Geometric thermodynamics of reaction-diffusion systems: Thermodynamic trade-off relations and optimal transport for pattern formation

Ryuna Nagayama [1,2,*] Kohei Yoshimura [1,2] Artemy Kolchinsky [3,2] and Sosuke Ito [1,2]
[1]*Department of Physics, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*
[2]*Universal Biology Institute, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*
[3]*ICREA-Complex Systems Lab, Pompeu Fabra University, 08003 Barcelona, Spain*

We establish universal relations between pattern formation and dissipation with a geometric approach to nonequilibrium thermodynamics of deterministic reaction-diffusion systems. We first provide a way to systematically decompose the entropy production rate based on the orthogonality of thermodynamic forces, thereby identifying the amount of dissipation caused by each factor. This enables us to extract the excess entropy production rate that genuinely contributes to the time evolution of patterns. We also show that a similar geometric method further decomposes the entropy production rate into detailed contributions, e.g., the dissipation from each point in real or wavenumber space. Second, we relate the excess entropy production rate to the details of the change in patterns through two types of thermodynamic trade-off relations for reaction-diffusion systems: thermodynamic speed limits and thermodynamic uncertainty relations. The former relates dissipation and the speed of pattern formation, and the latter bounds the excess entropy production rate with partial information on patterns, such as specific Fourier components of concentration distributions. In connection with the derivation of the thermodynamic speed limits, we also extend optimal transport theory to reaction-diffusion systems, which enables us to measure the speed of the time evolution. This extension of optimal transport also solves the minimization problem of the dissipation associated with the transition between two patterns, and constructs energetically efficient protocols for pattern formation. We numerically demonstrate our results using chemical traveling waves in the Fisher–Kolmogorov–Petrovsky–Piskunov equation and changes in symmetry in the Brusselator model. Our results apply to general reaction-diffusion systems and contribute to understanding the relations between pattern formation and unavoidable dissipation.

I. INTRODUCTION

A. Background and motivation

Reaction-diffusion systems (RDSs) have been used to study the formation of various spatiotemporal patterns [1–7] since the pioneering study of Turing [8]. In nature, pattern formation by RDSs achieves a variety of functions. For example, organisms use the reactions and diffusive dynamics of biomolecules for morphogenesis [4,5,9]. Cells also accelerate biochemical reactions through phase separation, which is interpretable as an RDS [10,11]. From an engineering perspective, we can also apply RDSs to biomimetic materials [12], computation [13–15], and information processing [16].

In achieving such functions, it is generally crucial to minimize costs associated with functions. One of the fundamental costs is energy dissipation, measured with the entropy production required to change a given pattern to a desired pattern through reactions and diffusion. In particular, biological systems must reduce energy dissipation to the extent that it does not alter the desired function, since they have access to only limited energy resources [17]. To consider such a minimization of entropy production, we clarify how dissipation relates to the spatiotemporal dynamics of patterns based on nonequilibrium thermodynamics of RDSs.

Most attempts to relate the time evolution of patterns to the dissipation are limited to qualitative observations of some specific systems [18–21] and are considerably less advanced. This may be because of the historical context: nonequilibrium thermodynamics of RDSs originated in the study of dissipative structures by Prigogine and his collaborators, who focused on steady-state patterns and stability rather than the nonstationary change in patterns [22,23]. Unfortunately, thermodynamics cannot predict stable steady-state patterns except in some special cases [22,24], and has been replaced by methods based on dynamical systems [25–27]. On the other hand, nonequilibrium thermodynamics itself has continued to develop away from dissipative structures. A major development was the establishment of stochastic thermodynamics, which deals with mesoscopic systems [28–30]. Stochastic thermodynamics has developed some methods to quantitatively connect time evolution with dissipation and has considered minimum dissipation problems for stochastic systems [31–46].

*Contact author: ryuna.nagayama@ubi.s.u-tokyo.ac.jp

More recently, developments from stochastic thermodynamics were imported into thermodynamics of chemical systems, including RDSs [47–49]. This has provided a thermodynamic framework for traveling waves [50] and phase separation [51] and reveals thermodynamic constraints on a particular class of RDSs that conserve mass [52]. Still, the universal relations between the time evolution of the patterns and the dissipation, which is valid in general RDSs, have not been found.

In this study, we reveal universal relations between the time evolution of patterns and the dissipation in RDSs by applying the wisdom of stochastic thermodynamics to general deterministic RDSs. Our analysis also extends to systems driven by particle exchange with the environment and external mechanical forces [53,54] and nonideal mixtures such as those describing phase separation [24,51,55–57]. Our first result is a decomposition of EPR according to different sources, which helps us to understand quantitatively where and how much dissipation occurs during the time evolution. One decomposition that has attracted particular attention is that of excess EPR, which becomes zero in steady state, versus housekeeping EPR, which remains positive even in the steady state [58]. Although such a decomposition is not unique [59–61], we mainly focus on the geometric decomposition because it enables us to extract the part that genuinely contributes to time evolution as excess EPR [62,63]. Our second result is a set of thermodynamic trade-off relations, universal inequalities that connect details of time evolution and dissipation. Although various types of trade-off relations may be derived, we focus specifically on thermodynamic speed limits (TSLs), that relate the speed of time evolution and dissipation [39–45,64–69], and thermodynamic uncertainty relations (TURs), that relate precision and dissipation [70,71].

In addition, we extend optimal transport theory, which deals with the transport between probability distributions [72,73], to RDSs, where the concentration distribution changes through reactions and diffusion. This extension is directly related to the problem of minimizing dissipation in pattern formation. In particular, we focus on the Wasserstein distance, a fundamental quantity in optimal transport theory [74–76] that also plays a central role in modern stochastic thermodynamics. For example, the Wasserstein distance determines the minimum dissipation [39–45,77] and has gained attention in various problems, such as optimal control of thermal engines [41,78,79] and information thermodynamics [41,44,80,81]. It also enables us to reinterpret the geometric excess/housekeeping decomposition from the perspective of optimal transport [45,62,63]. Furthermore, measuring the speed of the time evolution with the Wasserstein distance yields various TSLs [39–45,65–68] for stochastic systems.

We extend results from stochastic thermodynamics to deterministic RDSs by focusing on the common geometric structure of nonequilibrium thermodynamics, called *geometric thermodynamics* [82]. This geometric approach makes it possible to extend results for stochastic systems to deterministic systems as follows. The geometric excess/housekeeping decomposition was originally obtained by focusing on a geometry of thermodynamic forces for Langevin systems [62,63,82]. Because the geometry of thermodynamic forces is common to broad systems, the decomposition has been extended to deterministic CRNs [45] and fluid systems [83]. The

geometric interpretation of thermodynamic forces has also induced TURs for such deterministic systems [45,66,83,84]. In addition, generalization of the Wasserstein distance has led to the TSLs for deterministic CRNs by relating the excess EPR and the speed of time evolution [45,67].

B. Road map

All results in this paper are based on the similarity between Langevin systems and RDSs [Fig. 1(a)]. Before discussing the results for RDSs, we begin this paper by explaining the framework of geometric thermodynamics of Langevin systems and other relevant background, including optimal transport theory, in Sec. II. Although this part is essentially an aid to understanding the main result for RDSs, it also contains a new result reimported from RDSs, the wavenumber decomposition of the EPR (Sec. II C).

In Sec. III, we introduce nonequilibrium thermodynamics of RDSs. This preliminary section also provides vector notation, inner products, and generalizations of differential operators to simplify the description and calculation (Sec. III D). The notation helps us establish an analogy between stochastic systems and deterministic RDSs. In particular, the core of the subsequent results is that the EPR is expressed as the squared norm of the thermodynamic force.

We provide a way to decompose EPR using orthogonal decompositions of the thermodynamic force in Sec. IV. This geometric method enables us to systematically decompose the EPR into various contributions. In particular, extending the geometric excess/housekeeping decomposition [62,63,82] to RDSs [Fig. 1(b)] is important, since the excess part extracts the dissipation that truly contributes to the time evolution. It is obtained by projecting the thermodynamic force onto the space of conservative forces, which describe the relaxation (Sec. IV A). We also show that a similar approach based on orthogonality decomposes EPR into contributions from each point in the real or wavenumber space in Sec. IV B [Fig. 1(c)]. The latter, named wavenumber decomposition, has been unexplored even in stochastic thermodynamics. Finally, we numerically demonstrate the usefulness of the decompositions with the Fisher–Kolmogorov–Petrovsky–Piskunov (Fisher–KPP) equation and the Brusselator model, which show the appearance of a chemical traveling wave and change in the symmetry of patterns, respectively (Sec. IV C). The numerical demonstration reveals the difference in how the EPR and the excess EPR reflect the structure of the patterns and its time evolution by combining the excess/housekeeping decomposition with the local and wavenumber decompositions. In subsequent parts of this paper, we mainly study the details of the excess EPR to focus on the relations between the time evolution of patterns and the dissipation.

In Sec. V, we establish the optimal transport theory for RDSs and use it to establish TSLs [Fig. 1(d)] and minimum dissipation formulas for RDSs. In stochastic thermodynamics, two types of Wasserstein distances, the 1-Wasserstein and 2-Wasserstein distances, have provided different insights. We extend both of them to RDSs, which enables us to measure the distance between patterns and the speed of change in patterns. The extension of the 2-Wasserstein distance is an improvement over previous attempts [85,86] (Sec. V B), and
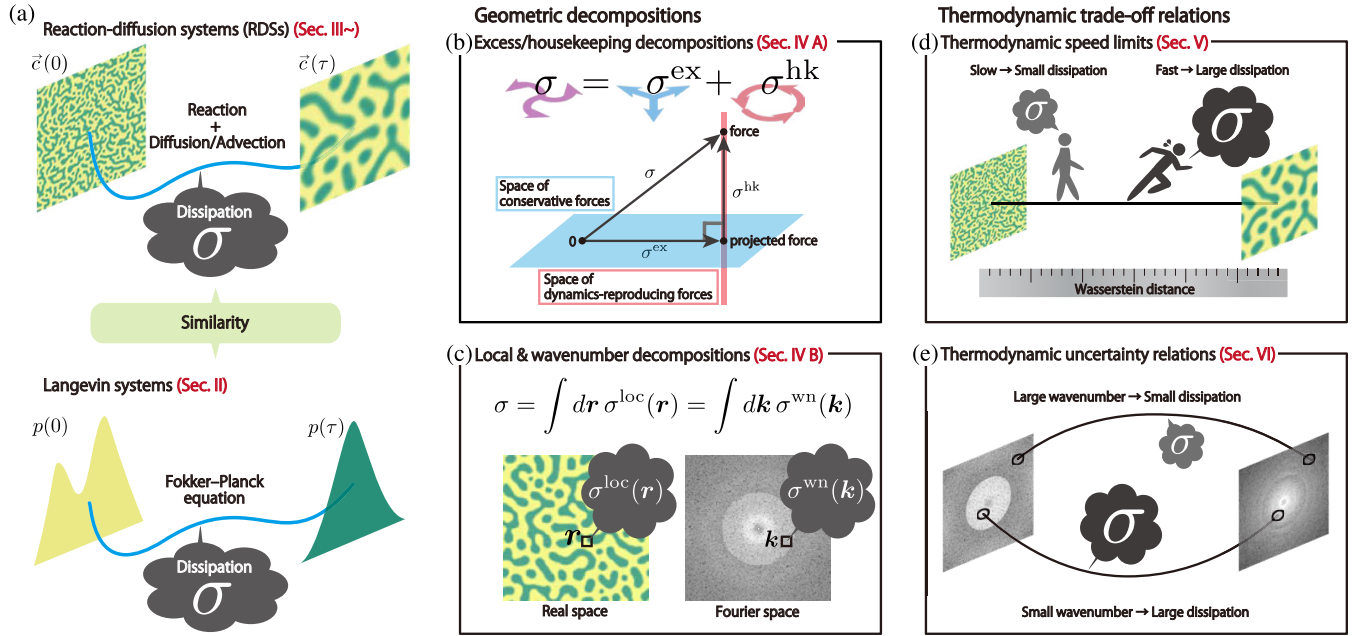
FIG. 1. Schematics of the results in this paper. The red letters in the figure indicate the corresponding chapters. (a) The similarity between RDSs and Langevin systems. For Langevin systems, we consider the time evolution from the probability $p(0)$ at time $t = 0$ to the probability $p(\tau)$ at time $t = \tau$. For RDSs, we consider the time evolution from the initial concentration $\vec{c}(0)$ to the final concentration $\vec{c}(\tau)$. Dissipation is quantified by the EPR σ for both systems. (b) The excess/housekeeping decomposition, the local decomposition, and the wavenumber decomposition. The EPR σ is decomposed into the excess EPR σ^{ex} and the housekeeping EPR σ^{hk}, which correspond to the contributions of the conservative force and the nonconservative force, respectively. This decomposition is introduced by projecting the force onto the space of conservative forces. (c) The EPR σ is also given by the integral of the local EPR $\sigma^{\mathrm{loc}}(\boldsymbol{r})$ in real space, and the integral of the wavenumber EPR $\sigma^{\mathrm{wn}}(\boldsymbol{k})$ in Fourier space. (d) TSLs refer to a trade-off between dissipation and speed. If the speed of the transition between initial and final states is slower, the dissipation can be smaller. The distance between initial and final states is measured by the Wasserstein distance and the minimal amount of dissipation also depends on this distance. (e) TURs refer to the trade-off between dissipation and the change in the spatial pattern. More dissipation is required to change the mode corresponding to smaller wavenumbers.

the extension of the 1-Wasserstein distance is essentially new (Sec. V A). The 2-Wasserstein distance leads to a series of TSLs, because it is closely related to excess EPR. It universally links the dissipation and the speed of the change in patterns. In addition, we obtain another series of TSLs based on the 1-Wasserstein distance by constructing a new quantity that measures the intensity of diffusion and reaction consistent with thermodynamics. It gives trade-off relations between three pieces of information: the speed of the time evolution, the dissipation, and the intensity of reaction and diffusion (Sec. V D). We show that the lower bounds given by the TSLs for the 1- and 2-Wasserstein distances are the minimum dissipation achievable under some different conditions (Sec. V E). We also provide protocols for achieving the minimum dissipation, confirming that generalized optimal transport theory can be applied to RDSs. Finally, we numerically demonstrate the TSLs and the minimum dissipation with the two systems in Sec. V F. In particular, we obtain several quantities related to the 1-Wasserstein distance analytically for the Fisher–KPP equation. The content of Sec. V integrates the two different approaches recently developed in stochastic thermodynamics based on optimal transport [44,45] within RDSs.

We also derive TURs, which bound the excess EPR using partial information about the system, in Sec. VI. In particular, we discover a myriad of TURs underlying the TUR for well-

mixed CRNs [84] by considering the Fourier transform of the concentration distribution (Sec. VI B). These TURs reveal that more dissipation is required to change the mode corresponding to smaller wavenumbers, i.e., lower spatial frequencies [Fig. 1(e)]. We numerically demonstrate the TURs using the Brusselator model, which shows a notable change in symmetry (Sec. VI C). It confirms that the TURs give lower bounds on the EPR, reflecting the change in the spatial structure of the patterns. In addition, we also compare the TURs with the wavenumber decomposition, which gives lower bounds on the EPR depending on the wavenumber.

II. BACKGROUND: GEOMETRIC THERMODYNAMICS FOR LANGEVIN SYSTEMS

Before proceeding to RDSs, we briefly introduce geometric thermodynamics for Langevin systems [82]. Although this section mainly consists of the existing results, it also includes a novel result, the wavenumber decomposition of the EPR in Sec. II C.

In the following, we consider a Brownian particle in a d-dimensional Euclidean space $\mathbb{R}^d$. We assume that the temperature is homogeneous and we set the temperature and Boltzmann's constant to unity for simplicity. The following Langevin equation describes the time evolution of the position

of the particle:

$$d_t \check{r}(t) = -D\nabla_r U(\check{r}(t); t) + DK^{\mathrm{nc}}(\check{r}(t); t) + \sqrt{2D}\xi(t), \quad (1)$$

where $d_t = d/dt$ stands for the time derivative, $\check{r}(t)$ indicates the position of the Brownian particle at time t, ∇_r is the differential operator for spatial coordinates $r \in \mathbb{R}^d$, $-\nabla_r U(\check{r}(t); t) := -\nabla_r U(r; t)|_{r=\check{r}(t)}$ indicates the potential force on the particle, $K^{\mathrm{nc}}(\check{r}(t); t)$ indicates the nonconservative mechanical force on the particle, D is the diffusion constant that is given by the mobility and temperature, and $\xi(t) = (\xi_i(t))_{i=1}^d$ is the white Gaussian noise satisfying $\langle \xi_i(t) \rangle = 0$ and $\langle \xi_i(t)\xi_j(t') \rangle = \delta_{ij}\delta(t - t')$. In later sections, we will use some of the symbols introduced in this section to represent their counterparts in RDSs.

A. Fokker–Planck equation and entropy production rate

The probability density $p(r; t)$ of the Brownian particle at position r at time t described by the Langevin equation (1) evolves according to the Fokker–Planck equation,

$$\partial_t p(r; t) = -\nabla_r \cdot J(r; t), \quad (2)$$

$$J(r; t) := Dp(r; t)[-\nabla_r(U(r; t) + \ln p(r; t)) + K^{\mathrm{nc}}(r; t)], \quad (3)$$

where $\partial_t = \partial/\partial t$ stands for the partial time derivative and $J(r; t)$ is the thermodynamic current. Here, $p(r; t)$ is the probability density and thus $\int_{\mathbb{R}^d} dr\, p(r; t) = 1$ and $p(r; t) \geqslant 0$ hold. In the following, we assume the following boundary conditions: $p(r; t)$ and its derivatives vanish as $\|r\| \to \infty$, where $\|\cdot\|$ is the Euclidean norm.

Defining thermodynamic force $F(r; t)$ as

$$F(r; t) := -\nabla_r(U(r; t) + \ln p(r; t)) + K^{\mathrm{nc}}(r; t), \quad (4)$$

the thermodynamic current J and force F satisfy a linear relation,

$$J(r; t) = \mathsf{M}(r; t)F(r; t), \quad (5)$$

where the $d \times d$ positive-definite matrix $\mathsf{M}_{ij}(r; t) := Dp(r; t)\delta_{ij}$ indicates the mobility tensor. We can rewrite the Fokker–Planck equation in Eq. (2) as $\partial_t p = -\nabla_r \cdot (\mathsf{M}F)$.

The EPR σ for Langevin systems is written as an inner product of J and F or a squared norm of F,

$$\sigma := \langle\langle J, F \rangle\rangle = \langle\langle \mathsf{M}F, F \rangle\rangle = \langle\langle F, F \rangle\rangle_{\mathsf{M}}$$

$$= \int_{\mathbb{R}^d} dr\, Dp(r; t)\|F(r; t)\|^2, \quad (6)$$

where the inner products $\langle\langle \cdot, \cdot \rangle\rangle$ and $\langle\langle \cdot, \cdot \rangle\rangle_{\mathsf{M}}$ are defined as $\langle\langle J', F' \rangle\rangle := \int_{\mathbb{R}^d} dr\, J'(r) \cdot F'(r)$ and $\langle\langle F', F'' \rangle\rangle_{\mathsf{M}} := \langle\langle \mathsf{M}F', F'' \rangle\rangle$ for all vector-valued functions $J'(r)$, $F'(r)$ and $F''(r)$ that take values at $\mathbb{R}^d$. The mobility tensor M can be regarded as the metric tensor because M is positive-definite. We also introduce the EP Σ_τ as $\Sigma_\tau := \int_0^\tau dt\, \sigma$.

B. Geometric excess/housekeeping decomposition of entropy production rate for Langevin systems

The Langevin system in Eq. (1) is driven by two contributions: one is the conservative contribution, which

causes relaxation to the equilibrium state corresponding to $U(r; t)$, and the other is the nonconservative contribution due to $K^{\mathrm{nc}}(r; t)$, which keeps the system out of equilibrium even in the steady state. The thermodynamic force F in Eq. (4) contains the two contributions, the conservative force $-\nabla_r(U(r; t) + \ln p(r; t))$, which is the gradient of $-(U(r; t) + \ln p(r; t))$, and the nonconservative force K^{nc}. Thus, the EPR $\sigma = \langle\langle F, F \rangle\rangle_{\mathsf{M}}$ quantifies these two contributions simultaneously.

To quantify the conservative and nonconservative contributions separately, we construct the geometric decomposition of the EPR by utilizing the generalized Pythagorean theorem for the force space with the inner product $\langle\langle, \rangle\rangle_{\mathsf{M}}$,

$$\langle\langle F, F \rangle\rangle_{\mathsf{M}} = \langle\langle F^*, F^* \rangle\rangle_{\mathsf{M}} + \langle\langle F - F^*, F - F^* \rangle\rangle_{\mathsf{M}}, \quad (7)$$

which is valid when we decompose F into two orthogonal parts F^* and $F - F^*$, satisfying

$$\langle\langle F^*, F - F^* \rangle\rangle_{\mathsf{M}} = 0. \quad (8)$$

The force F^* that allows for the Pythagorean theorem (7) is not unique [63]. We focus on F^*, which enables us to regard $\langle\langle F^*, F^* \rangle\rangle_{\mathsf{M}}$ and $\langle\langle F - F^*, F - F^* \rangle\rangle_{\mathsf{M}}$ as dissipation owing to conservative and nonconservative forces, respectively. For this purpose, we assume that $F^*(r; t)$ is the gradient of a potential as $F^*(r; t) = \nabla_r \phi^*(r; t)$, inspired by the original form of the conservative force $-\nabla_r(U(r; t) + \ln p(r; t))$. Then, we can derive the condition on ϕ^* as the sufficient condition for the orthogonality in Eq. (8),

$$\nabla_r \cdot (\mathsf{M}\nabla_r \phi^*) = \nabla_r \cdot (\mathsf{M}F), \quad (9)$$

which lets us determine F^* uniquely (see Appendix A 1 for details).

Using the decomposition of F into the conservative part F^* and the orthogonal part $F - F^*$, we define the excess and housekeeping EPRs as $\sigma^{\mathrm{ex}} := \langle\langle F^*, F^* \rangle\rangle_{\mathsf{M}}$ and $\sigma^{\mathrm{hk}} := \langle\langle F - F^*, F - F^* \rangle\rangle_{\mathsf{M}}$, respectively. These EPRs are nonnegative because they are represented by a squared norm. Then, the Pythagorean theorem (7) is a decomposition of EPR into the excess and the housekeeping EPRs,

$$\sigma = \sigma^{\mathrm{ex}} + \sigma^{\mathrm{hk}}. \quad (10)$$

Time integration gives a decomposition of the EP Σ_τ into the excess EP $\Sigma_\tau^{\mathrm{ex}}$ and the housekeeping EP $\Sigma_\tau^{\mathrm{hk}}$, $\Sigma_\tau = \Sigma_\tau^{\mathrm{ex}} + \Sigma_\tau^{\mathrm{hk}}$. Here, the excess EP $\Sigma_\tau^{\mathrm{ex}}$ and the housekeeping EP $\Sigma_\tau^{\mathrm{hk}}$ are defined as $\Sigma_\tau^{\mathrm{ex}} := \int_0^\tau dt\, \sigma^{\mathrm{ex}}$ and $\Sigma_\tau^{\mathrm{hk}} := \int_0^\tau dt\, \sigma^{\mathrm{hk}}$.

Since the geometric decomposition (10) is a generalized Pythagorean theorem, we can conceptualize the decomposition geometrically, as summarized in Fig. 2. Let us consider the geometric nature of the excess EPR. The conservative force $F^* = \nabla_r \phi^*$, whose squared norm provides the excess EPR, is uniquely given by the minimization problem

$$F^* = \underset{F' | \nabla_r \cdot (\mathsf{M}F') = \nabla_r \cdot (\mathsf{M}F)}{\arg \min} \langle\langle F', F' \rangle\rangle_{\mathsf{M}}, \quad (11)$$

which follows from condition Eq. (9) (see Appendix A 2 for details). As a result, we can rewrite the excess EPR σ^{ex} as the following variational problem

$$\sigma^{\mathrm{ex}} = \underset{F' | \nabla_r \cdot (\mathsf{M}F') = \nabla_r \cdot (\mathsf{M}F)}{\inf} \langle\langle F', F' \rangle\rangle_{\mathsf{M}}. \quad (12)$$
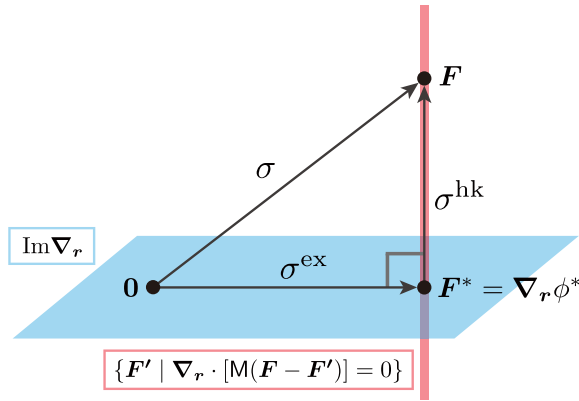
FIG. 2. The geometric decomposition of the EPR for Langevin systems. The blue plane indicates $\mathrm{Im}\boldsymbol{\nabla}_{\boldsymbol{r}}$, and the red line indicates $\{\boldsymbol{F}' \mid \boldsymbol{\nabla}_{\boldsymbol{r}} \cdot [\mathsf{M}(\boldsymbol{F} - \boldsymbol{F}')] = 0\}$. The thermodynamic force, whose squared norm is the EPR as $\sigma = \langle\!\langle \boldsymbol{F}, \boldsymbol{F} \rangle\!\rangle_{\mathsf{M}}$, is decomposed into the two orthogonal parts: $\boldsymbol{F}^*$ is the projection of $\boldsymbol{F}$ onto $\mathrm{Im}\boldsymbol{\nabla}_{\boldsymbol{r}}$, whose squared norm is the excess EPR as $\sigma^{\mathrm{ex}} = \langle\!\langle \boldsymbol{F}^*, \boldsymbol{F}^* \rangle\!\rangle_{\mathsf{M}}$, and the remaining part $\boldsymbol{F} - \boldsymbol{F}^*$, whose squared norm is the housekeeping EPR as $\sigma^{\mathrm{hk}} = \langle\!\langle \boldsymbol{F} - \boldsymbol{F}^*, \boldsymbol{F} - \boldsymbol{F}^* \rangle\!\rangle_{\mathsf{M}}$. We can regard $\boldsymbol{F}^*$ as the projection of $\boldsymbol{0}$ onto the red line or of $\boldsymbol{F}$ onto the blue plane.

The minimization problem (11) means that $\boldsymbol{F}^*$ is the closest point to the origin $\boldsymbol{0}$ in the affine subspace $\{\boldsymbol{F}' \mid \boldsymbol{\nabla}_{\boldsymbol{r}} \cdot [\mathsf{M}(\boldsymbol{F} - \boldsymbol{F}')] = 0\}$. The excess EPR can be seen as the shortest distance between this affine subspace and the origin.

We can also obtain a geometric interpretation of the housekeeping EPR because the set of conservative forces is the image of the gradient operator $\mathrm{Im}\,\boldsymbol{\nabla}_{\boldsymbol{r}} := \{\boldsymbol{F}' \mid \exists \phi, \ \boldsymbol{F}' = \boldsymbol{\nabla}_{\boldsymbol{r}}\phi\}$, and $\boldsymbol{F} - \boldsymbol{F}'$ for the element $\boldsymbol{F}' \in \{\boldsymbol{F}' \mid \boldsymbol{\nabla}_{\boldsymbol{r}} \cdot [\mathsf{M}(\boldsymbol{F} - \boldsymbol{F}')] = 0\}$ is in the orthogonal complement of $\mathrm{Im}\,\boldsymbol{\nabla}_{\boldsymbol{r}}$ with respect to the inner product $\langle\!\langle \cdot, \cdot \rangle\!\rangle_{\mathsf{M}}$. Consequently, the geometric decomposition (10) can also be treated from the viewpoint of the projection onto the subspace $\mathrm{Im}\boldsymbol{\nabla}_{\boldsymbol{r}}$: the conservative force $\boldsymbol{F}^*$ is given by the minimization problem

$$\boldsymbol{F}^* = \underset{\boldsymbol{F}' \in \mathrm{Im}\boldsymbol{\nabla}_{\boldsymbol{r}}}{\arg\min} \langle\!\langle \boldsymbol{F} - \boldsymbol{F}', \boldsymbol{F} - \boldsymbol{F}' \rangle\!\rangle_{\mathsf{M}}. \tag{13}$$

This is derived similarly to Eq. (11) by using the condition in Eq. (9) (see Appendix A 2 for details). Then, we can rewrite the housekeeping EPR σ^{hk} as the variational problem

$$\sigma^{\mathrm{hk}} = \inf_{\phi} \langle\!\langle \boldsymbol{F} - \boldsymbol{\nabla}_{\boldsymbol{r}}\phi, \boldsymbol{F} - \boldsymbol{\nabla}_{\boldsymbol{r}}\phi \rangle\!\rangle_{\mathsf{M}}, \tag{14}$$

which shows that the housekeeping EPR is the squared distance between the actual force and the subspace of conservative forces.

In addition to geometric interpretations, the constraint in Eq. (12) allows us discuss the physical meaning of the decomposition. The constraint with the Fokker–Planck equation $\partial_t p(\boldsymbol{r}; t) = -\boldsymbol{\nabla}_{\boldsymbol{r}} \cdot (\mathsf{M}\boldsymbol{F})$ yields the relation

$$\partial_t p = -\boldsymbol{\nabla}_{\boldsymbol{r}} \cdot (\mathsf{M}\boldsymbol{F}^*), \tag{15}$$

so we can interpret the excess EPR as the minimum dissipation required to reproduce the original dynamics. In contrast, the housekeeping EPR reflects the dissipation caused by the cyclic current that does not affect the dynamics, because $\boldsymbol{J}^{\mathrm{hk}} := \mathsf{M}(\boldsymbol{F} - \boldsymbol{\nabla}_{\boldsymbol{r}}\phi^*)$ satisfies $\boldsymbol{\nabla}_{\boldsymbol{r}} \cdot \boldsymbol{J}^{\mathrm{hk}} = 0$.

If the nonconservative force $\boldsymbol{K}^{\mathrm{nc}}$ in Eq. (4) is absent, then the housekeeping EPR always vanishes, and the optimal potential ϕ^* is given by $\phi^* = -U - \ln p$. Conversely, the excess EPR vanishes when the system is in steady state since the condition in Eq. (9) reduces to $\boldsymbol{\nabla}_{\boldsymbol{r}} \cdot (\mathsf{M}\boldsymbol{F}^*) = 0$ in steady state, which $\boldsymbol{F}^* = \boldsymbol{0}$ solves.

C. Local decomposition and wavenumber decomposition of entropy production rate for Langevin systems

The EPR is the volume integral of the positive quantity $\boldsymbol{J}(\boldsymbol{r}; t) \cdot \boldsymbol{F}(\boldsymbol{r}; t)$. From this viewpoint, we can decompose the dissipation at each spatial location. Similarly, it is expected that we can identify the dissipation at each wavenumber in Fourier space. In this section, we introduce two new geometric decompositions of the EPR. One is a decomposition of the EPR into contributions from each spatial location, and the other is a decomposition into contributions from each wavenumber.

Local decomposition.—We define the local EPR as

$$\sigma^{\mathrm{loc}}(\boldsymbol{r}) := \boldsymbol{J}(\boldsymbol{r}; t) \cdot \boldsymbol{F}(\boldsymbol{r}; t) \geqslant 0, \tag{16}$$

which satisfies

$$\sigma = \int_{\mathbb{R}^d} d\boldsymbol{r}\, \sigma^{\mathrm{loc}}(\boldsymbol{r}). \tag{17}$$

The local EPR $\sigma^{\mathrm{loc}}(\boldsymbol{r})$ is nonnegative and indicates dissipation at location $\boldsymbol{r}$.

We can also decompose the excess and housekeeping EPRs as $\sigma^{\mathrm{ex}} = \int_{\mathbb{R}^d} d\boldsymbol{r}\, \sigma^{\mathrm{ex,loc}}(\boldsymbol{r})$, and $\sigma^{\mathrm{hk}} = \int_{\mathbb{R}^d} d\boldsymbol{r}\, \sigma^{\mathrm{hk,loc}}(\boldsymbol{r})$, where the local excess and housekeeping EPRs are defined as

$$\sigma^{\mathrm{ex,loc}}(\boldsymbol{r}) := \boldsymbol{\nabla}_{\boldsymbol{r}}\phi^*(\boldsymbol{r}; t) \cdot \mathsf{M}(\boldsymbol{r}; t)\boldsymbol{\nabla}_{\boldsymbol{r}}\phi^*(\boldsymbol{r}; t)$$
$$= \boldsymbol{F}^*(\boldsymbol{r}; t) \cdot \mathsf{M}(\boldsymbol{r}; t)\boldsymbol{F}^*(\boldsymbol{r}; t), \tag{18}$$

$$\sigma^{\mathrm{hk,loc}}(\boldsymbol{r}) := [\boldsymbol{F}(\boldsymbol{r}; t) - \boldsymbol{\nabla}_{\boldsymbol{r}}\phi^*(\boldsymbol{r}; t)]$$
$$\cdot \mathsf{M}(\boldsymbol{r}; t)[\boldsymbol{F}(\boldsymbol{r}; t) - \boldsymbol{\nabla}_{\boldsymbol{r}}\phi^*(\boldsymbol{r}; t)]. \tag{19}$$

The local excess and housekeeping EPRs are nonnegative because the mobility tensor M is positive definite for all $\boldsymbol{r} \in V$. Note that ϕ^* is a solution of the partial differential equation in Eq. (9), which means that we need global information to obtain the local excess and housekeeping EPRs.

Because these local excess and housekeeping EPRs are not introduced by the geometric decomposition for the local EPR σ^{loc}, the geometric excess/housekeeping decomposition can be locally violated as $\sigma^{\mathrm{loc}}(\boldsymbol{r}) \neq \sigma^{\mathrm{ex,loc}}(\boldsymbol{r}) + \sigma^{\mathrm{hk,loc}}(\boldsymbol{r})$. In other words, there may be a nonzero cross-term $\sigma^{\mathrm{cross}}(\boldsymbol{r}) := \sigma^{\mathrm{loc}}(\boldsymbol{r}) - \sigma^{\mathrm{ex,loc}}(\boldsymbol{r}) - \sigma^{\mathrm{hk,loc}}(\boldsymbol{r}) = 2\boldsymbol{\nabla}_{\boldsymbol{r}}\phi^* \cdot \mathsf{M}[\boldsymbol{F} - \boldsymbol{\nabla}_{\boldsymbol{r}}\phi^*]$. The cross-term may be negative or positive, but it satisfies $\int_{\mathbb{R}^d} d\boldsymbol{r}\, \sigma^{\mathrm{cross}}(\boldsymbol{r}) = 0$, thereby guaranteeing that the geometric excess/housekeeping decomposition holds globally as $\sigma = \sigma^{\mathrm{ex}} + \sigma^{\mathrm{hk}}$.

Wavenumber decomposition.—Next, we decompose the EPR into nonnegative wavenumber components using Parseval's identity. Because Parseval's identity can be regarded as a generalization of the Pythagorean theorem, we can regard this decomposition as another kind of geometric decomposition.

We define the wavenumber EPR as

$$\sigma^{\mathrm{wn}}(\boldsymbol{k}) := \frac{1}{(2\pi)^d}\|\hat{\boldsymbol{F}}(\boldsymbol{k};t)\|^2 \geqslant 0, \quad (20)$$

where we introduced the weighted Fourier transform of a vector field $\boldsymbol{F}'$ with a weight $\sqrt{\mathsf{M}_{ii}} = \sqrt{Dp(\boldsymbol{r};t)}$ defined as

$$\hat{F}_i'(\boldsymbol{k};t) := \int_{\mathbb{R}^d} d\boldsymbol{r}\sqrt{Dp(\boldsymbol{r};t)}F_i'(\boldsymbol{r};t)\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}. \quad (21)$$

Note that $\hat{\boldsymbol{F}}(\boldsymbol{k};t)$ is a complex vector and its Euclidean norm is defined as $\|\hat{\boldsymbol{F}}\| := \sqrt{\sum_{i=1}^d \overline{\hat{F}_i}\hat{F}_i}$ with the overline indicating complex conjugate. The wavenumber EPR $\sigma^{\mathrm{wn}}(\boldsymbol{k})$ provides a decomposition of EPR as

$$\sigma = \int_{\mathbb{R}^d} d\boldsymbol{k}\, \sigma^{\mathrm{wn}}(\boldsymbol{k}). \quad (22)$$

Although this can be understood as a consequence of Parseval's identity, we can show it directly as

$$\int_{\mathbb{R}^d} d\boldsymbol{k}\, \sigma^{\mathrm{wn}}(\boldsymbol{k}) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} d\boldsymbol{k}\|\hat{\boldsymbol{F}}(\boldsymbol{k};t)\|^2.$$

$$= \int_{\mathbb{R}^d\times\mathbb{R}^d} d\boldsymbol{r}d\boldsymbol{r}'\left[\left\{\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} d\boldsymbol{k}\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}')}\right\}\right.$$

$$\left.\times D\sqrt{p(\boldsymbol{r};t)p(\boldsymbol{r}';t)}\boldsymbol{F}(\boldsymbol{r};t)\cdot\boldsymbol{F}(\boldsymbol{r}';t)\right]$$

$$= \int_{\mathbb{R}^d} d\boldsymbol{r}\, Dp(\boldsymbol{r};t)\|\boldsymbol{F}(\boldsymbol{r};t)\|^2 = \sigma. \quad (23)$$

Here, we use the Fourier transform of the delta function $\delta(\boldsymbol{r}-\boldsymbol{r}') = \int_{\mathbb{R}^d} d\boldsymbol{k}\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}')}/(2\pi)^d$.

As we did for local EPR, we can also decompose the excess and housekeeping EPRs into wavenumber contributions as $\sigma^{\mathrm{ex}} = \int_{\mathbb{R}^d} d\boldsymbol{k}\, \sigma^{\mathrm{ex,wn}}(\boldsymbol{k})$, and $\sigma^{\mathrm{hk}} = \int_{\mathbb{R}^d} d\boldsymbol{k}\, \sigma^{\mathrm{hk,wn}}(\boldsymbol{k})$ by defining the wavenumber excess and housekeeping EPRs as

$$\sigma^{\mathrm{ex,wn}}(\boldsymbol{k}) := \frac{1}{(2\pi)^d}\|\hat{\boldsymbol{F}}^*(\boldsymbol{k};t)\|^2 \quad (24)$$

$$\sigma^{\mathrm{hk,wn}}(\boldsymbol{k}) := \frac{1}{(2\pi)^d}\|\hat{\boldsymbol{F}}(\boldsymbol{k};t)-\hat{\boldsymbol{F}}^*(\boldsymbol{k};t)\|^2. \quad (25)$$

We remark that the geometric excess/housekeeping decomposition can also be violated at each wavenumber as $\sigma^{\mathrm{wn}}(\boldsymbol{k}) \neq \sigma^{\mathrm{ex,wn}}(\boldsymbol{k}) + \sigma^{\mathrm{hk,wn}}(\boldsymbol{k})$, and there is a nonzero cross term $\sigma^{\mathrm{cross,wn}}(\boldsymbol{k}) := \sigma^{\mathrm{wn}}(\boldsymbol{k}) - \sigma^{\mathrm{ex,wn}}(\boldsymbol{k}) + \sigma^{\mathrm{hk,wn}}(\boldsymbol{k})$, which satisfies $\int_{\mathbb{R}^d} d\boldsymbol{k}\, \sigma^{\mathrm{cross,wn}}(\boldsymbol{k}) = 0$.

The wavenumber decomposition is based on the orthonormality of the Fourier basis. Therefore, it may be possible to generalize the geometric decomposition of the EPR using an orthonormal basis other than the Fourier basis, e.g., a wavelet basis [87,88]. It may also be interesting to consider the spectral decomposition of the EPR [89] based on the Harada–Sasa relation [90] in terms of our wavenumber decomposition.

D. Wasserstein distance

The excess EPR obtained in the previous section can be interpreted as a geometric quantity using the Wasserstein geometry developed in optimal transport theory [63,72]. Here, we briefly review the Wasserstein distance and its dynamical reformulation, which is intrinsically important in thermodynamics.

The q-Wasserstein distance for a positive number $q \geqslant 1$ between two probability distributions p_A and p_B is defined as

$$W_q(p_A, p_B)$$

$$:= \left(\inf_{\pi\in\Pi(p_A,p_B)}\int_{\mathbb{R}^d\times\mathbb{R}^d} d\boldsymbol{r}d\boldsymbol{r}'\|\boldsymbol{r}-\boldsymbol{r}'\|^q\pi(\boldsymbol{r},\boldsymbol{r}')\right)^{\frac{1}{q}}, \quad (26)$$

where $\Pi(p_A, p_B)$ is the set of joint probability distributions with marginals p_A and p_B,

$$\Pi(p_A, p_B) := \left\{\pi \,\middle|\, p_A(\boldsymbol{r}) = \int_{\mathbb{R}^d} d\boldsymbol{r}'\pi(\boldsymbol{r},\boldsymbol{r}'),\right.$$

$$\left. p_B(\boldsymbol{r}') = \int_{\mathbb{R}^d} d\boldsymbol{r}\,\pi(\boldsymbol{r},\boldsymbol{r}'),\ \pi(\boldsymbol{r},\boldsymbol{r}')\geqslant 0\right\}.$$

Here, we assume that the moments up to the qth order are finite for the two probability distributions p_A and p_B. We can confirm that W_q satisfies the axioms of distance. We can also prove

$$W_q(p_A, p_B) \leqslant W_{q'}(p_A, p_B)\ \text{for}\ q \leqslant q', \quad (27)$$

by Hölder's inequality [72].

For every q, we can reformulate the q-Wasserstein distance as an optimization problem related to the dynamics of a probability distribution subject to a continuity equation. In particular, we can obtain the square of the 2-Wasserstein distance by the minimization problem

$$W_2(p_A, p_B)^2 = \inf_{p,\boldsymbol{F}'}\tau\int_0^\tau dt\int_{\mathbb{R}^d} d\boldsymbol{r}\|D\boldsymbol{F}'(\boldsymbol{r};t)\|^2 p(\boldsymbol{r};t) \quad (28)$$

with the following three constraints:

$$p(\cdot;0) = p_A(\cdot),\ p(\cdot;\tau) = p_B(\cdot),\ \partial_t p = -\boldsymbol{\nabla}_{\boldsymbol{r}}\cdot(pD\boldsymbol{F}'). \quad (29)$$

In other words, we minimize the right-hand side in Eq. (28) over trajectories of probability distributions that start and end on p_A and p_B and satisfy a continuity equation. This reformulation for the case $q = 2$ was initially made by Benamou and Brenier, so the equation (28) is called the Benamou–Brenier formula in optimal transport theory [77]. We can also consider an extension of the Benamou–Brenier formula for general q [91–93]. For the special case of $q = 1$, we can express the Benamou–Brenier formula as an optimization problem of the current, which is known as the Beckmann problem [94]

$$W_1(p_A, p_B) = \inf_{\boldsymbol{J}'}\int_0^\tau dt\int_{\mathbb{R}^d} d\boldsymbol{r}\|\boldsymbol{J}'(\boldsymbol{r};t)\|, \quad (30)$$

where we impose the following condition on $\boldsymbol{J}'$: there exists a time series of probability distribution p' satisfying

$$p'(\cdot;0) = p_A(\cdot),\ p'(\cdot;\tau) = p_B(\cdot),\ \partial_t p' = -\boldsymbol{\nabla}_{\boldsymbol{r}}\cdot\boldsymbol{J}'. \quad (31)$$

Both expressions of the 1-Wasserstein distance in the original definition (26) and the Beckmann problem (30) are reduced to an expression known as the Kantorovich–Rubinstein duality,

$$W_1(p_A, p_B) = \sup_{\phi\in\mathrm{Lip}^1}\left\{\int_{\mathbb{R}^d} d\boldsymbol{r}\,\phi(p_B - p_A)\right\}, \quad (32)$$

where the set of 1-Lipschitz functions is denoted by

$$\text{Lip}^1 := \{\phi | \|\nabla_{\boldsymbol{r}} \phi\| \leqslant 1\}. \tag{33}$$

The derivation of Eq. (32) from the original definition of the 1-Wasserstein distance in Eq. (26) is well known and based on the method of Lagrange multipliers [72]. The Beckmann problem (30) can be directly obtained from Kantorovich–Rubinstein duality (32) by again using the method of Lagrange multipliers [93,95].

E. Wasserstein geometry and thermodynamic trade-off relations

Considering a trajectory of probability distribution $\{p(t)\}_{t \in [0,\tau]}$ obeying the Fokker–Planck equation (2), we can define the length of the trajectory using the q-Wasserstein distance as

$$l_{q,\tau} := \int_0^\tau dt \, v_q(t), \tag{34}$$

with $v_q(t)$ defined as

$$v_q(t) := \lim_{\Delta t \to 0} \frac{W_q(p(t), p(t + \Delta t))}{\Delta t}. \tag{35}$$

This quantity $v_q(t)$ indicates the speed of the dynamics of $p(t)$ in the space of probability distributions. The form of the mobility tensor M, the Benamou–Brenier formula for the 2-Wasserstein distance (28), and the variational form of the excess EPR (12) lead to

$$\sigma^{\text{ex}} = \frac{v_2(t)^2}{D}, \tag{36}$$

which means the square root of the excess EPR at time t is proportional to the speed of the dynamics of the probability distribution.

The relation between σ^{ex} and $v_2(t)$ leads to the hierarchy of TSLs as [41]

$$\frac{W_2(p(0), p(\tau))^2}{D} \leqslant \frac{l_{2,\tau}^2}{D} \leqslant \tau \Sigma_\tau^{\text{ex}} \leqslant \tau \Sigma_\tau. \tag{37}$$

The inequality $l_{2,\tau} \geqslant W_2(p(0), p(\tau))$, which comes from the triangle inequality, yields the first inequality in Eq. (37). This inequality reflects the fact that $W_2(p(0), p(\tau))$ is the geodesic length between $p(0)$ and $p(\tau)$. The second inequality is derived from the Cauchy–Schwarz inequality $[\int_0^\tau dt \, v_2(t)]^2 \leqslant [\int_0^\tau dt][\int_0^\tau dt \, v_2(t)^2]$ and the relation between v_2 and σ^{ex} (36). The third inequality is a consequence of the nonnegativity of the housekeeping EP $\Sigma_\tau^{\text{hk}} \geqslant 0$ and the decomposition of the EP. Overall, the inequalities in Eq. (37) tell us that transitioning to a more distant distribution in less time requires more dissipation.

