* (Dover Publications, United States, 2012).