The inequality between Wasserstein distances in Eq. (27) leads to another hierarchy of TSLs as

$$\frac{W_1(p(0), p(\tau))^2}{D} \leqslant \frac{l_{1,\tau}^2}{D} \leqslant \tau \Sigma_\tau^{\text{ex}} \leqslant \tau \Sigma_\tau. \tag{38}$$

These lower bounds on the EPs are weaker than those in Eq. (37) because Eq. (27) shows

$$\frac{l_{1,\tau}^2}{D} \leqslant \frac{l_{2,\tau}^2}{D} \leqslant \tau \Sigma_\tau,$$

$$\frac{W_1(p(0), p(\tau))^2}{D} \leqslant \frac{W_2(p(0), p(\tau))^2}{D} \leqslant \tau \Sigma_\tau, \tag{39}$$

where we used the fact that Eq. (27) leads to $v_2(t) \geqslant v_1(t)$ and $l_{2,\tau} \geqslant l_{1,\tau}$. The generalization of the TSLs discussed here from Langevin dynamics to Markov jump processes (MJPs) is rather complicated, and we only note that there are multiple ways of generalizing [44,45].

Using the excess EPR, we can obtain a TUR for time-independent observable $\varphi(\boldsymbol{r})$ as

$$\left(d_t \langle \varphi \rangle_{p_t}\right)^2 \leqslant D \langle \|\nabla_{\boldsymbol{r}} \varphi\|^2 \rangle_{p_t} \sigma^{\text{ex}}, \tag{40}$$

where the bracket indicates the average over the probability distribution $p(\boldsymbol{r}; t)$, $\langle \varphi \rangle_{p_t} := \int_{\mathbb{R}^d} d\boldsymbol{r} \, p(\boldsymbol{r}; t) \varphi(\boldsymbol{r})$ [62]. The TUR represents a trade-off relation between the excess EPR σ^{ex}, the speed of the observable $d_t \langle \varphi \rangle_{p_t}$, and the average squared magnitude of the gradient of the observable $\langle \|\nabla_{\boldsymbol{r}} \varphi\|^2 \rangle_{p_t}$. In other words, we need more dissipation to make a flatter observable change faster. It is derived from the Cauchy–Schwarz inequality, $\langle\!\langle \nabla_{\boldsymbol{r}} \varphi, \nabla_{\boldsymbol{r}} \phi^* \rangle\!\rangle_{\mathsf{M}}^2 \leqslant \langle\!\langle \nabla_{\boldsymbol{r}} \varphi, \nabla_{\boldsymbol{r}} \varphi \rangle\!\rangle_{\mathsf{M}} \langle\!\langle \nabla_{\boldsymbol{r}} \phi^*, \nabla_{\boldsymbol{r}} \phi^* \rangle\!\rangle_{\mathsf{M}}$, and the fact that ϕ^* reproduces the dynamics as $\partial_t p = -\nabla_{\boldsymbol{r}} \cdot (\mathsf{M} \nabla_{\boldsymbol{r}} \phi^*)$. Here, the quantities appearing in the Cauchy–Schwarz inequality are given by $\langle\!\langle \nabla_{\boldsymbol{r}} \varphi, \nabla_{\boldsymbol{r}} \phi^* \rangle\!\rangle_{\mathsf{M}} = d_t \langle \varphi \rangle_{p_t}$, $\langle\!\langle \nabla_{\boldsymbol{r}} \varphi, \nabla_{\boldsymbol{r}} \varphi \rangle\!\rangle_{\mathsf{M}} = D \langle \|\nabla_{\boldsymbol{r}} \varphi\|^2 \rangle_{p_t}$ and $\langle\!\langle \nabla_{\boldsymbol{r}} \phi^*, \nabla_{\boldsymbol{r}} \phi^* \rangle\!\rangle_{\mathsf{M}} = \sigma^{\text{ex}}$, which proves the TUR.

We note our usage of the term TUR. Conventionally, a TUR refers to a lower bound on dissipation using the expectation value and the variance of a general current. Although the trade-off in Eq. (40) appears different from the conventional form of a TUR, we refer to it as a TUR. This is because the TUR (40) represents a specific case of the short-time limit of the conventional TUR for any initial state [96–98], as discussed in detail in Ref. [62].

We can also interpret the TUR from the viewpoint of the Wasserstein geometry by rewriting Eq. (40) as

$$v_\varphi(t) := \frac{|d_t \langle \varphi \rangle_{p_t}|}{\sqrt{\langle \|\nabla_{\boldsymbol{r}} \varphi\|^2 \rangle_{p_t}}} \leqslant v_2(t), \tag{41}$$

using the relation between $v_2(t)$ and σ^{ex} in Eq. (36). Here, we define $v_\varphi(t)$ as the speed of the observable φ normalized by the spatial fluctuation of φ. Therefore, the TUR means that the normalized speed of an observable is less than the speed of the probability distribution moving on the manifold of distributions equipped with the Wasserstein metric. This is similar to the Cramér–Rao bound [99] with parameter t, called the information geometric speed limit [100,101], written as

$$v_\varphi^I(t) := \frac{|d_t \langle \varphi \rangle_{p_t}|}{\sqrt{\text{Var}[\varphi]}} \leqslant v_I(t), \tag{42}$$

where $v_I(t) := \sqrt{\langle (d_t \ln p(t))^2 \rangle_{p_t}}$ is the square root of the Fisher information and $\text{Var}[\varphi]$ is the variance defined as $\text{Var}[\varphi] = \langle \varphi^2 \rangle_{p_t} - \langle \varphi \rangle_{p_t}^2$. From an information geometric point of view, we can regard $v_I(t)$ as the speed of the

TABLE I. Summary of essential quantities appearing in the total RDSs, the diffusion part, and the reaction part. Forces and currents take values in the second row at each spatial and temporal point. Operators $\overset{\leftrightarrow}{\mathsf{M}}$, m, and $\mathcal{M}$ map forces to the corresponding currents locally. Gradient operators can generate diffusion and reaction forces from a single potential function as $\vec{\boldsymbol{F}} = \nabla_{\boldsymbol{r}}\vec{\phi}$ and $\boldsymbol{f} = \nabla_{\mathsf{s}}\vec{\phi}$.

	Total	Diffusion	Reaction
Value	$\mathbb{R}^{N \times d} \oplus \mathbb{R}^M$	$\mathbb{R}^{N \times d}$	$\mathbb{R}^M$
Force	$\mathcal{F} = (\vec{\boldsymbol{F}}, \boldsymbol{f})$	$\vec{\boldsymbol{F}} = [\boldsymbol{F}_{(\alpha)}]_{\alpha=1}^N$	$\boldsymbol{f} = (f_\rho)_{\rho=1}^M$
Current	$\mathcal{J} = (\vec{\boldsymbol{J}}, \boldsymbol{j})$	$\vec{\boldsymbol{J}} = [\boldsymbol{J}_{(\alpha)}]_{\alpha=1}^N$	$\boldsymbol{j} = (j_\rho)_{\rho=1}^M$
Mobility	$\mathcal{M} = \overset{\leftrightarrow}{\mathsf{M}} \oplus \mathsf{m}$	$\overset{\leftrightarrow}{\mathsf{M}} = [\mathsf{M}_{(\alpha\beta)}]$	$\mathsf{m} = (\delta_{\rho\rho'}m_\rho)$
Gradient	$\nabla = \nabla_{\boldsymbol{r}} \oplus \nabla_{\mathsf{s}}$	$\nabla_{\boldsymbol{r}}$	∇_{s}
Potential	$\vec{\phi}(\boldsymbol{r}) \in \mathbb{R}^N$		

probability distribution, $p(t)$, on the manifold equipped with the information-geometric (Fisher) metric.

We can derive the inequality $v_1(t) \leqslant v_2(t)$, which provides the hierarchy of TSLs (39), from the TUR by taking the optimal potential of the Kantorovich–Rubinstein duality (32) as observable in Eq. (41).

III. THERMODYNAMICS OF REACTION-DIFFUSION SYSTEMS

Hereafter, we will focus on reaction-diffusion systems (RDSs). We will refer to the geometric framework reviewed above and extend it with additional notions to generalize it. We will need to use many kinds of symbols, which we summarize in Table I.

This section introduces the two basics of RDSs: dynamics and thermodynamics. We begin with a class of RDSs called closed systems in Sec. III A. A closed RDS does not exchange molecules with the outside, as opposed to an open RDS, the other class of RDSs, as explained in Sec. III B. Section III C introduces the thermodynamics of RDSs in terms of thermodynamic forces and the EPR. To simplify the discussion in subsequent sections, we unify quantities associated with reaction and diffusion by introducing appropriate vector fields and operators in Sec. III D. We also introduce the concept of conservative and nonconservative thermodynamic forces for RDSs, which play a central role in the geometric excess/housekeeping decomposition of EPR in Sec. III E.

A. Closed reaction-diffusion systems

We consider an RDS that describes the time evolution of a concentration distribution of N chemical species in the d-dimensional area $V \subseteq \mathbb{R}^d$ due to reactions, advection, and diffusion. If no particles interact with the outside of the system, the system is called a closed RDS. Let $\mathcal{S}$ index the chemical species as $\mathcal{S} := \{1, 2, \dots, N\}$. We write the αth chemical species ($\alpha \in \mathcal{S}$) as Z_α, and its concentration distribution at location $\boldsymbol{r} \in V$ and time t as $c_\alpha(\boldsymbol{r}; t)$. We note that $c_\alpha(\boldsymbol{r}; t)$ is not a probability density, and $\int d\boldsymbol{r} c_\alpha(\boldsymbol{r}; t)$ is not necessarily equal to 1 or any other constant.

The time evolution of the concentration of the αth chemical species under the assumption of Fick's law are usually given by the reaction-diffusion (RD) equation,

$$\partial_t c_\alpha(\boldsymbol{r}; t) = D_\alpha \nabla_{\boldsymbol{r}}^2 c_\alpha(\boldsymbol{r}; t) + R_\alpha(\boldsymbol{r}; t), \qquad (43)$$

where D_α indicates the diffusion constant of Z_α and $R_\alpha(\boldsymbol{r}; t)$ represents the effect of reactions. Note that the reaction term R_α can depend on the concentration distribution. We can rewrite the first term to $-\nabla_{\boldsymbol{r}} \cdot \boldsymbol{J}_{(\alpha)}^{\text{Fick}}(\boldsymbol{r}; t)$, where $\boldsymbol{J}_{(\alpha)}^{\text{Fick}}(\boldsymbol{r}; t) := -D_\alpha \nabla_{\boldsymbol{r}} c_\alpha(\boldsymbol{r}; t)$ is the current obeying Fick's law.

However, assuming Fick's law is optional to our discussion. The diffusion current given by Fick's law sometimes fails to describe the dynamics of chemical species, for example, under an electric field, which causes advection, or in a nondilute solution. To include a wider range of phenomena, we deal with the following more general RD equation for the αth chemical species,

$$\partial_t c_\alpha(\boldsymbol{r}; t) = -\nabla_{\boldsymbol{r}} \cdot \boldsymbol{J}_{(\alpha)}(\boldsymbol{r}; t) + R_\alpha(\boldsymbol{r}; t), \qquad (44)$$

where $\boldsymbol{J}_{(\alpha)}(\boldsymbol{r}; t) = [J_{(\alpha)i}(\boldsymbol{r}; t)]_{i=1,\dots,d}$ is the general diffusion current for the αth species. The diffusion currents can depend on the concentration distribution as in the case of Fick's law. Note that the general RD equation (44) reproduces the usual one (43) if the diffusion current obeys Fick's law as $\boldsymbol{J}_{(\alpha)}(\boldsymbol{r}; t) = \boldsymbol{J}_{(\alpha)}^{\text{Fick}}(\boldsymbol{r}; t)$.

In this paper, we assume one of the following three boundary conditions on diffusion currents. Note that the boundary conditions on diffusion currents constrain concentration distributions through the dependence of diffusion currents on concentration distributions. The first one is the no-flux boundary condition, $\boldsymbol{J}_{(\alpha)}(\boldsymbol{r}; t) \cdot \boldsymbol{n}(\boldsymbol{r}) = 0$ for all $\alpha \in \mathcal{S}, t$, and $\boldsymbol{r} \in \partial V$, where ∂V indicates the boundary of V, and $\boldsymbol{n}(\boldsymbol{r})$ is the unit normal vector of the surface at $\boldsymbol{r} \in \partial V$. This condition corresponds to a chemical reaction system in a container, where the exchange of particles via diffusion with the outside never happens. The second one is the periodic boundary condition: when the space V has a periodic structure, like a supercube, we may assume that all quantities depending on $\boldsymbol{r}$ satisfy the periodic boundary condition. Note that this rule applies not only to currents but also to other quantities. The third one is the fast decay of the diffusion currents at infinity, which we consider when $V = \mathbb{R}^d$. These conditions can be combined. For example, considering an infinitely long pipe with a square cross section, we may assume the no-flux boundary or the periodic boundary on the sides of the pipe, while supposing the fast decay for the current at infinity.

We can consider the thermodynamic structure of the RDS by rewriting the reaction term $R_\alpha(\boldsymbol{r}; t)$ with reaction currents based on the details of the reactions, as explained below. We consider M reactions indexed by $\mathcal{R} := \{1, 2, \dots, M\}$. We write the ρth reaction ($\rho \in \mathcal{R}$) as

$$\sum_{\alpha \in \mathcal{S}} v_{\alpha\rho}^+ Z_\alpha \rightleftharpoons \sum_{\alpha \in \mathcal{S}} v_{\alpha\rho}^- Z_\alpha, \qquad (45)$$

where $v_{\alpha\rho}^\pm$ indicates the number of Z_α consumed (+) or produced (−) by the ρth reaction. We assume $v_{\alpha\rho}^\pm$ is independent of location $\boldsymbol{r}$ and time t for all α and ρ. We write the forward (+) and reverse (−) fluxes of the ρth reaction at $(\boldsymbol{r}; t)$ as $j_\rho^+(\boldsymbol{r}; t)$ and $j_\rho^-(\boldsymbol{r}; t)$, respectively. All the fluxes are always assumed to be positive, $j_\rho^\pm(\boldsymbol{r}; t) > 0$. They can depend on space and time $(\boldsymbol{r}; t)$ directly and/or via concentration

distributions (e.g., assuming mass action kinetics, a flux is given by $j_\rho^\pm(\boldsymbol{r};t) = \kappa_\rho^\pm(\boldsymbol{r};t)\prod_\alpha[c_\alpha(\boldsymbol{r};t)]^{\nu_{\alpha\rho}^\pm}$, where $\kappa_\rho^\pm(\boldsymbol{r};t)$ is the reaction rate constant for the forward/reverse reaction). The reaction current of the ρth reaction $j_\rho(\boldsymbol{r};t)$ is given by $j_\rho(\boldsymbol{r};t) = j_\rho^+(\boldsymbol{r};t) - j_\rho^-(\boldsymbol{r};t)$. Using these reaction currents, we can rewrite $R_\alpha(\boldsymbol{r};t)$ as

$$R_\alpha(\boldsymbol{r};t) = \sum_{\rho\in\mathcal{R}} S_{\alpha\rho}\, j_\rho(\boldsymbol{r};t), \qquad (46)$$

where $S_{\alpha\rho} := \nu_{\alpha\rho}^- - \nu_{\alpha\rho}^+$ is the (α,ρ)th element of the stoichiometric matrix, which denotes the net increase of Z_α through the ρth reaction.

Finally, we rewrite the RD equation as a continuity equation, which will be convenient for future calculations. Introducing the vector notation $\vec{c} = (c_1,\dots,c_N)^\top$, $\vec{J} = (\boldsymbol{J}_{(1)},\dots,\boldsymbol{J}_{(N)})^\top$, and $\vec{R} = (R_1,\dots,R_N)^\top$ with transposition $^\top$, the RD equation (44) is rewritten in a simpler form $\partial_t\vec{c} = -\boldsymbol{\nabla}_{\boldsymbol{r}}\cdot\vec{J} + \vec{R}$. Here, we make the dependence on $(\boldsymbol{r};t)$ implicit to simplify the notation. Rewriting the reaction term as $\vec{R} = \nabla_{\mathsf{s}}^\top \boldsymbol{j}$ with the vector of the reaction current $\boldsymbol{j} := (j_1,\dots,j_M)^\top$ and the matrix $(\nabla_{\mathsf{s}})_{\rho\alpha} := S_{\alpha\rho}$, the RD equation reduces to a continuity equation

$$\partial_t\vec{c} = -\boldsymbol{\nabla}_{\boldsymbol{r}}\cdot\vec{J} + \nabla_{\mathsf{s}}^\top \boldsymbol{j}. \qquad (47)$$

Here, the matrix $\nabla_{\mathsf{s}}^\top$ is the stoichiometric matrix.

B. Open reaction-diffusion systems

We generally deal with an open RDS, where some of the N species can be exchanged with the outside. We classify the species into two categories: internal species, which are not exchanged with the outside, and external species, which are exchanged with the outside. Let $N_\mathcal{X}(\leqslant N)$ denote the number of internal species. We index the internal and external species as $\mathcal{X} := \{1,\dots,N_\mathcal{X}\}$ and $\mathcal{Y} := \{N_\mathcal{X}+1,\dots,N\} = \mathcal{S}\setminus\mathcal{X}$, respectively.

Taking into account the decomposition of $\mathcal{S}$ into $\mathcal{X}$ and $\mathcal{Y}$, we introduce the following notation. Let $\vec{e} = (e_1,\dots,e_N)^\top$ be an arbitrary vector consisting of N elements such as $\vec{c}$, $\vec{J}$, or $\vec{R}$. We define $\vec{e}_\mathcal{X}$ and $\vec{e}_\mathcal{Y}$ as the vectors of the first $N_\mathcal{X}$ elements and the last $N - N_\mathcal{X}$ elements of $\vec{e}$; therefore, they decompose $\vec{e}$ as $\vec{e} = (\vec{e}_\mathcal{X}^\top, \vec{e}_\mathcal{Y}^\top)^\top$.

We also define a subset of $\mathcal{R}$, $\mathcal{R}_\mathcal{X}$, as the set of the indexes of reactions that change the concentrations of internal chemical species,

$$\mathcal{R}_\mathcal{X} := \{\rho\in\mathcal{R}\,|\,\exists\alpha\in\mathcal{X},\, S_{\alpha\rho}\neq 0\}. \qquad (48)$$

Reactions whose index belongs to $\mathcal{R}\setminus\mathcal{R}_\mathcal{X}$ only change the concentrations of external species, while reactions whose index belongs to $\mathcal{R}_\mathcal{X}$ may also change the concentrations of external chemical species.

The exchange of external species can be modeled in various ways. For example, we can describe the interaction with the outside by fixing the concentrations of the external species on the boundary. We can also assume that the concentration distributions of the external species are homogeneous both on the boundary and in the bulk. In addition, we can control the concentration distribution of the external species with external currents.

Nonetheless, the time evolution of the internal species is essential for further discussion. With the notation already introduced in this section, they can be written as

$$\partial_t\vec{c}_\mathcal{X} = -\boldsymbol{\nabla}_{\boldsymbol{r}}\cdot\vec{J}_\mathcal{X} + (\nabla_{\mathsf{s}}^\top \boldsymbol{j})_\mathcal{X}. \qquad (49)$$

We impose the same boundary conditions on the diffusion current corresponding to the internal species in open systems as in closed systems, e.g., the no-flux boundary condition, the periodic boundary condition, or the fast decay of diffusion currents. In the following results, we only consider the time evolution of internal species, which is described by the continuity-equation-like form (49), so that we do not need to consider how to describe the interaction with the outside of the system. The only exception is a generalization of the 2-Wasserstein distance, where we have to assume the homogeneity of the external species, as discussed in Sec. V B.

C. Thermodynamic force and entropy production rate

Here, we introduce the thermodynamic forces and the EPR. Corresponding to the diffusion and reaction currents, two kinds of forces, diffusion and reaction forces, are introduced. In the following, we assume that the temperature is homogeneous in V, and we choose units so that the product of the temperature T and the gas constant R_{gas} is equal to 1 [47] for simplicity.

We also assume that the chemical potential can be defined for each species at each location. We let $\mu_\alpha(\boldsymbol{r};t)$ denote the chemical potential of the αth species at $(\boldsymbol{r};t)$. The chemical potential can depend on the concentration distribution and its spatial derivatives. For example, the chemical potential in an ideal dilute solution without applied mechanical forces is given by $\mu_\alpha^{\mathrm{id}} = \mu_\alpha^\circ + \ln c_\alpha$, with the standard chemical potential μ_α°, which is independent of t and $\boldsymbol{r}$. In the Cahn–Hilliard equation [55], the chemical potential also depends on the concentration gradient as $\mu_\alpha = c_\alpha^3 - c_\alpha - \gamma\boldsymbol{\nabla}_{\boldsymbol{r}}^2 c_\alpha$ with a constant γ.

Diffusion force.—We introduce the diffusion force. We write the diffusion force for the αth species at $(\boldsymbol{r};t)$ as $\boldsymbol{F}_{(\alpha)}(\boldsymbol{r};t) = [F_{(\alpha)i}(\boldsymbol{r};t)]_{i=1,\dots,d}$. The diffusion force is defined by using the chemical potential as

$$\boldsymbol{F}_{(\alpha)}(\boldsymbol{r};t) := -\boldsymbol{\nabla}_{\boldsymbol{r}}\mu_\alpha(\boldsymbol{r};t) + \boldsymbol{K}_{(\alpha)}^{\mathrm{nc}}(\boldsymbol{r};t), \qquad (50)$$

where $\boldsymbol{K}_{(\alpha)}^{\mathrm{nc}}$ denotes the nonconservative mechanical force on particles of the αth chemical species. Note that all mechanical forces on particles of the αth chemical species that can be represented by the gradient of a potential, e.g., gravity and Coulomb force, are included in the gradients of the chemical potential $-\boldsymbol{\nabla}_{\boldsymbol{r}}\mu_\alpha(\boldsymbol{r};t)$. We introduce the vector notation as $\vec{F} := (\boldsymbol{F}_{(1)},\dots,\boldsymbol{F}_{(N)})^\top$, $\vec{\mu} := (\mu_1,\dots,\mu_N)^\top$, and $\vec{K}^{\mathrm{nc}} := (\boldsymbol{K}_{(1)}^{\mathrm{nc}},\dots,\boldsymbol{K}_{(N)}^{\mathrm{nc}})^\top$, which let us rewrite Eq. (50) as

$$\vec{F} := -\boldsymbol{\nabla}_{\boldsymbol{r}}\vec{\mu} + \vec{K}^{\mathrm{nc}}. \qquad (51)$$

Here, the first term in the right-hand side, $\boldsymbol{\nabla}_{\boldsymbol{r}}\vec{\mu}$, indicates $(\boldsymbol{\nabla}_{\boldsymbol{r}}\mu_1,\dots,\boldsymbol{\nabla}_{\boldsymbol{r}}\mu_N)^\top$.

We assume a linear relation between the diffusion current and the diffusion force as

$$\vec{J}(\boldsymbol{r};t) = \overset{\leftrightarrow}{\mathsf{M}}(\boldsymbol{r};t)\vec{F}(\boldsymbol{r};t), \qquad (52)$$

where $\overset{\leftrightarrow}{\mathsf{M}}(\boldsymbol{r};t) = [\mathsf{M}_{(\alpha\beta)}(\boldsymbol{r};t)]_{\alpha,\beta\in\mathcal{S}}$ is the mobility tensor, each of whose elements is a $d \times d$ matrix as $\mathsf{M}_{(\alpha\beta)} = [M_{(\alpha\beta)ij}]_{i,j=1,\dots,d}$. It can be rewritten with the elements of current, force, and mobility tensor as $J_{(\alpha)i}(\boldsymbol{r};t) = \sum_{\beta\in\mathcal{S}}\sum_{j=1}^{d} M_{(\alpha\beta)ij}(\boldsymbol{r};t)F_{(\beta)j}(\boldsymbol{r};t)$. We further assume that $\overset{\leftrightarrow}{\mathsf{M}}(\boldsymbol{r};t)$ is symmetric and positive-definite: $M_{(\alpha\beta)ij} = M_{(\beta\alpha)ji}$ holds for all $\alpha, \beta \in \mathcal{S}$ and for all $1 \leqslant i, j \leqslant d$, and $\vec{F}'^{\top}\overset{\leftrightarrow}{\mathsf{M}}\vec{F}' = \sum_{\alpha,\beta\in\mathcal{S}}\sum_{i,j=1}^{d} F'_{(\alpha)i}M_{(\alpha\beta)ij}F'_{(\beta)j} > 0$ holds for all $\vec{F}' \neq \vec{\mathbf{0}}$ ($\vec{\mathbf{0}}$ indicates a diffusion force or current all of whose elements are zero).

The mobility tensor possibly depends on the concentration distribution. In the special case where the diffusion current obeys Fick's law, $\boldsymbol{J}_{(\alpha)}(\boldsymbol{r};t) = \boldsymbol{J}_{(\alpha)}^{\text{Fick}}(\boldsymbol{r};t)$ and the force is given by the chemical potential as $\boldsymbol{F}_{(\alpha)}(\boldsymbol{r};t) = -\nabla_{\boldsymbol{r}}\mu_{\alpha}^{\text{id}}(\boldsymbol{r};t)$, the mobility tensor becomes

$$\mathsf{M}_{(\alpha\beta)}(\boldsymbol{r};t) = D_{\alpha}c_{\alpha}(\boldsymbol{r};t)\delta_{\alpha\beta}\mathsf{I}, \tag{53}$$

where $\delta_{\alpha\beta}$ is the Kronecker delta and I is the $d \times d$ identity matrix. In general, $\mathsf{M}_{(\alpha\alpha)}$ will not be proportional to the identity matrix if the mobility of the αth species is not isotropic. The off-diagonal entries $\mathsf{M}_{(\alpha\beta)}$ can also be nonzero matrices, which represent inter-species effects on the diffusion currents by the forces [85,102].

Reaction force.—We next define the reaction force $\boldsymbol{f} = (f_1, \dots, f_M)^{\top}$. The ρth element $f_{\rho}(\boldsymbol{r};t)$, which gives the reaction force on the ρth reaction at $(\boldsymbol{r};t)$, is defined in terms of the chemical potential as

$$f_{\rho}(\boldsymbol{r};t) := -\sum_{\alpha\in\mathcal{S}} S_{\alpha\rho}\mu_{\alpha}(\boldsymbol{r};t). \tag{54}$$

Using the vector notations, we can rewrite Eq. (54) as

$$\boldsymbol{f} := -\nabla_{\mathsf{s}}\vec{\mu}. \tag{55}$$

We also impose the condition of local detailed balance,

$$f_{\rho}(\boldsymbol{r};t) = \ln\frac{j_{\rho}^{+}(\boldsymbol{r};t)}{j_{\rho}^{-}(\boldsymbol{r};t)}, \tag{56}$$

on the reaction force everywhere for all $\rho \in \mathcal{R}$. It relates the reaction force and the fluxes. In particular, when the fluxes obey mass action kinetics and the chemical potential is μ_{α}^{id}, the local detailed balance condition reduces to $\ln(\kappa_{\rho}^{+}/\kappa_{\rho}^{-}) = -\nabla_{\mathsf{s}}\vec{\mu}^{\circ}$ with the vector $\vec{\mu}^{\circ} := (\mu_1^{\circ}, \dots, \mu_N^{\circ})^{\top}$. Note that we can utilize the local detailed balance condition in Eq. (56) as the definition of the reaction force even for systems where the chemical potential cannot be defined or where some unrecognized chemical species are present.

The local detailed balance condition (56) establishes a relationship between the reaction force and the current, analogous to the one for diffusion force (52), as

$$\boldsymbol{j}(\boldsymbol{r};t) = \mathsf{m}(\boldsymbol{r};t)\boldsymbol{f}(\boldsymbol{r};t), \tag{57}$$

with a diagonal matrix $[\mathsf{m}(\boldsymbol{r};t)]_{\rho\rho'} := \delta_{\rho\rho'}m_{\rho}(\boldsymbol{r};t)$. Here, the diagonal elements of $\mathsf{m}(\boldsymbol{r};t)$ are defined as

$$m_{\rho}(\boldsymbol{r};t) := \begin{cases} \dfrac{j_{\rho}^{+}(\boldsymbol{r};t) - j_{\rho}^{-}(\boldsymbol{r};t)}{\ln j_{\rho}^{+}(\boldsymbol{r};t) - \ln j_{\rho}^{-}(\boldsymbol{r};t)} & (j_{\rho}^{+}(\boldsymbol{r};t) \neq j_{\rho}^{-}(\boldsymbol{r};t)), \\ j_{\rho}^{+}(\boldsymbol{r};t) & (j_{\rho}^{+}(\boldsymbol{r};t) = j_{\rho}^{-}(\boldsymbol{r};t)). \end{cases} \tag{58}$$

In the following, we refer to m as the edgewise Onsager coefficient matrix, inspired by the Onsager coefficient that provides the linear relation between the force and the current in a steady state [103]. We remark that the edgewise Onsager coefficient matrix can depend on the concentration distribution owing to the $\vec{c}$ dependence of the fluxes $j_{\rho}^{\pm}$. In this sense, it differs from the mobility tensor, which does not depend on the diffusion current. Despite its dependence on fluxes, m is a physically fruitful quantity, as shown in previous studies [44,45] and in what follows in this paper.

As the mobility tensor $\overset{\leftrightarrow}{\mathsf{M}}$, the edgewise Onsager coefficient matrix m is positive definite. This positive definiteness follows from the mathematical fact that the logarithmic mean between two positive numbers $a, b > 0$, $(a - b)/\ln(a/b)$, is always positive. This is because the diagonal element of the matrix m_{ρ} is the logarithmic mean between the forward and reverse fluxes $j_{\rho}^{\pm} > 0$. Actually, the edgewise Onsager coefficient m_{ρ} can be regarded as a kind of activity, that is, an indicator of the back-and-forth intensity of each reaction measured by the average of the speed of the forward and reverse reactions. For example, activity is conventionally evaluated by (double) the arithmetic mean $j_{\rho}^{+}(\boldsymbol{r};t) + j_{\rho}^{-}(\boldsymbol{r};t)$ and the geometric mean $\sqrt{j_{\rho}^{+}(\boldsymbol{r};t)j_{\rho}^{-}(\boldsymbol{r};t)}$, which are called the dynamical [104] and frenetic activity [105], respectively. The edgewise Onsager coefficient is intermediate between these conventional activities because the inequality between means, $\sqrt{ab} \leqslant (a - b)/\ln(a/b) \leqslant (a + b)/2$, yields

$$\sqrt{j_{\rho}^{+}(\boldsymbol{r};t)j_{\rho}^{-}(\boldsymbol{r};t)} \leqslant m_{\rho}(\boldsymbol{r};t) \leqslant \frac{j_{\rho}^{+}(\boldsymbol{r};t) + j_{\rho}^{-}(\boldsymbol{r};t)}{2}. \tag{59}$$

Entropy production rate.—The product between the forces and the currents defines the EPR σ as

$$\sigma := \int_V d\boldsymbol{r} \left(\sum_{\alpha\in\mathcal{S}} \boldsymbol{J}_{(\alpha)} \cdot \boldsymbol{F}_{(\alpha)} + \sum_{\rho\in\mathcal{R}} j_{\rho}f_{\rho} \right)$$

$$= \int_V d\boldsymbol{r} \, (\vec{J}^{\top}\vec{F} + \boldsymbol{j}^{\top}\boldsymbol{f}), \tag{60}$$

whose time integration provides the EP during time duration $[0, \tau]$ as $\Sigma_{\tau} := \int_0^{\tau} dt \, \sigma$. The definition of the EPR (60) derives the second law of thermodynamics, since the positivity of $\overset{\leftrightarrow}{\mathsf{M}}$ and the local detailed balance condition (56) lead to $\vec{J}^{\top}\vec{F} = \vec{F}^{\top}\overset{\leftrightarrow}{\mathsf{M}}\vec{F} \geqslant 0$ and $\boldsymbol{j}^{\top}\boldsymbol{f} \geqslant 0$, respectively. The equality of the second law $\sigma = 0$ is satisfied if and only if the system is in equilibrium, i.e., $j_{\rho}(\boldsymbol{r};t) = 0$ and $\boldsymbol{J}_{(\alpha)}(\boldsymbol{r};t) = \mathbf{0}$ hold for all $\alpha \in \mathcal{S}$, $\rho \in \mathcal{R}$, and $\boldsymbol{r} \in V$. It indicates that the EPR is a measure of the irreversibility of the system.

When the temperature is constant, we can regard the EP as dissipated work, which is the difference between the work done on the system and the increased free energy [47,106]. Hence, a smaller EP means that the state of the system changes with less work, that is, it is energetically efficient. From this viewpoint, today's nonequilibrium thermodynamics, especially stochastic thermodynamics, usually considers minimum dissipation problems [31–46] to control systems optimally. We later see that some results in this paper also directly relate to such minimization problems, which have not been previously discussed for RDSs.

Since RDSs consist of diffusive dynamics and reactions, the EPR (60) accounts for the dissipation arising from multiple factors. Thus, it is essential to decompose the EPR into contributions from different factors to understand the thermodynamic properties of a system. One of the simplest decompositions is into the EPR from diffusion σ^{diff} and from reactions σ^{react},

$$\sigma^{\text{diff}} := \int_V d\boldsymbol{r} \sum_{\alpha \in \mathcal{S}} \boldsymbol{J}_{(\alpha)} \cdot \boldsymbol{F}_{(\alpha)}, \ \sigma^{\text{react}} := \int_V d\boldsymbol{r} \sum_{\rho \in \mathcal{R}} j_\rho f_\rho. \tag{61}$$

We provide a unified way to perform decompositions into various factors in Sec. IV.

D. Unifying the diffusion and reaction

We consider the quantities associated with reaction and diffusion separately in the previous sections. However, treating these quantities together will be useful for further discussion. Here, we introduce two inner products and some operators to handle reactions and diffusion together.

Forces and currents.—We introduce the force $\mathcal{F}$ and the current $\mathcal{J}$ by unifying the diffusion force (current) and the reaction force (current) as

$$\mathcal{J} := (\vec{\boldsymbol{J}}, \boldsymbol{j}), \quad \mathcal{F} := (\vec{\boldsymbol{F}}, \boldsymbol{f}). \tag{62}$$

We refer to $\vec{\boldsymbol{F}}'$ and $\boldsymbol{f}'$ as the diffusion part and the reaction part of $\mathcal{F}' = (\vec{\boldsymbol{F}}', \boldsymbol{f}')$, respectively, and the same is true for the current. We also define the inner product of two vector fields $\mathcal{J}' = (\vec{\boldsymbol{J}}', \boldsymbol{j}')$ and $\mathcal{F}' = (\vec{\boldsymbol{F}}', \boldsymbol{f}')$ as

$$\langle\!\langle \mathcal{J}', \mathcal{F}' \rangle\!\rangle := \int_V d\boldsymbol{r} \left(\sum_{\alpha \in \mathcal{S}} \boldsymbol{J}'_\alpha \cdot \boldsymbol{F}'_\alpha + \sum_{\rho \in \mathcal{R}} j'_\rho f'_\rho \right)$$

$$= \int_V d\boldsymbol{r} \, (\vec{\boldsymbol{J}}'^\top \vec{\boldsymbol{F}}' + \boldsymbol{j}'^\top \boldsymbol{f}'), \tag{63}$$

which immediately leads to a new expression of the EPR as

$$\sigma = \langle\!\langle \mathcal{J}, \mathcal{F} \rangle\!\rangle. \tag{64}$$

Potentials and concentrations.—We introduce the inner product of two vector fields with N elements, e.g., the chemical potential $\vec{\mu}$, the concentration distribution $\vec{c}$, and its time derivative $\partial_t \vec{c}$, as

$$\langle \vec{\phi}, \vec{\psi} \rangle := \int_V d\boldsymbol{r} \sum_{\alpha \in \mathcal{S}} \phi_\alpha \psi_\alpha = \int_V d\boldsymbol{r} \, \vec{\phi}^\top \vec{\psi}, \tag{65}$$

with $\vec{\phi} = (\phi_1, \dots, \phi_N)^\top$ and $\vec{\psi} = (\psi_1, \dots, \psi_N)^\top$. We refer to $\vec{\psi}_\mathcal{X}$ and $\vec{\psi}_\mathcal{Y}$ as the internal part and the external part of $\vec{\psi}$, e.g., $\partial_t \vec{c}_\mathcal{X} = (\partial_t c_1, \dots, \partial_t c_{N_\mathcal{X}})^\top$ is the internal part of $\partial_t \vec{c} = (\partial_t c_1, \dots, \partial_t c_N)^\top$.

Generalized gradient and divergence operators.—We define the generalized gradient operator ∇, taking a potential $\vec{\phi} = (\phi_1, \dots, \phi_N)^\top$ to a force, as

$$\nabla \vec{\phi} := (\boldsymbol{\nabla}_{\boldsymbol{r}} \vec{\phi}, \nabla_{\mathsf{s}} \vec{\phi}), \tag{66}$$

where $\boldsymbol{\nabla}_{\boldsymbol{r}} \vec{\phi}$ is defined as $\boldsymbol{\nabla}_{\boldsymbol{r}} \vec{\phi} := (\boldsymbol{\nabla}_{\boldsymbol{r}} \phi_1, \dots, \boldsymbol{\nabla}_{\boldsymbol{r}} \phi_N)^\top$. We remark that $\nabla_{\mathsf{s}} \vec{\phi}(\boldsymbol{r})$ is an M-dimensional vector because ∇_{s} is an $M \times N$ matrix and $\vec{\phi}(\boldsymbol{r})$ is an N-dimensional vector.

The generalized gradient operator $\nabla = \boldsymbol{\nabla}_{\boldsymbol{r}} \oplus \nabla_{\mathsf{s}}$ is regarded as the direct sum between $\boldsymbol{\nabla}_{\boldsymbol{r}}$ and ∇_{s}, where $\oplus$ represents the direct sum. The generalized gradient ∇ enables us to unify the relations between the diffusion and reaction forces and the chemical potential in Eqs. (51) and (55) as

$$\mathcal{F} = -\nabla \vec{\mu} + \mathcal{K}^{\text{nc}}, \tag{67}$$

where we define the nonconservative mechanical force vector $\mathcal{K}^{\text{nc}}$ as $\mathcal{K}^{\text{nc}} := (\vec{\boldsymbol{K}}^{\text{nc}}, \boldsymbol{0})$.

We also define an operator $\nabla^\dagger$ that maps a current to the time evolution caused by the current as

$$\nabla^\dagger \mathcal{J}' := -\boldsymbol{\nabla}_{\boldsymbol{r}} \cdot \vec{\boldsymbol{J}}' + \nabla_{\mathsf{s}}^\top \boldsymbol{j}', \tag{68}$$

where $\mathcal{J}' = (\vec{\boldsymbol{J}}', \boldsymbol{j}')$ is a vector field with the reaction and diffusion parts, and $\nabla^\dagger \mathcal{J}'$ is a vector field with N elements. This definition means $-\nabla^\dagger$ is a generalized divergence operator. The operator $\nabla^\dagger$ simplifies the form of the time evolution (49) as

$$\partial_t \vec{c}_\mathcal{X} = (\nabla^\dagger \mathcal{J})_\mathcal{X}, \tag{69}$$

which emphasizes that a continuity equation gives the time evolution of the internal species.

The operator $\nabla^\dagger$ is conjugate to the generalized gradient operator ∇ in terms of the two inner products $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ and $\langle \cdot, \cdot \rangle$ as

$$\langle\!\langle \mathcal{J}', \nabla \vec{\phi} \rangle\!\rangle = \langle \nabla^\dagger \mathcal{J}', \vec{\phi} \rangle, \tag{70}$$

where $\vec{\phi}$ is a potential whose external part is the zero vector, and $\mathcal{J}'$ is a current whose diffusion part corresponds to the internal species, $\boldsymbol{J}'_{(\alpha)}$ for $\alpha \in \mathcal{X}$, satisfies the boundary conditions on the system. To derive the conjugation relations between ∇ and $\nabla^\dagger$, we do partial integration and use Gauss's theorem to calculate $\langle \nabla^\dagger \mathcal{J}', \vec{\phi} \rangle$ as follows:

$$\langle \nabla^\dagger \mathcal{J}', \vec{\phi} \rangle = \int_V d\boldsymbol{r} \sum_{\alpha \in \mathcal{S}} \{-\boldsymbol{\nabla}_{\boldsymbol{r}} \cdot \boldsymbol{J}'_{(\alpha)} + (\nabla_{\mathsf{s}}^\top \boldsymbol{j}')_\alpha\} \phi_\alpha$$

$$= \int_V d\boldsymbol{r} \left\{ \sum_{\alpha \in \mathcal{S}} (\boldsymbol{J}'_{(\alpha)} \cdot \boldsymbol{\nabla}_{\boldsymbol{r}} \phi_\alpha) + \boldsymbol{j}'^\top \nabla_{\mathsf{s}} \vec{\phi} \right\}$$

$$\quad - \int_V d\boldsymbol{r} \sum_{\alpha \in \mathcal{S}} \boldsymbol{\nabla}_{\boldsymbol{r}} \cdot (\phi_\alpha \boldsymbol{J}'_{(\alpha)})$$

$$= \langle\!\langle \mathcal{J}', \nabla \vec{\phi} \rangle\!\rangle - \sum_{\alpha \in \mathcal{S}} \int_{\partial V} d\boldsymbol{n} \cdot (\phi_\alpha \boldsymbol{J}'_{(\alpha)}). \tag{71}$$

Here, the summand in the second term of the last line, $\int_{\partial V} d\boldsymbol{n} \cdot (\phi_\alpha \boldsymbol{J}'_{(\alpha)})$, vanishes for all $\alpha \in \mathcal{S}$ because $\phi_\alpha = 0$ holds for all $\alpha \in \mathcal{Y}$, and $\boldsymbol{J}'_{(\alpha)}$ satisfies the boundary conditions for all $\alpha \in \mathcal{X}$. Thus, we obtain Eq. (70). We remind the reader that we also impose the periodic boundary condition on every field containing $\vec{\phi}$ if we impose it on the system.

Onsager operator.—Unifying the mobility tensor and the edgewise Onsager coefficient matrix, we introduce the Onsager operator $\mathcal{M} := \overleftrightarrow{\mathsf{M}} \oplus \mathsf{m}$ as the direct sum between $\overleftrightarrow{\mathsf{M}}$ and m, which maps forces to currents as

$$\mathcal{M} \mathcal{F}' = (\overleftrightarrow{\mathsf{M}} \vec{\boldsymbol{F}}', \mathsf{m} \boldsymbol{f}') \tag{72}$$

The Onsager operator $\mathcal{M}$ possibly depends on the concentration distribution in the same way that the mobility tensor and the edgewise Onsager coefficient matrix do.

The Onsager operator allows us to unify the linear relations between the diffusion and reaction forces and currents in Eqs. (52) and (57) as

$$\mathcal{J}(\boldsymbol{r};t) = \mathcal{M}(\boldsymbol{r};t)\mathcal{F}(\boldsymbol{r};t), \tag{73}$$

and the positive-definiteness of $\overset{\leftrightarrow}{\mathsf{M}}$ and m makes it invertible. The linear relation (73) lets us rewrite the dynamics of the internal species as

$$\partial_t \vec{c}_{\mathcal{X}} = (\nabla^\dagger \mathcal{M}\mathcal{F})_{\mathcal{X}}. \tag{74}$$

Since $\overset{\leftrightarrow}{\mathsf{M}}$ and m are symmetric, we obtain $\langle\!\langle \mathcal{M}\mathcal{F}', \mathcal{F}'' \rangle\!\rangle = \langle\!\langle \mathcal{F}', \mathcal{M}\mathcal{F}'' \rangle\!\rangle$ for any forces $\mathcal{F}'$ and $\mathcal{F}''$, which means that $\mathcal{M}$ is a self-adjoint operator. This property and the positive definiteness of $\mathcal{M}$ let us define a new inner product $\langle\!\langle \cdot, \cdot \rangle\!\rangle_{\mathcal{M}}$ as $\langle\!\langle \mathcal{F}', \mathcal{F}'' \rangle\!\rangle_{\mathcal{M}} := \langle\!\langle \mathcal{M}\mathcal{F}', \mathcal{F}'' \rangle\!\rangle = \langle\!\langle \mathcal{F}', \mathcal{M}\mathcal{F}'' \rangle\!\rangle$. The positive-definiteness also guarantees that the inner product is nondegenerate, i.e., $\langle\!\langle \mathcal{F}', \mathcal{F}' \rangle\!\rangle_{\mathcal{M}} > 0$ holds for any $\mathcal{F}' \neq (\vec{\mathbf{0}}, \mathbf{0})$. This inner product induced by $\mathcal{M}$ rewrites the EPR as the squared norm of the force as

$$\sigma = \langle\!\langle \mathcal{J}, \mathcal{F} \rangle\!\rangle = \langle\!\langle \mathcal{M}\mathcal{F}, \mathcal{F} \rangle\!\rangle = \langle\!\langle \mathcal{F}, \mathcal{F} \rangle\!\rangle_{\mathcal{M}}. \tag{75}$$

Now, the second law of thermodynamics is given by the nonnegativity of the norm, $\langle\!\langle \mathcal{F}, \mathcal{F} \rangle\!\rangle_{\mathcal{M}} \geqslant 0$.

E. Conservative and nonconservative forces

RDSs are driven by two types of forces: one is the force solely due to the chemical potential of the internal species, and the other is the force owing to the interaction with the outside of the system, i.e., the chemical potential of the external species and the nonconservative mechanical force $\vec{\boldsymbol{K}}^{\mathrm{nc}}$.

From this viewpoint, we can rewrite the force (67) as

$$\mathcal{F} = -\nabla \begin{pmatrix} \vec{\mu}_{\mathcal{X}} \\ \vec{0}_{\mathcal{Y}} \end{pmatrix} - \nabla \begin{pmatrix} \vec{0}_{\mathcal{X}} \\ \vec{\mu}_{\mathcal{Y}} \end{pmatrix} + \mathcal{K}^{\mathrm{nc}}. \tag{76}$$

Here, the first term is determined solely by the chemical potential of the internal species. The remaining two terms are the contributions from the chemical potential of the external species and the nonconservative mechanical force.

Inspired by the form in Eq. (76), we can decompose the force $\mathcal{F}$ into two parts as

$$\mathcal{F} = \nabla\vec{\phi} + \mathcal{F}^{\mathrm{nc}}, \tag{77}$$

where $\vec{\phi}$ in the first term is a potential whose external part is the zero vector, $\vec{\phi}_{\mathcal{Y}} = \vec{0}_{\mathcal{Y}}$. Here, the second term $\mathcal{F}^{\mathrm{nc}}$ is the remainder $\mathcal{F} - \nabla\vec{\phi}$. We refer to $\nabla\vec{\phi}$ and $\mathcal{F}^{\mathrm{nc}}$ as the conservative force and the nonconservative force, respectively. We remark that such a decomposition of the force into conservative and nonconservative forces is not unique. The representation with the chemical potential in Eq. (76) corresponds to the case where $\vec{\phi} = -(\vec{\mu}_{\mathcal{X}}^\top, \vec{0}_{\mathcal{Y}}^\top)^\top$ in Eq. (77). It indicates that $-\vec{\phi}$ may be easier to interpret thermodynamically than $\vec{\phi}$.

Using the decomposition in Eq. (77), we can rewrite the EPR $\sigma = \langle\!\langle \mathcal{J}, \mathcal{F} \rangle\!\rangle$ as

$$\sigma = \langle\!\langle \mathcal{J}, \nabla\vec{\phi} + \mathcal{F}^{\mathrm{nc}} \rangle\!\rangle$$
$$= \langle \partial_t \vec{c}, \vec{\phi} \rangle + \langle\!\langle \mathcal{J}, \mathcal{F}^{\mathrm{nc}} \rangle\!\rangle, \tag{78}$$

where we use the conjugation relation (70) and the assumption that $\phi_\alpha = 0$ for all $\alpha \in \mathcal{Y}$. If the system is driven solely by the conservative force $\nabla\vec{\phi}$, the second term in Eq. (78) vanishes so that the EPR becomes zero at a steady state, i.e., the system is in equilibrium at the steady state. The conservative force $\nabla\vec{\phi}$ drives relaxation to a state corresponding to $\vec{\phi}$ (see also Appendix C 2). On the other hand, the second term in Eq. (78) arising from the nonconservative force $\mathcal{F}^{\mathrm{nc}}$ may not be zero at the steady state. The nonconservative force $\mathcal{F}^{\mathrm{nc}}$ maintains the system out of equilibrium even at the steady state.

IV. GEOMETRIC DECOMPOSITIONS OF ENTROPY PRODUCTION RATE FOR REACTION-DIFFUSION SYSTEMS

One way to understand the thermodynamics of RDSs is to decompose dissipation into contributions from different causes. These decompositions can be performed by focusing on the geometry of the forces as follows. Since the EPR is given by the squared norm of the force as $\sigma = \langle\!\langle \mathcal{F}, \mathcal{F} \rangle\!\rangle_{\mathcal{M}}$, we can decompose it by using the Pythagorean theorem as

$$\langle\!\langle \mathcal{F}, \mathcal{F} \rangle\!\rangle_{\mathcal{M}} = \langle\!\langle \mathcal{F}^\perp, \mathcal{F}^\perp \rangle\!\rangle_{\mathcal{M}} + \langle\!\langle \mathcal{F} - \mathcal{F}^\perp, \mathcal{F} - \mathcal{F}^\perp \rangle\!\rangle_{\mathcal{M}}, \tag{79}$$

with the force $\mathcal{F}^\perp$ satisfying the orthogonality $\langle\!\langle \mathcal{F}^\perp, \mathcal{F} - \mathcal{F}^\perp \rangle\!\rangle_{\mathcal{M}} = 0$. For example, the decomposition into the contributions from diffusion and reactions (61) is rewritten using the Pythagorean theorem as

$$\langle\!\langle \mathcal{F}, \mathcal{F} \rangle\!\rangle_{\mathcal{M}} = \langle\!\langle \mathcal{F}^{\mathrm{diff}}, \mathcal{F}^{\mathrm{diff}} \rangle\!\rangle_{\mathcal{M}} + \langle\!\langle \mathcal{F}^{\mathrm{react}}, \mathcal{F}^{\mathrm{react}} \rangle\!\rangle_{\mathcal{M}}, \tag{80}$$

where the forces $\mathcal{F}^{\mathrm{diff}} := (\vec{\boldsymbol{F}}, \mathbf{0})$ and $\mathcal{F}^{\mathrm{react}} := (\vec{\mathbf{0}}, \boldsymbol{f})$ provide an orthogonal decomposition of the force as $\mathcal{F} = \mathcal{F}^{\mathrm{diff}} + \mathcal{F}^{\mathrm{react}}$ with $\langle\!\langle \mathcal{F}^{\mathrm{diff}}, \mathcal{F}^{\mathrm{react}} \rangle\!\rangle_{\mathcal{M}} = 0$.

In this section, we introduce more complicated decompositions, using the same geometric method. In Sec. IV A, we derive the geometric excess/housekeeping decomposition of the EPR by constructing an orthogonal decomposition of the force into the conservative and nonconservative parts. Notably, the excess EPR plays a central role in relating time evolution and dissipation, since it extracts the part of the dissipation rate that contributes to the change in the pattern. We also develop the local decomposition and the wavenumber decomposition, which enable us to identify the dissipation at each point in real space and Fourier space, in Sec. IV B. We demonstrate our geometric decompositions using two models, which are simple but present typical behaviors of RDSs, in Sec. IV C. We remark that we can further decompose the decompositions obtained in this section into contributions from reaction and diffusion in the same way as in Eq. (61).

A. Excess and housekeeping entropy production rate for reaction-diffusion systems

The EPR $\sigma = \langle\!\langle \mathcal{F}, \mathcal{F} \rangle\!\rangle_{\mathcal{M}}$ includes contributions from both conservative and nonconservative sources, as shown in Eq. (77). To quantify these two contributions separately, we construct the geometric excess/housekeeping decomposition of EPR for RDSs by using the Pythagorean theorem,

$$\langle\!\langle \mathcal{F}, \mathcal{F} \rangle\!\rangle_{\mathcal{M}} = \langle\!\langle \mathcal{F}^*, \mathcal{F}^* \rangle\!\rangle_{\mathcal{M}} + \langle\!\langle \mathcal{F} - \mathcal{F}^*, \mathcal{F} - \mathcal{F}^* \rangle\!\rangle_{\mathcal{M}}, \quad (81)$$

where $\mathcal{F}^*$ is a conservative force satisfying the orthogonality $\langle\!\langle \mathcal{F}^*, \mathcal{F} - \mathcal{F}^* \rangle\!\rangle$. In contrast to the simple case (80), it is difficult to find $\mathcal{F}^*$ directly. Instead, we obtain it by projecting the force $\mathcal{F}$ onto the conservative force space, which is defined as

$$\mathrm{Im}_{\mathcal{X}} \nabla := \{\nabla\vec{\phi} \,|\, \forall\alpha \in \mathcal{Y}, \, \phi_\alpha = 0\}. \quad (82)$$

Here, we use that conservative forces can be written as the generalized gradient of a potential whose external part is the zero vector, as discussed in Sec. III E. As we will confirm shortly afterward, the conservative force $\mathcal{F}^*$ obtained by the projection satisfies the orthogonality and allows us to use the Pythagorean theorem (81) to decompose EPR into contributions from conservative and nonconservative forces.

Projection of the force onto the conservative force space.— The projection of the force onto the conservative force space is defined by a variational problem

$$\mathcal{F}^* := \underset{\mathcal{F}' \in \mathrm{Im}_{\mathcal{X}}\nabla}{\arg\min} \langle\!\langle \mathcal{F} - \mathcal{F}', \mathcal{F} - \mathcal{F}' \rangle\!\rangle_{\mathcal{M}}, \quad (83)$$

where we impose the same boundary condition on the diffusion part of $\mathcal{M}\mathcal{F}'$ as we do on $\mathcal{J} = \mathcal{M}\mathcal{F}$. For example, if we consider a system with the no-flux boundary condition on the diffusion currents of the internal species, then we also impose the same condition on $(\overleftrightarrow{\mathbf{M}}\vec{F}')_{\mathcal{X}}$ in the minimization problem. By definition, $\mathcal{F}^*$ can be given as $\mathcal{F}^* = \nabla\vec{\phi}^*$ with

$$\vec{\phi}^* \in \underset{\vec{\phi}|\phi_{\mathcal{Y}}=\vec{0}_{\mathcal{Y}}}{\arg\min} \langle\!\langle \mathcal{F} - \nabla\vec{\phi}, \mathcal{F} - \nabla\vec{\phi} \rangle\!\rangle_{\mathcal{M}}. \quad (84)$$

Here, we also impose the same boundary condition on the diffusion part of $\mathcal{M}\nabla\vec{\phi}$ as we do on $\mathcal{J} = \mathcal{M}\mathcal{F}$ in the minimization problem. The minimization problem in Eq. (84) reduces to solving the partial differential equation, which is obtained as the Euler–Lagrange equation,

$$(\nabla^\dagger \mathcal{M}\mathcal{F})_{\mathcal{X}} = (\nabla^\dagger \mathcal{M}\nabla\vec{\phi}^*)_{\mathcal{X}}, \quad (85)$$

with the same boundary condition on $(\overleftrightarrow{\mathbf{M}}\boldsymbol{\nabla}_{\boldsymbol{r}}\vec{\phi}^*)_{\mathcal{X}}$ as we imposed on the original dynamics of internal species. Note that we can uniquely determine $\mathcal{F}^*$ since the inner product $\langle\!\langle \cdot, \cdot \rangle\!\rangle_{\mathcal{M}}$ is nondegenerate (see Appendix B 1 for details).

A notable property of the projected conservative force $\mathcal{F}^* = \nabla\vec{\phi}^*$ is that it preserves the time evolution of the original dynamics of internal species, since Eq. (85) provides the same time evolution of $\vec{c}_{\mathcal{X}}$:

$$\partial_t \vec{c}_{\mathcal{X}} = (\nabla^\dagger \mathcal{M}\mathcal{F})_{\mathcal{X}} = (\nabla^\dagger \mathcal{M}\nabla\vec{\phi}^*)_{\mathcal{X}}. \quad (86)$$

From this dynamics-conservation viewpoint, we may come up with another representation of $\mathcal{F}^*$ as

$$\mathcal{F}^* = \underset{\mathcal{F}'|(\nabla^\dagger\mathcal{M}\mathcal{F}')_{\mathcal{X}}=(\nabla^\dagger\mathcal{M}\mathcal{F})_{\mathcal{X}}}{\arg\min} \langle\!\langle \mathcal{F}', \mathcal{F}' \rangle\!\rangle_{\mathcal{M}}, \quad (87)$$
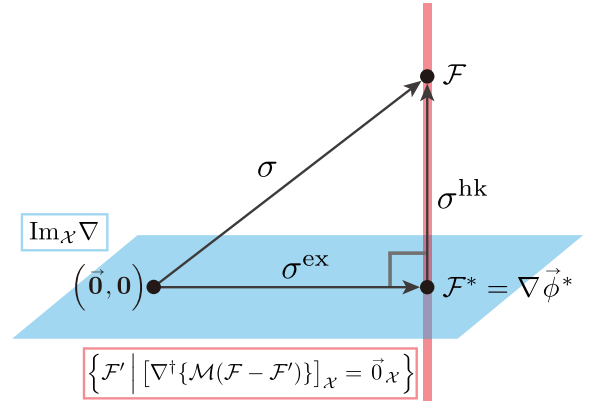

FIG. 3. The geometric decomposition of the EPR for RDSs. Here, the blue plane indicates $\mathrm{Im}_{\mathcal{X}}\nabla$, and the red line indicates $\{\mathcal{F}'|[\nabla^\dagger\{\mathcal{M}(\mathcal{F} - \mathcal{F}')\}]_{\mathcal{X}} = \vec{0}_{\mathcal{X}}\}$. The thermodynamic force, whose squared norm is the EPR $\sigma = \langle\!\langle \mathcal{F}, \mathcal{F} \rangle\!\rangle_{\mathcal{M}}$, is decomposed into the two orthogonal parts: the projection of $\mathcal{F}$ onto $\mathrm{Im}_{\mathcal{X}}\nabla$, $\mathcal{F}^*$, whose squared norm is the excess EPR $\sigma^{\mathrm{ex}} = \langle\!\langle \mathcal{F}^*, \mathcal{F}^* \rangle\!\rangle_{\mathcal{M}}$, and the remaining part, $\mathcal{F} - \mathcal{F}^*$, whose squared norm is the housekeeping EPR $\sigma^{\mathrm{hk}} = \langle\!\langle \mathcal{F} - \mathcal{F}^*, \mathcal{F} - \mathcal{F}^* \rangle\!\rangle_{\mathcal{M}}$. This description is parallel to the case of the Langevin systems shown in Fig. 2.

with the same boundary condition as Eq. (83). We can actually check that the minimization problem in Eq. (87) leads to the same Euler–Lagrange equation as Eq. (85). Note that the potential $\vec{\phi}$, whose external part is the zero vector, is introduced as the Lagrange multiplier for the constraint $(\nabla^\dagger\mathcal{M}\mathcal{F}')_{\mathcal{X}} = (\nabla^\dagger\mathcal{M}\mathcal{F})_{\mathcal{X}}$ in Eq. (87) (see Appendix B 2).

As already mentioned, the projected conservative force $\mathcal{F}^*$ is orthogonal to $\mathcal{F} - \mathcal{F}^*$ with respect to the inner product $\langle\!\langle \cdot, \cdot \rangle\!\rangle_{\mathcal{M}}$ as $\langle\!\langle \mathcal{F} - \mathcal{F}^*, \mathcal{F}^* \rangle\!\rangle_{\mathcal{M}} = \langle\!\langle \mathcal{M}(\mathcal{F} - \mathcal{F}^*), \nabla\vec{\phi}^* \rangle\!\rangle = \langle \nabla^\dagger\mathcal{M}(\mathcal{F} - \mathcal{F}^*), \vec{\phi}^* \rangle = 0$. Here, we used the boundary condition on $(\overleftrightarrow{\mathbf{M}}\boldsymbol{\nabla}_{\boldsymbol{r}}\vec{\phi}^*)_{\mathcal{X}}$, the condition that $\phi_\alpha^* = 0$ for all $\alpha \in \mathcal{Y}$, and Eq. (70) in the second transformation. We also used the condition $\phi_\alpha^* = 0$ for all $\alpha \in \mathcal{Y}$ and the Euler–Lagrange equation (85) in the third transformation. This orthogonality finally leads to the Pythagorean theorem (81), which enables us to define a decomposition of the EPR.

Excess/housekeeping decomposition of EPR.—Defining the excess and housekeeping EPRs (see also Fig. 3) as

$$\sigma^{\mathrm{ex}} := \langle\!\langle \mathcal{F}^*, \mathcal{F}^* \rangle\!\rangle_{\mathcal{M}}, \quad \sigma^{\mathrm{hk}} := \langle\!\langle \mathcal{F} - \mathcal{F}^*, \mathcal{F} - \mathcal{F}^* \rangle\!\rangle_{\mathcal{M}}, \quad (88)$$

the Pythagorean theorem (81) leads to the geometric excess/housekeeping decomposition of the EPR

$$\sigma = \sigma^{\mathrm{ex}} + \sigma^{\mathrm{hk}}. \quad (89)$$

Here, the nonnegativity of the EPRs is ensured, since each term is the squared norm of the corresponding force. The time integration of the decomposition (89) immediately yields the geometric excess/housekeeping decomposition of the EP as

$$\Sigma_\tau = \Sigma_\tau^{\mathrm{ex}} + \Sigma_\tau^{\mathrm{hk}}. \quad (90)$$

The physical meaning of the excess and housekeeping EPRs is explained based on the properties of $\mathcal{F}^*$ in the following.

We focus on the excess EPR, which plays a central role in the following sections. The minimization problems, Eq. (83)

and Eq. (87), imply that the excess EPR σ^{ex} is the minimum dissipation incurred by any conservative force that induces the original dynamics of the internal species. Indeed, the property of $\mathcal{F}^*$ (87) gives the excess EPR by the following optimization problem,

$$\sigma^{\mathrm{ex}} = \inf_{\mathcal{F}'|(\nabla^\dagger \mathcal{M}\mathcal{F}')_{\mathcal{X}} = (\nabla^\dagger \mathcal{M}\mathcal{F})_{\mathcal{X}}} \langle\!\langle \mathcal{F}', \mathcal{F}' \rangle\!\rangle_{\mathcal{M}}, \quad (91)$$

which means that the excess EPR extracts the unavoidable dissipation due to the time evolution of the pattern. The variational expression (91) also shows that the excess EPR vanishes at the steady state when the system is closed, since the zero function $\mathcal{F}^* = (\vec{\mathbf{0}}, \mathbf{0})$ then satisfies the constraints. If the system is open, the excess EPR will be zero as long as the concentrations of the internal species are stationary, even if those of the external species are not. We also remark that we can rewrite σ^{ex} as the time derivative of a quantity, defined in terms of the conservative force, that drives relaxation [45] (see also Appendix C).

We emphasize that the excess EPR defined here is different from Prigogine's one for RDSs [22]. The latter is defined only near the steady state and introduced to predict the stability of the steady state by its sign. On the other hand, the former is always nonnegative and can be defined far from the steady state or even in systems without steady states. In essence, it identifies the minimum dissipation required for the time evolution. The common feature of both is that they become zero in the steady state.

We also reveal the physical meaning of the remaining contribution, namely, the housekeeping EPR

$$\sigma^{\mathrm{hk}} = \inf_{\mathcal{F}' \in \mathrm{Im}_{\mathcal{X}} \nabla} \langle\!\langle \mathcal{F} - \mathcal{F}', \mathcal{F} - \mathcal{F}' \rangle\!\rangle_{\mathcal{M}}$$
$$= \inf_{\vec{\phi}|\vec{\phi}_{\mathcal{Y}} = \vec{0}_{\mathcal{Y}}} \langle\!\langle \mathcal{F} - \nabla\vec{\phi}, \mathcal{F} - \nabla\vec{\phi} \rangle\!\rangle_{\mathcal{M}}. \quad (92)$$

Although this paper mainly focuses on the excess EPR, the remaining housekeeping EPR also has an important physical meaning: it is the dissipation caused by the current that maintains the pattern without changing it. To show this, we rewrite the Euler–Lagrange equation (85) with the current corresponding to the projected conservative force $\mathcal{J}^* = \mathcal{M}\mathcal{F}^*$ as $(\nabla^\dagger \mathcal{J})_{\mathcal{X}} = (\nabla^\dagger \mathcal{J}^*)_{\mathcal{X}}$. This indicates that $\mathcal{J} - \mathcal{J}^*$ is a *cyclic current*, which does not affect the dynamics of the internal species as

$$\partial_t \vec{c}_{\mathcal{X}} = (\nabla^\dagger \mathcal{J}^*)_{\mathcal{X}} + (\nabla^\dagger [\mathcal{J} - \mathcal{J}^*])_{\mathcal{X}} = (\nabla^\dagger \mathcal{J}^*)_{\mathcal{X}}. \quad (93)$$

It is worth noting that while the cyclic currents are cyclic in terms of internal species, physically they are driven by external factors, such as external species or nonconservative mechanical forces. This is also the case with homogeneous CRNs without detailed balance, where external species are often made implicit. If no external species exist, the system will be detailed balanced, and no cyclic motion will be observed.

We also remark that the cyclic current $\mathcal{J} - \mathcal{J}^*$ can affect the concentrations of the external species. Observe that $\partial_t \vec{c}_{\mathcal{Y}} = (\nabla^\dagger \mathcal{J}^*)_{\mathcal{Y}} + (\nabla^\dagger [\mathcal{J} - \mathcal{J}^*])_{\mathcal{Y}}$, where the second term on the right-hand side does not vanish generally. In other words, the housekeeping EPR consists of the cyclic contribution plus the contribution from the diffusion of external species. Suppose that the mobility tensor has no direct

interaction terms between internal and external species, $\mathsf{M}_{(\alpha\beta)} = \mathsf{O}$ if $(\alpha, \beta) \in \mathcal{X} \times \mathcal{Y}$ or $\mathcal{Y} \times \mathcal{X}$, where O is the zero matrix. Then, the interpretation is clearly depicted by the decomposition $\sigma^{\mathrm{hk}} = \sigma_{\mathcal{X}}^{\mathrm{cyc}} + \sigma_{\mathcal{Y}}^{\mathrm{diff}}$ with two nonnegative contributions

$$\sigma_{\mathcal{X}}^{\mathrm{cyc}} := \int_V d\mathbf{r} \Bigg[(\mathbf{f} - \nabla_{\mathsf{s}}\vec{\phi}^*)^\top \mathsf{m}(\mathbf{f} - \nabla_{\mathsf{s}}\vec{\phi}^*)$$
$$+ \sum_{\alpha,\beta \in \mathcal{X}} (\mathbf{F}_{(\alpha)} - \nabla_{\mathbf{r}}\phi_\alpha^*)^\top \mathsf{M}_{(\alpha\beta)}(\mathbf{F}_{(\beta)} - \nabla_{\mathbf{r}}\phi_\beta^*) \Bigg] \geqslant 0, \quad (94)$$

$$\sigma_{\mathcal{Y}}^{\mathrm{diff}} := \int_V d\mathbf{r} \sum_{\alpha \in \mathcal{Y}} \mathbf{J}_{(\alpha)} \cdot \mathbf{F}_{(\alpha)} \geqslant 0. \quad (95)$$

The first term $\sigma_{\mathcal{X}}^{\mathrm{cyc}}$ reflects the cyclic motion of the internal species, arising from the cyclic current $\mathcal{J} - \mathcal{J}^*$. The remainder reflects the dissipation stemming from diffusion of the external species. This term is particular to RDSs, as the housekeeping EPR in a homogeneous CRN can be written only with cyclic contributions [45]. The decomposition is straightforwardly proved using the definition of σ^{hk} and the fact that $\mathbf{F}_{(\alpha)} - \nabla_{\mathbf{r}}\phi_\alpha^* = \mathbf{F}_{(\alpha)}$ for any $\alpha \in \mathcal{Y}$ because $\phi_\alpha^* = 0$ for any $\alpha \in \mathcal{Y}$ with the assumption $\mathsf{M}_{(\alpha\beta)} = \mathsf{O}$ if $(\alpha, \beta) \in \mathcal{X} \times \mathcal{Y}$ or $(\alpha, \beta) \in \mathcal{Y} \times \mathcal{X}$.

B. Local decomposition and wavenumber decomposition of entropy production rate

In this section, we discuss the local and wavenumber decomposition of EPR for RDSs, analogous to the decomposition for Langevin systems from Sec. II C. The local and wavenumber EPRs enable us to quantify the dissipation at each point in the real and wavenumber spaces, respectively. Thus, we can detect local dissipation caused by pattern formation via the local and wavenumber EPRs for RDSs.

Local decomposition.—We define the local EPR at the position $\mathbf{r}$ as

$$\sigma^{\mathrm{loc}}(\mathbf{r}) := \vec{\mathbf{J}}^\top(\mathbf{r}; t)\vec{\mathbf{F}}(\mathbf{r}; t) + \mathbf{j}^\top(\mathbf{r}; t)\mathbf{f}(\mathbf{r}; t). \quad (96)$$

The nonnegativity $\sigma^{\mathrm{loc}}(\mathbf{r}) \geqslant 0$ holds locally as it is locally that $\overleftrightarrow{\mathsf{M}}$ is assumed to be positive-definite and $\mathbf{j}$ and $\mathbf{f}$ have the same sign. The volume integral gives the EPR,

$$\sigma = \int_V d\mathbf{r}\, \sigma^{\mathrm{loc}}(\mathbf{r}). \quad (97)$$

This equality (97) indicates a decomposition of the EPR into the local EPRs, the dissipation at each location.

We also decompose the excess and housekeeping EPRs into the contributions from each location as

$$\sigma^{\mathrm{ex,loc}}(\mathbf{r}) := (\nabla_{\mathbf{r}}\vec{\phi}^*)^\top \overleftrightarrow{\mathsf{M}}\nabla_{\mathbf{r}}\vec{\phi}^* + (\nabla_{\mathsf{s}}\vec{\phi}^*)^\top \mathsf{m}\nabla_{\mathsf{s}}\vec{\phi}^*, \quad (98)$$
$$\sigma^{\mathrm{hk,loc}}(\mathbf{r}) := (\vec{\mathbf{F}} - \nabla_{\mathbf{r}}\vec{\phi}^*)^\top \overleftrightarrow{\mathsf{M}}(\vec{\mathbf{F}} - \nabla_{\mathbf{r}}\vec{\phi}^*)$$
$$+ (\mathbf{f} - \nabla_{\mathsf{s}}\vec{\phi}^*)^\top \mathsf{m}(\mathbf{f} - \nabla_{\mathsf{s}}\vec{\phi}^*), \quad (99)$$

which satisfy $\int_V d\mathbf{r}\, \sigma^{\mathrm{ex,loc}}(\mathbf{r}) = \sigma^{\mathrm{ex}}$ and $\int_V d\mathbf{r}\, \sigma^{\mathrm{hk,loc}}(\mathbf{r}) = \sigma^{\mathrm{hk}}$. The local excess and housekeeping EPRs are nonnegative because $\overleftrightarrow{\mathsf{M}}$ and m are positive-definite. Thus, we can interpret

$\sigma^{\mathrm{ex,loc}}(\boldsymbol{r})$ as the EPR due to the projected conservative force $\mathcal{F}^*$ at the location $\boldsymbol{r}$. We can also regard $\sigma^{\mathrm{hk,loc}}(\boldsymbol{r})$ as the EPR due to the cyclic current at the location $\boldsymbol{r}$. Note that $\sigma^{\mathrm{ex,loc}}(\boldsymbol{r}) = 0$ and $\sigma^{\mathrm{hk,loc}}(\boldsymbol{r}) = \sigma^{\mathrm{loc}}(\boldsymbol{r})$ hold for all $\boldsymbol{r} \in V$ if and only if $\nabla \vec{\phi}^* = (\vec{\mathbf{0}}, \mathbf{0})$. The time evolution of the internal species needs to be stationary to achieve this condition because $\partial_t \vec{c}_{\mathcal{X}} = (\nabla^\dagger \mathcal{M} \nabla \vec{\phi}^*)_{\mathcal{X}} = \vec{0}_{\mathcal{X}}$ holds when $\nabla \vec{\phi}^* = (\vec{\mathbf{0}}, \mathbf{0})$.

We remark that the definitions of the local excess and housekeeping EPRs include $\vec{\phi}^*$, which is defined in terms of global information about the RDS. For this reason, the local excess and housekeeping EPRs cannot be defined solely from local information. Reflecting this nonlocality, the local excess and housekeeping EPRs do not sum up to the local EPR $\sigma^{\mathrm{loc}}(\boldsymbol{r}) \neq \sigma^{\mathrm{ex,loc}}(\boldsymbol{r}) + \sigma^{\mathrm{hk,loc}}(\boldsymbol{r})$. The nonzero cross-term can be obtained as $\sigma^{\mathrm{cross}}(\boldsymbol{r}) := \sigma^{\mathrm{loc}}(\boldsymbol{r}) - \sigma^{\mathrm{ex,loc}}(\boldsymbol{r}) - \sigma^{\mathrm{hk,loc}}(\boldsymbol{r}) = 2(\nabla_{\boldsymbol{r}} \vec{\phi}^*)^\top \overleftrightarrow{\mathsf{M}} (\vec{F} - \nabla_{\boldsymbol{r}} \vec{\phi}^*) + 2(\nabla_{\mathsf{s}} \vec{\phi}^*)^\top \mathsf{m} (f - \nabla_{\mathsf{s}} \vec{\phi}^*)$, which may be positive or negative in sign. It also satisfies $\int_V d\boldsymbol{r} \, \sigma^{\mathrm{cross}}(\boldsymbol{r}) = 0$, which maintains the geometric decomposition globally as $\sigma = \sigma^{\mathrm{ex}} + \sigma^{\mathrm{hk}}$.

Wavenumber decomposition.—We also provide the wavenumber decomposition of the EPR using Parseval's identity. We define the weighted Fourier transform of forces $\hat{\vec{F}}'$ and $\hat{f}'$ as

$$\hat{F}'_{(\alpha)i}(\boldsymbol{k};t) := \int_V d\boldsymbol{r} \left[\overleftrightarrow{\mathsf{M}}^{\frac{1}{2}} \vec{F}' \right]_{(\alpha)i} e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}, \quad (100)$$

$$\hat{f}'_\rho(\boldsymbol{k};t) := \int_V d\boldsymbol{r} \left[\mathsf{m}^{\frac{1}{2}} f' \right]_\rho e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}. \quad (101)$$

Here, we use the square root of the mobility tensor $\overleftrightarrow{\mathsf{M}}^{\frac{1}{2}}$ and the edgewise Onsager coefficient matrix $\mathsf{m}^{\frac{1}{2}}$, which satisfy $\sum_{\gamma \in \mathcal{S}} \sum_{k=1}^d [\overleftrightarrow{\mathsf{M}}^{\frac{1}{2}}]_{(\alpha\gamma)ik} [\overleftrightarrow{\mathsf{M}}^{\frac{1}{2}}]_{(\gamma\beta)kj} = M_{(\alpha\beta)ij}$ and $[\mathsf{m}^{\frac{1}{2}}]_{\rho\rho'} = \sqrt{m_\rho} \delta_{\rho\rho'}$. The existence of these operators is guaranteed by the positive-definiteness of $\overleftrightarrow{\mathsf{M}}$ and m. Note that the elements of $\overleftrightarrow{\mathsf{M}}^{\frac{1}{2}}$ satisfy $[\overleftrightarrow{\mathsf{M}}^{\frac{1}{2}}]_{(\alpha\beta)ij}(\boldsymbol{r};t) = \sqrt{D_\alpha c_\alpha(\boldsymbol{r};t)} \delta_{\alpha\beta} \delta_{ij}$ if the mobility tensor has the simple form in Eq. (53).

Since periodic boundary conditions discretize the wavenumber, the details of the wavenumber decomposition differ slightly between cases where periodic boundary conditions are imposed and cases where they are not. If we do not impose periodic boundary conditions on the system, we define the wavenumber EPR as

$$\sigma^{\mathrm{wn}}(\boldsymbol{k}) := \frac{1}{(2\pi)^d} [\hat{\vec{F}}^\dagger(\boldsymbol{k};t) \hat{\vec{F}}(\boldsymbol{k};t) + \hat{f}^\dagger(\boldsymbol{k};t) \hat{f}(\boldsymbol{k};t)]$$

$$= \frac{1}{(2\pi)^d} \left[\sum_{\alpha \in \mathcal{S}} \sum_{i=1}^d \overline{\hat{F}_{(\alpha)i}} \hat{F}_{(\alpha)i} + \sum_\rho \overline{\hat{f}_\rho} \hat{f}_\rho \right] \geqslant 0,$$
$$(102)$$

where the superscript $\dagger$ indicates conjugate transpose. We can obtain the decomposition

$$\sigma = \int_{\mathbb{R}^d} d\boldsymbol{k} \, \sigma^{\mathrm{wn}}(\boldsymbol{k}), \quad (103)$$

by Parseval's identity. It follows from the Fourier transform of the delta function, as was the case for Langevin systems in Eq. (23) (see also Appendix D for the derivation). On the other hand, if we consider a system with periodic boundaries,

we define the wavenumber EPR as

$$\sigma^{\mathrm{wn}}(\boldsymbol{k}) := \frac{1}{|V|} [\hat{\vec{F}}^\dagger(\boldsymbol{k};t) \hat{\vec{F}}(\boldsymbol{k};t) + \hat{f}^\dagger(\boldsymbol{k};t) \hat{f}(\boldsymbol{k};t)]$$

$$= \frac{1}{|V|} \left[\sum_{\alpha \in \mathcal{S}} \sum_{i=1}^d \overline{\hat{F}_{(\alpha)i}} \hat{F}_{(\alpha)i} + \sum_\rho \overline{\hat{f}_\rho} \hat{f}_\rho \right] \geqslant 0, \quad (104)$$

Here, we let $|V|$ denote the volume of the space V. We consider discrete wavenumbers because of the periodic boundary conditions. With periodic boundary conditions, we can also obtain the decomposition

$$\sigma = \sum_{\boldsymbol{k}} \sigma^{\mathrm{wn}}(\boldsymbol{k}), \quad (105)$$

using the Fourier series expansion of the delta function, $\delta(\boldsymbol{r}) = \sum_{\boldsymbol{k}} e^{i\boldsymbol{k}\cdot\boldsymbol{r}}/|V|$ (see also Appendix D).

We can also define the wavenumber decomposition of the excess and housekeeping EPRs, $\sigma^{\mathrm{ex,wn}}$ and $\sigma^{\mathrm{hk,wn}}$, using $(\hat{\vec{F}}^*, \hat{f}^*)$ and $(\hat{\vec{F}} - \hat{\vec{F}}^*, \hat{f} - \hat{f}^*)$ instead of $(\hat{\vec{F}}, \hat{f})$, respectively. We can interpret $\sigma^{\mathrm{ex,wn}}(\boldsymbol{k})$ as the EPR owing to the projected conservative force $\mathcal{F}^*$ at the wavenumber $\boldsymbol{k}$. We can also regard $\sigma^{\mathrm{hk,wn}}(\boldsymbol{k})$ as the EPR caused by the cyclic current at the wavenumber $\boldsymbol{k}$. Note that the geometric excess/housekeeping decomposition can be violated at each wavenumber, $\sigma^{\mathrm{wn}}(\boldsymbol{k}) \neq \sigma^{\mathrm{ex,wn}}(\boldsymbol{k}) + \sigma^{\mathrm{hk,wn}}(\boldsymbol{k})$, as in the case for Langevin systems.

The local and wavenumber decompositions are based on the orthonormality of the basis. Therefore, it is also possible to decompose the EPR using an orthonormal basis other than the Fourier basis. For example, a wavelet basis [87,88] may allow us to quantify the dissipation corresponding to particular wave packets.

C. Numerical examples: Geometric decompositions

Here, we show numerical examples of the geometric decompositions for open RDSs. We discuss the Fisher–KPP equation and the Brusselator model in one dimension, $V_1 = [-0.5, 0.5]$. For numerical results here and in the following sections, we use r, k, ∂_r, $J_{(\alpha)}$ and $F_{(\alpha)}$ instead of $\boldsymbol{r}$, $\boldsymbol{k}$, $\nabla_{\boldsymbol{r}}$, $\boldsymbol{J}_{(\alpha)}$ and $\boldsymbol{F}_{(\alpha)}$ respectively, because we consider only one-dimensional systems. We use the same models with the same parameters for other numerical results in Secs. V F and VI C.

Fisher–KPP equation.—The Fisher–KPP equation consists of an internal species Z_1, an external species Z_2, and an autocatalytic reaction

$$Z_1 + Z_2 \underset{\kappa_1^-}{\overset{\kappa_1^+}{\rightleftharpoons}} 2Z_1. \quad (106)$$

Now, the index sets are $\mathcal{S} = \{1, 2\}$, $\mathcal{X} = \{1\}$, $\mathcal{Y} = \{2\}$, $\mathcal{R} = \{1\}$, and $\mathcal{R}_{\mathcal{X}} = \{1\} = \mathcal{R}$, and the vectors $\vec{c} = (c_1, c_2)^\top$, $\vec{c}_{\mathcal{X}} = (c_1)$, and $\vec{c}_{\mathcal{Y}} = (c_2)$. The stoichiometric matrix is

$$\nabla_{\mathsf{s}}^\top = \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \quad (107)$$

In this system, we assume that the concentration of the external species is kept homogeneous by the interaction with the outside: $c_2(r;t) = 1$ holds for all $r \in V_1$. We let the mobility tensor take the simple form in Eq. (53), and assume Fick's

law for the diffusion currents, $J_{(\alpha)} = -D_\alpha \partial_r c_\alpha$. Here, $J_{(2)} = 0$ because $c_2(r;t)$ is homogeneous and $\partial_r c_2 = 0$. We also assume mass action kinetics for the reaction fluxes: $j_1^+(r;t) = \kappa_1^+ c_1(r;t) c_2(r;t) = \kappa_1^+ c_1(r;t)$, $j_1^-(r;t) = \kappa_1^- c_1(r;t)^2$. Then, we can write the dynamics as

$$\partial_t c_1 = D_1 \partial_r^2 c_1 + \kappa_1^+ c_1 - \kappa_1^- c_1^2. \qquad (108)$$

We impose the no-flux boundary condition, $\partial_r c_1(r;t)|_{r=\pm 0.5} = 0$. We also use the parameters $D_1 = 10^{-4}$ and $(\kappa_1^+, \kappa_1^-) = (1, 1)$.

In the Fisher–KPP equation, we can explicitly write down the condition to determine the potential $\vec{\phi}^*$ (85) as

$$D_1 \partial_r^2 c_1 + \kappa_1^+ c_1 - \kappa_1^- c_1^2$$

$$= -\partial_r (D_1 c_1 \partial_r \phi_1^*) + \frac{\kappa_1^+ c_1 - \kappa_1^- c_1^2}{\ln(\kappa_1^+ c_1) - \ln(\kappa_1^- c_1^2)} \phi_1^*, \quad (109)$$

which $\phi_1^* = \ln(\kappa_1^+/(\kappa_1^- c_1))$ solves.

Therefore, the excess EPR is

$$\sigma^{\mathrm{ex}} = \int_{-0.5}^{0.5} dr \left[D_1 c_1 (\partial_r \phi_1^*)^2 + \frac{(\kappa_1^+ c_1 - \kappa_1^- c_1^2)(\phi_1^*)^2}{\ln(\kappa_1^+ c_1) - \ln(\kappa_1^- c_1^2)} \right]$$

$$= \int_{-0.5}^{0.5} dr \left[D_1 c_1 (\partial_r \ln c_1)^2 + (\kappa_1^+ c_1 - \kappa_1^- c_1^2) \ln \frac{\kappa_1^+}{\kappa_1^- c_1} \right].$$

This is the same as the total EPR σ, because $F_{(1)} = \partial_r \ln c_1$ and $f_1 = \ln(\kappa_1^+/(\kappa_1^- c_1))$.

In Fig. 4, we numerically show the time series of the local EPR σ^{loc} and the concentration distribution of internal species c_1. As is well known, in the Fisher–KPP equation, the area of high concentration of c_1 spreads over time [107]: Unlike standard diffusion, total concentration is not conserved. The reaction does not occur inside the high-concentration area because $c_1 = 1$ is the equilibrium concentration of the reaction in Eq. (106) with given parameters. No diffusion occurs inside the high-concentration area because the concentration gradient vanishes. As a result, the local EPR is larger at the boundary between the high- and low-concentration areas, with no dissipation inside the high-concentration area.

Brusselator model.—The Brusselator model consists of two internal species, Z_1 and Z_2, an external species Z_3, and three reactions,

$$Z_3 \underset{\kappa_1^-}{\overset{\kappa_1^+}{\rightleftharpoons}} Z_1, \quad Z_1 \underset{\kappa_2^-}{\overset{\kappa_2^+}{\rightleftharpoons}} Z_2, \quad 2Z_1 + Z_2 \underset{\kappa_3^-}{\overset{\kappa_3^+}{\rightleftharpoons}} 3Z_1, \quad (110)$$
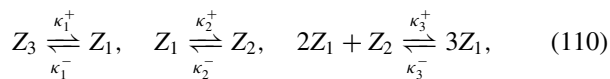
where we label the reactions $\rho = 1, 2, 3$ from left to right. The index sets are $\mathcal{S} = \{1, 2, 3\}$, $\mathcal{X} = \{1, 2\}$, $\mathcal{Y} = \{3\}$, $\mathcal{R} = \{1, 2, 3\}$, and $\mathcal{R}_\mathcal{X} = \{1, 2, 3\} = \mathcal{R}$, and the vectors $\vec{c} = (c_1, c_2, c_3)^\top$, $\vec{c}_\mathcal{X} = (c_1, c_2)^\top$, and $\vec{c}_\mathcal{Y} = (c_3)$. The stoichiometric matrix is

$$\nabla_s^\top = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \\ -1 & 0 & 0 \end{pmatrix}. \qquad (111)$$

The concentration of the external species is again assumed to be homogeneous due to the interaction with the outside: $c_3(r;t) = 1$ for all $r \in V_1$. We let the mobility tensor have the simple form in Eq. (53) and assume Fick's law for the diffusion currents, $J_{(\alpha)} = -D_\alpha \partial_r c_\alpha$. Here, $J_{(3)} = 0$ because
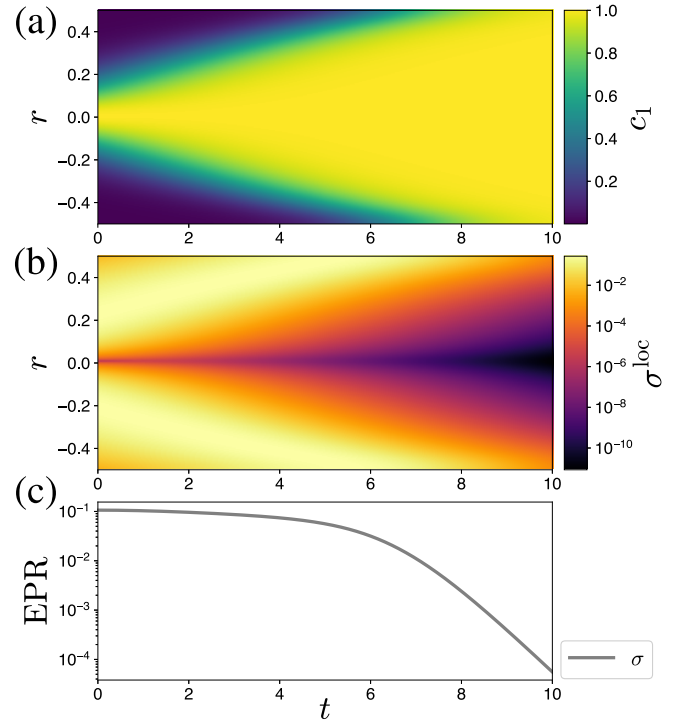

FIG. 4. (a) The time series of c_1 in the Fisher–KPP equation. The area with high c_1 (yellow) expands as time passes. (b) The local EPR σ^{loc} in the Fisher–KPP equation. The local EPR is colored on a logarithmic scale. Comparing (a) with (b), we can see that dissipation occurs at the boundary between high- and low-concentration areas. (c) The EPR σ in the Fisher–KPP equation. Reflecting the relaxation to the equilibrium state, the EPR is monotonically decreasing in time.

c_3 is homogeneous and $\partial_r c_3 = 0$. We also assume the mass action kinetics for the reaction fluxes: $j_1^+(r;t) = \kappa_1^+ c_3(r;t) = \kappa_1^+$, $j_1^-(r;t) = \kappa_1^- c_1(r;t)$, $j_2^+(r;t) = \kappa_2^+ c_1(r;t)$, $j_2^-(r;t) = \kappa_2^- c_2(r;t)$, $j_3^+(r;t) = \kappa_3^+ c_1(r;t)^2 c_2(r;t)$, and $j_3^-(r;t) = \kappa_3^- c_1(r;t)^3$. Then, the dynamics are given by

$$\begin{cases} \partial_t c_1 = D_1 \partial_r^2 c_1 + \kappa_1^+ - \kappa_1^- c_1 \\ \qquad\quad - \kappa_2^+ c_1 + \kappa_2^- c_2 + \kappa_3^+ c_1^2 c_2 - \kappa_3^- c_1^3 & (112) \\ \partial_t c_2 = D_2 \partial_r^2 c_2 + \kappa_2^+ c_1 - \kappa_2^- c_2 - \kappa_3^+ c_1^2 c_2 + \kappa_3^- c_1^3. & (113) \end{cases}$$

In the following numerical examples, we use the parameters as follows: $D_1 = 1.6 \times 10^{-4}$, $D_2 = 10^{-3}$, and $(\kappa_1^+, \kappa_1^-, \kappa_2^+, \kappa_2^-, \kappa_3^+, \kappa_3^-) = (1, 1, 10, 0.1, 1, 1)$. We also impose the periodic boundary conditions, which let each wavenumber be determined by an integer n as

$$k(n) = \frac{2n\pi}{|V_1|} = 2n\pi, \qquad (114)$$

where $|V_1| = 1$ is the system size. In the following, we abbreviate the wavenumber EPRs $\sigma^{\mathrm{wn}}(\boldsymbol{k})$ and $\sigma^{\mathrm{ex,wn}}(\boldsymbol{k})$ as $\sigma^{\mathrm{wn}}(n)$ and $\sigma^{\mathrm{ex,wn}}(n)$.

We demonstrate how the EPR and the excess EPR change quantitatively with the time evolution of the concentration distribution. Note that we numerically compute the excess EPR because the condition determining the potential ϕ^* (85) for the Brusselator model is difficult to solve analytically,
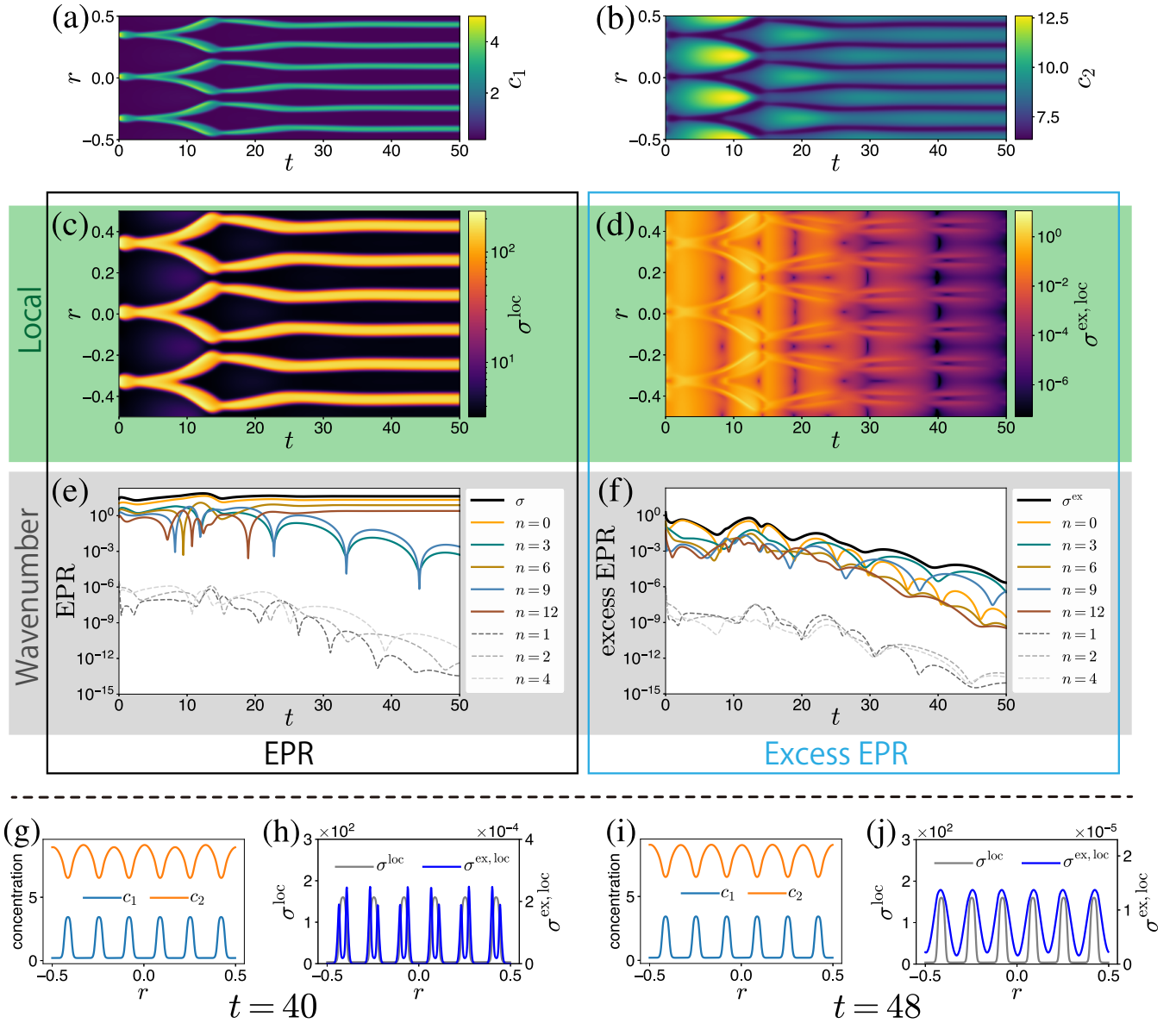
FIG. 5. The comparison of the EPR and the excess EPR in the Brusselator model. (a) The time series of c_1. (b) The time series of c_2. The symmetry of the pattern changes from threefold with time evolution. It reaches sixfold symmetry at around $t = 25$. (c) The local EPR σ^{loc} colored on a logarithmic scale. In contrast to the Fisher–KPP equation [Fig. 4(b)], σ^{loc} is large on the peaks of the pattern (areas where c_1 is high), not on the edges of the pattern. (d) The local excess EPR $\sigma^{\mathrm{ex,loc}}$ colored on a logarithmic scale. It tends to be larger on the edge of the pattern. (e) The EPR and the wavenumber EPRs for some modes. The EPR σ (black line) does not vanish to maintain the pattern. Because of the threefold symmetry of the pattern, the wavenumber EPR $\sigma^{\mathrm{wn}}(n)$ is small if $n \neq 3n'$ for any integer n'. Further, while the wavenumber EPRs of the form $\sigma^{\mathrm{wn}}(6n')$ do not decay, those with $n \neq 6n'$ decay even if n is a multiple of three. It reflects that the symmetry of the pattern becomes sixfold as the system approaches the steady state. (f) The excess EPR and the wavenumber excess EPRs for some modes. The excess EPR σ^{ex} (black line) is considerably smaller than the EPR and decreases as the system approaches the steady state. The threefold symmetry of the pattern makes the wavenumber excess EPR be small for n that are not multiples of three, as in the case of the wavenumber EPRs. In contrast to the EPR, there is a qualitative order $\sigma^{\mathrm{ex,wn}}(6n') \leqslant \sigma^{\mathrm{ex,wn}}(3n')$. This is caused by two facts: (1) the excess EPR reflects the change in pattern, and (2) the Fourier components corresponding to wavenumbers that are multiples of three and not six decay as the system approaches the steady state, while those corresponding to multiples of six barely change. (g) Concentrations $c_1(r;t)$ (cyan line) and $c_2(r;t)$ (orange line) as a function of spatial location r at $t = 40$. (h) Values of $\sigma^{\mathrm{loc}}(r)$ (gray line) and $\sigma^{\mathrm{ex,loc}}(r)$ (blue line) as a function of spatial location r at $t = 40$. The local EPR becomes large on the peaks of the pattern, while the local excess EPR has strong peaks at the edges of the pattern. (i) Values of $\vec{c}_{\mathcal{X}}(t)$ as a function of spatial location r at $t = 48$. (j) The values $\sigma^{\mathrm{loc}}(r)$ (gray line) and $\sigma^{\mathrm{ex,loc}}(r)$ (blue line) as a function of spatial location r at $t = 48$. Both of the local EPRs become large on the peaks of the pattern.

unlike the case of the Fisher–KPP equation. The time series of the concentration distribution and the EPRs are shown in Fig. 5.

We can see that the excess EPR [black line in Fig. 5(f)] decreases (nonmonotonically) as the system approaches the steady state. In addition, the excess EPR σ^{ex} is much smaller

than the EPR σ [black line in Fig. 5(e)] because the majority of the dissipation is the housekeeping EPR, which is caused by cyclic currents that do not affect the time evolution. In other words, for these parameter values, more dissipation is used to maintain the pattern than to change it.

The local EPRs show the difference between the EPR and the excess EPR in more detail. In contrast to the previous model, the local EPR σ^{loc} is larger on the pattern [areas where c_1 is high in Fig. 5(a)], rather than the edges of the pattern. This tendency reflects that the violation of the detailed balance requires dissipation to maintain the pattern, even when the change in the concentration is stationary. On the other hand, the local excess EPR $\sigma^{\mathrm{ex,loc}}$ is large at the edges of the pattern unless the system is close to the steady state. This reflects the following fact: When the system is far from the steady state, the time evolution tends to be faster at the edges of the pattern, where the gradients of concentrations are larger. We also show the details of these tendencies in Figs. 5(g)–5(j): at $t = 40$, the local excess EPR is large on the edge of the pattern, while the local EPR is large on the pattern, as shown in Fig. 5(h). On the other hand, both local EPRs take a large value on the pattern once the system approaches steady state ($t = 48$), as shown in Fig. 5(j).

The wavenumber decomposition shows the difference between the EPR and the excess EPR more explicitly. This is because the symmetry of the pattern is clear in this numerical example. Here, we focus on the change in the symmetry of the pattern from threefold to sixfold. Near the steady state, the wavenumber EPR $\sigma^{\mathrm{wn}}(n)$ is larger at n that are multiples of six than at other multiples of three [Fig. 5(e)]. This reflects that the symmetry of the pattern is sixfold at the steady state. On the other hand, near the steady state, the wavenumber excess EPR is larger at n that are multiples of three but not multiples of six [Fig. 5(f)]. This reflects the faster decay of the wavenumber components of the pattern corresponding to n that are multiples of three rather than multiples of six.

V. OPTIMAL TRANSPORT AND THERMODYNAMIC SPEED LIMITS FOR REACTION-DIFFUSION SYSTEMS

We can understand dissipation in RDSs from the perspective of thermodynamic trade-off relations, which quantify the minimum dissipation required to achieve an objective. In particular, we focus on the thermodynamic speed limits (TSLs), which are trade-off relations between the speed of the dynamics and dissipation. They are geometric relations since they typically use some measure of "distance" between the initial and final patterns to quantify the speed. We measure the distance between two patterns of an RDS with the Wasserstein distance, similar to Langevin systems and MJPs [39–45,65–68].

An RDS is a composite of chemical reactions and diffusive dynamics. Since some kinds of Wasserstein distance have been studied for both types of dynamics, we can generalize the 1-Wasserstein and 2-Wasserstein distances to RDSs to derive TSLs. Sections V A and V B are dedicated to the generalization of the Wasserstein distances, while some differences between these distances are discussed in Sec. V C. We derive TSLs with the 1- and 2-Wasserstein distances for RDSs using a connection between the 2-Wasserstein distance

and the excess EPR in Sec. V D. We revisit the TSLs in terms of minimum dissipation in Sec. V E. The final section V F provides numerical demonstrations of the TSLs.

Note that we need to specify boundary conditions to define the Wasserstein distance variationally (otherwise, it will not be well defined). We adopt here the boundary conditions discussed in Sec. III B for quantities that are considered as currents, e.g., a quantity obtained by acting the mobility tensor on a force.

A. 1-Wasserstein distance for reaction-diffusion systems

Fixing the boundary conditions and stoichiometry, we define the 1-Wasserstein distance by generalizing the Beckmann problem (30) as

$$W_{1,\mathcal{X}}(\vec{c}^A, \vec{c}^B) := \inf_{\mathcal{J}'} \left\{ \int_0^\tau dt |\mathcal{J}'|_{\mathrm{RD}} \right\}, \qquad (115)$$

with the norm of current $|\cdot|_{\mathrm{RD}}$ defined as

$$|\mathcal{J}'|_{\mathrm{RD}} := \int_V dr \left[\sum_{\alpha \in \mathcal{S}} \|\boldsymbol{J}'_{(\alpha)}\| + \sum_{\rho \in \mathcal{R}} |j'_\rho| \right]. \qquad (116)$$

Here, we impose the following condition on $\mathcal{J}'$: there exists a time series of concentration distributions $\vec{c}'$ satisfying

$$\partial_t \vec{c}'_{\mathcal{X}} = (\nabla^\dagger \mathcal{J}')_{\mathcal{X}}, \quad \vec{c}'_{\mathcal{X}}(0) = \vec{c}^A_{\mathcal{X}}, \quad \vec{c}'_{\mathcal{X}}(\tau) = \vec{c}^B_{\mathcal{X}}. \quad (117)$$

The condition means that the minimization is performed over all time series that obey the continuity equation with current $\mathcal{J}'$ connecting $\vec{c}^A$ and $\vec{c}^B$ only with respect to the internal species. Note that the concentrations of the external species are irrelevant in the formula. As a result, the 1-Wasserstein distance can be zero even if $\vec{c}^A \neq \vec{c}^B$ as long as $\vec{c}^A_{\mathcal{X}} = \vec{c}^B_{\mathcal{X}}$. Therefore, it should be regarded as a distance between concentration distributions of internal species rather than the full concentration profiles.

We can reduce the optimization for time in Eq. (115) as

$$W_{1,\mathcal{X}}(\vec{c}^A, \vec{c}^B) = \inf_{\mathcal{U}} |\mathcal{U}|_{\mathrm{RD}}. \qquad (118)$$

with the constraint that $\mathcal{U}$ satisfies

$$\vec{c}^B_{\mathcal{X}} - \vec{c}^A_{\mathcal{X}} = (\nabla^\dagger \mathcal{U})_{\mathcal{X}}, \qquad (119)$$

and the boundary condition on the diffusion part of $\mathcal{U}$ for the internal species. Using this reduced optimization problem, we can compute $W_{1,\mathcal{X}}$ numerically with less computational complexity. We provide the derivation of Eq. (118) and its geometric interpretation in Appendix E 1.

In addition, we can also generalize the Kantorovich–Rubinstein duality (32) to RDSs by considering the dual problem of the minimization problem in Eq. (118) (see the details in Appendix E 2) as

$$W_{1,\mathcal{X}}(\vec{c}^A, \vec{c}^B) = \sup_{\vec{\phi} \in \mathrm{Lip}^1_{\mathcal{X}}} \langle \vec{\phi}, \vec{c}^B - \vec{c}^A \rangle. \qquad (120)$$

Here, the set $\mathrm{Lip}^1_{\mathcal{X}}$ appearing in the conditions of optimization is a generalization of the set of 1-Lipschitz functions Lip^1 in

Eq. (33), and defined as

$$\text{Lip}^1_{\mathcal{X}} := \left\{ \vec{\phi} \,\middle|\, \forall \alpha \in \mathcal{X}, \ \|\boldsymbol{\nabla}_{\boldsymbol{r}} \phi_\alpha\| \leqslant 1; \right.$$

$$\forall \alpha \in \mathcal{Y}, \phi_\alpha = 0;$$

$$\left. \forall \rho \in \mathcal{R}_{\mathcal{X}}, \ |(\nabla_{\mathsf{s}} \vec{\phi})_\rho| \leqslant 1 \right\}. \quad (121)$$

This representation supports the justification for generalizing the 1-Wasserstein distance as the form in Eq. (115). It was also used to obtain an analytical form of the 1-Wasserstein distance in the numerical example that appears in Sec. V F.

B. 2-Wasserstein distance for reaction-diffusion systems

We define the 2-Wasserstein distance between concentration distributions of the internal species while fixing the boundary conditions, the stoichiometry, the $\vec{c}$ dependence of the Onsager operator, and the concentration distribution of the external species as

$$W_{2,\mathcal{X}}(\vec{c}^A, \vec{c}^B | \vec{b}_{\mathcal{Y}})^2 := \inf_{\vec{c}', \mathcal{F}'} \left\{ \tau \int_0^\tau dt \, \langle\!\langle \mathcal{F}', \mathcal{F}' \rangle\!\rangle_{\mathcal{M}_{\vec{c}'}} \right\}, \quad (122)$$

where we impose the conditions

$$\partial_t \vec{c}'_{\mathcal{X}} = (\nabla^\dagger \mathcal{M}_{\vec{c}'} \mathcal{F}')_{\mathcal{X}}, \quad (123)$$

$$\vec{c}'_{\mathcal{X}}(0) = \vec{c}^A_{\mathcal{X}}, \ \vec{c}'_{\mathcal{X}}(\tau) = \vec{c}^B_{\mathcal{X}}, \quad (124)$$

and

$$\vec{c}'_{\mathcal{Y}}(t) = \vec{c}^A_{\mathcal{Y}} = \vec{c}^B_{\mathcal{Y}} = \vec{b}_{\mathcal{Y}}, \quad (125)$$

on $\vec{c}'$ and $\mathcal{F}'$. In particular, the conditions in Eq. (125) correspond to the concentration of the external species being fixed. Here, we write $\mathcal{M}_{\vec{c}'}$ instead of $\mathcal{M}$ because we want to emphasize its dependence on concentration $\vec{c}'$. In addition, we assume that $\overleftrightarrow{\mathsf{M}}$ and m depend on time only through concentration and not explicitly, so that $W_{2,\mathcal{X}}$ is invariant to changes of the parameter τ. Note that the 2-Wasserstein distance $W_{2,\mathcal{X}}(\vec{c}^A, \vec{c}^B | \vec{b}_{\mathcal{Y}})$ is not a distance between concentration distributions of all species, but only those of the internal species, as in the case of the 1-Wasserstein distance. This is because the condition (125) fixes the concentration distribution of the external species to $\vec{b}_{\mathcal{Y}}$.

This definition generalizes the Benamou–Brenier formula of the 2-Wasserstein distance for Langevin systems in Eq. (28), MJPs [108], and CRNs [45]. This 2-Wasserstein distance also extends the dissipation distance [85,86], defined in the context of the gradient flow structure of detailed-balanced RDSs, to general open systems.

We can also rewrite Eq. (122) in the form of an optimization problem not for the force $\mathcal{F}'$ but for the potential $\vec{\phi}$ as

$$W_{2,\mathcal{X}}(\vec{c}^A, \vec{c}^B | \vec{b}_{\mathcal{Y}})^2 := \inf_{\vec{\phi}, \vec{c}'} \left\{ \tau \int_0^\tau dt \, \langle\!\langle \nabla \vec{\phi}, \nabla \vec{\phi} \rangle\!\rangle_{\mathcal{M}_{\vec{c}'}} \right\}, \quad (126)$$

with the conditions

$$\partial_t \vec{c}'_{\mathcal{X}} = (\nabla^\dagger \mathcal{M}_{\vec{c}'} \nabla \vec{\phi})_{\mathcal{X}}, \quad \phi_{\mathcal{Y}} = \vec{0}_{\mathcal{Y}}, \quad (127)$$

and the same conditions as in Eqs. (124) and (125) (see also Appendices E 3 and E 4). This representation originates from the following two facts: the definition of the 2-Wasserstein distance corresponds to the minimization problem of EP, and

such the minimization is achieved by a conservative force in short-time limit (91).

C. Features of 1- and 2-Wasserstein distances

The 1- and 2-Wasserstein distances are not distances between concentration distributions of all the species but between those of the internal species, as mentioned in earlier sections. In fact, we can prove that the axioms of distance hold for $W_{1,\mathcal{X}}$ and $W_{2,\mathcal{X}}$ as distances between concentration distributions of the internal species (see also Appendix E 5).

In addition, these distances cannot be defined between arbitrary concentration distributions. The constraints on dynamics in Eq. (117) or Eq. (123) let us define the 1- and 2-Wasserstein distances between the concentration distributions $\vec{c}^A$ and $\vec{c}^B$ satisfying

$$\vec{c}^B - \vec{c}^A \in \text{Im}_{\mathcal{X}} \nabla^\dagger, \quad (128)$$

where the set $\text{Im}_{\mathcal{X}} \nabla^\dagger$ is defined as

$$\text{Im}_{\mathcal{X}} \nabla^\dagger := \{ \vec{c}' \,|\, \exists \mathcal{J}', \vec{c}'_{\mathcal{X}} = (\nabla^\dagger \mathcal{J}')_{\mathcal{X}} \}. \quad (129)$$

Here, we impose the boundary condition that was imposed on the RDS on the diffusion part of $\mathcal{J}'$ in Eq. (129). In other words, we can define $W_{1,\mathcal{X}}(\vec{c}^A, \vec{c}^B)$ and $W_{2,\mathcal{X}}(\vec{c}^A, \vec{c}^B | \vec{b}_{\mathcal{Y}})$ only for $\vec{c}^B \in \vec{c}^A + \text{Im}_{\mathcal{X}} \nabla^\dagger := \{ \vec{c}' \,|\, \vec{c}' - \vec{c}^A \in \text{Im}_{\mathcal{X}} \nabla^\dagger \}$, which is a generalization of the concentration space of a CRN restricted by stoichiometry. We call this affine space the *stoichiometric manifold* following Ref. [109] (it is also called the stoichiometric compatibility class [110]). Thus, the form of the operator $\nabla^\dagger$ and the boundary conditions determine whether the Wasserstein distances between a given pair of concentration distributions are well defined. Note that the Wasserstein distances between concentration distributions belonging to the same time series obtained by the time evolution according to the RDSs are always well defined.

The 1- and 2-Wasserstein distances require different information to compute. We need two distributions of the internal species $\vec{c}^A_{\mathcal{X}}$ and $\vec{c}^B_{\mathcal{X}}$, the operator ∇ (or $\nabla^\dagger$) and the boundary conditions to obtain the 1-Wasserstein distance. In contrast, we need the time-independent concentration distribution of the external species $\vec{b}_{\mathcal{Y}}$ and the form of $\mathcal{M}$ as a functional of the concentration distributions, in addition to them, to obtain the 2-Wasserstein distance. We can regard ∇ and the boundary conditions as having information on the topology of the CRN and the topology of the space where diffusion occurs, and $\mathcal{M}$ as having information on the kinetic aspect of diffusion and reactions. Therefore, the topology determines the 1-Wasserstein distance, and the topology and the kinetic properties determine the 2-Wasserstein distance.

Because of the difference in the information required to define the 1- and 2-Wasserstein distances, there is no simple inequality between $W_{1,\mathcal{X}}$ and $W_{2,\mathcal{X}}$, as seen in Eq. (27). To compare the 1- and 2-Wasserstein distances, we define the following functional of the Onsager operator, which represents the intensity of reactions and diffusive dynamics,

$$|\mathcal{M}|^{\text{tot}}_{\mathcal{X}} := \int_V d\boldsymbol{r} \left[M^{\max}_{\mathcal{X}}(\boldsymbol{r};t) + \sum_{\rho \in \mathcal{R}_{\mathcal{X}}} m_\rho(\boldsymbol{r};t) \right]. \quad (130)$$

Here, the intensity of diffusive dynamics of the internal species $M_{\mathcal{X}}^{\max}(\boldsymbol{r};t)$ is defined by evaluating the magnitude of the mobility tensor as

$$M_{\mathcal{X}}^{\max}(\boldsymbol{r};t) := \max_{\vec{\boldsymbol{F}}'} \vec{\boldsymbol{F}}'^{\top} \overset{\leftrightarrow}{\mathsf{M}}(\boldsymbol{r};t) \vec{\boldsymbol{F}}', \tag{131}$$

under the following conditions:

$$\forall \alpha \in \mathcal{X}, \|\boldsymbol{F}'_{(\alpha)}\| = 1, \ \forall \alpha \in \mathcal{Y}, \boldsymbol{F}'_{(\alpha)} = \boldsymbol{0}. \tag{132}$$

These conditions allow for the equal incorporation of contributions from each internal chemical species when measuring the intensity of diffusion. When we consider the simple mobility tensor (53), our definition of $M_{\mathcal{X}}^{\max}(\boldsymbol{r};t)$ reduces to

$$M_{\mathcal{X}}^{\max}(\boldsymbol{r};t) = \sum_{\alpha \in \mathcal{X}} D_\alpha c_\alpha(\boldsymbol{r};t). \tag{133}$$

This form confirms that $M_{\mathcal{X}}^{\max}(\boldsymbol{r};t)$ represents the intensity of diffusion, since it becomes large at $(\boldsymbol{r};t)$ where the concentration of species with a large diffusion coefficient is high. We also regard m_ρ as the intensity of the ρth reaction since it measures the average reaction rate of the forward and reverse fluxes with the logarithmic mean.

We consider the following situation: (i) the $\vec{c}$ dependence of the Onsager operator is fixed, and (ii) the t dependence of the Onsager operator is based only on the t dependence of $\vec{c}$. We represent its dependence as $\mathcal{M}_{\vec{c}}$. The intensity of the Onsager operator enables us to obtain an inequality between the 1- and 2-Wasserstein distances as

$$\frac{W_{1,\mathcal{X}}(\vec{c}^A, \vec{c}^B)^2}{\langle |\mathcal{M}_{\vec{c}^\star}|_{\mathcal{X}}^{\text{tot}} \rangle_\tau} \leqslant W_{2,\mathcal{X}}(\vec{c}^A, \vec{c}^B | \vec{b}_{\mathcal{Y}})^2 \tag{134}$$

under the same topology, namely, the same boundary conditions and the same stoichiometry (see Appendix E 6 for the proof). Here, the concentration distribution $\vec{c}^\star(t)$ is the optimizer for the 2-Wasserstein distance (122), which satisfies $\vec{c}_{\mathcal{X}}^\star(0) = \vec{c}_{\mathcal{X}}^A, \vec{c}_{\mathcal{X}}^\star(\tau) = \vec{c}_{\mathcal{X}}^B$, and $\vec{c}_{\mathcal{Y}}^\star(t) = \vec{b}_{\mathcal{Y}}$. We also let $\langle \cdots \rangle_\tau$ denote the time average as $\langle \cdots \rangle_\tau = (1/\tau) \int_0^\tau dt \cdots$.

D. Thermodynamic speed limits with Wasserstein distances

We can find a relation between the excess EPR and the 2-Wasserstein distance for RDSs as in the case of Langevin systems, MJPs, and CRNs [45,62,63]. The relation leads to TSLs based on the 2-Wasserstein distance. Moreover, the inequality between the 1- and 2-Wasserstein distances enables us to obtain TSLs based on the 1-Wasserstein distance. We focus on the 2-Wasserstein distance. We define the path length between the initial and final concentration distributions of internal species induced by the 2-Wasserstein distance as

$$l_{2,\tau} := \int_0^\tau dt \, v_2(t), \tag{135}$$

with the speed of the dynamics of $\vec{c}(t)$ on the set $\vec{c}(0) + \text{Im}_{\mathcal{X}} \nabla^\dagger$ (see also Fig. 6),

$$v_2(t) := \lim_{\Delta t \to 0} \frac{W_{2,\mathcal{X}}(\vec{c}(t), \vec{c}(t + \Delta t) | \vec{c}_{\mathcal{Y}}(t))}{\Delta t}. \tag{136}$$

Unlike the case of the 2-Wasserstein distance, we can define $v_2(t)$ even if $\mathcal{M}$ explicitly depends on time by fixing $\mathcal{M}$ as the value at time t.
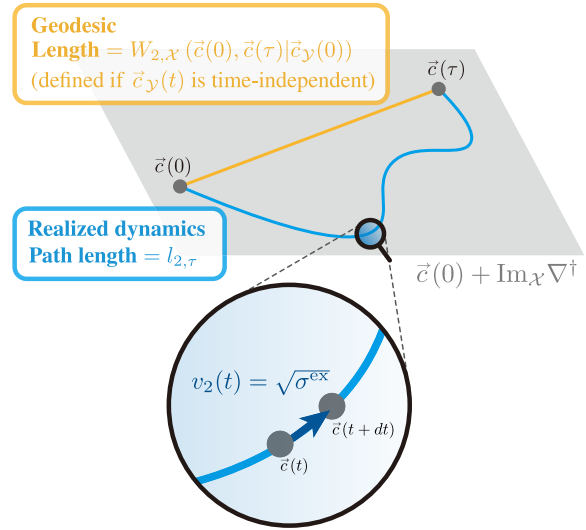

FIG. 6. Schematic illustration of the relation between the 2-Wasserstein distance and the excess EPR. The trajectory of the realized dynamics during the period $[0, \tau]$ (light-blue curve) is a curve with length $l_{2,\tau}$ on the stoichiometric manifold $\vec{c}(0) + \text{Im}_{\mathcal{X}} \nabla^\dagger$ (gray space). We can define the geodesic between the initial and final concentration distributions, $\vec{c}(0)$ and $\vec{c}(\tau)$ (yellow line) if the concentrations of the external species are constant in time as $\vec{c}_{\mathcal{Y}}(t) = \vec{c}_{\mathcal{Y}}(0)$. The geodesic also lies on the stoichiometric manifold and has length $W_{2,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau) | \vec{c}_{\mathcal{Y}}(0))$. The speed of the concentration distribution moving on the stoichiometric manifold $v_2(t)$ equals the square root of the excess EPR, $\sqrt{\sigma^{\text{ex}}}$.

The speed of the dynamics squared equals the excess EPR,

$$\sigma^{\text{ex}} = v_2(t)^2. \tag{137}$$

We can prove this relation as follows. Taking $\Delta t \ll 1$, the definition of the 2-Wasserstein distance in Eq. (122) and the constraints in Eqs. (123)–(125) lead to

$$W_{2,\mathcal{X}}(\vec{c}(t), \vec{c}(t + \Delta t) | \vec{c}_{\mathcal{Y}}(t))^2$$
$$= \Delta t^2 \inf_{\mathcal{F}'} \langle\!\langle \mathcal{F}', \mathcal{F}' \rangle\!\rangle_{\mathcal{M}_{\vec{c}(t)}} + o(\Delta t^2), \tag{138}$$

with the constraint

$$\vec{c}_{\mathcal{X}}(t + \Delta t) - \vec{c}_{\mathcal{X}}(t) = \Delta t (\nabla^\dagger \mathcal{M}_{\vec{c}(t)} \mathcal{F}')_{\mathcal{X}} + o(\Delta t). \tag{139}$$

We can obtain

$$\lim_{\Delta t \to 0} \frac{W_{2,\mathcal{X}}(\vec{c}(t), \vec{c}(t + \Delta t) | \vec{c}_{\mathcal{Y}}(t))^2}{\Delta t^2}$$
$$= \inf_{\mathcal{F}' | \partial_t \vec{c}_{\mathcal{X}}(t) = (\nabla^\dagger \mathcal{M}_{\vec{c}(t)} \mathcal{F}')_{\mathcal{X}}} \langle\!\langle \mathcal{F}', \mathcal{F}' \rangle\!\rangle_{\mathcal{M}_{\vec{c}(t)}} \tag{140}$$

by taking limit $\Delta t \to 0$ after dividing both sides of Eq. (138) by Δt^2 and both sides of Eq. (139) by Δt. Then we derive the relation between the speed of the dynamics and the excess EPR (137) by comparing the right-hand side of Eq. (140) and a form of the excess EPR (91) because $\partial_t \vec{c}_{\mathcal{X}}(t) = (\nabla^\dagger \mathcal{M}_{\vec{c}(t)} \mathcal{F})_{\mathcal{X}}$ holds for the force $\mathcal{F}$ in the original dynamics.

This relation between the speed of the dynamics and the excess EPR leads to the TSL,

$$l_{2,\tau}^2 \leqslant \tau \Sigma_\tau^{\text{ex}} \leqslant \tau \Sigma_\tau. \tag{141}$$

The first inequality is derived from the Cauchy–Schwarz inequality $[\int_0^\tau dt \, v_2(t)]^2 \leqslant [\int_0^\tau dt][\int_0^\tau dt \, v_2(t)^2]$ and the property of the excess EPR in Eq. (137). We remark that the equality of the first inequality holds if and only if $v_2(t)$ (or equivalently σ^{ex}) is independent of time, which follows from the conditions for the equality of the Cauchy–Schwarz inequality to hold.

We also obtain a lower bound determined from more limited information, only the initial and final distributions,

$$W_{2,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau)|\vec{c}_{\mathcal{Y}}(0))^2 \leqslant l_{2,\tau}^2 \leqslant \tau \Sigma_\tau^{\text{ex}} \leqslant \tau \Sigma_\tau, \quad (142)$$

if we consider the system where the concentrations of all the external species and the concentration dependence of $\mathcal{M}$ are independent of time. Here, the first inequality is a consequence of the triangle inequality for the 2-Wasserstein distance, which is proved in Appendix E 5. The TSL for $l_{2,\tau}$ is tighter than the one for $W_{2,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau))$ reflecting the fact that the path $\{\vec{c}_\mathcal{X}\}_{t \in [0,\tau]}$ is generally not the geodesic, whose length is $W_{2,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau))$.

The TSL for RDSs (141) implies a trade-off between the dissipation due to pattern formation and the change speed of the pattern because the TSL in Eq. (141) is rewritten as $(l_{2,\tau}/\tau)^2 \leqslant (\Sigma_\tau^{\text{ex}}/\tau)$, where $l_{2,\tau}/\tau$ means the change speed from the initial pattern at $t = 0$ to the final pattern at $t = \tau$, and $\Sigma_\tau^{\text{ex}}/\tau = \langle \sigma^{\text{ex}} \rangle_\tau$ means the time average of dissipation due to the time evolution of the pattern. This trade-off relation means that the slower the speed of pattern formation, the smaller the dissipation can be.

The 1-Wasserstein distance provides a different series of TSLs. Similar to Eq. (135), we define the path length between the initial and final distributions with the 1-Wasserstein distance as

$$l_{1,\tau} := \int_0^\tau dt \, v_1(t), \quad (143)$$

where the integrand $v_1(t)$ indicates the speed of dynamics measured with $W_{1,\mathcal{X}}$,

$$v_1(t) := \lim_{\Delta t \to 0} \frac{W_{1,\mathcal{X}}(\vec{c}(t), \vec{c}(t + \Delta t))}{\Delta t}. \quad (144)$$

The inequality between the 1- and 2-Wasserstein distances (134) and this speed v_1 provide a lower bound of the excess EPR,

$$\sigma^{\text{ex}} = v_2(t)^2 \geqslant \frac{v_1(t)^2}{|\mathcal{M}|_\mathcal{X}^{\text{tot}}}. \quad (145)$$

This inequality leads to the TSLs based on the 1-Wasserstein distance (see the derivations in Appendix E 7),

$$\frac{W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau))^2}{\langle |\mathcal{M}|_\mathcal{X}^{\text{tot}} \rangle_\tau} \leqslant \frac{l_{1,\tau}^2}{\langle |\mathcal{M}|_\mathcal{X}^{\text{tot}} \rangle_\tau} \leqslant \tau \Sigma_\tau^{\text{ex}} \leqslant \tau \Sigma_\tau, \quad (146)$$

which generalize TSLs obtained for MJPs [44] to RDSs and tighten them by using the excess EP and the line length $l_{1,\tau}$. As in the case of MJPs, equality between the leftmost and rightmost sides is achievable as $\tau \Sigma_\tau = W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau))^2/\langle |\mathcal{M}|_\mathcal{X}^{\text{tot}} \rangle_\tau$. We discuss this equality from the viewpoint of minimum dissipation in the next section.

In contrast to the TSLs for the 2-Wasserstein distance (142), we can always consider the leftmost term in Eq. (146).

This is because we can define the 1-Wasserstein distance even when $\vec{c}_\mathcal{Y}$ and the $\vec{c}$ dependence of $\mathcal{M}$ change in time. We also remark that the TSL for $l_{1,\tau}$ is tighter than the one for $W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau))$. This reflects the fact that the path $\{\vec{c}_\mathcal{X}\}_{t \in [0,\tau]}$ is generally not the geodesic, whose length is $W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau))$, as in the case of the 2-Wasserstein distance (142).

The TSL for the 1-Wasserstein distance provides a more detailed physical insight than the one with the 2-Wasserstein distance because the 1-Wasserstein distance lets us separately treat the kinetic parameters and the speed of the time-evolution, which are merged in the case of the 2-Wasserstein distance. To obtain this physical insight, we rewrite the TSL $l_{1,\tau}^2/\langle |\mathcal{M}|_\mathcal{X}^{\text{tot}} \rangle_\tau \leqslant \tau \Sigma_\tau^{\text{ex}}$ in Eq. (146) as $[(l_{1,\tau}/\tau)^2/\langle |\mathcal{M}|_\mathcal{X}^{\text{tot}} \rangle_\tau] \leqslant (\Sigma_\tau^{\text{ex}}/\tau)$, where $l_{1,\tau}/\tau$ is the change speed of the pattern measured with the 1-Wasserstein distance, and $\langle |\mathcal{M}|_\mathcal{X}^{\text{tot}} \rangle_\tau$ indicates the time average of intensity of reactions and diffusive dynamics. This rewriting allows us to regard the TSLs as trade-off relations between the dissipation due to pattern formation, speed of pattern change, and intensity of reactions and diffusion. Simply put, smaller dissipation is possible when the speed of pattern change is slower, or when the intensity of reactions and diffusion is greater.

We also remark that it is not obvious which lower bound of $\tau \Sigma_\tau^{\text{ex}}$ is tighter: $W_{2,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau)|\vec{c}_\mathcal{Y}(0))^2$ in Eq. (142) or $W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau))^2/\langle |\mathcal{M}|_\mathcal{X}^{\text{tot}} \rangle_\tau$ in Eq. (146), even though there is the hierarchy in short-time limit (145). This is because the denominator in the left-hand side of the inequality between $W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau))$ and $W_{2,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau))$ (134) is different from the denominator on the left-hand side of the TSL for $W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau))$ (146): the former refers to the optimal time series of concentration distributions for the 2-Wasserstein distance, $\vec{c}^\star$, while the latter refers to the original time series, $\vec{c}$. Similarly, it is not clear which lower bound is tighter: $l_{2,\tau}^2$ in Eq. (142) or $l_{1,\tau}^2/\langle |\mathcal{M}|_\mathcal{X}^{\text{tot}} \rangle_\tau$ in Eq. (146) (see also Appendix E 7).

E. Minimum dissipation and optimal transport

We consider the minimum dissipation required to evolve the concentration distribution of the internal species from the initial distribution $\vec{c}_\mathcal{X}(0)$ to the final distribution $\vec{c}_\mathcal{X}(\tau)$ over time τ. Here, we regard the dissipation as the functional of the force $\mathcal{F}'$ and the Onsager operator $\mathcal{M}'$ as

$$\Sigma_\tau[\mathcal{M}', \mathcal{F}'] = \int_0^\tau dt \, \langle\!\langle \mathcal{F}', \mathcal{F}' \rangle\!\rangle_{\mathcal{M}'}, \quad (147)$$

based on the form of the EPR (75).

In fact, the minimum dissipation can be made to vanish by letting the intensity of reactions and diffusive dynamics $|\mathcal{M}|_\mathcal{X}^{\text{tot}}$ be infinite. Therefore, minimum dissipation must be considered under constraints on the Onsager operator $\mathcal{M}$. Here, we show that the 1- and 2-Wasserstein distances provide the minimum dissipation under different physical constraints on the Onsager operator $\mathcal{M}$.

We can relate the 2-Wasserstein distance to the minimum dissipation as follows. We consider the following minimization problem

$$\inf_{\vec{c}', \mathcal{F}'} \Sigma_\tau[\mathcal{M}_{\vec{c}'}, \mathcal{F}'] \quad (148)$$

under four conditions:

(i) the internal part of the concentration distribution $\vec{c}\,'_{\mathcal{X}}$ satisfies

$$\vec{c}\,'_{\mathcal{X}}(0) = \vec{c}_{\mathcal{X}}(0), \ \ \vec{c}\,'_{\mathcal{X}}(\tau) = \vec{c}_{\mathcal{X}}(\tau), \tag{149}$$

(ii) the time evolution of $\vec{c}\,'_{\mathcal{X}}$ is given by

$$\partial_t \vec{c}\,'_{\mathcal{X}} = (\nabla^\dagger \mathcal{M}_{\vec{c}'} \mathcal{F}'), \tag{150}$$

(iii) the external part of the concentration distribution $\vec{c}\,'_{\mathcal{Y}}$ is independent of time, i.e., it satisfies

$$\vec{c}\,'_{\mathcal{Y}}(t) = \vec{c}_{\mathcal{Y}}(0), \tag{151}$$

for all $t \in [0, \tau]$, and

(iv) the Onsager operator $\mathcal{M}_{\vec{c}'}$ depends on time only through $\vec{c}\,'$. This minimization is related to the 2-Wasserstein distance as follows:

$$\inf_{\vec{c}', \mathcal{F}'} \Sigma_\tau[\mathcal{M}_{\vec{c}'}, \mathcal{F}'] = \frac{W_{2,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau)|\vec{c}_{\mathcal{Y}}(0))^2}{\tau}. \tag{152}$$

It is directly derived from the definition of the 2-Wasserstein distance (122).

The optimal force in the minimization problem (152) is given by the one in the minimization problem in the definition of the 2-Wasserstein distance (122) and it is conservative. The conservativeness is verified by the fact that the 2-Wasserstein distance corresponds to the minimization problem over conservative forces (126). When the system is driven by the optimal force, the concentration distribution moves on the geodesic determined by the 2-Wasserstein distance with constant speed,

$$v_2 = \frac{W_{2,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau)|\vec{c}_{\mathcal{Y}}(0))}{\tau}. \tag{153}$$

This is verified by considering the optimizer of the 2-Wasserstein distance (see also Appendix E 4).

The minimum dissipation (152) ensures that the equalities of the TSLs for the 2-Wasserstein distance in Eq (142) are achievable. It lets us regard the TSL for the 2-Wasserstein distance as an achievable lower bound of dissipation. The properties of the optimizer of the minimum dissipation (152) yield the following consequences when the system achieves the equality of the TSL for the 2-Wasserstein distance: The conservativeness of the optimal force makes the housekeeping EPR disappear, and the constant speed (153) makes the EPR independent of time.

We can also relate the 1-Wasserstein distance to the minimum dissipation when the dependence of the Onsager operator on the concentration distribution is controllable. We consider the following minimization problem

$$\inf_{\mathcal{M}', \mathcal{F}'} \Sigma_\tau[\mathcal{M}', \mathcal{F}'] \tag{154}$$

under two conditions:

(i) the current $\mathcal{M}'\mathcal{F}'$ changes the concentration distribution from the initial to the final distribution,

$$\vec{c}_{\mathcal{X}}(\tau) - \vec{c}_{\mathcal{X}}(0) = \left(\nabla^\dagger \int_0^\tau dt\, \mathcal{M}'\mathcal{F}' \right)_{\mathcal{X}}, \tag{155}$$

(ii) the intensity of reactions and diffusive dynamics has an upper bound M_0,

$$\langle |\mathcal{M}'|_{\mathcal{X}}^{\mathrm{tot}} \rangle_\tau \leqslant M_0. \tag{156}$$

This minimization is related to the 1-Wasserstein distance as follows (see Appendix E 8 a for the proof),

$$\inf_{\mathcal{M}', \mathcal{F}'} \Sigma_\tau[\mathcal{M}', \mathcal{F}'] = \frac{W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau))^2}{\tau M_0}. \tag{157}$$

As an optimizer of the minimization problem in Eq. (157), we can take an Onsager operator $\mathcal{M}^\diamond = \overleftrightarrow{\mathsf{M}}^\diamond \oplus \mathsf{m}^\diamond$ and a force $\mathcal{F}^\diamond = (\vec{F}^\diamond, f^\diamond)$ that are independent of time and satisfy $|\tau \mathcal{M}^\diamond \mathcal{F}^\diamond|_{\mathrm{RD}} = W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau))$ (see also Appendix E 8 a). Thus, the concentration distribution moves on the geodesic determined by the 1-Wasserstein distance with the constant speed,

$$v_1 = \frac{W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau))}{\tau}, \tag{158}$$

when the system is driven by the optimizer $(\mathcal{M}^\diamond, \mathcal{F}^\diamond)$. We remark that the time independence of the optimizer explicitly provides the geodesic as follows: let $\vec{c}^\diamond$ denote the concentration distribution whose internal part evolves according to $\partial_t \vec{c}_{\mathcal{X}}^\diamond = (\nabla^\dagger \mathcal{M}^\diamond \mathcal{F}^\diamond)_{\mathcal{X}}$, then the internal part of $\vec{c}^\diamond$ satisfies

$$\vec{c}_{\mathcal{X}}^\diamond(t) = \left(1 - \frac{t}{\tau} \right) \vec{c}_{\mathcal{X}}(0) + \frac{t}{\tau} \vec{c}_{\mathcal{X}}(\tau), \tag{159}$$

since $\mathcal{M}^\diamond \mathcal{F}^\diamond$ does not depend on time.

We can also reinterpret the optimizer $(\mathcal{M}^\diamond, \mathcal{F}^\diamond)$ from an operational viewpoint. We recombine the reaction part of the optimizer $\mathsf{m}^\diamond$ and $f^\diamond$ into the forward and reverse fluxes as below. Let us define the forward and reverse fluxes $j_\rho^{\pm\diamond}$ as

$$j_\rho^{+\diamond} := \frac{e^{f_\rho^\diamond}}{e^{f_\rho^\diamond} - 1} \mathsf{m}_\rho^\diamond f_\rho^\diamond, \quad j_\rho^{-\diamond} := \frac{1}{e^{f_\rho^\diamond} - 1} \mathsf{m}_\rho^\diamond f_\rho^\diamond. \tag{160}$$

Then, these satisfy the following relations: $f_\rho^\diamond = \ln(j_\rho^{+\diamond}/j_\rho^{-\diamond})$ and $\mathsf{m}_\rho^\diamond = (j_\rho^{+\diamond} - j_\rho^{-\diamond})/\ln(j_\rho^{+\diamond}/j_\rho^{-\diamond})$. In particular, if the reactions obey mass action kinetics, we can obtain the reaction rate constants that realize the fluxes $j_\rho^{\pm\diamond}$. Similarly, if the mobility tensor has the simple form (53) and the system is an ideal dilute solution, we can also obtain the diffusion coefficients and the mechanical forces that realize $\overleftrightarrow{\mathsf{M}}^\diamond$ and $\vec{F}^\diamond$ (see also Appendix E 8 b).

The minimum dissipation (157) also ensures that the equalities of the TSLs for the 1-Wasserstein distance in Eq. (146) are achievable by setting $M_0 = \langle |\mathcal{M}|_{\mathcal{X}}^{\mathrm{tot}} \rangle_\tau$. It provides the TSL for the 1-Wasserstein distance with physical meaning, as an achievable lower bound on dissipation. When the system achieves the equality of the TSL for the 1-Wasserstein distance, the EPR is independent of time, since the optimizer of the minimum dissipation with the 1-Wasserstein distance (157) does not depend on time. We also remark that the EP equals the excess EP under the optimal protocol since the leftmost and rightmost sides in the TSLs (146) coincide. It implies that we can choose a conservative force as an optimizer of the minimum dissipation problem (157), which is verified with the Kantorovich–Rubinstein duality (see also Appendix E 8 c).
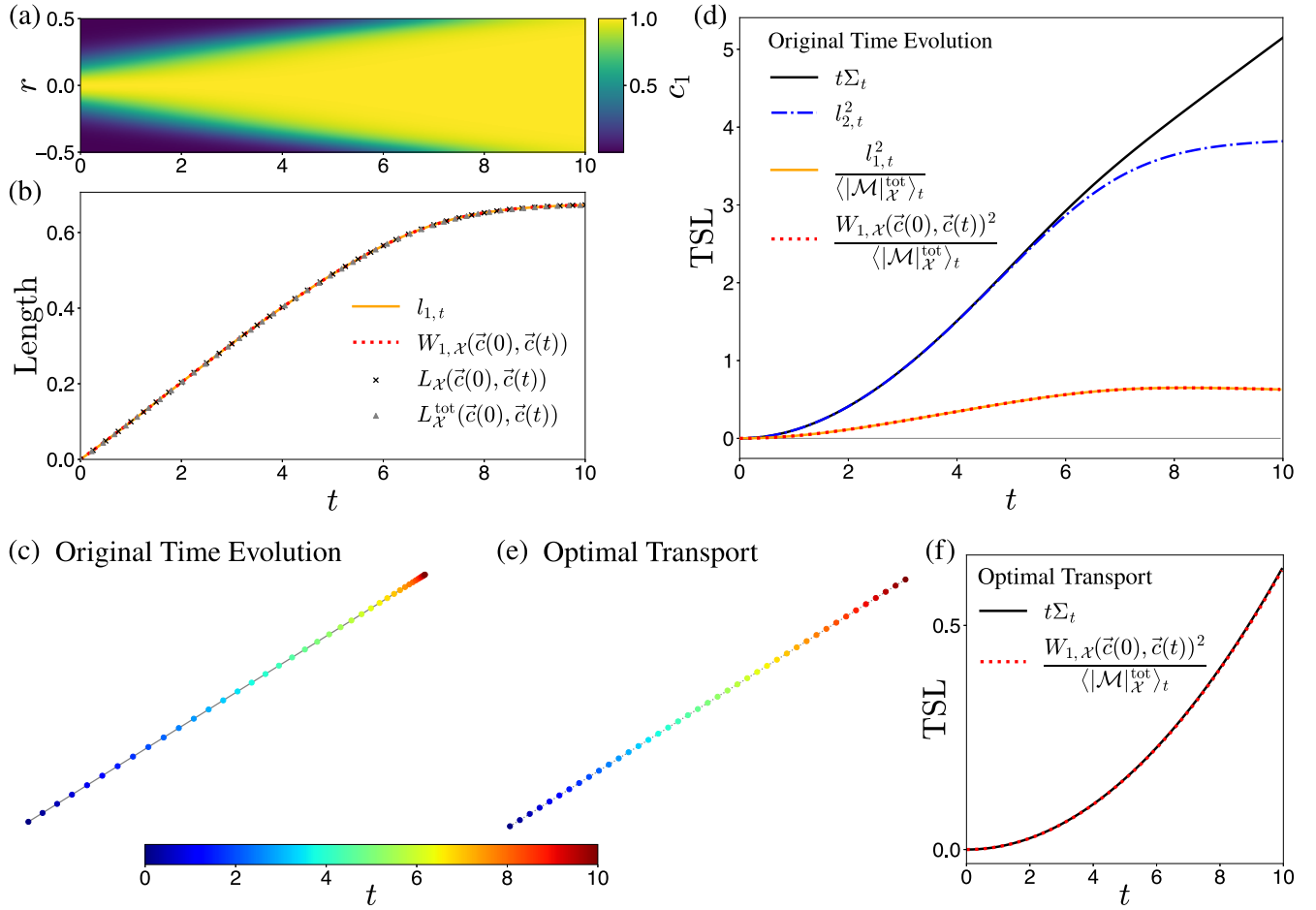
FIG. 7. The TSLs and optimal transport in the Fisher–KPP equation. (a) The time series of c_1. The figure is the same as Fig. 4(a). (b) Various lengths between $\vec{c}(0)$ and $\vec{c}(t)$: $l_{1,t}$, $W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(t))$, $L_{\mathcal{X}}(\vec{c}(0), \vec{c}(t))$, and $L_{\mathcal{X}}^{\text{tot}}(\vec{c}(0), \vec{c}(t))$. They have the same value for all time in the time series. (c) The original time series of the concentration distribution embedded in a two-dimensional plane by multidimensional scaling, which preserves the 1-Wasserstein distance between distributions as much as possible. Each point represents the concentration distribution at the time specified by the color. In this case, the time series is placed strictly on a straight line since $l_{1,t} = W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(t))$ holds for all t. (d) The TSLs. All the TSLs are confirmed to hold, and the one based on the path length with the 2-Wasserstein distance $l_{2,t}$ is relatively tighter. The gray line indicates zero. (e) The time series generated by the optimal protocol that achieves the minimum dissipation given by the 1-Wasserstein distance (157). We use $\tau = 10$ and $M_0 = \langle |\mathcal{M}|_{\mathcal{X}}^{\text{tot}} \rangle_\tau$, which is obtained by the original time evolution. It is embedded in the same two-dimensional plane as (c). The optimal protocol moves the pattern along the shortest path from the initial distribution $\vec{c}(0)$ to the destination $\vec{c}(\tau)$ with constant speed. (f) The achievement of the equality of the TSL for the 1-Wasserstein distance by the same optimal protocol as the one used in (e). The black line and the red-dotted line are $t\Sigma_t$ and the TSL for the 1-Wasserstein distance under the optimal protocol, respectively.

F. Numerical examples: Optimal transport and thermodynamic speed limits

We show numerical results for the TSLs [Eqs. (146) and (141)] by using the same time series as in Sec. IV C. We show the results for the Fisher–KPP equation in Fig. 7, and those for the Brusselator in Fig. 8. We compute the 1-Wasserstein distance by the primal-dual algorithm [111,112]. We utilize the multidimensional scaling [113] to embed the orbit of the time evolution on the stoichiometric manifold into a two-dimensional plane. This method keeps the pairwise 1-Wasserstein distances between the concentration distributions at any two times as close as possible (for details, see Appendix E 9). To compare with the behavior of the 1-Wasserstein distance, we use the L^1 distance between

concentration distributions of the internal species,

$$L_{\mathcal{X}}(\vec{c}(0), \vec{c}(\tau)) := \sum_{\alpha \in \mathcal{X}} \int_V d\boldsymbol{r} |c_\alpha(\boldsymbol{r}; \tau) - c_\alpha(\boldsymbol{r}; 0)|, \quad (161)$$

and the L^1 distance between total concentrations of the internal species,

$$L_{\mathcal{X}}^{\text{tot}}(\vec{c}(0), \vec{c}(\tau)) := \sum_{\alpha \in \mathcal{X}} \left| \int_V d\boldsymbol{r}\, c_\alpha(\boldsymbol{r}; \tau) - \int_V d\boldsymbol{r}\, c_\alpha(\boldsymbol{r}; 0) \right|.$$

$$(162)$$

Note that $L_{\mathcal{X}}$ accounts for not only changes in total concentrations but also changes in the shape of the pattern, which is not taken into account in $L_{\mathcal{X}}^{\text{tot}}$. The triangle inequality implies
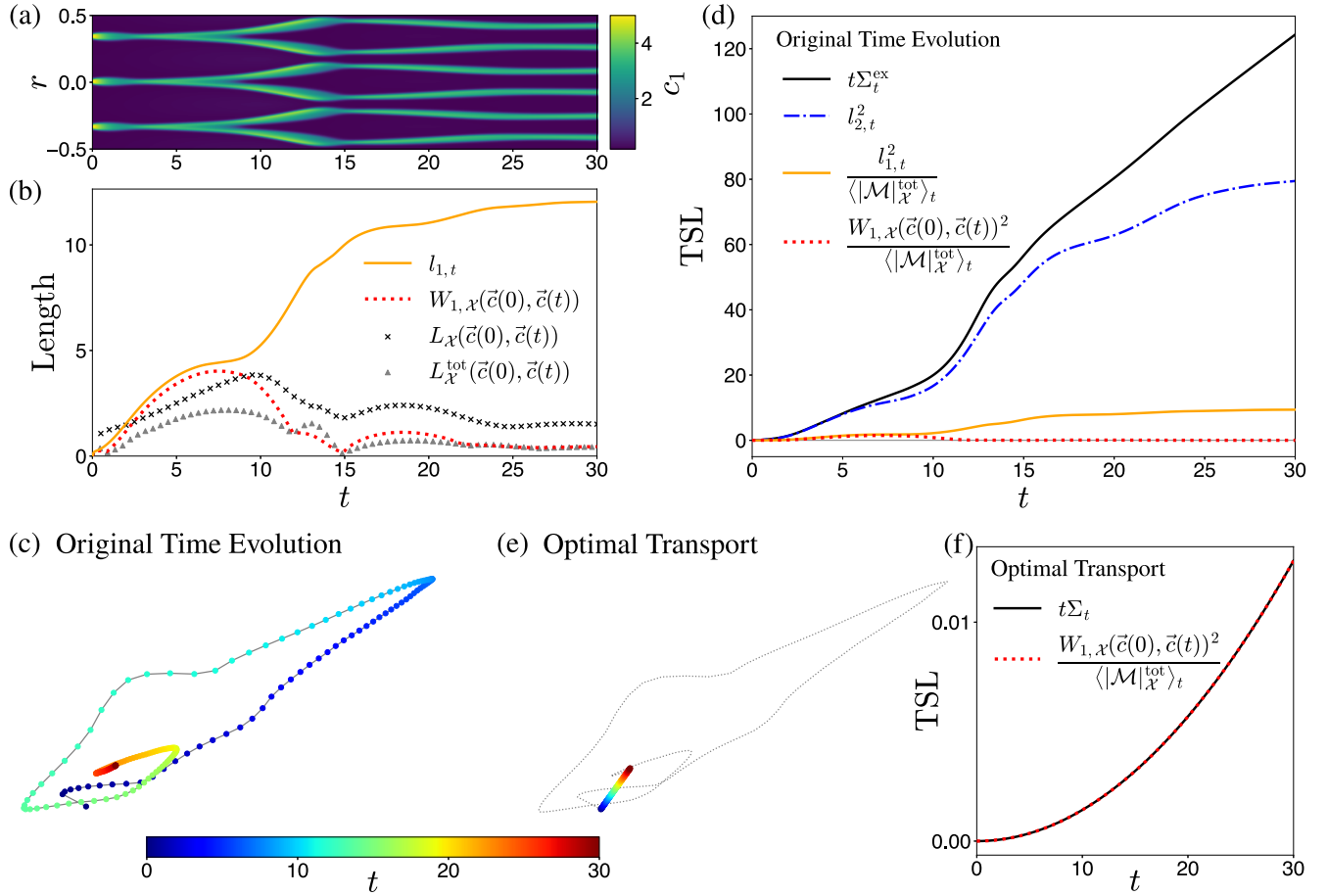
FIG. 8. The TSLs and optimal transport in the Brusselator model. (a) The time series of c_1. The figure shows the time interval $[0,30]$ from Fig. 5(a). (b) Various lengths between $\vec{c}(0)$ and $\vec{c}(t)$: $l_{1,t}$, $W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(t))$, $L_{\mathcal{X}}(\vec{c}(0), \vec{c}(t))$, and $L_{\mathcal{X}}^{\text{tot}}(\vec{c}(0), \vec{c}(t))$. We can see that the 1-Wasserstein distance satisfies the triangle inequality $l_{1,t} \geqslant W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(t))$ and has no obvious relationship with the other distances. (c) The original time series of the concentration distribution embedded in a two-dimensional plane by multidimensional scaling, which preserves the 1-Wasserstein distance between distributions as much as possible. Each point represents the concentration distribution at the time specified by the color. The concentration first moves away from the initial distribution and later returns toward it. (d) The TSLs. All the TSLs are confirmed to hold, and the one based on the path length with the 2-Wasserstein distance $l_{2,t}$ is relatively tight. The gray line indicates zero. (e) The time series generated by the optimal protocol that achieves the minimum dissipation given by the 1-Wasserstein distance (157). We use $\tau = 30$ and $M_0 = \langle |\mathcal{M}|_{\mathcal{X}}^{\text{tot}} \rangle_\tau$, which is obtained from the original time evolution. It is embedded in the same two-dimensional plane as (c). The optimal protocol moves the pattern along the shortest path from the initial distribution $\vec{c}(0)$ to the destination $\vec{c}(\tau)$ with a constant speed. (f) The achievement of the equality of the TSL for the 1-Wasserstein distance by the same optimal protocol as the one used in (e). The black line and the red dotted line are $t\Sigma_t$ and the TSL for the 1-Wasserstein distance under the optimal protocol, respectively.

$L_{\mathcal{X}}(\vec{c}(0), \vec{c}(\tau)) \geqslant L_{\mathcal{X}}^{\text{tot}}(\vec{c}(0), \vec{c}(\tau))$. If the system is well mixed, these distances become equivalent.

Fisher–KPP equation.—In Fig. 7(b), we show the four lengths between $\vec{c}(0)$ and $\vec{c}(t)$ for the Fisher–KPP equation, $l_{1,t}$, $W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(t))$, $L_{\mathcal{X}}(\vec{c}(0), \vec{c}(t))$, and $L_{\mathcal{X}}^{\text{tot}}(\vec{c}(0), \vec{c}(t))$. We can see that the lengths have the same value for all t. This equivalence of lengths is due to the following two conditions: (i) the system consists of only one internal species Z_1 and only one reaction, and (ii) the concentration is monotonically increasing for all locations and time, i.e., $\partial_t c_1 \geqslant 0$ holds for all $\boldsymbol{r}$ and t (see Appendix E 10 a for the proof). From the viewpoint of the path on the stoichiometric manifold, the equivalence $W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(t)) = l_{1,t}$ implies that the pattern evolves along the geodesic as shown in Fig. 7(c). We also remark that the

speed of the time evolution is not constant, i.e., it slows down over time.

Reflecting the monotonic increase of the area with a high concentration of c_1 shown in Fig. 7(a), the lengths between $\vec{c}(0)$ and $\vec{c}(t)$ increase monotonically over time. In particular, the lengths increase approximately in proportion to t when the concentration of Z_1 is not saturated (roughly $0 \leqslant t \leqslant 5$). We remark that the lengths are strictly proportional to t if the concentration distribution is a traveling wave solution with one wavefront, which is a simpler case than the numerical example (see also Appendix E 10 b).

In Fig. 7(d), we demonstrate the TSLs using the Fisher–KPP equation. We use the EP Σ_t instead of the excess EP Σ_t^{ex} because $\Sigma_t = \Sigma_t^{\text{ex}}$ holds as explained in Sec. IV C. The

squared length $l_{2,t}^2$ bounds $t\Sigma_t$ especially tight when the concentration of Z_1 is not saturated (roughly $0 \leqslant t \leqslant 5$). This is because the EPR for the traveling wave solution of the Fisher–KPP equation is independent of time [50] so that $v_2(t)$ satisfies the condition for the equality of the TSL. However, in the time region where the change in the concentration distribution is small, all of the TSLs become looser because $t\Sigma_t$ keeps increasing in proportion to time while the lengths saturate as the system approaches the steady state.

We can construct an optimal protocol that reduces the thermodynamic cost of changing the pattern from $\vec{c}_{\mathcal{X}}(0)$ to $\vec{c}_{\mathcal{X}}(\tau)$ to the minimum value provided by the TSL for the 1-Wasserstein distance as discussed in Sec. V E. It is obtained by solving the minimum dissipation problem (157) with $M_0 = \langle |\mathcal{M}|_{\mathcal{X}}^{\text{tot}} \rangle_\tau$. In the following, let us fix $\tau = 10$. From a geometric point of view, this protocol moves the pattern from $\vec{c}(0)$ to $\vec{c}(\tau)$ along the geodesic (159) as in the original time evolution; the difference with the original time evolution is that the speed on the stoichiometric manifold is uniform, as shown in Fig. 7(e).

In Fig. 7(f), we show that $t\Sigma_t$ and the TSL for the 1-Wasserstein distance coincide under the optimal protocol. This is verified as follows. Since the optimizer of the minimization problem in Eq. (157) is independent of time, we obtain

$$t\Sigma_t = \left(\frac{t}{\tau}\right)^2 \tau \Sigma_\tau. \tag{163}$$

The time-independence of the optimizer also lets $\langle |\mathcal{M}|_{\mathcal{X}}^{\text{tot}} \rangle_t = \langle |\mathcal{M}|_{\mathcal{X}}^{\text{tot}} \rangle_\tau$ hold for all t. We also have $W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(t)) = (t/\tau)W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau))$ since the concentration distribution moves on the geodesic with a constant speed. Thus, the lower bound provided by the TSL for the 1-Wasserstein distance satisfies

$$\frac{W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(t))^2}{\langle |\mathcal{M}|_{\mathcal{X}}^{\text{tot}} \rangle_t} = \left(\frac{t}{\tau}\right)^2 \frac{W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau))^2}{\langle |\mathcal{M}|_{\mathcal{X}}^{\text{tot}} \rangle_\tau}. \tag{164}$$

Combining Eq. (163) and Eq. (164) concludes that $t\Sigma_t = W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(t))^2/\langle |\mathcal{M}|_{\mathcal{X}}^{\text{tot}} \rangle_t$ holds for all t since the optimal protocol achieves the minimum dissipation (157) as $\tau \Sigma_\tau = W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau))^2/\langle |\mathcal{M}|_{\mathcal{X}}^{\text{tot}} \rangle_\tau$.

Brusselator model.—The time series used in the following is the same as that used in Sec. IV C. We will focus on the time interval $t \in [0, 30]$ as shown in Fig. 8(a), where the concentrations undergo significant changes.

In Fig. 8(b), we show the four lengths between $\vec{c}(0)$ and $\vec{c}(t)$ of the Brusselator. Unlike the Fisher–KPP equation, they behave differently, with only $l_{1,t}$ increasing monotonically. There is no definite order between $W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(t))$ and $L_{\mathcal{X}}(\vec{c}(0), \vec{c}(t))$ or between $W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(t))$ and $L_{\mathcal{X}}^{\text{tot}}(\vec{c}(0), \vec{c}(t))$. In particular, the 1-Wasserstein distance decreases to almost zero at time $t = 15$ after increasing. This is because the total concentrations at $t = 0$ and $t = 15$ are very close, which is evident from $L_{\mathcal{X}}^{\text{tot}}(\vec{c}(0), \vec{c}(15)) \simeq 0$. From the viewpoint of the path on the stoichiometric manifold, the behavior of the 1-Wasserstein distance implies that the pattern goes back near the initial state after once moving away from the initial state. The time series of the concentration distribution embedded in a

two-dimensional plane [Fig. 8(c)] verifies this behavior. The path of the pattern is not a geodesic of the 1-Wasserstein distance, since $l_{1,t}$ and $W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(t))$ are different.

In Fig. 8(d), we demonstrate the TSLs using the Brusselator. As in the case of the Fisher–KPP equation, the TSL for $l_{2,t}$ is tighter than the TSLs for $l_{1,t}$ and $W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(t))$. Moreover, the TSL for $l_{1,t}$ is tighter than the one with $W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(t))$. This is because the path of the time series $\{\vec{c}(t)\}_{t \in [0,30]}$ is not a geodesic as mentioned above. As in the case of the Fisher–KPP equation, we can also see that all of the TSLs become looser when the system approaches the steady state. This is caused by the fact that the increase proportional to t in $t\Sigma_t^{\text{ex}}$ persists while the pattern stops changing.

As in the case of the Fisher–KPP equation, we can construct an optimal protocol that reduces the thermodynamic cost to the minimum value provided by the TSL for the 1-Wasserstein distance. In the following, we take $\tau = 30$ as the final time in Eq. (157). Since the TSL for the 1-Wasserstein distance is weak, the optimal protocol can significantly decrease the dissipation required to change the pattern. This reduction of dissipation is due to the following circumstances: as already discussed, the time series of the concentration distribution takes a detour in the present system [Fig. 8(c)]. In contrast, the optimal protocol moves the pattern along the shortest path to the destination, which is close to the initial distribution, as shown in Fig. 8(e). As a result, the optimal protocol significantly reduces the total change in the pattern, which leads to low dissipation. In Fig. 8(f), we also show the TSL for the 1-Wasserstein distance under the optimal protocol. It behaves the same way as in the case of the Fisher–KPP equation: equality always holds, and $t\Sigma_t$ and the bound are proportional to t^2.

VI. THERMODYNAMIC UNCERTAINTY RELATIONS FOR REACTION-DIFFUSION SYSTEMS

We now provide another thermodynamic trade-off relation, the thermodynamic uncertainty relation (TUR). We refer to the lower bound of the instantaneous (excess) EPR by partial information about the pattern and system fluctuations as the TUR following previous studies [45,66,84]. This is in contrast to the TSLs, which bound the (excess) EP arising in finite time using information about changes in the entire pattern. After providing the TUR for general observables in Sec. VI A, we obtain the TUR for the Fourier coefficients of the concentration distribution as a special case of the general TUR in Sec. VI B. The TUR for the Fourier coefficients generalizes the TUR for CRNs found in previous study [84] by reflecting the spatial structure of the pattern.

A. General thermodynamic uncertainty relation for reaction-diffusion systems

Here, we derive the TUR for a general observable, which can take complex values. The term *observable* is used to refer to complex-valued, time-independent vector functions whose external parts are zero; i.e., an observable belongs to the set $\{\vec{\varphi} : \mathbb{R}^d \to \mathbb{C}^N \mid \forall \alpha \in \mathcal{Y}, \ \varphi_\alpha = 0\}$. We impose periodic boundary condition on $\vec{\varphi}$ if we consider systems with

periodic boundaries. Since we extend the scope of treatment to complex-valued functions, we also extend the inner products $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ and $\langle \cdot, \cdot \rangle$ to complex-valued functions by taking conjugate transpose, denoted by $\dagger$, as

$$\langle\!\langle \mathcal{J}', \mathcal{F}' \rangle\!\rangle := \int_V d\boldsymbol{r} \, (\vec{\boldsymbol{J}}'^{\dagger} \vec{\boldsymbol{F}}' + \boldsymbol{j}'^{\dagger} \boldsymbol{f}'), \tag{165}$$

and

$$\langle \vec{\phi}, \vec{\psi} \rangle := \int_V d\boldsymbol{r} \, \vec{\phi}^{\dagger} \vec{\psi}. \tag{166}$$

Under these preparations, we obtain the TUR for RDSs as a generalization of Eq. (40),

$$|d_t \langle \vec{c}, \vec{\varphi} \rangle|^2 \leqslant \mathfrak{D}_{\vec{\varphi}} \sigma^{\mathrm{ex}}, \tag{167}$$

where the indicator of the fluctuation of $\vec{\varphi}$ is defined as

$$\mathfrak{D}_{\vec{\varphi}} := \int_V d\boldsymbol{r} \, [(\boldsymbol{\nabla_r} \vec{\varphi})^{\dagger} \overleftrightarrow{\mathsf{M}} \boldsymbol{\nabla_r} \vec{\varphi} + \vec{\varphi}^{\dagger} \check{D} \vec{\varphi}]. \tag{168}$$

Here, the first term on the right-hand side of Eq. (168) originates from the spatial inhomogeneity of $\vec{\varphi}$ and the mobility of the system. The matrix in the second term $\check{D}$ is the scaled diffusion coefficient matrix defined as

$$\check{D}_{\alpha\beta}(\boldsymbol{r};t) := \sum_{\rho \in \mathcal{R}} \frac{j_{\rho}^+(\boldsymbol{r};t) + j_{\rho}^-(\boldsymbol{r};t)}{2} S_{\alpha\rho} S_{\beta\rho} \tag{169}$$

$$= \sum_{\rho \in \mathcal{R}} a_{\rho}(\boldsymbol{r};t) S_{\alpha\rho} S_{\beta\rho}. \tag{170}$$

This matrix appears in the diffusion coefficient matrix of the chemical Langevin equation, which is a microscopic description of chemical reactions. In the chemical Langevin equation, the diffusion coefficient matrix is inversely proportional to the system size, and $\check{D}$ is its proportional coefficient. This implies that the scaled diffusion coefficient reflects the intrinsic fluctuations of chemical reactions [84]. Thus, the second term in $\mathfrak{D}_{\vec{\varphi}}$ (168) corresponds to the fluctuation of the reactions.

The TUR (167) is derived as follows. We can obtain an inequality between the scaled diffusion coefficient and the edgewise Onsager coefficient as

$$\check{D}_{\alpha\beta} \geqslant \sum_{\rho \in \mathcal{R}} m_{\rho} S_{\alpha\rho} S_{\beta\rho} = \sum_{\rho, \rho' \in \mathcal{R}} [\mathsf{m}]_{\rho\rho'} S_{\alpha\rho} S_{\beta\rho'}, \tag{171}$$

by using the inequality in Eq. (59). This yields

$$\mathfrak{D}_{\vec{\varphi}} \geqslant \langle\!\langle \nabla \vec{\varphi}, \nabla \vec{\varphi} \rangle\!\rangle_{\mathcal{M}}. \tag{172}$$

Then, we can derive the TUR as

$$\begin{aligned}
|d_t \langle \vec{c}, \vec{\varphi} \rangle|^2 &= |\langle \nabla^{\dagger} \mathcal{M} \mathcal{F}, \vec{\varphi} \rangle|^2 = |\langle \nabla^{\dagger} \mathcal{M} \mathcal{F}^*, \vec{\varphi} \rangle|^2 \\
&= |\langle\!\langle \mathcal{M} \mathcal{F}^*, \nabla \vec{\varphi} \rangle\!\rangle|^2 = |\langle\!\langle \mathcal{F}^*, \nabla \vec{\varphi} \rangle\!\rangle_{\mathcal{M}}|^2 \\
&\leqslant \langle\!\langle \nabla \vec{\varphi}, \nabla \vec{\varphi} \rangle\!\rangle_{\mathcal{M}} \langle\!\langle \mathcal{F}^*, \mathcal{F}^* \rangle\!\rangle_{\mathcal{M}} \\
&\leqslant \mathfrak{D}_{\vec{\varphi}} \sigma^{\mathrm{ex}},
\end{aligned}$$

where we use Eqs. (69) and (85) in the first line, the boundary condition for $\mathcal{F}^*$ between the first and second line, the Cauchy–Schwarz inequality for the inner product $\langle\!\langle \cdot, \cdot \rangle\!\rangle_{\mathcal{M}}$ between the second and third line, and Eq. (172) in the last transformation.

We can also interpret the TUR from the viewpoint of the Wasserstein geometry, as we did for the TUR for Langevin systems in Eq. (41). By rewriting Eq. (167), we obtain

$$v_{\vec{\varphi}} := \frac{|d_t \langle \vec{c}, \vec{\varphi} \rangle|}{\sqrt{\mathfrak{D}_{\vec{\varphi}}}} \leqslant v_2, \tag{173}$$

where we used $\sigma^{\mathrm{ex}} = v_2^2$. This inequality means that the speed of the observable change $v_{\vec{\varphi}}$, normalized by the indicator of fluctuation $\mathfrak{D}_{\vec{\varphi}}$, can always be upper bounded by the speed of the concentration distribution v_2 measured with the 2-Wasserstein distance.

B. Thermodynamic uncertainty relations for Fourier component of concentration distribution

Here, we prove a TUR for the Fourier transform of a concentration distribution by properly choosing the observable in the general TUR (167). In the following part, we only consider internal species, $\alpha \in \mathcal{X}$. We define the Fourier transform of the concentration distribution as

$$\tilde{c}_{\alpha}(\boldsymbol{k};t) := \int_V d\boldsymbol{r} \, c_{\alpha}(\boldsymbol{r};t) \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}, \tag{174}$$

where we regard $c_{\alpha}(\boldsymbol{r};t)$ as zero outside of V. If we consider the system with the periodic boundary condition, we obtain $c_{\alpha}(\boldsymbol{r};t) = \sum_{\boldsymbol{k}} \tilde{c}_{\alpha}(\boldsymbol{k};t) \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}/|V|$. In this case, the vector $\boldsymbol{k}$ takes discrete values depending on the details of V.

Fixing $\alpha \in \mathcal{X}$ and letting $\vec{\varphi}$ satisfy $(\vec{\varphi}(\boldsymbol{r}))_{\beta} = \delta_{\alpha\beta} \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}$ in Eq. (167) yields a TUR for the Fourier transform,

$$\frac{|d_t \tilde{c}_{\alpha}(\boldsymbol{k};t)|^2}{\boldsymbol{k} \cdot \mathsf{M}^{\mathrm{tot}}_{(\alpha\alpha)}(t) \boldsymbol{k} + \check{D}^{\mathrm{tot}}_{\alpha\alpha}(t)} \leqslant \sigma^{\mathrm{ex}}, \tag{175}$$

where $\mathsf{M}^{\mathrm{tot}}_{(\alpha\alpha)}(t) := \int_V d\boldsymbol{r} \, \mathsf{M}_{(\alpha\alpha)}(\boldsymbol{r};t)$ and $\check{D}^{\mathrm{tot}}_{\alpha\alpha}(t) := \int_V d\boldsymbol{r} \, \check{D}_{\alpha\alpha}(\boldsymbol{r};t)$. This TUR generalizes the one for well-mixed CRNs studied in the previous work [84] because it reduces to the same form as the TUR for CRNs,

$$\frac{\left|d_t c^{\mathrm{tot}}_{\alpha}(t)\right|^2}{\check{D}^{\mathrm{tot}}_{\alpha\alpha}(t)} \leqslant \sigma^{\mathrm{ex}}, \tag{176}$$

if $\boldsymbol{k}$ is $\boldsymbol{0}$. Here, we let $c^{\mathrm{tot}}_{\alpha}(t)$ denote the total concentration of αth species at time t as $c^{\mathrm{tot}}_{\alpha}(t) := \int_V d\boldsymbol{r} c_{\alpha}(\boldsymbol{r};t)$.

If the mobility tensor has the simple form in Eq. (53), we obtain a simpler form of the TUR as

$$\frac{|d_t \tilde{c}_{\alpha}(\boldsymbol{k};t)|^2}{\|\boldsymbol{k}\|^2 D_{\alpha} c^{\mathrm{tot}}_{\alpha}(t) + \check{D}^{\mathrm{tot}}_{\alpha\alpha}(t)} \leqslant \sigma^{\mathrm{ex}}. \tag{177}$$

The first term of the denominator in Eq. (177) monotonically increases for $\|\boldsymbol{k}\|$, while the second term is independent of $\|\boldsymbol{k}\|$. Thus, the TUR (177) indicates that we can realize large $|d_t \tilde{c}(\boldsymbol{k};t)|$ with small dissipation if $\|\boldsymbol{k}\|$ is large. We can also detect which mode and which species is dominant in the time evolution by considering the tightness of the TUR (177) depending on wavenumber $\boldsymbol{k}$.

In the context of pattern formation, the trade-off relation in Eq. (177) means that spatially global pattern formation, which is given by the mode change $|d_t \tilde{c}_{\alpha}(\boldsymbol{k};t)|$ corresponding to smaller wavenumber $\boldsymbol{k}$, is more dissipative rather than spatially local pattern formation, which is given by the mode
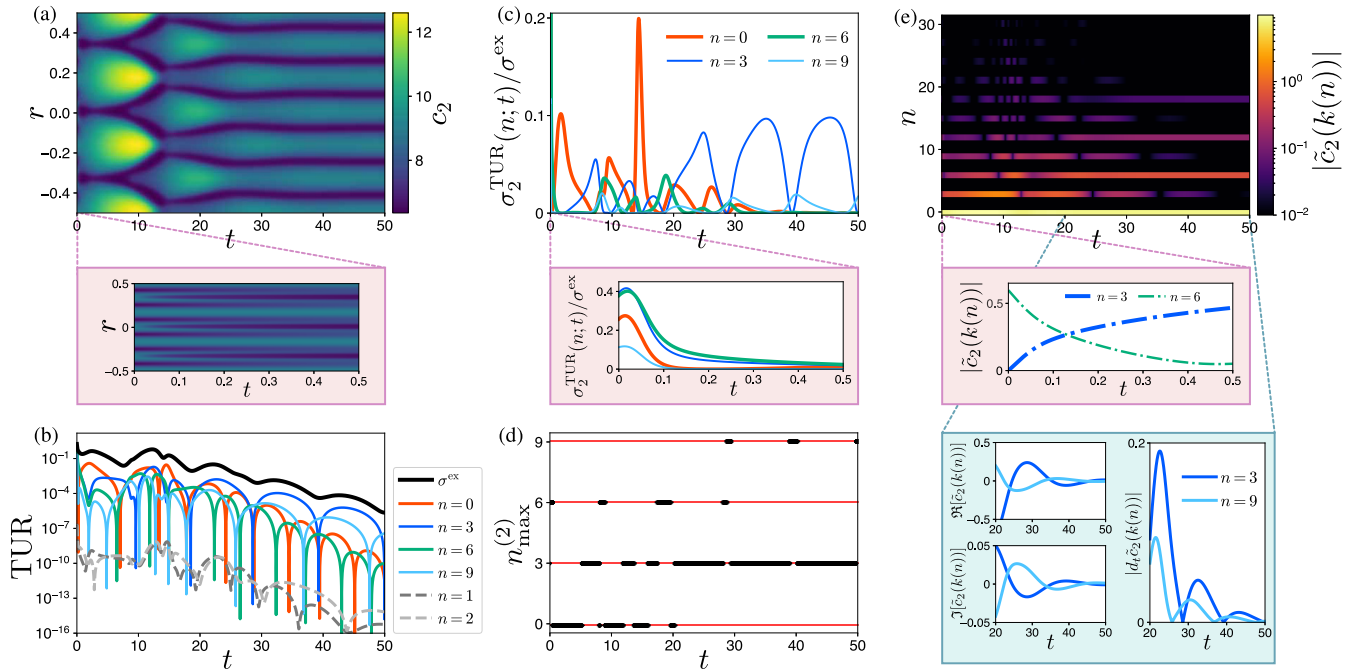
FIG. 9. The TUR (177) in the Brusselator model. (a) The time series of c_2 for comparison. In the early stage of the time evolution ($t \in [0, 0.5]$), the symmetry of the pattern instantly changes from sixfold to threefold, as shown in the pink panel. After that, the pattern goes to the steady state with sixfold symmetry. (b) Semilog plot of the excess EPR (black line) and the lower bounds $\sigma_2^{\mathrm{TUR}}(n; t)$ for various n. The lower bounds corresponding to multiples of three ($n = 0, 3, 6, 9$) are tighter than the bounds corresponding to $n = 1, 2$. (c) The ratio of $\sigma_2^{\mathrm{TUR}}(n; t)$ to the excess EPR for $n = 0, 3, 6, 9$. In the early stage ($t \in [0, 0.5]$), $\sigma_2^{\mathrm{TUR}}(3; t)$ or $\sigma_2^{\mathrm{TUR}}(6; t)$ provides the tightest bound as shown in the pink panel. (d) $n_{\max}^{(2)}(t)$ (black dots). Reflecting the symmetry of the pattern, $n_{\max}^{(2)}(t)$ are multiples of three for all time t (red lines). Near the stationary pattern, $n_{\max}^{(2)}(t) = 3$ for almost all t. (e) The time series of $|\tilde{c}_2(k(n))|$. We omit $|\tilde{c}_2(k(n))|$ for $n \geqslant 32$ because they are sufficiently small. The threefold symmetry of the pattern lets $|\tilde{c}_2(k(n))|$ have large values when $n = 3n'$ for any integer n'. Since the symmetry of the pattern transitions to sixfold as approaching the steady state, $|\tilde{c}_2(k(n))|$ decays if n is a multiple of three but not a multiple of six at large t. In contrast, the magnitude $|\tilde{c}_2(k(6))|$ decays while $|\tilde{c}_2(k(3))|$ increases in the early stage of the pattern formation (pink panel) reflecting the instant transition from the pattern with sixfold symmetry to the pattern with threefold symmetry. The light-blue panel shows the asynchronous oscillations of the Fourier components corresponding to $n = 3, 9$ in $t \in [20, 50]$. Here, we let $\mathfrak{R}[z]$ and $\mathfrak{I}[z]$ denote the real part and imaginary part of the complex number z, respectively. The oscillations of the two modes are not synchronized. Due to the asynchronous nature of the oscillations, the times when $|d_t \tilde{c}_2(k(3); t)|$ vanishes differ from the times when $|d_t \tilde{c}_2(k(9); t)|$ vanishes.

change $|d_t \tilde{c}_\alpha(\boldsymbol{k}'; t)|$ corresponding to larger wavenumber $\boldsymbol{k}'$ ($\|\boldsymbol{k}'\| > \|\boldsymbol{k}\|$). Thus, the TUR (177) quantifies a required dissipation to form spatial patterns according to its spatial structure.

We can also obtain a TSL-like thermodynamic trade-off relation

$$\frac{|\tilde{c}_\alpha(\boldsymbol{k}; \tau) - \tilde{c}_\alpha(\boldsymbol{k}; 0)|^2}{\boldsymbol{k} \cdot \langle \mathsf{M}_{(\alpha\alpha)}^{\mathrm{tot}} \rangle_\tau \boldsymbol{k} + \langle \check{D}_{\alpha\alpha}^{\mathrm{tot}} \rangle_\tau} \leqslant \tau \Sigma_\tau^{\mathrm{ex}}, \tag{178}$$

which is derived by integrating the TUR in Eq. (175) over time and using the triangle inequality for time integration and the Cauchy–Schwarz inequality (see also Appendix F for details). Note that the initial time 0 can be arbitrarily set.

C. Numerical examples: Thermodynamic uncertainty relations

We demonstrate the TUR for the Fourier transform of a concentration distribution (177), using the same time series of the Brusselator as in Sec. IV C, shown in Fig. 5. Because of the periodic boundary condition imposed on the system, we can obtain the Fourier transform of the concentration distribution

as

$$\tilde{c}_\alpha(k(n); t) = \int_{-0.5}^{0.5} dr \, c_\alpha(r; t) \mathrm{e}^{-ik(n)r} \tag{179}$$

with discrete wavenumbers $k(n)$ (114). To discuss the property of the TUR, we introduce the mode $n_{\max}^{(\alpha)}(t)$ that provides the tightest lower bound on dissipation at time t in the sense of the TUR as

$$n_{\max}^{(\alpha)}(t) := \arg \max_n \sigma^{\mathrm{TUR}}(n; t), \tag{180}$$

with

$$\sigma_\alpha^{\mathrm{TUR}}(n; t) := \frac{|d_t \tilde{c}_\alpha(k(n); t)|^2}{k(n)^2 D_\alpha c_\alpha^{\mathrm{tot}}(t) + \check{D}_{\alpha\alpha}^{\mathrm{tot}}(t)}. \tag{181}$$

Here, we focus on the chemical species Z_2, as indexed by $\alpha = 2$ (see Appendix G for results for Z_1, corresponding to $\alpha = 1$) and discuss the relations between the pattern dynamics and the lower bound of dissipation.

In the following, we show that the lower bound in the TUR reflects the symmetry of the pattern (Fig. 9). In this numerical

example, the pattern always has a threefold symmetry. Thus, the pattern changes occur only in modes corresponding to multiples of three. This leads to large $\sigma_2^{\mathrm{TUR}}(n;t)$ for n that are multiples of three, as shown in Fig. 9(b). After the pattern with sixfold symmetry is formed ($t > 20$), $|\tilde{c}_2(k(n))|$ decays if n is a multiple of three but not a multiple of six Fig. 9(e). In particular, the decay of the mode $n = 3$ is significant. It leads to $n_{\max}^{(2)}(t) = 3$ for almost all $t > 20$ [Figs. 9(c) and 9(d)].

Rapid changes in the pattern have a significant impact on the TUR. In the early stages of the time evolution, the symmetry of the pattern instantly changes from sixfold to threefold, as shown in the pink panel in Fig. 9(a). This change is accompanied by a rapid decay of $|\tilde{c}_2(k(6))|$ and a rapid increase of $|\tilde{c}_2(k(3))|$ [the pink panel in Fig. 9(e)]. It makes $\sigma_2^{\mathrm{TUR}}(n;t)$ be tighter for $n = 3$ and $n = 6$ than other modes, as shown in the pink panel in Fig. 9(c).

The lower bound in the TUR also reflects oscillations in the Fourier components of the pattern. The speed of change of the Fourier component $|d_t\tilde{c}_2(k(n);t)|$ intermittently goes to zero owing to the oscillatory behavior of $\tilde{c}_2(k(n);t)$. This causes vanishing of the lower bounds $\sigma_2^{\mathrm{TUR}}(n;t)$, shown as sharp peaks in the semilog plot [Fig. 9(b)]. In particular, the asynchronous oscillations of $\tilde{c}_2(k(n);t)$ shown in the light-blue panel of Fig. 9(e) induce the temporary deviations from $n_{\max}^{(2)}(t) = 3$ in $t \geqslant 20$, shown in Fig. 9(d). For example, the light-blue panel of Fig. 9(e) shows that $|d_t\tilde{c}_2(k(3);t)|$ vanishes at $t \simeq 40$, while $|d_t\tilde{c}_2(k(9);t)|$ does not. It lets $n_{\max}^{(2)}(t)$ deviate to nine from three at $t \simeq 39$ in Fig. 9(d).

These observations imply that the new TUR determined by the change of each Fourier component (175) may provide a tighter bound of EPR than the TUR of the well-mixed CRN [84] if we consider chemical systems with pattern changes. Since the denominator of $\sigma_2^{\mathrm{TUR}}(n;t)$ is smaller for smaller n, the case $n = 0$ [corresponding simply to the TUR of the well-mixed CRN as Eq. (176)] gives the tightest bound if $|d_t\tilde{c}_2(k(n);t)|$ is independent of n. In actual pattern formation, however, the magnitude of $d_t\tilde{c}_2(k(n);t)$ is biased for each n, reflecting the spatiotemporal structure of the pattern. It allows $\sigma_2^{\mathrm{TUR}}(n;t)$ be tighter when n is not zero.

We also compare the lower bounds provided by the TUR and the wavenumber decomposition in Fig. 10. Since the wavenumber excess EPR is nonnegative at each mode, we obtain

$$\sigma^{\mathrm{ex}} = \sum_{n'} \sigma^{\mathrm{ex,wn}}(n') \geqslant \sigma^{\mathrm{ex,wn}}(n). \tag{182}$$

Thus, the value of the wavenumber excess EPR provides a lower bound of the excess EPR. The bound determined by wave number as well as the TUR. We can see that the bounds provided by the TUR are closer to the wavenumber excess EPR on the corresponding mode than the excess EPR itself, except for the points where $|d_t\tilde{c}_\alpha(n)|$ vanishes. It implies that the TUR can estimate the wavenumber excess EPR rather than the excess EPR. Note that, however, the lower bounds provided by the TUR possibly become larger than the wavenumber excess EPR at the corresponding mode. For instance, we show that both $\sigma_1^{\mathrm{TUR}}(3;t)$ and $\sigma_2^{\mathrm{TUR}}(3;t)$ are larger than $\sigma^{\mathrm{ex,wn}}(3)$ at small t, where the pattern changes rapidly, in Fig. 10(c).
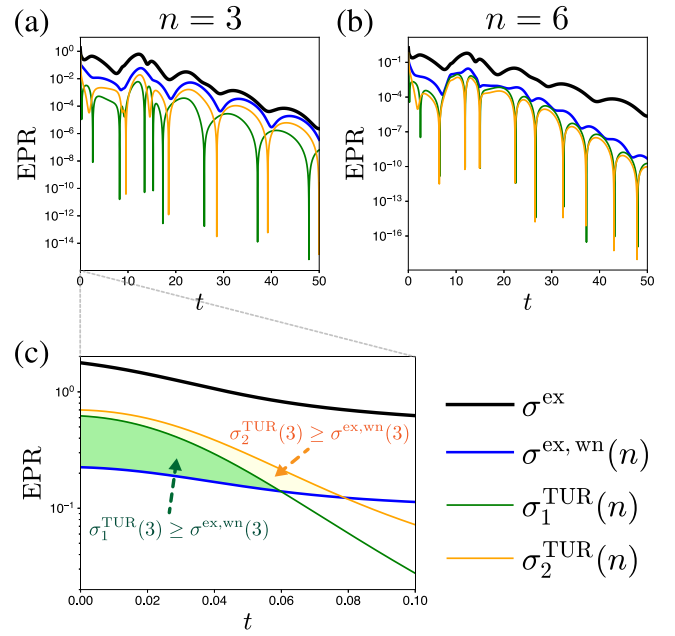

FIG. 10. The comparison of the TUR (175) and the wavenumber decomposition of the excess EPR. (a) The behavior of the excess EPR and its lower bounds ($n = 3$). The three bounds $\sigma^{\mathrm{ex,wn}}(3)$ (blue line), $\sigma_1^{\mathrm{TUR}}(3;t)$ (green line), and $\sigma_2^{\mathrm{TUR}}(3;t)$ (orange line) are smaller than the excess EPR (black line) for all t. In addition, the lower bounds provided by the TUR are close to the wavenumber excess EPR rather than the excess EPR. (b) The behavior of the excess EPR and its lower bounds ($n = 6$). We can observe a similar trend as in the case of $n = 3$. (c) The behavior of the excess EPR and its lower bounds near the initial time ($n = 3$). The areas filled in light green and light yellow indicate $\sigma_1^{\mathrm{TUR}}(3;t) \geqslant \sigma^{\mathrm{ex,wn}}(3)$ and $\sigma_2^{\mathrm{TUR}}(3;t) \geqslant \sigma^{\mathrm{ex,wn}}(3)$, respectively.

VII. DISCUSSION

In this paper, we established universal relations between the time evolution of patterns and dissipation by extending the framework of geometric thermodynamics [45,62,63,66,82] to RDSs. These relations clarify how dissipation constrains or reflects the spatiotemporal dynamics of patterns. Our results also enable the treatment of deterministic pattern dynamics, which previous geometric thermodynamics could not address.

In particular, we constructed the geometric excess/housekeeping decomposition of EPR and showed that the excess EPR can be understood as the minimum dissipation rate required to reproduce the original time evolution. We also proposed a geometric decomposition of the EPR, excess EPR, and housekeeping EPR into nonnegative contributions from different spatial locations and Fourier modes (wave numbers). Our decompositions were illustrated using detailed numerical examples.

In addition, the excess EPR was related to the details of the pattern and its time evolution by the newly derived TSLs and TURs for RDSs. TSLs are trade-off relations between the speed of the time evolution and the dissipation. These inequalities characterize the dissipation in a finite duration using information about the whole pattern. In contrast, TURs bound the instantaneous excess EPR using partial information about the pattern, such as the Fourier components of the concentration distribution. Our trade-off relations are

applicable to deterministic systems with large degrees of freedom; in this way they differ from the existing attempts that treat pattern formation in the presence of fluctuations [114–117].

Our results advance nonequilibrium thermodynamics of RDSs, suggesting engineering applications such as quantifying energy efficiency and optimizing control. In this way, we go beyond the traditional approach, which aim to predict stationary patterns and analyze their stability. This goal cannot be achieved in general cases [22,26,27], which may give the impression that thermodynamics is useless for understanding pattern formation. However, once RDSs are considered objects of control, thermodynamics becomes an essential tool because it quantifies the required dissipation to achieve a desired change in patterns. We have taken the first step in developing this tool, by developing inequalities and extending optimal transport theory so as to relate dissipation with the change in patterns. We believe this new direction will become increasingly important as RDSs undergo exploration by experimentalists and engineers.

Our results also provide insights into stochastic thermodynamics. For instance, we have introduced the wavenumber decomposition for Langevin systems. In addition, our results indicate that optimal transport theory, which in stochastic thermodynamics has been studied separately for systems with discrete and continuous variables, can be applied to general Markov processes with both kinds of variables. To be precise, the master equation for such a general system consists of diffusion terms for the continuous variables and Markov jump terms for the discrete variables. It is equivalent to an RD equation since the MJP is formally the same as an unimolecular CRN. Therefore, all of our results, derived here for RD equations, are also applicable to general master equations and bring new fundamental tools, e.g., decompositions of EPR and thermodynamic trade-off relations, for biological processes modeled by Markov processes with discrete and continuous variables, such as F1-ATPase [118].

One of the main contributions of this paper is to extend optimal transport theory by generalizing the 1- and 2-Wasserstein distances to RDSs. Optimal transport solves the problem of minimizing the dissipation necessary to transition between two given patterns, and it provides an operational way to quantify the speed of evolution, as required for the TSLs. Previously, the 2-Wasserstein distance was extended to deterministic chemical systems by using its consistency with gradient structures [45,85,86], but the generalizations do not treat open systems explicitly. The generalization of the 1-Wasserstein distance to deterministic chemical systems have not received as much attention even though the usefulness of the 1-Wasserstein distance in stochastic thermodynamics has been revealed [43,44]. We have extended the 1- and 2-Wasserstein distances to deterministic open RDSs and showed that they relate to the EP associated with pattern formation. The Wasserstein distances have different advantages: the 1-Wasserstein distance allows us to measure the distance between patterns even if the 2-Wasserstein distance cannot be defined, and it enables us to separately treat the distance and the kinetic information (130). On the other hand, the 2-Wasserstein distance gives the excess EPR a geometric interpretation and provides tight TSLs.

Another development relating to the 1-Wasserstein distance is the intensity of reactions and diffusive dynamics, which represents the kinetic information and links the distance to thermodynamics. The intensity of reactions measures the average rate of forward and reverse reactions with the logarithmic mean. It corresponds to the dynamical state mobility defined in MJPs [44]. The intensity of diffusion is defined by a variational problem, consistent with the diffusion coefficient in Langevin systems. This definition enables us to quantify the intensity of diffusion even for systems with spatial anisotropy and interactions between species.

We remark that our geometric decompositions were derived by expressing dissipation as the squared norm of the thermodynamic forces. In fact, it is possible to express dissipation in terms of other functionals, leading to alternative geometric decompositions. In particular, geometric decompositions may be derived by using the fact that information geometry and Hessian geometry generalize orthogonality and the Pythagorean theorem [119–121]. Previously, this was used to derive different extensions of the excess/housekeeping decomposition by Maes and Netočný from Langevin systems to MJPs and CRNs. One method involves fixing another mean of the fluxes instead of the logarithmic mean m_ρ and constructing a nonlinear relationship between the force and current based on Hessian geometry [109]. Another method uses information geometry to study one-way fluxes, instead of currents [66]. We may also generalize these methods to RDSs. However, if we adopt the first alternative, we can no longer regard the excess EPR as the minimum EPR required to realize the original instantaneous change in patterns. It is not in keeping with our original goal of clarifying the relations between the time evolution of patterns and the unavoidable dissipation. The second alternative avoids this problem, but does not directly link the excess EPR to the Wasserstein distance. Thus, we have taken the method based on the quadratic form of the thermodynamic forces [45].

We would like to end this paper with an introduction to some future directions. First, the wavenumber decomposition can be extended to bases other than the Fourier basis (e.g., the wavelet basis). Second, while our results are valid for general RDSs, they may be specialized in useful ways to individual systems; for instance, the TUR for a general observable (167) will yield interesting bounds on EPR by considering specific observables depending on the nature of the system. In addition, it may be possible to explore other expressions of minimum dissipation by appropriately tailoring the imposed constraints. This may allow quantifying minimum dissipation in specific systems, such as active phase separation [122–125] and systems used for computation [13–16], thus highlighting system-specific thermodynamics bounds on pattern formation. Third, it may be possible to utilize the lower bounds of EPR to estimate the dissipation of actual pattern formations as studied in the growing field of thermodynamic inference [98,126].

Although our framework already accommodates a broad class of systems, including nonideal mixtures and systems with complex boundary conditions, it would also be meaningful to further extend our results to more complicated situations. For example, we possibly generalize our results to RDSs on general spaces, such as curved surfaces [127,128]

or graphs [129,130]. This may be done by replacing the differential operator ∇_r with an appropriate counterpart for curved surfaces or graphs. This generalization will reveal trade-offs and help advance an energetic understanding of pattern formations in living systems, e.g., chemical waves on the cell membrane [131] and epidemics in metapopulation models [132]. In addition, we may also generalize our framework to systems involving additional physical fields, such as mechanical stresses appeared in mechanochemical feedback [133,134]. A promising direction may be to integrate our thermodynamic approach with the results on general hydrodynamic systems [83].

APPENDIX A: DETAILS OF THE PROJECTED FORCE IN LANGEVIN SYSTEMS

In this Appendix, we provide the details of the conservative force $\boldsymbol{F}^*$. We derive the condition in Eq. (9) from the orthogonality (8) and the uniqueness of $\boldsymbol{F}^*$ using the condition in Appendix A 1. We also derive the minimization problem in Eqs. (11) and (13) in Appendix A 2. In Appendix A 3, we provide the way to obtain Eq. (9) from the minimization problems in Eqs. (11) and (13).

1. Sufficiency of Eq. (9) for orthogonality, and the uniqueness of its solution

Letting $\boldsymbol{F}^*$ be the gradient of a potential as $\boldsymbol{F}^* = \nabla_r \phi^*$, we can rewrite the inner product between $\boldsymbol{F}^*$ and $\boldsymbol{F} - \boldsymbol{F}^*$ as

$$\langle\!\langle \boldsymbol{F}^*, \boldsymbol{F} - \boldsymbol{F}^* \rangle\!\rangle_{\mathsf{M}} = \langle\!\langle \nabla_r \phi^*, \boldsymbol{F} - \nabla_r \phi^* \rangle\!\rangle_{\mathsf{M}}$$
$$= \int_{\mathbb{R}^d} d\boldsymbol{r}\, \nabla_r \phi^* \cdot \mathsf{M}(\boldsymbol{F} - \nabla_r \phi^*)$$
$$= -\int_{\mathbb{R}^d} d\boldsymbol{r}\, \phi^* \nabla_r \cdot [\mathsf{M}(\boldsymbol{F} - \nabla_r \phi^*)], \quad (A1)$$

where we used the boundary condition for $p(\boldsymbol{r};t)$ to ignore the surface term in the third line. Then, the orthogonality in Eq. (8) reduces to

$$\int_{\mathbb{R}^d} d\boldsymbol{r}\, \phi^* \nabla_r \cdot [\mathsf{M}(\boldsymbol{F} - \nabla_r \phi^*)] = 0, \quad (A2)$$

which shows that Eq. (9) is a sufficient condition of the orthogonality.

Let ϕ^1 and ϕ^2 be solutions of Eq. (9), and $\boldsymbol{F}^1$ and $\boldsymbol{F}^2$ the corresponding gradient forces. The norm of the difference between these forces is zero because

$$\langle\!\langle \boldsymbol{F}^1 - \boldsymbol{F}^2, \boldsymbol{F}^1 - \boldsymbol{F}^2 \rangle\!\rangle_{\mathsf{M}}$$
$$= \int_{\mathbb{R}^d} d\boldsymbol{r}\, \nabla_r (\phi^1 - \phi^2) \cdot \mathsf{M}\nabla_r (\phi^1 - \phi^2)$$
$$= -\int_{\mathbb{R}^d} d\boldsymbol{r}\, (\phi^1 - \phi^2) \nabla_r \cdot [\mathsf{M}\{\nabla_r (\phi^1 - \phi^2)\}]$$
$$= \int_{\mathbb{R}^d} d\boldsymbol{r}\, (\phi^1 - \phi^2) \nabla_r \cdot [\mathsf{M}(\boldsymbol{F} - \boldsymbol{F})] = 0, \quad (A3)$$

where we used the boundary condition for $p(\boldsymbol{r};t)$ to ignore the surface term in the third line, and Eq. (9) in the fourth line. Therefore, the two forces are identical as the norm is nondegenerate. Note that the potential satisfying Eq. (9) is not unique because $\phi^* + C$, where C is a constant, is a solution of Eq. (9) when ϕ^* is a solution.

2. The derivation of minimization problems [Eqs. (11) and (13)]

The condition in Eq. (9) also leads to the orthogonality

$$\langle\!\langle \boldsymbol{F}^*, \boldsymbol{F}' - \boldsymbol{F}^* \rangle\!\rangle_{\mathsf{M}} = 0, \quad (A4)$$

where $\boldsymbol{F}'$ satisfies the condition in Eq. (11). We can check it by calculating in a similar way as in Eq. (A1). Using the orthogonality in Eq. (A4), we obtain the inequality

$$\langle\!\langle \boldsymbol{F}', \boldsymbol{F}' \rangle\!\rangle_{\mathsf{M}} = \langle\!\langle \boldsymbol{F}^*, \boldsymbol{F}^* \rangle\!\rangle_{\mathsf{M}} + \langle\!\langle \boldsymbol{F}' - \boldsymbol{F}^*, \boldsymbol{F}' - \boldsymbol{F}^* \rangle\!\rangle_{\mathsf{M}}$$
$$\geqslant \langle\!\langle \boldsymbol{F}^*, \boldsymbol{F}^* \rangle\!\rangle_{\mathsf{M}} \quad (A5)$$

for all $\boldsymbol{F}'$ satisfying the condition in Eq. (11). It immediately leads to the minimization problem in Eq. (11) because the equality in Eq. (A5) is achieved if and only if $\boldsymbol{F}' = \boldsymbol{F}^*$.

Similarly, the condition in Eq. (9) also leads to another orthogonality

$$\langle\!\langle \boldsymbol{F} - \boldsymbol{F}^*, \boldsymbol{F}^* - \boldsymbol{F}' \rangle\!\rangle_{\mathsf{M}} = 0, \quad (A6)$$

for all $\boldsymbol{F}' \in \mathrm{Im}\nabla_r$. It yields the inequality

$$\langle\!\langle \boldsymbol{F} - \boldsymbol{F}', \boldsymbol{F} - \boldsymbol{F}' \rangle\!\rangle_{\mathsf{M}}$$
$$= \langle\!\langle \boldsymbol{F} - \boldsymbol{F}^*, \boldsymbol{F} - \boldsymbol{F}^* \rangle\!\rangle_{\mathsf{M}} + \langle\!\langle \boldsymbol{F}^* - \boldsymbol{F}', \boldsymbol{F}^* - \boldsymbol{F}' \rangle\!\rangle_{\mathsf{M}}$$
$$\geqslant \langle\!\langle \boldsymbol{F} - \boldsymbol{F}^*, \boldsymbol{F} - \boldsymbol{F}^* \rangle\!\rangle_{\mathsf{M}}, \quad (A7)$$

which leads to the minimization problem in Eq. (13) because the equality in Eq. (A7) is achieved if and only if $\boldsymbol{F}' = \boldsymbol{F}^*$.

3. The derivation of the condition Eq. (9) as the Euler–Lagrange equation

We remark that the conditions Eq. (9) and $\boldsymbol{F}^* = \nabla_r \phi^*$ are conversely obtained from two variational problems Eqs. (12) and (14). By considering the action functionals

$$\mathcal{I}_{\mathrm{hk}}[\phi] := \frac{1}{2} \langle\!\langle \boldsymbol{F} - \nabla_r \phi, \boldsymbol{F} - \nabla_r \phi \rangle\!\rangle_{\mathsf{M}} = \int d\boldsymbol{r}\, I_{\mathrm{hk}}, \quad (A8)$$

$$\mathcal{I}_{\mathrm{ex}}[\boldsymbol{F}', \phi] := \frac{1}{2} \langle\!\langle \boldsymbol{F}', \boldsymbol{F}' \rangle\!\rangle_{\mathsf{M}} + \int d\boldsymbol{r}\, \phi\{\nabla_r \cdot [\mathsf{M}(\boldsymbol{F} - \boldsymbol{F}')]\}$$
$$= \int d\boldsymbol{r}\, I_{\mathrm{ex}} \quad (A9)$$

with

$$I_{\text{hk}} := \frac{1}{2}(\boldsymbol{F} - \boldsymbol{\nabla_r}\phi) \cdot \mathsf{M}(\boldsymbol{F} - \boldsymbol{\nabla_r}\phi), \qquad (A10)$$

$$I_{\text{ex}} := \frac{1}{2}\boldsymbol{F}' \cdot \mathsf{M}\boldsymbol{F}' - \boldsymbol{\nabla_r}\phi \cdot \mathsf{M}(\boldsymbol{F} - \boldsymbol{F}'), \qquad (A11)$$

the two variational problems, Eqs. (12) and (14), are solved by the Euler–Lagrange equations

$$\frac{\delta \mathcal{I}_{\text{hk}}[\phi]}{\delta \phi}\bigg|_{\phi=\phi^*} = \left[\frac{\partial I_{\text{hk}}}{\partial \phi} - \boldsymbol{\nabla_r} \cdot \frac{\partial I_{\text{hk}}}{\partial(\boldsymbol{\nabla_r}\phi)}\right]\bigg|_{\phi=\phi^*} = 0, \quad (A12)$$

$$\frac{\delta \mathcal{I}_{\text{ex}}[\boldsymbol{F}', \phi]}{\delta \boldsymbol{F}'}\bigg|_{\boldsymbol{F}'=\boldsymbol{F}^*, \phi=\phi^*} = \frac{\partial I_{\text{ex}}}{\partial \boldsymbol{F}'}\bigg|_{\boldsymbol{F}'=\boldsymbol{F}^*, \phi=\phi^*} = \boldsymbol{0}. \quad (A13)$$

Here, ϕ in Eq. (A9) is the Lagrange multiplier which gives the condition in Eq. (9) from the variation $\delta\mathcal{I}_{\text{ex}}[\boldsymbol{F}', \phi]/\delta\phi|_{\boldsymbol{F}'=\boldsymbol{F}^*, \phi=\phi^*} = 0$. The first Euler–Lagrange equation (A12) directly provides the condition in Eq. (9). On the other hand, the second Euler–Lagrange equation (A13) provides $\boldsymbol{F}^* = \boldsymbol{\nabla_r}\phi^*$. Substituting it into the constraint $\boldsymbol{\nabla_r} \cdot [\mathsf{M}(\boldsymbol{F} - \boldsymbol{F}')] = \boldsymbol{0}$, we obtain Eq. (9).

APPENDIX B: DETAILS OF THE PROJECTED FORCE IN REACTION-DIFFUSION SYSTEMS

Here, we derive the Euler–Lagrange equation in Eq. (85) from the optimization problem in Eq. (84) or (87). We also derive the uniqueness of $\mathcal{F}^*$. The derivation reduces to the case of Langevin systems in Appendix A 3 by considering a particular situation: There is only one species and no chemical reactions.

1. The Euler–Lagrange equation for the projection and the uniqueness of the projected conservative force

First, we derive Eq. (85) from Eq. (84). We define a functional to minimize as $\mathcal{I}_{\text{hk}}[\vec{\phi}] := \langle\!\langle \mathcal{F} - \nabla\vec{\phi}, \mathcal{F} - \nabla\vec{\phi} \rangle\!\rangle_{\mathcal{M}}/2 = \int_V d\boldsymbol{r} I_{\text{hk}}$ with

$$\begin{aligned} I_{\text{hk}} := &\frac{1}{2}[\vec{\boldsymbol{F}} - \boldsymbol{\nabla_r}\vec{\phi}]^\top \overset{\leftrightarrow}{\mathsf{M}}[\vec{\boldsymbol{F}} - \boldsymbol{\nabla_r}\vec{\phi}] \\ &+ \frac{1}{2}[\boldsymbol{f} - \nabla_{\mathsf{s}}\vec{\phi}]^\top \mathsf{m}[\boldsymbol{f} - \nabla_{\mathsf{s}}\vec{\phi}]. \end{aligned} \quad (B1)$$

The functional derivative of $\mathcal{I}_{\text{hk}}$ leads to the condition to be satisfied by the optimal potential $\vec{\phi}^*$ as

$$\begin{aligned} \frac{\delta \mathcal{I}_{\text{hk}}}{\delta \phi_\alpha}\bigg|_{\vec{\phi}=\vec{\phi}^*} &= \left(\frac{\partial I_{\text{hk}}}{\partial \phi_\alpha} - \boldsymbol{\nabla_r} \cdot \frac{\partial I_{\text{hk}}}{\partial[\boldsymbol{\nabla_r}\phi_\alpha]}\right)\bigg|_{\vec{\phi}=\vec{\phi}^*} \\ &= [-\nabla_{\mathsf{s}}^\top \mathsf{m}(\boldsymbol{f} - \nabla_{\mathsf{s}}\vec{\phi}^*) + \boldsymbol{\nabla_r} \cdot \{\overset{\leftrightarrow}{\mathsf{M}}(\vec{\boldsymbol{F}} - \boldsymbol{\nabla_r}\vec{\phi}^*)\}]_\alpha \\ &= [-\nabla^\dagger \mathcal{M}(\mathcal{F} - \nabla\vec{\phi}^*)]_\alpha = 0, \quad (B2) \end{aligned}$$

for all $\alpha \in \mathcal{X}$, while $\phi_\alpha^* = 0$ holds for all $\alpha \in \mathcal{Y}$. This condition (B2) is nothing more than the condition indicated in Eq. (85).

Second, we show that the projected conservative force $\mathcal{F}^*$ is unique, while $\vec{\phi}^*$ is not. If we consider a constant vector

$\vec{C}$ satisfying $\vec{C}_{\mathcal{Y}} = \vec{0}_{\mathcal{Y}}$, whose gradient $\nabla\vec{C}$ is a zero vector, $\vec{\phi}^* + \vec{C}$ can also be a solution of the minimization problem in Eq. (84) and thus $\vec{\phi}^*$ is not unique. Supposing Eq. (85) has two solutions, $\vec{\phi}^1$ and $\vec{\phi}^2$, which satisfy $\vec{\phi}_{\mathcal{Y}}^1 = \vec{\phi}_{\mathcal{Y}}^2 = \vec{0}_{\mathcal{Y}}$, under the boundary condition, we obtain

$$\begin{aligned} \langle\!\langle \mathcal{F}^1 - \mathcal{F}^2, \mathcal{F}^1 - \mathcal{F}^2 \rangle\!\rangle_{\mathcal{M}} &= \langle\!\langle \mathcal{M}\nabla(\vec{\phi}^1 - \vec{\phi}^2), \nabla(\vec{\phi}^1 - \vec{\phi}^2) \rangle\!\rangle \\ &= \langle \nabla^\dagger \mathcal{M}\nabla(\vec{\phi}^1 - \vec{\phi}^2), \vec{\phi}^1 - \vec{\phi}^2 \rangle \\ &= \langle \nabla^\dagger \mathcal{M}(\mathcal{F} - \mathcal{F}), \vec{\phi}^1 - \vec{\phi}^2 \rangle \\ &= 0, \quad (B3) \end{aligned}$$

where we define $\mathcal{F}^i := \nabla\vec{\phi}^i$ for $i = 1, 2$. The second transformation in Eq. (B3) is allowed because we are imposing the appropriate boundary condition, and the third line follows from the Euler–Lagrange equation (85). We also use the condition, $\phi_\alpha^1 - \phi_\alpha^2 = 0$ for all $\alpha \in \mathcal{Y}$, in these transformations. Therefore, the nondegenerateness of the inner product $\langle\!\langle \cdot, \cdot \rangle\!\rangle_{\mathcal{M}}$ concludes $\mathcal{F}^1 = \mathcal{F}^2$, that is, the uniqueness of the projected potential force.

2. The Euler–Lagrange equation for the minimum dissipation

We derive the condition (85) from the minimization problem in Eq. (87). To solve this constraint minimization problem, we execute the method of Lagrange multiplier with the multiplier $\vec{\phi}$, whose external part is the zero vector as $\vec{\phi}_{\mathcal{Y}} = \vec{0}_{\mathcal{Y}}$. Then, the functional to optimize is

$$\begin{aligned} \mathcal{I}_{\text{ex}}[\mathcal{F}', \vec{\phi}] &:= \frac{1}{2}\langle\!\langle \mathcal{F}', \mathcal{F}' \rangle\!\rangle_{\mathcal{M}} + \langle \vec{\phi}, \nabla^\dagger \mathcal{M}(\mathcal{F} - \mathcal{F}') \rangle \\ &= \frac{1}{2}\langle\!\langle \mathcal{F}', \mathcal{F}' \rangle\!\rangle_{\mathcal{M}} + \langle\!\langle \nabla\vec{\phi}, \mathcal{F} - \mathcal{F}' \rangle\!\rangle_{\mathcal{M}}. \quad (B4) \end{aligned}$$

We obtain the conditions to be satisfied by the optimizer $\mathcal{F}^*$ and $\vec{\phi}^*$ by taking the functional derivative of $\mathcal{I}_{\text{ex}}[\mathcal{F}', \vec{\phi}]$ as

$$\frac{\delta \mathcal{I}_{\text{ex}}[\mathcal{F}', \vec{\phi}]}{\delta \boldsymbol{F}'_{(\alpha)}}\bigg|_{\vec{\phi}=\vec{\phi}^*, \mathcal{F}'=\mathcal{F}^*} = \sum_{\beta\in\mathcal{S}} \mathsf{M}_{(\alpha\beta)}(\boldsymbol{F}^*_{(\beta)} - \boldsymbol{\nabla_r}\phi^*_\beta) = \boldsymbol{0},$$

$$(B5)$$

for all $\alpha \in \mathcal{S}$ and

$$\frac{\delta \mathcal{I}_{\text{ex}}[\mathcal{F}', \vec{\phi}]}{\delta f'_\rho}\bigg|_{\vec{\phi}=\vec{\phi}^*, \mathcal{F}'=\mathcal{F}^*} = m_\rho\{f^*_\rho - (\nabla_{\mathsf{s}}\vec{\phi}^*)_\rho\} = 0, \quad (B6)$$

for all $\rho \in \mathcal{R}$. The positive definiteness of $\overset{\leftrightarrow}{\mathsf{M}}$ and m enables us to obtain $\vec{\boldsymbol{F}}^* = \boldsymbol{\nabla_r}\vec{\phi}^*$ and $\boldsymbol{f}^* = \nabla_{\mathsf{s}}\vec{\phi}^*$ from Eq. (B5) and Eq. (B6). Unifying these results yields

$$\mathcal{F}^* = \nabla\vec{\phi}^*. \quad (B7)$$

We obtain the condition in Eq. (85) by substituting Eq. (B7) into the constraint $(\nabla^\dagger \mathcal{M}\mathcal{F})_{\mathcal{X}} = (\nabla^\dagger \mathcal{M}\mathcal{F}^*)_{\mathcal{X}}$.

APPENDIX C: GRADIENT FLOW AND RELAXATION IN REACTION-DIFFUSION SYSTEMS

Gradient flow is the flow that causes a specific function (functional) to become smaller. The gradient flow structure of RDSs [53,85] is deeply related to the thermodynamics of

RDSs, e.g., the relaxation to the equilibrium of the closed ideal dilute solution is described by the gradient flow to the equilibrium concentration distribution (Appendix C 1). We can also verify that the conservative force $\nabla \vec{\phi}$ provides a gradient flow, which describes the relaxation to a state determined by the potential $\vec{\phi}$ (Appendix C 2). Moreover, considering the potential $\vec{\phi}^*$, which provides the excess EPR, the time evolution of the concentration distribution of the internal species can be represented only by a gradient flow. The gradient flow structure allows us to rewrite the excess EPR in a similar form to the EPR of the ideal dilute solutions, which relax to equilibrium (Appendix C 3).

1. Relaxation to equilibrium and gradient flow of the ideal solutions

We consider relaxation to the equilibrium of a closed ideal dilute solution without mechanical forces, where the chemical potential is written as $\mu_\alpha^{\mathrm{id}} = \mu_\alpha^\circ + \ln c_\alpha$. Since the system is closed and there is no mechanical force applied to the system, there exists an equilibrium concentration distribution $\vec{c}^{\mathrm{eq}}$. The equilibrium concentration distribution provides zero force as $-\nabla \vec{\mu}^{\mathrm{id,eq}} = (\vec{0}, \mathbf{0})$, where $\vec{\mu}^{\mathrm{id,eq}}$ is the chemical potential of the equilibrium state $\mu_\alpha^{\mathrm{id,eq}} = \mu_\alpha^\circ + \ln c_\alpha^{\mathrm{eq}}$.

In this situation, the force $\mathcal{F}$ is written as

$$\mathcal{F} = -\nabla(\vec{\mu}^{\mathrm{id}} - \vec{\mu}^{\mathrm{id,eq}}). \tag{C1}$$

Focusing on the concrete form of $\vec{\mu}^{\mathrm{id}}$, we can rewrite the difference between the chemical potentials $\vec{\mu}^{\mathrm{id}} - \vec{\mu}^{\mathrm{id,eq}}$ as

$$\mu_\alpha^{\mathrm{id}} - \mu_\alpha^{\mathrm{id,eq}} = \ln \frac{c_\alpha}{c_\alpha^{\mathrm{eq}}} = \frac{\delta}{\delta c_\alpha} D_{\mathrm{KL}}(\vec{c} \| \vec{c}^{\mathrm{eq}}), \tag{C2}$$

where $D_{\mathrm{KL}}(\cdot \| \cdot)$ is the generalized Kullback–Leibler (KL) divergence,

$$D_{\mathrm{KL}}(\vec{c} \| \vec{c}') := \int_V d\mathbf{r} \sum_{\alpha \in \mathcal{S}} \left(c_\alpha \ln \frac{c_\alpha}{c_\alpha'} - c_\alpha + c_\alpha' \right). \tag{C3}$$

Introducing the vector $\delta/\delta \vec{c} = (\delta/\delta c_1, \ldots, \delta/\delta c_N)^\top$, we can rewrite the force as

$$\mathcal{F} = -\nabla \frac{\delta}{\delta \vec{c}} D_{\mathrm{KL}}(\vec{c} \| \vec{c}^{\mathrm{eq}}), \tag{C4}$$

which rewrites the RD equation as the gradient flow toward the equilibrium [53],

$$\partial_t \vec{c} = -\nabla^\dagger \mathcal{M} \nabla \frac{\delta}{\delta \vec{c}} D_{\mathrm{KL}}(\vec{c} \| \vec{c}^{\mathrm{eq}}). \tag{C5}$$

We can also rewrite the EPR as

$$\begin{aligned}
\sigma &= -\langle\!\langle \mathcal{J}, \nabla(\vec{\mu}^{\mathrm{id}} - \vec{\mu}^{\mathrm{id,eq}}) \rangle\!\rangle \\
&= -\langle \partial_t \vec{c}, \vec{\mu}^{\mathrm{id}} - \vec{\mu}^{\mathrm{id,eq}} \rangle \\
&= -\int_V d\mathbf{r} \sum_{\alpha \in \mathcal{S}} (\partial_t c_\alpha) \ln \frac{c_\alpha}{c_\alpha^{\mathrm{eq}}} \\
&= -\partial_t D_{\mathrm{KL}}(\vec{c} \| \vec{c}^{\mathrm{eq}}), \tag{C6}
\end{aligned}$$

where we can regard the KL divergence $D_{\mathrm{KL}}(\vec{c} \| \vec{c}^{\mathrm{eq}})$ as the Gibbs free energy difference between the state $\vec{c}$ and the equilibrium state $\vec{c}^{\mathrm{eq}}$.

2. Relaxation due to the conservative force

In Sec. III E, we decompose the force at time t as

$$\mathcal{F} = \nabla \vec{\phi}(t) + \mathcal{F}^{\mathrm{nc}}. \tag{C7}$$

Note that $\mathcal{F}$ and $\mathcal{F}^{\mathrm{nc}}$ also depend on time, and the external part of the potential $\vec{\phi}(t)$ is the zero vector.

Using the potential $\vec{\phi}(t)$, we can introduce a pseudocanonical distribution corresponding to $\vec{\phi}$ [45] as

$$c_\alpha^{\mathrm{pcan}}(\mathbf{r}; t) := c_\alpha(\mathbf{r}; t) \frac{e^{\phi_\alpha(\mathbf{r};t)}}{\Xi_\alpha(t)} \tag{C8}$$

Here, the parameter $\Xi_\alpha(t)$ can be freely chosen to satisfy the following conditions:

$$\forall \rho, \quad \sum_{\alpha \in \mathcal{X}} S_{\alpha\rho} \ln \Xi_\alpha(t) = 0 \tag{C9}$$

for $\alpha \in \mathcal{X}$, and $\Xi_\alpha(t) = 1$ for $\alpha \in \mathcal{Y}$. This can be interpreted as meaning that each $\ln \Xi(s) = (\ln \Xi_\alpha(s))$ at time s is a conservation law. That is, at any time t, we have

$$\begin{aligned}
&\partial_t \int_V d\mathbf{r} \sum_{\alpha \in \mathcal{X}} c_\alpha(\mathbf{r}; t) \ln \Xi_\alpha(s) \\
&= \int_V d\mathbf{r} \sum_{\alpha \in \mathcal{X}} \{-\nabla_{\mathbf{r}} \cdot \mathbf{J}_{(\alpha)} + (\nabla_{\mathsf{s}} \mathbf{j})_\alpha\} \ln \Xi_\alpha(s) \\
&= \int_V d\mathbf{r} \sum_{\rho \in \mathcal{R}} j_\rho \left(\sum_{\alpha \in \mathcal{X}} S_{\alpha\rho} \ln \Xi_\alpha(s) \right) = 0, \tag{C10}
\end{aligned}$$

hence, $\int_V d\mathbf{r} \sum_\alpha c_\alpha(\mathbf{r}; t) \ln \Xi_\alpha(s)$ does not depend on t. Here, we ignore the surface term by using the boundary condition on the current and use $\nabla_{\mathbf{r}} \ln \Xi_\alpha(s) = 0$ in the second line. We can use the parameter $\Xi_\alpha(t)$ to let conserved quantities, quantities kept constant over time evolution, have the same value in $\vec{c}$ and $\vec{c}^{\mathrm{pcan}} := (c_1^{\mathrm{pcan}}, \ldots c_N^{\mathrm{pcan}})^\top$. We may also choose $\Xi_\alpha(t) = 1$ simply for all $\alpha \in \mathcal{X}$. When we consider closed ideal solutions, the pseudocanonical distribution $\vec{c}^{\mathrm{pcan}}$ reduces to $\vec{c}^{\mathrm{eq}}$ using $\vec{\phi} = -(\vec{\mu}^{\mathrm{id}} - \vec{\mu}^{\mathrm{id,eq}})$ and $\Xi_\alpha(t) = 1$.

Using the pseudocanonical distribution, we obtain

$$\begin{aligned}
\frac{\delta}{\delta c_\alpha} D_{\mathrm{KL}}(\vec{c} \| \vec{c}^{\mathrm{pcan}}) &= \ln \frac{c_\alpha}{c_\alpha^{\mathrm{pcan}}} \\
&= -\phi_\alpha + \ln \Xi_\alpha, \tag{C11}
\end{aligned}$$

where we consider $\vec{c}$ and $\vec{c}^{\mathrm{pcan}}$ to be independent, i.e., we ignore c_α appearing in the definition of c_α^{pcan} in Eq. (C8) in the variational calculations. Note that $-\phi_\alpha + \ln \Xi_\alpha$ is zero for all $\alpha \in \mathcal{Y}$ because $\phi_\alpha = 0$ and $\Xi_\alpha = 1$ hold for all $\alpha \in \mathcal{Y}$. Then, we can rewrite the RD equation for the internal species as

$$\begin{aligned}
\partial_t \vec{c}_{\mathcal{X}} &= (\nabla^\dagger \mathcal{M} \nabla \vec{\phi} + \nabla^\dagger \mathcal{M} \mathcal{F}^{\mathrm{nc}})_{\mathcal{X}} \\
&= \left(-\nabla^\dagger \mathcal{M} \nabla \frac{\delta}{\delta \vec{c}} D_{\mathrm{KL}}(\vec{c} \| \vec{c}^{\mathrm{pcan}}) + \nabla^\dagger \mathcal{M} \mathcal{F}^{\mathrm{nc}} \right)_{\mathcal{X}}, \tag{C12}
\end{aligned}$$

where we use Eqs. (C11) and (C9), and $\mathbf{r}$ independence of $\Xi_\alpha(t)$. The rewritten form of the RD equation allows us to regard the conservative force $\nabla \vec{\phi}$ as driving the relaxation to the pseudocanonical distribution corresponding to $\vec{\phi}$, since the first term in the second line in Eq. (C12) is written as

the gradient flow. This gradient flow is similar to the one for the relaxation to the equilibrium in Eq. (C5).

3. The excess entropy production rate and gradient flow

To obtain the geometric excess/housekeeping EPR, we decompose the force as

$$\mathcal{F} = \nabla \vec{\phi}^* + (\mathcal{F} - \mathcal{F}^*), \tag{C13}$$

where $\mathcal{F}^* = \nabla \vec{\phi}^*$. In the following, $\vec{\phi}^*(t)$ indicates the potential for the excess EPR at time t. Using the pseudocanonical distribution corresponding to $\vec{\phi}^*(t)$, we can rewrite the RD equation for the internal species as the gradient flow,

$$\partial_t \vec{c}_{\mathcal{X}} = \left(-\nabla^\dagger \mathcal{M} \nabla \frac{\delta}{\delta \vec{c}} D_{\mathrm{KL}}(\vec{c} \| \vec{c}^{\mathrm{pcan}}) \right)_{\mathcal{X}}, \tag{C14}$$

where we use $[\nabla^\dagger \mathcal{M}(\mathcal{F} - \mathcal{F}^*)]_{\mathcal{X}} = \vec{0}_{\mathcal{X}}$. The gradient flow (C14) means that the time evolution of the concentration distribution of the internal species can be written solely in terms of the relaxation to the pseudocanonical distribution corresponding to $\vec{\phi}^*$ determined independently at each time.

Corresponding to Eq. (C14) that the time evolution of the concentration distribution of the internal species can be described only by the relaxation to different states at different times, the excess EPR has a similar informational geometric form to Eq. (C6),

$$\sigma^{\mathrm{ex}} = -\partial_t D_{\mathrm{KL}}(\vec{c}(t) \| \vec{c}^{\mathrm{pcan}}(s))|_{s=t}, \tag{C15}$$

where $\vec{c}^{\mathrm{pcan}}$ indicates the pseudocanonical distribution corresponding to $\vec{\phi}^*$. We can verify the representation of the excess EPR in Eq. (C15) as

$$-\partial_t D_{\mathrm{KL}}(\vec{c}(t) \| \vec{c}^{\mathrm{pcan}}(s))|_{s=t}$$
$$= \int_V d\mathbf{r} \sum_{\alpha \in \mathcal{X}} (\partial_t c_\alpha)(\phi_\alpha^* - \ln \Xi_\alpha)$$
$$= \int_V d\mathbf{r} \sum_{\alpha \in \mathcal{X}} (\partial_t c_\alpha)\phi_\alpha^*$$
$$= \langle \partial_t \vec{c}, \vec{\phi}^* \rangle = \langle \nabla^\dagger \mathcal{M} \nabla \vec{\phi}^*, \vec{\phi}^* \rangle$$
$$= \langle\!\langle \mathcal{M} \nabla \vec{\phi}^*, \nabla \vec{\phi}^* \rangle\!\rangle = \langle\!\langle \nabla \vec{\phi}^*, \nabla \vec{\phi}^* \rangle\!\rangle_{\mathcal{M}}$$
$$= \sigma^{\mathrm{ex}}. \tag{C16}$$

In the calculation, we use Eq. (C10) in the second line, the assumption $\phi_\alpha^* = 0$ for all $\alpha \in \mathcal{Y}$ in the third line, and the condition on $\vec{\phi}^*$ (85) in the fourth line.

APPENDIX D: THE DERIVATION OF THE WAVENUMBER DECOMPOSITION OF THE ENTROPY PRODUCTION RATE IN REACTION-DIFFUSION SYSTEMS

Here, we provide the derivation of the wavenumber decomposition of EPR in Eq. (103). For simplicity, we abbreviate the weighted Fourier transform in Eq. (101) as $\hat{\vec{F}} = \int_V d\mathbf{r} \overset{\leftrightarrow}{\mathsf{M}}^{\frac{1}{2}} \vec{F} e^{-i\mathbf{k}\cdot\mathbf{r}}$ and $\hat{f} = \int_V d\mathbf{r} \, \mathsf{m}^{\frac{1}{2}} f e^{-i\mathbf{k}\cdot\mathbf{r}}$. If we do not impose the periodic boundary condition on the system, the

Fourier transform of the delta function leads to

$$\int_{\mathbb{R}^d} d\mathbf{k} \, \sigma^{\mathrm{wn}}(\mathbf{k}) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} d\mathbf{k} [\hat{\vec{F}}^\dagger \hat{\vec{F}} + \hat{f}^\dagger \hat{f}]$$

$$= \int_{V \times V} d\mathbf{r} d\mathbf{r}' \int_{\mathbb{R}^d} d\mathbf{k} \frac{e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}}{(2\pi)^d}$$
$$\times \left[\vec{F}^\top(\mathbf{r}) \overset{\leftrightarrow}{\mathsf{M}}^{\frac{1}{2}}(\mathbf{r}) \overset{\leftrightarrow}{\mathsf{M}}^{\frac{1}{2}}(\mathbf{r}') \vec{F}(\mathbf{r}') \right.$$
$$\left. + f^\top(\mathbf{r}) \mathsf{m}^{\frac{1}{2}}(\mathbf{r}) \mathsf{m}^{\frac{1}{2}}(\mathbf{r}') f(\mathbf{r}') \right]$$

$$= \int_{V \times V} d\mathbf{r} d\mathbf{r}' \delta(\mathbf{r} - \mathbf{r}')$$
$$\times \left[\vec{F}^\top(\mathbf{r}) \overset{\leftrightarrow}{\mathsf{M}}^{\frac{1}{2}}(\mathbf{r}) \overset{\leftrightarrow}{\mathsf{M}}^{\frac{1}{2}}(\mathbf{r}') \vec{F}(\mathbf{r}') \right.$$
$$\left. + f^\top(\mathbf{r}) \mathsf{m}^{\frac{1}{2}}(\mathbf{r}) \mathsf{m}^{\frac{1}{2}}(\mathbf{r}') f(\mathbf{r}') \right]$$

$$= \int_V d\mathbf{r} [\vec{F}^\top \overset{\leftrightarrow}{\mathsf{M}} \vec{F} + f^\top \mathsf{m} f] = \sigma, \tag{D1}$$

where we omit the argument t. We also obtain Eq. (105) by replacing $[1/(2\pi)^d] \int_{\mathbb{R}^d} d\mathbf{k}$ with $[1/|V|] \sum_{\mathbf{k}}$ and using $\sum_{\mathbf{k}} e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}/|V| = \delta(\mathbf{r} - \mathbf{r}')$ in the case of the systems with periodic boundaries.

APPENDIX E: THE DETAILS OF THE WASSERSTEIN GEOMETRY

In this Appendix, we provide more details on the Wasserstein distances for RDSs. First, we justify the reformulation of the 1-Wasserstein distance in Eq. (118) in Appendix E 1. We also generalize the Kantorovich–Rubinstein duality (32) in Appendix E 2. Second, we discuss the properties of the optimizer of the 2-Wasserstein distance in Appendix E 3. We introduce two reformulations of the 2-Wasserstein distance in Appendix E 4. We also confirm that the Wasserstein distances satisfy the axioms of distance in Appendix E 5. We derive the inequality between the Wasserstein distances (134) and the TSL based on the 1-Wasserstein distance (146) in Appendices E 6 and E 7. We provide details of the minimum dissipation determined by the 1-Wasserstein distance in Appendix E 8. Finally, we see the property of the 1-Wasserstein distance for the Fisher–KPP equation in Appendix E 10.

In the following, we often use $\vec{c}(0)$ and $\vec{c}(\tau)$ instead of $\vec{c}^A$ and $\vec{c}^B$ since we need to solve the optimization problem for time series $\vec{c} = \{\vec{c}(t)\}_{t \in [0,\tau]}$ such that $\vec{c}_{\mathcal{X}}^A = \vec{c}_{\mathcal{X}}(0)$ and $\vec{c}_{\mathcal{X}}^B = \vec{c}_{\mathcal{X}}(\tau)$ hold from the constraints.

1. Reduction of computational complexity of the 1-Wasserstein distance

Here, we derive the reduced form of the 1-Wasserstein distance (118) from the original definition (115). In the following, $\mathcal{J}^\diamond = (\vec{J}^\diamond, j^\diamond)$ denotes an optimizer of Eq. (115). Letting $\mathcal{U}^\diamond = (\vec{U}^\diamond, u^\diamond)$ be the optimizer of the right-hand side

in Eq. (118), the following inequality holds:

$$
\inf_{\mathcal{U}} |\mathcal{U}|_{\mathrm{RD}} = |\mathcal{U}^{\diamond}|_{\mathrm{RD}}
$$

$$
= \int_0^{\tau} dt \int_V d\boldsymbol{r} \left[\sum_{\alpha \in \mathcal{S}} \left\| \frac{\boldsymbol{U}_{(\alpha)}^{\diamond}}{\tau} \right\| + \sum_{\rho \in \mathcal{R}} \left| \frac{u_{\rho}^{\diamond}}{\tau} \right| \right]
$$

$$
= \int_0^{\tau} dt \left| \frac{\mathcal{U}^{\diamond}}{\tau} \right|_{\mathrm{RD}}
$$

$$
\geqslant \inf_{\mathcal{J}'} \int_0^{\tau} dt |\mathcal{J}'|_{\mathrm{RD}}
$$

$$
= W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau)), \qquad \text{(E1)}
$$

where we use that $\mathcal{U}^{\diamond}/\tau$ satisfies the conditions imposed on the minimization problem (115) because we can rewrite the condition in Eq. (119) to $\partial_t \vec{c}_{\mathcal{X}} = [\nabla^{\dagger}(\mathcal{U}^{\diamond}/\tau)]_{\mathcal{X}}$. On the other hand, we can obtain the inequality in the opposite direction by using $\mathcal{J}^{\diamond}$ as

$$
W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau))
$$

$$
= \int_0^{\tau} dt |\mathcal{J}^{\diamond}|_{\mathrm{RD}}
$$

$$
= \int_0^{\tau} dt \int_V d\boldsymbol{r} \left[\sum_{\alpha \in \mathcal{S}} \| \boldsymbol{J}_{(\alpha)}^{\diamond} \| + \sum_{\rho \in \mathcal{R}} | j_{\rho}^{\diamond} | \right]
$$

$$
\geqslant \int_V d\boldsymbol{r} \left[\sum_{\alpha \in \mathcal{S}} \left\| \int_0^{\tau} dt\, \boldsymbol{J}_{(\alpha)}^{\diamond} \right\| + \sum_{\rho \in \mathcal{R}} \left| \int_0^{\tau} dt\, j_{\rho}^{\diamond} \right| \right]
$$

$$
= \left| \int_0^{\tau} dt\, \mathcal{J}^{\diamond} \right|_{\mathrm{RD}} \geqslant \inf_{\mathcal{U}} |\mathcal{U}|_{\mathrm{RD}}, \qquad \text{(E2)}
$$

where we define time-integrated current as

$$
\int_0^{\tau} dt\, \mathcal{J}^{\diamond} := \left(\int_0^{\tau} dt\, \vec{\boldsymbol{J}}^{\diamond}, \int_0^{\tau} dt\, \boldsymbol{j}^{\diamond} \right),
$$

with

$$
\left(\int_0^{\tau} dt\, \vec{\boldsymbol{J}}^{\diamond} \right)_{(\alpha),i} := \int_0^{\tau} dt\, J_{(\alpha)i}^{\diamond}
$$

$$
\left(\int_0^{\tau} dt\, \boldsymbol{j}^{\diamond} \right)_{\rho} := \int_0^{\tau} dt\, j_{\rho}^{\diamond}.
$$

We also use that time integrated current (E1) satisfies the condition in Eq. (119) to obtain Eq. (118). Unifying the two inequalities, Eqs. (E1) and (E2) leads to the new form of $W_{1,\mathcal{X}}$ (118). Such a rewrite means that the optimizer of Eq. (115) is not unique because we can take

$$
\boldsymbol{J}_{(\alpha)}^{\diamond}(t) = \Theta_{\alpha}(t) \boldsymbol{U}_{(\alpha)}^{\diamond}, \ \ j_{\rho}^{\diamond}(t) = \vartheta_{\rho}(t) u_{\rho}^{\diamond} \qquad \text{(E3)}
$$

as the optimal current of Eq. (115) with the arbitrary nonnegative-valued functions $\Theta_{\alpha}(t)$ and $\vartheta_{\rho}(t)$ that satisfy

$$
\int_0^{\tau} dt\, \Theta_{\alpha}(t) = 1, \ \ \int_0^{\tau} dt\, \vartheta_{\rho}(t) = 1. \qquad \text{(E4)}
$$

The optimization problem in Eq. (118) provides a geometric interpretation of the 1-Wasserstein distance. To confirm it, we rewrite the reduced form of the 1-Wasserstein distance

(118) with the original time series of current $\mathcal{J}$, which satisfies $\vec{c}_{\mathcal{X}}(\tau) - \vec{c}_{\mathcal{X}}(0) = (\nabla^{\dagger} \int_0^{\tau} dt\, \mathcal{J})_{\mathcal{X}}$, as follows:

$$
W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau)) = \inf_{\mathcal{U}} |\mathcal{U}|_{\mathrm{RD}}
$$

$$
= \inf_{\mathcal{U}} \left| \int_0^{\tau} dt\, \mathcal{J} - \left(\int_0^{\tau} dt\, \mathcal{J} - \mathcal{U} \right) \right|_{\mathrm{RD}}
$$

$$
= \inf_{\mathcal{U}^{\mathrm{cyc}}} \left| \int_0^{\tau} dt\, \mathcal{J} - \mathcal{U}^{\mathrm{cyc}} \right|_{\mathrm{RD}}, \qquad \text{(E5)}
$$

where infimum at the third line is taken among all $\mathcal{U}^{\mathrm{cyc}}$ satisfying the condition $(\nabla^{\dagger} \mathcal{U}^{\mathrm{cyc}})_{\mathcal{X}} = \vec{0}_{\mathcal{X}}$, and $\mathcal{U}$ at the first and second lines satisfies Eq. (119). Here, we use the identity $|\mathcal{U}|_{\mathrm{RD}} = | \int_0^{\tau} dt\, \mathcal{J} - (\int_0^{\tau} dt\, \mathcal{J} - \mathcal{U})|_{\mathrm{RD}}$ in the first transform, and let $\mathcal{U}^{\mathrm{cyc}}$ denote $\int_0^{\tau} dt\, \mathcal{J} - \mathcal{U}$. Then, the condition on $\mathcal{U}$, namely, $\vec{c}_{\mathcal{X}}(\tau) - \vec{c}_{\mathcal{X}}(0) = (\nabla^{\dagger} \mathcal{U})_{\mathcal{X}}$, yields $(\nabla^{\dagger} \mathcal{U}^{\mathrm{cyc}})_{\mathcal{X}} = (\nabla^{\dagger} [\int_0^{\tau} dt\, \mathcal{J} - \mathcal{U}])_{\mathcal{X}} = \vec{0}_{\mathcal{X}}$. This implies that the optimization in the last line in Eq. (E5) is performed on all currents that are cyclic in the sense explained in Sec. IV A, i.e., the candidate $\mathcal{U}^{\mathrm{cyc}}$ does not affect the time evolution of the internal species. Obviously, the optimizer of Eq. (E5) is the projection of the original time-integrated current $\int_0^{\tau} dt\, \mathcal{J}$ onto the space of cyclic (time-integrated) currents with respect to the norm $| \cdot |_{\mathrm{RD}}$. It indicates that the 1-Wasserstein distance is the distance between the original time-integrated current and the space of cyclic (time-integrated) currents measured with $| \cdot |_{\mathrm{RD}}$.

In the following, we consider the characteristics of the optimizer $\mathcal{U}^{\diamond}$. We can rewrite the reduced form of the 1-Wasserstein distance (118) with Lagrange multiplier $\vec{\phi}$, whose external part is the zero vector, as

$$
W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau)) = \inf_{\mathcal{U}} \sup_{\vec{\phi}|\vec{\phi}_{\mathcal{Y}}=\vec{0}_{\mathcal{Y}}} \mathcal{I}_{1,\mathcal{X}}[\mathcal{U}, \vec{\phi}], \qquad \text{(E6)}
$$

where the functional $\mathcal{I}_{1,\mathcal{X}}[\mathcal{U}, \vec{\phi}]$ is defined as

$$
\mathcal{I}_{1,\mathcal{X}}[\mathcal{U}, \vec{\phi}] := |\mathcal{U}|_{\mathrm{RD}} + \langle \vec{\phi}, \vec{c}(\tau) - \vec{c}(0) - \nabla^{\dagger} \mathcal{U} \rangle. \qquad \text{(E7)}
$$

Here, we consider the supremum over the Lagrange multiplier $\vec{\phi}$ under the condition $\vec{\phi}_{\mathcal{Y}} = \vec{0}_{\mathcal{Y}}$ because the term $\langle \vec{\phi}, \vec{c}(\tau) - \vec{c}(0) - \nabla^{\dagger} \mathcal{U} \rangle$ only gives a contribution for the constraint on internal species (119) when $\vec{\phi}_{\mathcal{Y}} = \vec{0}_{\mathcal{Y}}$. The boundary condition imposed on $\mathcal{U}$ lets us transform Eq. (E7) by partial integration to

$$
\mathcal{I}_{1,\mathcal{X}}[\mathcal{U}, \vec{\phi}] = |\mathcal{U}|_{\mathrm{RD}} + \langle \vec{\phi}, \vec{c}(\tau) - \vec{c}(0) \rangle - \langle\!\langle \nabla \vec{\phi}, \mathcal{U} \rangle\!\rangle. \qquad \text{(E8)}
$$

Then, its functional derivative and $\vec{\phi}_{\mathcal{Y}} = \vec{0}_{\mathcal{Y}}$ lead to

$$
\boldsymbol{U}_{(\alpha)}^{\diamond} = \boldsymbol{0}, \ u_{\rho}^{\diamond} = 0 \qquad \text{(E9)}
$$

for all $\alpha \in \mathcal{Y}$ and $\rho \in \mathcal{R} \setminus \mathcal{R}_{\mathcal{X}}$, and

$$
\boldsymbol{U}_{(\alpha)}^{\diamond} = \|\boldsymbol{U}_{(\alpha)}^{\diamond}\| \nabla_{\boldsymbol{r}} \phi_{\alpha}^{\diamond},
$$

$$
u_{\rho}^{\diamond} = |u_{\rho}^{\diamond}| (\nabla_{\mathsf{s}} \vec{\phi}^{\diamond})_{\rho}, \qquad \text{(E10)}
$$

for all $\alpha \in \mathcal{X}$ and $\rho \in \mathcal{R}_{\mathcal{X}}$. Thus, for all $\alpha \in \mathcal{X}$, the optimal potential $\vec{\phi}^{\diamond}$ satisfies $\|\nabla_{\boldsymbol{r}} \phi_{\alpha}^{\diamond}\| = 1$ unless $\boldsymbol{U}_{(\alpha)}^{\diamond} = \boldsymbol{0}$. Similarly, for all $\rho \in \mathcal{R}_{\mathcal{X}}$, $|(\nabla_{\mathsf{s}} \vec{\phi}^{\diamond})_{\rho}| = 1$ holds unless $u_{\rho}^{\diamond} = 0$. Note that these conditions indicate that the gradient of a potential determines the direction of the optimal current. In addition,

these conditions lead to a new expression,

$$W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau)) = |\mathcal{U}^\diamond|_{\mathrm{RD}}$$

$$= \int_V d\mathbf{r} \left[\sum_{\alpha \in \mathcal{X}} \|\mathbf{U}^\diamond_{(\alpha)}\| + \sum_{\rho \in \mathcal{R}_\mathcal{X}} |u^\diamond_\rho| \right]$$

$$= \int_V d\mathbf{r} \left[\sum_{\alpha \in \mathcal{X}} \nabla_\mathbf{r} \phi^\diamond_\alpha \cdot \mathbf{U}^\diamond_{(\alpha)} + \sum_{\rho \in \mathcal{R}_\mathcal{X}} (\nabla_\mathbf{s} \vec{\phi}^\diamond)_\rho u^\diamond_\rho \right]. \quad (E11)$$

This expression shows that the value of $\nabla_\mathbf{r} \phi_\alpha(\mathbf{r})$ does not affect $W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau))$ if $\mathbf{U}^\diamond_{(\alpha)}(\mathbf{r}; t) = \mathbf{0}$ at the position $\mathbf{r}$. The value $[\nabla_\mathbf{s} \vec{\phi}^\diamond(\mathbf{r})]_\rho$ also does not matter where $u^\diamond_\rho(\mathbf{r}; t) = 0$.

2. Kantorovich–Rubinstein duality of the 1-Wasserstein distance

We here verify the Kantorovich–Rubinstein duality,

$$\langle \vec{\phi}^\bullet, \vec{c}(\tau) - \vec{c}(0) \rangle = |\mathcal{U}^\diamond|_{\mathrm{RD}}, \quad (E12)$$

where $\mathcal{U}^\diamond = (\vec{U}^\diamond, u^\diamond)$ and $\vec{\phi}^\bullet$ denote the optimizer of Eqs. (118) and (120), respectively. To obtain Eq. (E12), we show the inequalities

$$\langle \vec{\phi}^\bullet, \vec{c}(\tau) - \vec{c}(0) \rangle \leqslant |\mathcal{U}^\diamond|_{\mathrm{RD}}, \quad (E13)$$

and

$$\langle \vec{\phi}^\bullet, \vec{c}(\tau) - \vec{c}(0) \rangle \geqslant |\mathcal{U}^\diamond|_{\mathrm{RD}}. \quad (E14)$$

The first inequality is easily obtained as

$$\langle \vec{\phi}^\bullet, \vec{c}(\tau) - \vec{c}(0) \rangle$$

$$= \langle \vec{\phi}^\bullet, \nabla^\dagger \mathcal{U}^\diamond \rangle = \langle\langle \nabla \vec{\phi}^\bullet, \mathcal{U}^\diamond \rangle\rangle$$

$$= \int_V d\mathbf{r} \left[\sum_{\alpha \in \mathcal{X}} \nabla_\mathbf{r} \phi^\bullet_\alpha \cdot \mathbf{U}^\diamond_{(\alpha)} + \sum_{\rho \in \mathcal{R}_\mathcal{X}} (\nabla_\mathbf{s} \vec{\phi}^\bullet)_\rho u^\diamond_\rho \right]$$

$$\leqslant \int_V d\mathbf{r} \left[\sum_{\alpha \in \mathcal{X}} \|\mathbf{U}^\diamond_{(\alpha)}\| + \sum_{\rho \in \mathcal{R}_\mathcal{X}} |u^\diamond_\rho| \right] = |\mathcal{U}^\diamond|_{\mathrm{RD}}, \quad (E15)$$

where we use the generalized 1-Lipschitz continuity (121), $\|\nabla_\mathbf{r} \phi^\bullet_\alpha\| \leqslant 1$ for $\alpha \in \mathcal{X}$ and $|(\nabla_\mathbf{s} \vec{\phi}^\bullet)_\rho| \leqslant 1$ for $\rho \in \mathcal{R}_\mathcal{X}$, and the conditions on $\mathcal{U}^\diamond$ in Eq. (E9).

To verify the second inequality, we consider the optimization problem

$$\sup_{\vec{\psi}} \langle \vec{\psi}, \vec{c}(\tau) - \vec{c}(0) \rangle, \quad (E16)$$

with the following conditions: $\vec{\psi} \in \mathrm{Lip}^1_\mathcal{X}$ and the diffusion part of $\nabla \vec{\psi}$ satisfies the same boundary conditions as the one imposed on $\vec{U}^\diamond$. Using another representation of the conditions for the generalized 1-Lipschitz continuity (121),

$$\begin{aligned} \forall \alpha \in \mathcal{X}, \quad & \|\nabla_\mathbf{r} \psi_\alpha\|^2 - 1 \leqslant 0 \\ \forall \rho \in \mathcal{R}_\mathcal{X}, \quad & |(\nabla_\mathbf{s} \vec{\psi})_\rho|^2 - 1 \leqslant 0 \end{aligned}, \quad (E17)$$

we can rewrite the optimization problem (E16) with the Lagrange multipliers $\{U^\mathrm{L}_\alpha \geqslant 0\}_{\alpha \in \mathcal{X}}$ and $\{u^\mathrm{L}_\rho \geqslant 0\}_{\rho \in \mathcal{R}_\mathcal{X}}$ as

$$\sup_{\vec{\psi}} \inf_{\substack{\{U^\mathrm{L}_\alpha \geqslant 0\}_{\alpha \in \mathcal{X}}, \\ \{u^\mathrm{L}_\rho \geqslant 0\}_{\rho \in \mathcal{R}_\mathcal{X}}}} \mathcal{I}_\mathrm{KR}\left[\vec{\psi}, \{U^\mathrm{L}_\alpha\}_{\alpha \in \mathcal{X}}, \{u^\mathrm{L}_\rho\}_{\rho \in \mathcal{R}_\mathcal{X}} \right], \quad (E18)$$

with the following conditions: $\vec{\psi}_Y = \vec{0}_Y$ and the diffusion part of $\nabla \vec{\psi}$ satisfies the same boundary conditions as the one imposed on $\vec{U}^\diamond$. Here, the functional $\mathcal{I}_\mathrm{KR}$ is defined as

$$\mathcal{I}_\mathrm{KR} := \langle \vec{\psi}, \vec{c}(\tau) - \vec{c}(0) \rangle$$

$$- \frac{1}{2} \int_V d\mathbf{r} \sum_{\alpha \in \mathcal{X}} U^\mathrm{L}_\alpha (\|\nabla_\mathbf{r} \psi_\alpha\|^2 - 1)$$

$$- \frac{1}{2} \int_V d\mathbf{r} \sum_{\rho \in \mathcal{R}_\mathcal{X}} u^\mathrm{L}_\rho (|(\nabla_\mathbf{s} \vec{\psi})_\rho|^2 - 1). \quad (E19)$$

Computing the functional derivatives of Eq. (E19), we can verify that the optimizer of Eq. (E18), $(\vec{\psi}^\bullet, \{U^{\mathrm{L}\bullet}_\alpha\}_{\alpha \in \mathcal{X}}, \{u^{\mathrm{L}\bullet}_\rho\}_{\rho \in \mathcal{R}_\mathcal{X}})$, satisfies

$$c_\alpha(\tau) - c_\alpha(0)$$

$$= -\nabla_\mathbf{r} \cdot (U^{\mathrm{L}\bullet}_\alpha \nabla_\mathbf{r} \psi^\bullet_\alpha) + \sum_{\rho \in \mathcal{R}_\mathcal{X}} S_{\alpha\rho} u^{\mathrm{L}\bullet}_\rho (\nabla_\mathbf{s} \vec{\psi}^\bullet)_\rho, \quad (E20)$$

for all $\alpha \in \mathcal{X}$ as the Euler–Lagrange equation and

$$\begin{aligned} \forall \alpha \in \mathcal{X}, \quad & U^{\mathrm{L}\bullet}_\alpha (\|\nabla_\mathbf{r} \psi^\bullet_\alpha\|^2 - 1) = 0 \\ \forall \rho \in \mathcal{R}_\mathcal{X}, \quad & u^{\mathrm{L}\bullet}_\rho (|(\nabla_\mathbf{s} \vec{\psi}^\bullet)_\rho|^2 - 1) = 0 \end{aligned}, \quad (E21)$$

as the complementary slackness conditions. These conditions lead to

$$\langle \vec{\psi}^\bullet, \vec{c}(\tau) - \vec{c}(0) \rangle = \frac{1}{2} \int_V d\mathbf{r} \sum_{\alpha \in \mathcal{X}} U^{\mathrm{L}\bullet}_\alpha (\|\nabla_\mathbf{r} \psi^\bullet_\alpha\|^2 + 1)$$

$$+ \frac{1}{2} \int_V d\mathbf{r} \sum_{\rho \in \mathcal{R}_\mathcal{X}} u^{\mathrm{L}\bullet}_\rho (|(\nabla_\mathbf{s} \vec{\psi}^\bullet)_\rho|^2 + 1)$$

$$= \int_V d\mathbf{r} \left[\sum_{\alpha \in \mathcal{X}} U^{\mathrm{L}\bullet}_\alpha + \sum_{\rho \in \mathcal{R}_\mathcal{X}} u^{\mathrm{L}\bullet}_\rho \right], \quad (E22)$$

where we substitute the Euler–Lagrange equation (E20) into Eq. (E18) in the first transform and use the complementary slackness conditions (E21) in the second transform. The nonnegativity of the Lagrange multipliers and the complementary slackness conditions (E21) allow us to rewrite the right-hand side of Eq. (E22) as $|\mathcal{U}^\bullet|_{\mathrm{RD}}$, with a new time-integrated current $\mathcal{U}^\bullet = (\vec{U}^\bullet, u^\bullet)$ defined as

$$\mathbf{U}^\bullet_{(\alpha)} := \begin{cases} U^{\mathrm{L}\bullet}_\alpha \nabla_\mathbf{r} \psi^\bullet_\alpha & (\alpha \in \mathcal{X}) \\ \mathbf{0} & (\alpha \in \mathcal{Y}) \end{cases},$$

$$(u^\bullet)_\rho := \begin{cases} u^{\mathrm{L}\bullet}_\rho (\nabla_\mathbf{s} \vec{\psi}^\bullet)_\rho & (\rho \in \mathcal{R}_\mathcal{X}) \\ 0 & (\rho \in \mathcal{R} \setminus \mathcal{R}_\mathcal{X}) \end{cases}. \quad (E23)$$

Thus, we obtain

$$\langle \vec{\psi}^\bullet, \vec{c}(\tau) - \vec{c}(0) \rangle = |\mathcal{U}^\bullet|_{\mathrm{RD}} \geqslant |\mathcal{U}^\diamond|_{\mathrm{RD}}, \quad (E24)$$

where we use the fact that $\mathcal{U}^\bullet$ becomes a candidate of the optimization problem in Eq. (118). This is because the conditions on the diffusion part of $\nabla\vec{\psi}^\bullet$ let $\mathcal{U}^\bullet$ satisfy the boundary conditions imposed in the optimization problem in Eq. (118), and the Euler–Lagrange equation (E20) leads to $\vec{c}_\mathcal{X}(\tau) - \vec{c}_\mathcal{X}(0) = (\nabla^\dagger \mathcal{U}^\bullet)_\mathcal{X}$. Finally, we can verify the second inequality (E14) by combining this inequality (E24) and the consequence of adding the conditions in the optimization in Eq. (E16),

$$\langle \vec{\phi}^\bullet, \vec{c}(\tau) - \vec{c}(0) \rangle \geqslant \langle \vec{\psi}^\bullet, \vec{c}(\tau) - \vec{c}(0) \rangle. \tag{E25}$$

We remark that the equality (E12) implies that $\mathcal{U}^\bullet$ is also an optimizer for the 1-Wasserstein distance (118).

3. The optimizer of the 2-Wasserstein distance

We can rewrite the optimization problem in Eq. (122) using Lagrange multiplier $\vec{\phi}$ as

$$W_{2,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau) | \vec{b}_\mathcal{Y})^2 = \inf_{\vec{c}, \mathcal{F}} \sup_{\vec{\phi} | \vec{\phi}_\mathcal{Y} = \vec{0}_\mathcal{Y}} \tau \mathcal{I}_{2,\mathcal{X}}[\vec{c}, \mathcal{F}, \vec{\phi}], \tag{E26}$$

where $\mathcal{I}_{2,\mathcal{X}}$ is the functional defined as

$$\mathcal{I}_{2,\mathcal{X}}[\vec{c}, \mathcal{F}, \vec{\phi}] := \int_0^\tau dt \big[\langle\!\langle \mathcal{F}, \mathcal{F} \rangle\!\rangle_{\mathcal{M}_{\vec{c}}} + 2\langle \vec{\phi}, \partial_t \vec{c} - \nabla^\dagger \mathcal{M}_{\vec{c}} \mathcal{F} \rangle \big]. \tag{E27}$$

Here, we consider the supremum over the Lagrange multiplier $\vec{\phi}$ under the condition $\vec{\phi}_\mathcal{Y} = \vec{0}_\mathcal{Y}$ because the term $\langle \vec{\phi}, \partial_t \vec{c} - \nabla^\dagger \mathcal{M}_{\vec{c}} \mathcal{F} \rangle$ only gives a contribution for the constraint on internal species Eq. (123) when $\vec{\phi}_\mathcal{Y} = \vec{0}_\mathcal{Y}$. Because we fix $\vec{c}(0)$ and $\vec{c}(\tau)$, impose boundary conditions on $\overleftrightarrow{\mathsf{M}}\vec{F}$ for internal species, and let $\vec{\phi}$ satisfy $\phi_\alpha = 0$ for all $\alpha \in \mathcal{Y}$, partial integration yields

$$\mathcal{I}_{2,\mathcal{X}}[\vec{c}, \mathcal{F}, \vec{\phi}] = 2\langle \vec{\phi}(\tau), \vec{c}(\tau) \rangle - 2\langle \vec{\phi}(0), \vec{c}(0) \rangle$$
$$+ \int_0^\tau dt [\langle\!\langle \mathcal{F} - 2\nabla\vec{\phi}, \mathcal{F} \rangle\!\rangle_{\mathcal{M}_{\vec{c}}} - 2\langle \partial_t \vec{\phi}, \vec{c} \rangle]. \tag{E28}$$

In the following, we write the optimizer of the right-hand side of Eq. (E26) as $(\vec{c}^\star, \mathcal{F}^\star, \vec{\phi}^\star)$. As in case in Appendix B, the functional derivative of Eq. (E28) leads to the conditions to be satisfied by $(\vec{c}^\star, \mathcal{F}^\star, \vec{\phi}^\star)$,

$$\sum_{\beta \in \mathcal{S}} \mathsf{M}^\star_{(\alpha\beta)}(F^\star_{(\beta)} - \nabla_r \phi^\star_\beta) = 0 \tag{E29}$$

for all $\alpha \in \mathcal{S}$, and

$$m^\star_\rho \{ f^\star_\rho - (\nabla_s \vec{\phi}^\star)_\rho \} = 0 \tag{E30}$$

for all $\rho \in \mathcal{R}$, where $\mathcal{F}^\star = (\vec{F}^\star, f^\star)$, and $\mathsf{M}^\star_{(\alpha\beta)} = [\overleftrightarrow{\mathsf{M}}^\star]_{(\alpha\beta)}$ and $m^\star_\rho = [\mathsf{m}^\star]_{\rho\rho}$ are given by $\mathcal{M}_{\vec{c}^\star} = \overleftrightarrow{\mathsf{M}}^\star \oplus \mathsf{m}^\star$. $\overleftrightarrow{\mathsf{M}}^\star$ and $m^\star$ indicate the mobility tensor and the edgewise Onsager coefficient matrix for the optimal concentration distribution $\vec{c}^\star$. These results and positive-definiteness of $\overleftrightarrow{\mathsf{M}}^\star$ and $\mathsf{m}^\star$ make the optimal force be the gradient of potential,

$$\mathcal{F}^\star = \nabla\vec{\phi}^\star. \tag{E31}$$

This condition means that in order to minimize EP, we should drive the system by the conservative thermodynamic force corresponding to the potential $\vec{\phi}^\star$.

4. Reformulations of the 2-Wasserstein distance

Here, we introduce two reformulations of the 2-Wasserstein distance. The condition for the optimal force in Eq. (E31) let us rewrite the 2-Wasserstein distance as the form in Eq. (126),

$$W_{2,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau) | \vec{b}_\mathcal{Y})^2 = \inf_{\vec{c}, \vec{\phi} | \vec{\phi}_\mathcal{Y} = \vec{0}_\mathcal{Y}} \tau \int_0^\tau dt \langle\!\langle \nabla\vec{\phi}, \nabla\vec{\phi} \rangle\!\rangle_{\mathcal{M}_{\vec{c}}}, \tag{E32}$$

with the conditions

$$\partial_t \vec{c}_\mathcal{X} = (\nabla^\dagger \mathcal{M}_{\vec{c}} \nabla\vec{\phi})_\mathcal{X}, \quad \vec{c}_\mathcal{Y}(t) = \vec{b}_\mathcal{Y}. \tag{E33}$$

We can also rewrite the 2-Wasserstein distance as

$$W_{2,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau) | \vec{b}_\mathcal{Y}) = \inf_{\vec{c}, \vec{\phi} | \vec{\phi}_\mathcal{Y} = \vec{0}_\mathcal{Y}} \int_0^\tau dt \sqrt{\langle\!\langle \nabla\vec{\phi}, \nabla\vec{\phi} \rangle\!\rangle_{\mathcal{M}_{\vec{c}}}}, \tag{E34}$$

with the same conditions as Eq. (E33).

To prove the equivalence between Eqs. (E32) and (E34), we consider the optimizer of Eq. (E34). Let $(\vec{c}^\sharp, \vec{\phi}^\sharp)$ denote the optimizer of Eq. (E34). For the derivation of Eq. (E34), it is sufficient to confirm

$$W_{2,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau) | \vec{b}_\mathcal{Y})^2 \geqslant \left(\int_0^\tau dt \sqrt{\langle\!\langle \nabla\vec{\phi}^\sharp, \nabla\vec{\phi}^\sharp \rangle\!\rangle_{\mathcal{M}_{\vec{c}^\sharp}}} \right)^2, \tag{E35}$$

and

$$W_{2,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau) | \vec{b}_\mathcal{Y})^2 \leqslant \left(\int_0^\tau dt \sqrt{\langle\!\langle \nabla\vec{\phi}^\sharp, \nabla\vec{\phi}^\sharp \rangle\!\rangle_{\mathcal{M}_{\vec{c}^\sharp}}} \right)^2. \tag{E36}$$

We can easily show the first inequality (E35)

$$W_{2,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau) | \vec{b}_\mathcal{Y})^2$$
$$= \tau \int_0^\tau dt \langle\!\langle \nabla\vec{\phi}^\star, \nabla\vec{\phi}^\star \rangle\!\rangle_{\mathcal{M}_{\vec{c}^\star}}$$
$$= \left(\int_0^\tau dt \right) \left(\int_0^\tau dt \langle\!\langle \nabla\vec{\phi}^\star, \nabla\vec{\phi}^\star \rangle\!\rangle_{\mathcal{M}_{\vec{c}^\star}} \right)$$
$$\geqslant \left(\int_0^\tau dt \sqrt{\langle\!\langle \nabla\vec{\phi}^\star, \nabla\vec{\phi}^\star \rangle\!\rangle_{\mathcal{M}_{\vec{c}^\star}}} \right)^2$$
$$\geqslant \left(\int_0^\tau dt \sqrt{\langle\!\langle \nabla\vec{\phi}^\sharp, \nabla\vec{\phi}^\sharp \rangle\!\rangle_{\mathcal{M}_{\vec{c}^\sharp}}} \right)^2, \tag{E37}$$

where we used the Cauchy–Schwarz inequality and the fact that $(\vec{c}^\sharp, \vec{\phi}^\sharp)$ denote the optimizer of Eq. (E34).

To derive the second inequality (E36), we use arc-length reparametrization by referring to the literatures [91,135]. Introducing a function $s_\epsilon(t)$ for $t \in [0, \tau]$ with sufficiently small $\epsilon > 0$ as

$$s_\epsilon(t) := \int_0^t dt' \sqrt{\epsilon + \langle\!\langle \nabla\vec{\phi}^\sharp(t'), \nabla\vec{\phi}^\sharp(t') \rangle\!\rangle_{\mathcal{M}_{\vec{c}^\sharp(t')}}}, \tag{E38}$$

we can define the inverse function $t_\epsilon := s_\epsilon^{-1}$ because $d_t s_\epsilon(t) > 0$ holds so that $s_\epsilon(t)$ is an increasing function of t. The inverse function $t_\epsilon(s)$ satisfies

$$d_s t_\epsilon(s)|_{s=s_\epsilon(t)} = (d_t s_\epsilon)^{-1} = \frac{1}{\sqrt{\epsilon + \langle\!\langle \nabla\vec{\phi}^\sharp, \nabla\vec{\phi}^\sharp \rangle\!\rangle_{\mathcal{M}_{\vec{c}^\sharp}}}}. \quad (E39)$$

We define the reparametrized concentration distribution $\vec{\omega}(s)$ as $\vec{\omega}(s) := \vec{c}^\sharp(t_\epsilon(s))$, and a potential $\vec{\zeta}(s)$ as $\vec{\zeta}(s) := (d_s t_\epsilon(s))\vec{\phi}^\sharp(t_\epsilon(s))$. These quantities satisfy

$$\vec{\omega}(0) = \vec{c}^\sharp(0) = \vec{c}(0), \ \ \vec{\omega}(s_\epsilon(\tau)) = \vec{c}^\sharp(\tau) = \vec{c}(\tau), \quad (E40)$$

$$\vec{\omega}_{\mathcal{Y}}(s) = \vec{c}^\sharp_{\mathcal{Y}}(t_\epsilon(s)) = \vec{b}_{\mathcal{Y}}, \quad (E41)$$

and

$$
\begin{aligned}
\partial_s \vec{\omega}_{\mathcal{X}}(s) &= (d_s t_\epsilon(s)) \, \partial_t \vec{c}^\sharp_{\mathcal{X}}(t)\big|_{t=t_\epsilon(s)} \\
&= (d_s t_\epsilon(s))\big(\nabla^\dagger \mathcal{M}_{\vec{c}^\sharp(t_\epsilon(s))} \nabla\vec{\phi}^\sharp(t_\epsilon(s))\big)_{\mathcal{X}} \\
&= (\nabla^\dagger \mathcal{M}_{\vec{\omega}(s)} \nabla\vec{\zeta}(s))_{\mathcal{X}}. \quad (E42)
\end{aligned}
$$

These conditions in Eqs. (E40)–(E42) are the same as the conditions imposed on the optimization problem in Eq. (E32) if we replace the time duration τ with $s_\epsilon(\tau)$. From the definition in Eqs. (E32) and (E39), we thus obtain

$$
\begin{aligned}
&W_{2,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau)|\vec{b}_{\mathcal{Y}})^2 \\
&\leqslant s_\epsilon(\tau) \int_0^{s_\epsilon(\tau)} ds \langle\!\langle \nabla\vec{\zeta}, \nabla\vec{\zeta} \rangle\!\rangle_{\mathcal{M}_{\vec{\omega}}} \\
&= s_\epsilon(\tau) \int_0^\tau dt [d_t s_\epsilon(t)]\big[d_s t_\epsilon(s)|_{s=s_\epsilon(t)}\big]^2 \langle\!\langle \nabla\vec{\phi}^\sharp, \nabla\vec{\phi}^\sharp \rangle\!\rangle_{\mathcal{M}_{\vec{c}^\sharp}} \\
&= s_\epsilon(\tau) \int_0^\tau dt\, \frac{\langle\!\langle \nabla\vec{\phi}^\sharp, \nabla\vec{\phi}^\sharp \rangle\!\rangle_{\mathcal{M}_{\vec{c}^\sharp}}}{\sqrt{\epsilon + \langle\!\langle \nabla\vec{\phi}^\sharp, \nabla\vec{\phi}^\sharp \rangle\!\rangle_{\mathcal{M}_{\vec{c}^\sharp}}}} \\
&\leqslant \left(\int_0^\tau dt\, \sqrt{\epsilon + \langle\!\langle \nabla\vec{\phi}^\sharp, \nabla\vec{\phi}^\sharp \rangle\!\rangle_{\mathcal{M}_{\vec{c}^\sharp}}}\right)^2. \quad (E43)
\end{aligned}
$$

Taking the limit $\epsilon \to 0$ in Eq. (E43) leads to the inequality we need to derive (E36).

The optimizer of Eq. (E32) $(\vec{c}^\star, \vec{\phi}^\star)$ is one of the optimizers of Eq. (E34) because we can easily derive

$$W_{2,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau)|\vec{b}_{\mathcal{Y}}) = \int_0^\tau dt\, \sqrt{\langle\!\langle \nabla\vec{\phi}^\star, \nabla\vec{\phi}^\star \rangle\!\rangle_{\mathcal{M}_{\vec{c}^\star}}}, \quad (E44)$$

by repeating the same argument above using

$$s_\epsilon^\star(t) := \int_0^t dt'\, \sqrt{\epsilon + \langle\!\langle \nabla\vec{\phi}^\star(t'), \nabla\vec{\phi}^\star(t') \rangle\!\rangle_{\mathcal{M}_{\vec{c}^\star(t')}}}, \quad (E45)$$

instead of $s_\epsilon(t)$. The form of the 2-Wasserstein distance in Eq. (E44) means that the optimizer $(\vec{c}^\star, \vec{\phi}^\star)$ satisfies the equality condition of the Cauchy–Schwarz inequality in Eq. (E37) so that the condition

$$\frac{W_{2,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau)|\vec{b}_{\mathcal{Y}})}{\tau} = \sqrt{\langle\!\langle \nabla\vec{\phi}^\star, \nabla\vec{\phi}^\star \rangle\!\rangle_{\mathcal{M}_{\vec{c}^\star}}} \quad (E46)$$

holds for all $t \in [0, \tau]$. This condition means that the geodesic in Fig. 6 has the constant speed.

5. Axioms of distance

Here, we confirm that the Wasserstein distances satisfy the axioms of distance: nondegenerateness, symmetry, and the triangle inequality.

First, we prove the nondegenerateness of the Wasserstein distances. The nondegenerateness of the 1-Wasserstein distance is $W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau)) = 0 \Leftrightarrow \vec{c}_{\mathcal{X}}(0) = \vec{c}_{\mathcal{X}}(\tau)$. Letting $W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau)) = 0$ hold, all the elements of the optimal current are zero. Then, the constraint $\partial_t \vec{c}_{\mathcal{X}} = (\nabla^\dagger \mathcal{J}')_{\mathcal{X}}$ leads to $\vec{c}_{\mathcal{X}}(0) = \vec{c}_{\mathcal{X}}(\tau)$. Conversely, assuming $\vec{c}_{\mathcal{X}}(0) = \vec{c}_{\mathcal{X}}(\tau)$, the current with zero elements satisfies the constraints imposed on the optimization problem for the 1-Wasserstein distance so that $W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau)) = 0$ holds. We can also prove the nondegenerateness of the 2-Wasserstein distance by replacing the current with the force and arguing similarly.

Second, we prove the symmetry of the Wasserstein distance. The symmetry of the 1-Wasserstein distance is $W_{1,\mathcal{X}}(\vec{c}^A, \vec{c}^B) = W_{1,\mathcal{X}}(\vec{c}^B, \vec{c}^A)$. Letting $\vec{c}^\diamond$ and $\mathcal{J}^\diamond$ denote the optimizer for $W_{1,\mathcal{X}}(\vec{c}^A, \vec{c}^B)$, the time reversal quantities $\vec{c}'(t) := \vec{c}^\diamond(\tau - t)$ and $\mathcal{J}'(t) := -\mathcal{J}^\diamond(\tau - t)$ satisfy the constraints imposed on the optimization problem for $W_{1,\mathcal{X}}(\vec{c}^B, \vec{c}^A)$. Then, we obtain an inequality,

$$
\begin{aligned}
W_{1,\mathcal{X}}(\vec{c}^A, \vec{c}^B) &= \int_0^\tau dt |\mathcal{J}^\diamond(t)|_{\mathrm{RD}} \\
&= \int_0^\tau dt |-\mathcal{J}'(\tau - t)|_{\mathrm{RD}} \\
&= \int_0^\tau dt |\mathcal{J}'(t)|_{\mathrm{RD}} \\
&\geqslant W_{1,\mathcal{X}}(\vec{c}^B, \vec{c}^A). \quad (E47)
\end{aligned}
$$

Similarly, we can obtain the inequality in the opposite direction, $W_{1,\mathcal{X}}(\vec{c}^A, \vec{c}^B) \leqslant W_{1,\mathcal{X}}(\vec{c}^B, \vec{c}^A)$. We can also prove the symmetry of the 2-Wasserstein distance in the same way.

Third, we prove the triangle inequality of the 1-Wasserstein distance, $W_{1,\mathcal{X}}(\vec{c}^A, \vec{c}^B) + W_{1,\mathcal{X}}(\vec{c}^B, \vec{c}^C) \geqslant W_{1,\mathcal{X}}(\vec{c}^A, \vec{c}^C)$. Letting the optimizer for $W_{1,\mathcal{X}}(\vec{c}^A, \vec{c}^B)$ and $W_{1,\mathcal{X}}(\vec{c}^B, \vec{c}^C)$ be $(\vec{c}^{\diamond,1}, \mathcal{J}^{\diamond,1})$ and $(\vec{c}^{\diamond,2}, \mathcal{J}^{\diamond,2})$, respectively, a new concentration distributions $\vec{c}'$ and a new current $\mathcal{J}'$ defined as

$$\vec{c}'(t) := \begin{cases} \vec{c}^{\diamond,1}(2t) & \left(0 \leqslant t < \frac{\tau}{2}\right) \\ \vec{c}^{\diamond,2}(2t - \tau) & \left(\frac{\tau}{2} \leqslant t \leqslant \tau\right) \end{cases}, \quad (E48)$$

$$\mathcal{J}'(t) := \begin{cases} 2\mathcal{J}^{\diamond,1}(2t) & \left(0 \leqslant t < \frac{\tau}{2}\right) \\ 2\mathcal{J}^{\diamond,2}(2t - \tau) & \left(\frac{\tau}{2} \leqslant t \leqslant \tau\right) \end{cases}, \quad (E49)$$

satisfy the constraints imposed on the optimization problem for $W_{1,\mathcal{X}}(\vec{c}^A, \vec{c}^C)$: $\vec{c}'(0) = \vec{c}^A$, $\vec{c}'(\tau) = \vec{c}^C$, and $\partial_t \vec{c}'_{\mathcal{X}} = (\nabla^\dagger \mathcal{J}')_{\mathcal{X}}$. Thus, we obtain the triangle inequality as

$$
\begin{aligned}
W_{1,\mathcal{X}}(\vec{c}^A, \vec{c}^C) &\leqslant \int_0^\tau dt |\mathcal{J}'(t)|_{\mathrm{RD}} \\
&= \int_0^{\frac{\tau}{2}} dt\, 2|\mathcal{J}^{\diamond,1}(2t)|_{\mathrm{RD}} + \int_{\frac{\tau}{2}}^\tau dt\, 2|\mathcal{J}^{\diamond,2}(2t - \tau)|_{\mathrm{RD}} \\
&= \int_0^\tau dt |\mathcal{J}^{\diamond,1}(t)|_{\mathrm{RD}} + \int_0^\tau dt |\mathcal{J}^{\diamond,2}(t)|_{\mathrm{RD}} \\
&= W_{1,\mathcal{X}}(\vec{c}^A, \vec{c}^B) + W_{1,\mathcal{X}}(\vec{c}^B, \vec{c}^C). \quad (E50)
\end{aligned}
$$

Finally, we prove the triangle inequality of the 2-Wasserstein distance, $W_{2,\mathcal{X}}(\vec{c}^A, \vec{c}^B|\vec{b}_\mathcal{Y}) + W_{2,\mathcal{X}}(\vec{c}^B, \vec{c}^C|\vec{b}_\mathcal{Y}) \geqslant W_{2,\mathcal{X}}(\vec{c}^A, \vec{c}^C|\vec{b}_\mathcal{Y})$. We need to use the reformulation of the 2-Wasserstein distance in Eq. (E34) to derive the triangle inequality [135]. Letting the optimizer of the optimization problem in Eq. (E34) for $W_{2,\mathcal{X}}(\vec{c}^A, \vec{c}^B|\vec{b}_\mathcal{Y})$ and $W_{2,\mathcal{X}}(\vec{c}^B, \vec{c}^C|\vec{b}_\mathcal{Y})$ be $(\vec{c}^{\sharp,1}, \vec{\phi}^{\sharp,1})$ and $(\vec{c}^{\sharp,2}, \vec{\phi}^{\sharp,2})$, respectively, a new concentration distributions $\vec{c}'$ and a new potential $\vec{\phi}'$ defined as

$$\vec{c}'(t) := \begin{cases} \vec{c}^{\sharp,1}(2t) & \left(0 \leqslant t < \dfrac{\tau}{2}\right) \\ \vec{c}^{\sharp,2}(2t-\tau) & \left(\dfrac{\tau}{2} \leqslant t \leqslant \tau\right) \end{cases}, \qquad (E51)$$

$$\vec{\phi}'(t) := \begin{cases} 2\vec{\phi}^{\sharp,1}(2t) & \left(0 \leqslant t < \dfrac{\tau}{2}\right) \\ 2\vec{\phi}^{\sharp,2}(2t-\tau) & \left(\dfrac{\tau}{2} \leqslant t \leqslant \tau\right) \end{cases}, \qquad (E52)$$

satisfy the constraints imposed on the optimization problem for $W_{2,\mathcal{X}}(\vec{c}^A, \vec{c}^C|\vec{b}_\mathcal{Y})$: $\vec{c}'_\mathcal{X}(0) = \vec{c}^A_\mathcal{X}$, $\vec{c}'_\mathcal{X}(\tau) = \vec{c}^C_\mathcal{X}$, $\vec{c}'_\mathcal{Y}(t) = \vec{b}_\mathcal{Y}$, and $\partial_t \vec{c}'_\mathcal{X} = (\nabla^\dagger \mathcal{M}_{\vec{c}'} \nabla \vec{\phi}')_\mathcal{X}$. Thus, we obtain the triangle inequality as

$$W_{2,\mathcal{X}}(\vec{c}^A, \vec{c}^C|\vec{b}_\mathcal{Y}) \leqslant \int_0^\tau dt \sqrt{\langle\!\langle \nabla\vec{\phi}', \nabla\vec{\phi}'\rangle\!\rangle_{\mathcal{M}_{\vec{c}'}}}$$

$$= \int_0^{\frac{\tau}{2}} dt \, 2\sqrt{\langle\!\langle \nabla\vec{\phi}^{\sharp,1}(2t), \nabla\vec{\phi}^{\sharp,1}(2t)\rangle\!\rangle_{\mathcal{M}_{\vec{c}^{\sharp,1}(2t)}}}$$

$$+ \int_{\frac{\tau}{2}}^\tau dt \, 2\sqrt{\langle\!\langle \nabla\vec{\phi}^{\sharp,2}(2t-\tau), \nabla\vec{\phi}^{\sharp,2}(2t-\tau)\rangle\!\rangle_{\mathcal{M}_{\vec{c}^{\sharp,2}(2t-\tau)}}}$$

$$= \int_0^\tau dt \sqrt{\langle\!\langle \nabla\vec{\phi}^{\sharp,1}(t), \nabla\vec{\phi}^{\sharp,1}(t)\rangle\!\rangle_{\mathcal{M}_{\vec{c}^{\sharp,1}(t)}}}$$

$$+ \int_0^\tau dt \sqrt{\langle\!\langle \nabla\vec{\phi}^{\sharp,2}(t), \nabla\vec{\phi}^{\sharp,2}(t)\rangle\!\rangle_{\mathcal{M}_{\vec{c}^{\sharp,2}(t)}}}$$

$$= W_{2,\mathcal{X}}(\vec{c}^A, \vec{c}^B|\vec{b}_\mathcal{Y}) + W_{2,\mathcal{X}}(\vec{c}^B, \vec{c}^C|\vec{b}_\mathcal{Y}). \qquad (E53)$$

6. Derivation of the inequality between the Wasserstein distances

Here, we derive the inequality between the 1- and 2-Wasserstein distances (134). In the following, we let $(\vec{c}^\star, \mathcal{F}^\star)$ denote the optimizer of the minimization problem for the 2-Wasserstein distance (122). The corresponding current $\mathcal{J}^\star = (\vec{J}^\star, j^\star)$ is defined as $\mathcal{J}^\star = \mathcal{M}_{\vec{c}^\star}\mathcal{F}^\star$. We use a new current $\mathcal{J}^\times = (\vec{J}^\times, j^\times)$ and a new force $\mathcal{F}^{\mathrm{uni}} = (\vec{F}^{\mathrm{uni}}, f^{\mathrm{uni}})$, defined as

$$\vec{J}^\times_{(\alpha)} := \begin{cases} \vec{J}^\star_{(\alpha)} & (\alpha \in \mathcal{X}) \\ \mathbf{0} & (\alpha \in \mathcal{Y}) \end{cases}, \quad j^\times_\rho := \begin{cases} j^\star_\rho & (\rho \in \mathcal{R}_\mathcal{X}) \\ 0 & (\rho \in \mathcal{R} \setminus \mathcal{R}_\mathcal{X}) \end{cases},$$

and

$$\vec{F}^{\mathrm{uni}}_{(\alpha)} := \begin{cases} \dfrac{\vec{J}^\star_{(\alpha)}}{\|\vec{J}^\star_{(\alpha)}\|} & (\alpha \in \mathcal{X}, \|\vec{J}^\star_{(\alpha)}\| \neq 0) \\ \mathbf{0} & (\text{otherwise}) \end{cases},$$

$$f^{\mathrm{uni}}_\rho := \begin{cases} \dfrac{j^\star_\rho}{|j^\star_\rho|} & (\rho \in \mathcal{R}_\mathcal{X}, |j^\star_\rho| \neq 0) \\ 0 & (\text{otherwise}) \end{cases}.$$

The current $\mathcal{J}^\times$ satisfies

$$\int_0^\tau dt |\mathcal{J}^\times|_{\mathrm{RD}} \geqslant W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau)), \qquad (E54)$$

since the conditions imposed on $(\vec{c}^\star, \mathcal{F}^\star)$ let $\mathcal{J}^\times$ be a candidate of the optimization problem in the definition of the 1-Wasserstein distance (115). The force $\mathcal{F}^{\mathrm{uni}}$ also satisfies

$$\langle\!\langle \mathcal{F}^{\mathrm{uni}}, \mathcal{F}^{\mathrm{uni}} \rangle\!\rangle_{\mathcal{M}_{\vec{c}^\star}} \leqslant |\mathcal{M}_{\vec{c}^\star}|^{\mathrm{tot}}_\mathcal{X}. \qquad (E55)$$

We obtain this using the following two facts: the diffusion part $\vec{F}^{\mathrm{uni}}(\mathbf{r};t)$ is a candidate of the maximization problem in the definition of $M^{\max}_\mathcal{X}(\mathbf{r};t)$ (131), and the reaction part satisfies $|f^{\mathrm{uni}}_\rho(\mathbf{r};t)|^2 \leqslant 1$ for all $\rho \in \mathcal{R}_\mathcal{X}$ and $f^{\mathrm{uni}}_\rho(\mathbf{r};t) = 0$ for all $\rho \in \mathcal{R} \setminus \mathcal{R}_\mathcal{X}$. Using the notation $\mathcal{M}_{\vec{c}^\star} = \overset{\leftrightarrow}{\mathsf{M}}{}^\star \oplus \mathsf{m}^\star$, we can derive Eq. (E55) from these facts as $\langle\!\langle \mathcal{F}^{\mathrm{uni}}, \mathcal{F}^{\mathrm{uni}} \rangle\!\rangle_{\mathcal{M}_{\vec{c}^\star}} = \int_V d\mathbf{r}[\vec{F}^{\mathrm{uni}\top} \overset{\leftrightarrow}{\mathsf{M}}{}^\star \vec{F}^{\mathrm{uni}} + \sum_{\rho \in \mathcal{R}} m^\star_\rho (f^{\mathrm{uni}}_\rho)^2] \leqslant \int_V d\mathbf{r}[M^{\max}_\mathcal{X} + \sum_{\rho \in \mathcal{R}_\mathcal{X}} m^\star_\rho] = |\mathcal{M}_{\vec{c}^\star}|^{\mathrm{tot}}_\mathcal{X}$. The force $\mathcal{F}^{\mathrm{uni}}$ also relates to $|\mathcal{J}^\times|_{\mathrm{RD}}$ as

$$|\mathcal{J}^\times|_{\mathrm{RD}} = \langle\!\langle \mathcal{F}^{\mathrm{uni}}, \mathcal{F}^\star \rangle\!\rangle_{\mathcal{M}_{\vec{c}^\star}}, \qquad (E56)$$

which is verified as

$$|\mathcal{J}^\times|_{\mathrm{RD}} = \int_V d\mathbf{r} \left[\sum_{\alpha \in \mathcal{S}} \|\vec{J}^\times_{(\alpha)}\| + \sum_{\rho \in \mathcal{R}} |j^\times_\rho| \right]$$

$$= \int_V d\mathbf{r} \left[\sum_{\alpha \in \mathcal{X}} \|\vec{J}^\star_{(\alpha)}\| + \sum_{\rho \in \mathcal{R}_\mathcal{X}} |j^\star_\rho| \right]$$

$$= \int_V d\mathbf{r} \left[\sum_{\alpha \in \mathcal{X}} \vec{F}^{\mathrm{uni}}_{(\alpha)} \cdot \vec{J}^\star_{(\alpha)} + \sum_{\rho \in \mathcal{R}_\mathcal{X}} f^{\mathrm{uni}}_\rho j^\star_\rho \right]$$

$$= \langle\!\langle \mathcal{F}^{\mathrm{uni}}, \mathcal{J}^\star \rangle\!\rangle = \langle\!\langle \mathcal{F}^{\mathrm{uni}}, \mathcal{F}^\star \rangle\!\rangle_{\mathcal{M}_{\vec{c}^\star}}. \qquad (E57)$$

To obtain the desired inequality (134), we use the Cauchy–Schwarz inequality,

$$\langle\!\langle \mathcal{F}^{\mathrm{uni}}, \mathcal{F}^\star \rangle\!\rangle_{\mathcal{M}_{\vec{c}^\star}} \leqslant \sqrt{\langle\!\langle \mathcal{F}^{\mathrm{uni}}, \mathcal{F}^{\mathrm{uni}} \rangle\!\rangle_{\mathcal{M}_{\vec{c}^\star}} \langle\!\langle \mathcal{F}^\star, \mathcal{F}^\star \rangle\!\rangle_{\mathcal{M}_{\vec{c}^\star}}}$$

$$\leqslant \sqrt{|\mathcal{M}_{\vec{c}^\star}|^{\mathrm{tot}}_\mathcal{X} \langle\!\langle \mathcal{F}^\star, \mathcal{F}^\star \rangle\!\rangle_{\mathcal{M}_{\vec{c}^\star}}}. \qquad (E58)$$

Here, we also use the property of $\mathcal{F}^{\mathrm{uni}}$ in Eq. (E55). Integrating both sides of this inequality (E58), we obtain

$$\int_0^\tau dt \sqrt{|\mathcal{M}_{\vec{c}^\star}|^{\mathrm{tot}}_\mathcal{X} \langle\!\langle \mathcal{F}^\star, \mathcal{F}^\star \rangle\!\rangle_{\mathcal{M}_{\vec{c}^\star}}} \geqslant \int_0^\tau dt \langle\!\langle \mathcal{F}^{\mathrm{uni}}, \mathcal{F}^\star \rangle\!\rangle_{\mathcal{M}_{\vec{c}^\star}}$$

$$= \int_0^\tau dt |\mathcal{J}^\times|_{\mathrm{RD}}$$

$$\geqslant W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau)), \quad (E59)$$

where we use Eq. (E56) in the first transform and Eq. (E54) in the second transform. The Cauchy–Schwarz inequality also provides the following inequality:

$$\left(\int_0^\tau dt \sqrt{|\mathcal{M}_{\vec{c}^\star}|^{\mathrm{tot}}_\mathcal{X} \langle\!\langle \mathcal{F}^\star, \mathcal{F}^\star \rangle\!\rangle_{\mathcal{M}_{\vec{c}^\star}}} \right)^2$$

$$\leqslant \left(\int_0^\tau dt \, |\mathcal{M}_{\vec{c}^\star}|^{\mathrm{tot}}_\mathcal{X} \right) \left(\int_0^\tau dt \langle\!\langle \mathcal{F}^\star, \mathcal{F}^\star \rangle\!\rangle_{\mathcal{M}_{\vec{c}^\star}} \right)$$

$$= \left(\frac{1}{\tau} \int_0^\tau dt\, |\mathcal{M}_{\vec{c}^\star}|_{\mathcal{X}}^{\mathrm{tot}} \right) \left(\tau \int_0^\tau dt\, \langle\!\langle \mathcal{F}^\star, \mathcal{F}^\star \rangle\!\rangle_{\mathcal{M}_{\vec{c}^\star}} \right)$$

$$= \langle |\mathcal{M}_{\vec{c}^\star}|_{\mathcal{X}}^{\mathrm{tot}} \rangle_\tau\, W_{2,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau)|\vec{c}_\mathcal{Y}(t))^2. \qquad (\mathrm{E}60)$$

Here, $\vec{c}_\mathcal{Y}(t)$ does not depend on time. Combining Eqs. (E59) and (E60) leads to the desired inequality (134).

7. Derivation of thermodynamic speed limit based on the 1-Wasserstein distance

Here, we derive the TSLs in Eqs. (145) and (146) from the inequality between the Wasserstein distances in Eq. (134). Substituting $\vec{c}^A = \vec{c}(t)$, $\vec{c}^B = \vec{c}(t + \Delta t)$, and $\vec{b}_\mathcal{Y} = \vec{c}_\mathcal{Y}(t)$ with $\Delta t \ll 1$ into Eq. (134), we obtain

$$\frac{W_{1,\mathcal{X}}(\vec{c}(t), \vec{c}(t + \Delta t))^2}{\frac{1}{\Delta t} \int_t^{t+\Delta t} ds\, |\mathcal{M}|_{\mathcal{X}}^{\mathrm{tot}}} \leqslant W_{2,\mathcal{X}}(\vec{c}(t), \vec{c}(t + \Delta t)|\vec{c}_\mathcal{Y}(t))^2. \qquad (\mathrm{E}61)$$

Expanding the denominator of the left-hand side with respect to Δt, we can rewrite this inequality as

$$\frac{W_{1,\mathcal{X}}(\vec{c}(t), \vec{c}(t + \Delta t))^2}{|\mathcal{M}|_{\mathcal{X}}^{\mathrm{tot}}} + o(\Delta t^2)$$

$$\leqslant W_{2,\mathcal{X}}(\vec{c}(t), \vec{c}(t + \Delta t)|\vec{c}_\mathcal{Y}(t))^2. \qquad (\mathrm{E}62)$$

Then, dividing both sides by Δt^2 and taking the limit $\Delta t \to 0$ yields Eq. (145). Rewriting Eq. (145) as $v_1 \leqslant \sqrt{|\mathcal{M}|_{\mathcal{X}}^{\mathrm{tot}} \sigma^{\mathrm{ex}}}$ and using the Cauchy–Schwarz inequality, we obtain a part of the TSLs in Eq. (146) as

$$l_{1,\tau}^2 = \left[\int_0^\tau dt\, v_1(t) \right]^2$$

$$\leqslant \left[\int_0^\tau dt\, \sqrt{|\mathcal{M}|_{\mathcal{X}}^{\mathrm{tot}} \sigma^{\mathrm{ex}}} \right]^2$$

$$\leqslant \left(\int_0^\tau dt\, |\mathcal{M}|_{\mathcal{X}}^{\mathrm{tot}} \right) \left(\int_0^\tau dt\, \sigma^{\mathrm{ex}} \right)$$

$$= \langle |\mathcal{M}|_{\mathcal{X}}^{\mathrm{tot}} \rangle_\tau\, \tau \Sigma_\tau^{\mathrm{ex}}. \qquad (\mathrm{E}63)$$

We also obtain $l_{1,\tau} \geqslant W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau))$ as a direct consequence of the triangle inequality for the 1-Wasserstein distance. Unifying these results, we reach the TSLs in Eq. (146).

Although it is not obvious which is tighter, $l_{2,\tau}^2$ or $l_{1,\tau}^2 / \langle |\mathcal{M}|_{\mathcal{X}}^{\mathrm{tot}} \rangle_\tau$ in the TSLs, we can obtain the following inequality directly from Eq. (145),

$$l_{2,\tau}^2 \geqslant \frac{l_{1,\tau}^2}{\max_{t \in [0,\tau]} |\mathcal{M}|_{\mathcal{X}}^{\mathrm{tot}}}. \qquad (\mathrm{E}64)$$

8. Details of the minimum dissipation formula with the 1-Wasserstein distance (157)

a. Derivation of the minimum dissipation formula with the 1-Wasserstein distance (157)

Here, we prove the minimum dissipation formula with the 1-Wasserstein distance (157).

First, we verify that the right-hand side in Eq. (157) provides a lower bound of the EP under the conditions in

Eqs. (155) and (156) as

$$\Sigma_\tau[\mathcal{M}', \mathcal{F}'] \geqslant \frac{W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau))^2}{\tau \langle |\mathcal{M}'|_{\mathcal{X}}^{\mathrm{tot}} \rangle_\tau} \geqslant \frac{W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau))^2}{\tau M_0}. \qquad (\mathrm{E}65)$$

Here, the first inequality is obtained in the same way as in the case of Eq. (134). It is enough to use $(\mathcal{M}', \mathcal{F}')$ instead of $(\mathcal{M}_{\vec{c}^\star}, \mathcal{F}^\star)$ in Appendix E 6. The second inequality is due to the inequality (156).

Second, we construct an optimizer $(\mathcal{M}^\diamond, \mathcal{F}^\diamond)$, which achieves

$$\Sigma_\tau[\mathcal{M}^\diamond, \mathcal{F}^\diamond] = \frac{W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau))^2}{\tau M_0}. \qquad (\mathrm{E}66)$$

Let $\mathcal{U}^\diamond = (\vec{U}^\diamond, u^\diamond)$ be the optimizer of the right-hand side in Eq. (118), which satisfies $W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau)) = |\mathcal{U}^\diamond|_{\mathrm{RD}}$ and $\vec{c}_\mathcal{X}(\tau) - \vec{c}_\mathcal{X}(0) = (\nabla^\dagger \mathcal{U}^\diamond)_\mathcal{X}$. Using $\mathcal{U}^\diamond$, we can obtain the optimal Onsager operator $\mathcal{M}^\diamond = \overset{\leftrightarrow}{\mathsf{M}}^\diamond \oplus \mathsf{m}^\diamond$ as

$$\mathsf{M}_{(\alpha\beta)}^\diamond := \frac{M_0 \|\vec{U}_{(\alpha)}^\diamond\|}{|\mathcal{U}^\diamond|_{\mathrm{RD}}} \delta_{\alpha\beta} \mathsf{I}, \quad m_\rho^\diamond := \frac{M_0 |u_\rho^\diamond|}{|\mathcal{U}^\diamond|_{\mathrm{RD}}}, \qquad (\mathrm{E}67)$$

and the optimal force $\mathcal{F}^\diamond = (\vec{F}^\diamond, f^\diamond)$ as

$$\vec{F}_{(\alpha)}^\diamond := \begin{cases} \dfrac{|\mathcal{U}^\diamond|_{\mathrm{RD}}}{\tau M_0} \dfrac{\vec{U}_{(\alpha)}^\diamond}{\|\vec{U}_{(\alpha)}^\diamond\|} & (\vec{U}_{(\alpha)}^\diamond \neq \mathbf{0}) \\ \mathbf{0} & (\vec{U}_{(\alpha)}^\diamond = \mathbf{0}) \end{cases}, \qquad (\mathrm{E}68)$$

$$f_\rho^\diamond := \begin{cases} \dfrac{|\mathcal{U}^\diamond|_{\mathrm{RD}}}{\tau M_0} \dfrac{u_\rho^\diamond}{|u_\rho^\diamond|} & (u_\rho^\diamond \neq 0) \\ 0 & (u_\rho^\diamond = 0) \end{cases}. \qquad (\mathrm{E}69)$$

The optimizer $(\mathcal{M}^\diamond, \mathcal{F}^\diamond)$ satisfies the constraint (155) because $\mathcal{M}^\diamond \mathcal{F}^\diamond = \mathcal{U}^\diamond/\tau$ holds. We can also verify the remaining condition (156) and the equality (E66) by direct calculation.

b. Operational interpretation of the optimizer of the minimization problem in Eq. (157)

The optimization problem for $\mathcal{M}'$ and $\mathcal{F}'$ in Eq. (157) can be reformulated in terms of of control parameters. This allows us to interpret the optimizer of the minimization problem in Eq. (157) that is constructed in the previous section, $(\mathcal{M}^\diamond, \mathcal{F}^\diamond)$, from an operational viewpoint. Here, we provide such an operational interpretation of the optimizer $(\mathcal{M}^\diamond, \mathcal{F}^\diamond)$.

We consider the ideal dilute solution with the two assumptions: the mobility tensor has the simple form (53), and the reactions obey the mass action law. Under these assumptions, the time evolution of the concentration distribution of the internal species $\alpha \in \mathcal{X}$ is given by

$$\partial_t c_\alpha = \nabla_{\boldsymbol{r}} \cdot \left\{ D_\alpha c_\alpha \left(\nabla_{\boldsymbol{r}} \ln c_\alpha - \boldsymbol{F}_{(\alpha)}^{\mathrm{mech}} \right) \right\}$$

$$+ \sum_{\rho \in \mathcal{R}_\mathcal{X}} S_{\alpha\rho} \left(\kappa_\rho^+ \prod_{\alpha \in \mathcal{S}} c_\alpha^{\nu_{\alpha\rho}^+} - \kappa_\rho^- \prod_{\alpha \in \mathcal{S}} c_\alpha^{\nu_{\alpha\rho}^-} \right), \qquad (\mathrm{E}70)$$

where we let $\boldsymbol{F}_{(\alpha)}^{\mathrm{mech}}$ denote the force acting on the αth species.

Let us regard the diffusion coefficient D_α, the force $\boldsymbol{F}_{(\alpha)}^{\mathrm{mech}}$, and the reaction rate constant $\kappa_\rho^\pm$ in Eq. (E70) as the controllable parameters. Then, using the concentration distribution

$\vec{c}^{\diamond}$ introduced in Sec. V E, we can obtain the optimal diffusion coefficient and the optimal force as

$$D_{\alpha}^{\diamond}(t) := \frac{M_0 \|\boldsymbol{U}_{(\alpha)}^{\diamond}\|}{|\mathcal{U}^{\diamond}|_{\mathrm{RD}} c_{\alpha}^{\diamond}(t)}, \qquad \text{(E71)}$$

and

$$\boldsymbol{F}_{(\alpha)}^{\mathrm{mech}\diamond}(t) := \boldsymbol{F}_{(\alpha)}^{\diamond} + \boldsymbol{\nabla}_{\boldsymbol{r}} \ln c_{\alpha}^{\diamond}(t), \qquad \text{(E72)}$$

for all $\alpha \in \mathcal{S}$. We can also construct the optimal reaction rate constant as

$$\kappa_{\rho}^{+\diamond}(t) := \frac{1}{\prod_{\alpha \in \mathcal{S}} \{c_{\alpha}^{\diamond}(t)\}^{\nu_{\alpha\rho}^+}} \frac{e^{f_{\rho}^{\diamond}}}{e^{f_{\rho}^{\diamond}} - 1} \frac{u_{\rho}^{\diamond}}{\tau}$$

$$\kappa_{\rho}^{-\diamond}(t) := \frac{1}{\prod_{\alpha \in \mathcal{S}} \{c_{\alpha}^{\diamond}(t)\}^{\nu_{\alpha\rho}^-}} \frac{1}{e^{f_{\rho}^{\diamond}} - 1} \frac{u_{\rho}^{\diamond}}{\tau}, \qquad \text{(E73)}$$

for all $\rho \in \mathcal{R}$. We can easily verify that these optimal parameters reproduce $(\mathcal{M}^{\diamond}, \mathcal{F}^{\diamond})$ as

$$D_{\alpha}^{\diamond} c_{\alpha}^{\diamond} \delta_{\alpha\beta} \mathsf{I} = \frac{M_0 \|\boldsymbol{U}_{(\alpha)}^{\diamond}\|}{|\mathcal{U}^{\diamond}|_{\mathrm{RD}}} \delta_{\alpha\beta} \mathsf{I} = \mathsf{M}_{(\alpha\beta)}^{\diamond}, \qquad \text{(E74)}$$

$$\frac{\kappa_{\rho}^{+\diamond} \prod_{\alpha \in \mathcal{S}} c_{\alpha}^{\diamond \nu_{\alpha\rho}^+} - \kappa_{\rho}^{-\diamond} \prod_{\alpha \in \mathcal{S}} c_{\alpha}^{\diamond \nu_{\alpha\rho}^-}}{\ln \kappa_{\rho}^{+\diamond} \prod_{\alpha \in \mathcal{S}} c_{\alpha}^{\diamond \nu_{\alpha\rho}^+} - \ln \kappa_{\rho}^{-\diamond} \prod_{\alpha \in \mathcal{S}} c_{\alpha}^{\diamond \nu_{\alpha\rho}^-}} = \frac{M_0 |u_{\rho}^{\diamond}|}{|\mathcal{U}^{\diamond}|_{\mathrm{RD}}} = m_{\rho}^{\diamond},$$

$$\text{(E75)}$$

$$\boldsymbol{F}_{(\alpha)}^{\mathrm{mech}\diamond}(t) - \boldsymbol{\nabla}_{\boldsymbol{r}} \ln c_{\alpha}^{\diamond}(t) = \boldsymbol{F}_{(\alpha)}^{\diamond}, \qquad \text{(E76)}$$

and

$$\ln \frac{\kappa_{\rho}^{+\diamond} \prod_{\alpha \in \mathcal{S}} c_{\alpha}^{\diamond \nu_{\alpha\rho}^+}}{\kappa_{\rho}^{-\diamond} \prod_{\alpha \in \mathcal{S}} c_{\alpha}^{\diamond \nu_{\alpha\rho}^-}} = f_{\rho}^{\diamond}. \qquad \text{(E77)}$$

Note that the dependence of the optimal parameters on time is completely determined by $\vec{c}^{\diamond}(t)$ because the optimizer $(\mathcal{M}^{\diamond}, \mathcal{F}^{\diamond})$ is independent of time.

c. Conservative force as an optimizer of the minimization problem in Eq. (157)

We remark that we can take a conservative force as an optimal force of the minimization problem (157) by using $\mathcal{U}^{\bullet}$ introduced in Appendix E 2 instead of $\mathcal{U}^{\diamond}$. In this case, a new optimal Onsager operator $\mathcal{M}^{\bullet} = \overset{\leftrightarrow}{\mathsf{M}}^{\bullet} \oplus \mathsf{m}^{\bullet}$ is given by

$$\mathsf{M}_{(\alpha\beta)}^{\bullet} := \frac{M_0 \|\boldsymbol{U}_{(\alpha)}^{\bullet}\|}{|\mathcal{U}^{\bullet}|_{\mathrm{RD}}} \delta_{\alpha\beta} \mathsf{I}, \quad m_{\rho}^{\bullet} := \frac{M_0 |u_{\rho}^{\bullet}|}{|\mathcal{U}^{\bullet}|_{\mathrm{RD}}}, \qquad \text{(E78)}$$

and we can take the conservative force

$$\mathcal{F}^{\bullet} := \nabla \left(\frac{|\mathcal{U}^{\bullet}|_{\mathrm{RD}}}{\tau M_0} \vec{\psi}^{\bullet} \right), \qquad \text{(E79)}$$

as a new optimal force. They satisfy $\mathcal{M}^{\bullet} \mathcal{F}^{\bullet} = \mathcal{U}^{\bullet}/\tau$ due to the complementary slackness condition (E21) and the definition of $\mathcal{U}^{\bullet}$ (E23). To derive $\mathcal{M}^{\bullet} \mathcal{F}^{\bullet} = \mathcal{U}^{\bullet}/\tau$, we used $\|\boldsymbol{U}_{(\alpha)}^{\bullet}\| = U_{\alpha}^{\mathrm{L}\bullet}$ for $\alpha \in \mathcal{X}$ and $\|u_{\rho}^{\bullet}\| = u_{\rho}^{\mathrm{L}\bullet}$ for $\rho \in \mathcal{R}_{\mathcal{X}}$. Since the time-integrated current $\mathcal{U}^{\bullet}$ is an optimizer of the 1-Wasserstein distance as remarked in Appendix E 2, the Onsager operator $\mathcal{M}^{\bullet}$ and the force $\mathcal{F}^{\bullet}$ achieve the equality

$$\Sigma_{\tau}[\mathcal{M}^{\bullet}, \mathcal{F}^{\bullet}] = \frac{W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau))^2}{\tau M_0}, \qquad \text{(E80)}$$

and satisfy the constraints in Eqs. (155) and (156).

9. Embedding time series of concentration distributions into Euclidean space by multidimensional scaling

We introduce the procedure of the multidimensional scaling [113]. This method is used to visualize the similarity of a data set in a low-dimensional Euclidean space. Let $[0, \tau]$ be the time interval of the reaction-diffusion dynamics. We used the multidimensional scaling for $n + 1$ samples of the concentration $\vec{c}(t_a)$ where $t_a = (a/n)\tau$ and $a \in \{0, 1, \ldots, n\}$. We aim to find the points $\{\boldsymbol{r}^{(0)}, \boldsymbol{r}^{(1)}, \ldots, \boldsymbol{r}^{(n)}\}$ in the d_0-dimensional Euclidean space such that $\|\boldsymbol{r}^{(a)} - \boldsymbol{r}^{(b)}\| \simeq W_{1,\mathcal{X}}(\vec{c}(t_a), \vec{c}(t_b))$ is approximately satisfied for all $(a, b) \in \{0, 1, \ldots, n\}^2$. In the numerical examples, we use $(d_0, n) = (2, 40)$ for the Fisher–KPP equation and $(d_0, n) = (2, 200)$ for the Brusselator model. The detailed procedure is as follows.

We define a $(n + 1) \times (n + 1)$ matrix W as

$$(\mathsf{W})_{ab} := W_{1,\mathcal{X}}(\vec{c}(t_a), \vec{c}(t_b))^2, \qquad \text{(E81)}$$

for $(a, b) \in \{0, 1, \ldots, n\}^2$. In general, when determining coordinates solely from the distance structure between each point, there is an ambiguity in the choice of the origin. To determine the origin, we apply the double centering to W, i.e., we transform W with a matrix $(\mathsf{C})_{ab} := \delta_{ab} - 1/(n + 1)$ as

$$\mathsf{W}_{\mathrm{C}} := -\frac{1}{2} \mathsf{CWC}. \qquad \text{(E82)}$$

We note that W_{C} satisfies the following two properties: (i) $\sum_a (\mathsf{W}_{\mathrm{C}})_{ab} = 0$ and $\sum_b (\mathsf{W}_{\mathrm{C}})_{ab} = 0$, and (ii) the square of the distance is recovered as

$$(\mathsf{W}_{\mathrm{C}})_{aa} + (\mathsf{W}_{\mathrm{C}})_{bb} - 2(\mathsf{W}_{\mathrm{C}})_{ab} = (\mathsf{W})_{ab}$$

$$= W_{1,\mathcal{X}}(\vec{c}(t_a), \vec{c}(t_b))^2. \quad \text{(E83)}$$

Here, the property (i) corresponds to the center of gravity being the origin, and the property (ii) corresponds to considering each component $(\mathsf{W}_{\mathrm{C}})_{ab}$ as an inner product between vectors from the origin to each point. We can verify these properties using

$$(\mathsf{W}_{\mathrm{C}})_{ab} = -\frac{1}{2} \left[(\mathsf{W})_{ab} + \frac{\sum_{a',b'} (\mathsf{W})_{a'b'}}{(n + 1)^2} \right]$$

$$+ \frac{\sum_{a'} (\mathsf{W})_{a'b} + \sum_{b'} (\mathsf{W})_{ab'}}{2(n + 1)}, \qquad \text{(E84)}$$

$(\mathsf{W})_{aa} = 0$, and $(\mathsf{W})_{ab} = (\mathsf{W})_{ba}$.

Since W_{C} is symmetric, it is diagonalizable with an orthogonal matrix Q as

$$\mathsf{W}_{\mathrm{C}} = \mathsf{Q}\Lambda\mathsf{Q}^{\top}, \qquad \text{(E85)}$$

with $\Lambda := \mathrm{diag}(\lambda_0, \lambda_1, \ldots, \lambda_n)$. Here, we assume that the eigenvalues of W_{C} are in descending order, i.e., $\lambda_0 \geqslant \lambda_1 \geqslant \cdots \geqslant \lambda_n$. We note that some eigenvalues can be negative because the 1-Wasserstein distance is non-Euclidean. We can obtain $\boldsymbol{r}^{(a)}$ if λ_{d_0-1} is nonnegative as

$$(\boldsymbol{r}^{(a)})_i = (\mathsf{Q})_{ai} \sqrt{\lambda_i}, \qquad \text{(E86)}$$

for $0 \leqslant i \leqslant d_0 - 1$. Note that we cannot embed the time series in the d_0 dimensional Euclidean space if λ_{d_0-1} is negative.

We also discuss the accuracy of the multidimensional scaling. To do so, we define the estimation error of the distance between $\vec{c}(t_a)$ and $\vec{c}(t_b)$ as

$$\epsilon_{ab} := |W_{1,\mathcal{X}}(\vec{c}(t_a), \vec{c}(t_b))^2 - \|\boldsymbol{r}^{(a)} - \boldsymbol{r}^{(b)}\|^2|. \qquad \text{(E87)}$$

The definition of $\boldsymbol{r}^{(a)}$ (E86) yields

$$\|\boldsymbol{r}^{(a)} - \boldsymbol{r}^{(b)}\|^2$$

$$= \sum_{i=0}^{d_0-1} \lambda_i (\mathsf{Q})_{ai}^2 + \sum_{i=0}^{d_0-1} \lambda_i (\mathsf{Q})_{bi}^2 - 2 \sum_{i=0}^{d_0-1} \lambda_i (\mathsf{Q})_{ai} (\mathsf{Q})_{bi}$$

$$= (\mathsf{W}_\mathrm{C})_{aa} + (\mathsf{W}_\mathrm{C})_{bb} - 2(\mathsf{W}_\mathrm{C})_{ab} - \sum_{i=d_0}^{n} \lambda_i ((\mathsf{Q})_{ai} - (\mathsf{Q})_{bi})^2$$

$$= W_{1,\mathcal{X}}(\vec{c}(t_a), \vec{c}(t_b))^2 - \sum_{i=d_0}^{n} \lambda_i ((\mathsf{Q})_{ai} - (\mathsf{Q})_{bi})^2 \quad \text{(E88)}$$

where we use the element of Eq. (E85), i.e., $(\mathsf{W}_\mathrm{C})_{ab} = \sum_{i=0}^{n} \lambda_i (\mathsf{Q})_{ai} (\mathsf{Q})_{bi}$, in the second line. We also use Eq. (E83) in the last transform. Thus, we obtain

$$\epsilon_{ab} = \left| \sum_{i=d_0}^{n} \lambda_i ((\mathsf{Q})_{ai} - (\mathsf{Q})_{bi})^2 \right|. \quad \text{(E89)}$$

We can derive the upper bound of ϵ_{ab} for $a \neq b$ as

$$\epsilon_{ab} \leqslant \left(\max_{d_0 \leqslant i \leqslant n} |\lambda_i| \right) \sum_{i=d_0}^{n} ((\mathsf{Q})_{ai} - (\mathsf{Q})_{bi})^2$$

$$\leqslant \left(\max_{d_0 \leqslant i \leqslant n} |\lambda_i| \right) \sum_{i=0}^{n} ((\mathsf{Q})_{ai} - (\mathsf{Q})_{bi})^2$$

$$= 2 \max_{d_0 \leqslant i \leqslant n} |\lambda_i|, \quad \text{(E90)}$$

where we use the orthogonality of Q, namely, $\sum_{i=0}^{n} (\mathsf{Q})_{ai} (\mathsf{Q})_{bi} = \delta_{ab}$, in the last transform. This upper bound implies that the eigenvalues truncated in the embedding determine the error. The error becomes small if the following two conditions are satisfied: (i) the dimension d_0 is sufficiently large to express the trajectory that we embed, and (ii) the absolute values of the negative eigenvalues are enough small, i.e., the effect of non-Euclideanity is enough small.

10. The property of the 1-Wasserstein distance for the Fisher–KPP equation

a. The equivalence of the lengths in the Fisher–KPP equation and other simple reaction-diffusion systems

In this section, we consider an RDS in $V \subset \mathbb{R}^d$, which satisfies $\mathcal{X} = \{1\}$ and $\mathcal{R}_\mathcal{X} = \{1\}$. The dynamics of the concentration distribution of the internal species Z_1 is given by a simple RD equation,

$$\partial_t c_1(\boldsymbol{r}; t) = -\nabla_{\boldsymbol{r}} \cdot \boldsymbol{J}_{(1)} + j_1. \quad \text{(E91)}$$

The way to interact with the outside of the system and the boundary conditions can be whatever is appropriate. The Fisher–KPP equation is an example of such a system.

Here, we prove that each monotonic increase,

$$\partial_t c_1(\boldsymbol{r}; t) \geqslant 0 \quad \text{for all } \boldsymbol{r} \in V \text{ and } t \in [0, \tau], \quad \text{(E92)}$$

or decrease,

$$\partial_t c_1(\boldsymbol{r}; t) \leqslant 0 \quad \text{for all } \boldsymbol{r} \in V \text{ and } t \in [0, \tau], \quad \text{(E93)}$$

of the concentration leads to the equivalence of the lengths,

$$l_{1,\tau} = W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau))$$

$$= L_\mathcal{X}(\vec{c}(0), \vec{c}(\tau)) = L_\mathcal{X}^{\mathrm{tot}}(\vec{c}(0), \vec{c}(\tau)). \quad \text{(E94)}$$

We first verify the hierarchy of lengths,

$$l_{1,\tau} \geqslant W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau))$$

$$\geqslant L_\mathcal{X}(\vec{c}(0), \vec{c}(\tau)) \geqslant L_\mathcal{X}^{\mathrm{tot}}(\vec{c}(0), \vec{c}(\tau)), \quad \text{(E95)}$$

to obtain the equivalence (E94). It is enough to show the second inequality in Eq. (E95) since we can easily derive the first and third inequalities from the triangle inequality. To obtain the desired inequality, we take a potential $\phi_1'(\boldsymbol{r})$ as $\phi_1'(\boldsymbol{r}) = 1$ if the concentration is monotonically increasing (E92) or $\phi_1'(\boldsymbol{r}) = -1$ if the concentration is monotonically decreasing (E93). This potential $(\phi_1'(\boldsymbol{r}), \vec{0}_\mathcal{Y}^\top)^\top =: \vec{\phi}'$ belongs to the set $\mathrm{Lip}_\mathcal{X}^1$ (121) because $\|\nabla_{\boldsymbol{r}} \phi_1'\| \leqslant 1$ and $|\phi_1'| \leqslant 1$ hold. Thus, the potential $\vec{\phi}'$ is a candidate for the Kantorovich–Rubinstein duality (120), which leads

$$W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau)) = \sup_{\vec{\phi} \in \mathrm{Lip}_\mathcal{X}^1} \langle \vec{\phi}, \vec{c}(\tau) - \vec{c}(0) \rangle$$

$$\geqslant \langle \vec{\phi}', \vec{c}(\tau) - \vec{c}(0) \rangle$$

$$= \int_V d\boldsymbol{r} \, \phi_1'(\boldsymbol{r}) \{ c_1(\boldsymbol{r}; \tau) - c_1(\boldsymbol{r}; 0) \}$$

$$= \int_V d\boldsymbol{r} \, |c_1(\boldsymbol{r}; \tau) - c_1(\boldsymbol{r}; 0)|$$

$$= L_\mathcal{X}(\vec{c}(0), \vec{c}(\tau)) \quad \text{(E96)}$$

Second, we show that the maximum length in the hierarchy equals the minimum one as

$$l_{1,\tau} = L_\mathcal{X}^{\mathrm{tot}}(\vec{c}(0), \vec{c}(\tau)). \quad \text{(E97)}$$

The minimization problem for the 1-Wasserstein distance (118) provides the variational form of $v_1(t)$,

$$v_1(t) = \inf_{\boldsymbol{J}_{(1)}', j_1'} \int_V d\boldsymbol{r} \, \{ \|\boldsymbol{J}_{(1)}'\| + |j_1'| \}, \quad \text{(E98)}$$

with the following condition: the currents $(\boldsymbol{J}_{(1)}', j_1')$ satisfies

$$\partial_t c(\boldsymbol{r}; t) = -\nabla_{\boldsymbol{r}} \cdot \boldsymbol{J}_{(1)}'(\boldsymbol{r}; t) + j_1'(\boldsymbol{r}; t). \quad \text{(E99)}$$

We can obtain the lower bound of v_1 as

$$v_1(t) \geqslant \int_V d\boldsymbol{r} \, |\partial_t c_1(\boldsymbol{r}; t)|, \quad \text{(E100)}$$

by replacing $(0, \tau)$ with $(t, t + \Delta t)$ in Eq. (E96), dividing this equation by Δt, and taking the limit $\Delta t \to 0$. This lower bound is achievable by taking $\boldsymbol{J}_{(1)}' = \boldsymbol{0}$ and $j_1' = \partial_t c_1$ in Eq. (E98). Here, we can easily verify that $\boldsymbol{J}_{(1)}' = \boldsymbol{0}$ and $j_1' = \partial_t c_1$ satisfy the condition (E99). Thus, we obtain

$$v_1(t) = \int_V d\boldsymbol{r} \, |\partial_t c_1(\boldsymbol{r}; t)|. \quad \text{(E101)}$$

Integrating Eq. (E101) leads to the desired equality (E97) as

$$
\begin{aligned}
l_{1,\tau} &= \int_0^\tau dt\, v_1(t) \\
&= \int_0^\tau dt \int_V d\boldsymbol{r}\, |\partial_t c_1(\boldsymbol{r};t)| \\
&= \left| \int_0^\tau dt \int_V d\boldsymbol{r}\, \partial_t c_1(\boldsymbol{r};t) \right| \\
&= \left| \int_V d\boldsymbol{r}\, \{ c_1(\boldsymbol{r};\tau) - c_1(\boldsymbol{r};0) \} \right| \\
&= L_{\mathcal{X}}^{\text{tot}}(\vec{c}(0), \vec{c}(\tau))
\end{aligned}
\tag{E102}
$$

where we used the fact that the sign of $\partial_t c_1$ is invariant to obtain the third line. Combining Eqs. (E95) and (E97) yields the desired result in Eq. (E94).

b. The 1-Wasserstein distance for the traveling wave solution in the Fisher–KPP equation

In this section, we focus on the Fisher–KPP equation in the one-dimensional Euclidean space $\mathbb{R}$. Letting the concentration of the external species $c_2(\boldsymbol{r};t) = c_2$ be homogeneous, the time evolution of c_1 is given by

$$
\partial_t c_1 = D_1 \partial_r^2 c_1 + (\kappa_1^+ c_2) c_1 \left(1 - \frac{c_1}{c_1^{\text{eq}}} \right),
\tag{E103}
$$

where c_1^{eq} is the equilibrium concentration given by $c_1^{\text{eq}} = \kappa_1^+ c_2 / \kappa_1^-$. For each wave speed $v_{\text{wave}} \geqslant 2\sqrt{D_1 \kappa_1^+ c_2}$, the equation admits traveling wave solutions [107], whose form is given by

$$
c_1(r;t) = c_{\text{wave}}(r - v_{\text{wave}} t) = c_{\text{wave}}(x).
\tag{E104}
$$

The function $c_{\text{wave}}(x)$ is a monotonically decreasing function of x, and thus

$$
\partial_x c_{\text{wave}}(x) \leqslant 0,
\tag{E105}
$$

holds for all $x \in \mathbb{R}$. The function $c_{\text{wave}}(x)$ also satisfies the boundary conditions

$$
\lim_{x \to -\infty} c_{\text{wave}}(x) = c_1^{\text{eq}}, \quad \lim_{x \to \infty} c_{\text{wave}}(x) = 0.
\tag{E106}
$$

For these traveling wave solutions $c_1(r;t) = c_{\text{wave}}(r - v_{\text{wave}} t)$, we can obtain the closed form of the lengths,

$$
\begin{aligned}
l_{1,\tau} &= W_{1,\mathcal{X}}(\vec{c}(0), \vec{c}(\tau)) \\
&= L_{\mathcal{X}}(\vec{c}(0), \vec{c}(\tau)) = L_{\mathcal{X}}^{\text{tot}}(\vec{c}(0), \vec{c}(\tau)) \\
&= c_1^{\text{eq}} v_{\text{wave}} \tau.
\end{aligned}
\tag{E107}
$$

We can verify it as follows. Combining Eqs. (E104) and (E105) leads to

$$
\partial_t c_1(r;t) = -v_{\text{wave}} \partial_x c_{\text{wave}}(x)|_{x = r - v_{\text{wave}} t} \geqslant 0,
\tag{E108}
$$

for all $r \in \mathbb{R}$ and $t \in [0, \tau]$, which lets the traveling wave solutions satisfy the monotonicity (E92). Thus, the result of the previous section (E94) confirms the first, second, and third

equalities in Eq. (E107). We can also verify the remaining part,

$$
l_{1,\tau} = c_1^{\text{eq}} v_{\text{wave}} \tau,
\tag{E109}
$$

using Eq. (E101) as

$$
\begin{aligned}
l_{1,\tau} &= \int_0^\tau dt \int_{\mathbb{R}} dr\, |\partial_t c_1(\boldsymbol{r};t)| \\
&= -v_{\text{wave}} \int_0^\tau dt \int_{-\infty}^{\infty} dx\, \partial_x c_{\text{wave}}(x) \\
&= c_1^{\text{eq}} v_{\text{wave}} \tau,
\end{aligned}
\tag{E110}
$$

where we used Eq. (E106) in the space integration.

APPENDIX F: THE DERIVATION OF THE TRADE-OFF RELATION IN EQ. (178)

We provide the derivation of Eq. (178). The triangle inequality for the time integration leads to

$$
\begin{aligned}
|\tilde{c}_\alpha(\boldsymbol{k};\tau) - \tilde{c}_\alpha(\boldsymbol{k};0)| &= \left| \int_0^\tau dt [d_t \tilde{c}_\alpha(\boldsymbol{k};t)] \right| \\
&\leqslant \int_0^\tau dt |d_t \tilde{c}_\alpha(\boldsymbol{k};t)|.
\end{aligned}
\tag{F1}
$$

The right-hand side of Eq. (F1) has an upper bound from the TUR (175) as

$$
\int_0^\tau dt |d_t \tilde{c}_\alpha(\boldsymbol{k};t)| \leqslant \int_0^\tau dt \sqrt{\sigma^{\text{ex}}} \sqrt{\boldsymbol{k} \cdot \mathsf{M}_{(\alpha\alpha)}^{\text{tot}}(t) \boldsymbol{k} + \check{D}_{\alpha\alpha}^{\text{tot}}(t)}.
\tag{F2}
$$

The Cauchy–Schwarz inequality provides an upper bound for the right-hand side of Eq. (F2) as

$$
\begin{aligned}
&\int_0^\tau dt \sqrt{\sigma^{\text{ex}}} \sqrt{\boldsymbol{k} \cdot \mathsf{M}_{(\alpha\alpha)}^{\text{tot}}(t) \boldsymbol{k} + \check{D}_{\alpha\alpha}^{\text{tot}}(t)} \\
&\leqslant \sqrt{\int_0^\tau dt\, \sigma^{\text{ex}}} \sqrt{\int_0^\tau dt \left[\boldsymbol{k} \cdot \mathsf{M}_{(\alpha\alpha)}^{\text{tot}}(t) \boldsymbol{k} + \check{D}_{\alpha\alpha}^{\text{tot}}(t) \right]} \\
&= \sqrt{\tau \Sigma_\tau^{\text{ex}} \left[\boldsymbol{k} \cdot \langle \mathsf{M}_{(\alpha\alpha)}^{\text{tot}} \rangle_\tau \boldsymbol{k} + \langle \check{D}_{\alpha\alpha}^{\text{tot}} \rangle_\tau \right]}.
\end{aligned}
\tag{F3}
$$

We obtain Eq. (178) by unifying Eqs. (F1)–(F3).

APPENDIX G: COMPLEMENTARY NUMERICAL DEMONSTRATION OF THE THERMODYNAMIC UNCERTAINTY RELATION FOR $\tilde{c}_1(\boldsymbol{k};t)$ IN THE BRUSSELATOR

In Fig. 11, we show the TURs $\sigma_1^{\text{TUR}}(n;t)$ for various n and scatter plots of $n_{\max}^{(1)}(t)$, corresponding to the chemical species Z_1 ($\alpha = 1$). Compared to the case of $\alpha = 2$, the mode $n_{\max}^{(1)}$ tends to be larger than $n_{\max}^{(2)}$. This is probably because the pattern of c_1 has more extreme peaks than the pattern of c_2, resulting in a large time variation of the mode corresponding to a large n.
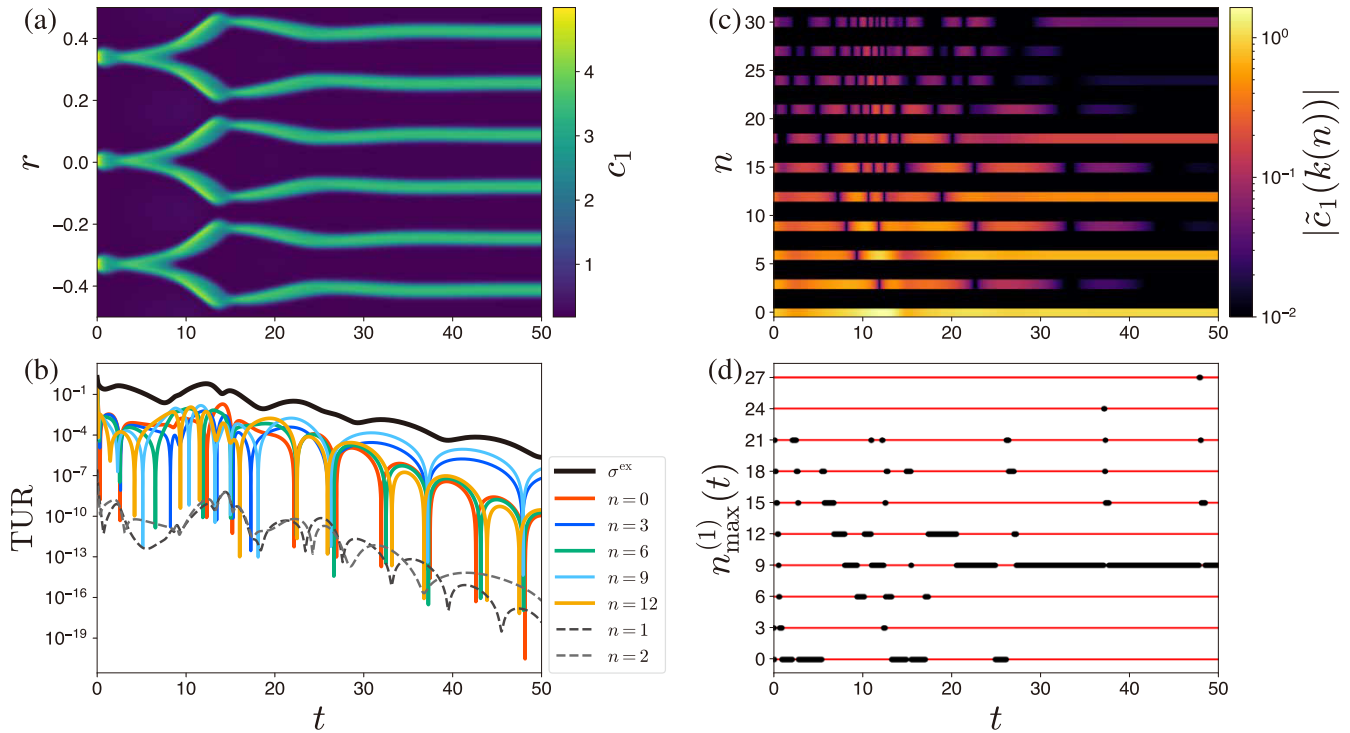
FIG. 11. (a) The time series of c_1. The symmetry of the pattern changes from threefold to sixfold. (b) The excess EPR (black line) and its lower bounds $\sigma_1^{\text{TUR}}(n;t)$ for various n. (c) The time series of $|\tilde{c}_1(k(n))|$. We omit $|\tilde{c}_1(k(n))|$ for $n \geqslant 32$. Since the symmetry of the pattern goes from threefold to sixfold, $|\tilde{c}_1(k(n))|$ decays if n is a multiple of three but not a multiple of six. (d) $n_{\max}^{(1)}(t)$ (black dots). Reflecting the symmetry of the pattern, $n_{\max}^{(1)}(t)$ is multiples of three (red lines) for all time t. Near the stationary pattern ($t > 20$), $n_{\max}^{(1)}(t)$ is nine for almost all t.

[1] J. E. Pearson, Complex patterns in a simple system, Science **261**, 189 (1993).

[2] B. P. Belousov, A periodic reaction and its mechanism, Ref. Radiats. Med. (1958).

[3] K. C. Huang, Y. Meir, and N. S. Wingreen, Dynamic structures in *Escherichia coli*: Spontaneous formation of mine rings and mind polar zones, Proc. Natl. Acad. Sci. USA **100**, 12724 (2003).

[4] S. Kondo and T. Miura, Reaction-diffusion model as a framework for understanding biological pattern formation, Science **329**, 1616 (2010).

[5] J. D. Murray, *Mathematical Biology: II: Spatial Models and Biomedical Applications* (Springer, New York, 2003).

[6] R. S. Cantrell and C. Cosner, *Spatial Ecology via Reaction-Diffusion Equations* (John Wiley & Sons, Hoboken, NJ, 2004).

[7] R. A. Fisher, The wave of advance of advantageous genes, Ann. Eugenics **7**, 355 (1937).

[8] A. M. Turing, The chemical basis of morphogenesis, Philos. Trans. R. Soc. London, Ser. B **237**, 37 (1952).

[9] L. Wolpert, Positional information and the spatial pattern of cellular differentiation, J. Theor. Biol. **25**, 1 (1969).

[10] S. F. Banani, H. O. Lee, A. A. Hyman, and M. K. Rosen, Biomolecular condensates: Organizers of cellular biochemistry, Nat. Rev. Mol. Cell Biol. **18**, 285 (2017).

[11] M. Castellana, M. Z. Wilson, Y. Xu, P. Joshi, I. M. Cristea, J. D. Rabinowitz, Z. Gitai, and N. S. Wingreen, Enzyme clustering accelerates processing of intermediates through metabolic channeling, Nat. Biotechnol. **32**, 1011 (2014).

[12] R. Yoshida, Self-oscillating gels driven by the Belousov–Zhabotinsky reaction as novel smart materials, Adv. Mater. **22**, 3463 (2010).

[13] A. Adamatzky, B. D. L. Costello, and T. Asai, *Reaction-Diffusion Computers* (Elsevier, Amsterdam, 2005).

[14] J. Gorecki, K. Gizynski, J. Guzowski, J. N. Gorecka, P. Garstecki, G. Gruenert, and P. Dittrich, Chemical computing with reaction–diffusion processes, Philos. Trans. R. Soc. A **373**, 20140219 (2015).

[15] J. M. Parrilla-Gutierrez, A. Sharma, S. Tsuda, G. J. Cooper, G. Aragon-Camarasa, K. Donkers, and L. Cronin, A programmable chemical computer with memory and pattern recognition, Nat. Commun. **11**, 1442 (2020).

[16] Y. Chen and T. Pock, Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration, IEEE Trans. Pattern Anal. Mach. Intell. **39**, 1256 (2016).

[17] H. Qian, Phosphorylation energy hypothesis: Open chemical systems and their biological functions, Annu. Rev. Phys. Chem. **58**, 113 (2007).

[18] B. R. Irvin and J. Ross, Calculation of the rate of entropy production for a model chemical reaction, J. Chem. Phys. **89**, 1064 (1988).

[19] M. P. Hanson, Spatial structures in dissipative systems, J. Chem. Phys. **60**, 3210 (1974).

[20] H. Mahara, N. J. Suematsu, T. Yamaguchi, K. Ohgane, Y. Nishiura, and M. Shimomura, Three-variable reversible Gray–Scott model, J. Chem. Phys. **121**, 8968 (2004).

[21] H. Mahara and T. Yamaguchi, Entropy balance in distributed reversible Gray–Scott model, Physica D **239**, 729 (2010).

[22] P. Glansdorff and I. Prigogine, *Thermodynamic Theory of Structure, Stability and Fluctuations* (John Wiley and Sons, New York, 1971).

[23] G. Nicolis and I. Prigogine, *Self-Organization in Nonequilibrium Systems: From Dissipative Structures to Order Through Fluctuations* (Wiley-Blackwell, London, 1977).

[24] F. Avanzini, T. Aslyamov, É. Fodor, and M. Esposito, Nonequilibrium thermodynamics of non-ideal reaction-diffusion systems: Implications for active self-organization, J. Chem. Phys. **161**, 174108 (2024).

[25] Y. Kuramoto, *Chemical Oscillations, Waves, and Turbulence* (Springer, Berlin, 1984).

[26] M. C. Cross and P. C. Hohenberg, Pattern formation outside of equilibrium, Rev. Mod. Phys. **65**, 851 (1993).

[27] A. Rupe and J. P. Crutchfield, On principles of emergent organization, Phys. Rep. **1071**, 1 (2024).

[28] C. Jarzynski, Nonequilibrium equality for free energy differences, Phys. Rev. Lett. **78**, 2690 (1997).

[29] K. Sekimoto, *Stochastic Energetics* (Springer, Berlin, 2010).

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

[31] K. Proesmans, Y. Dreher, M. Gavrilov, J. Bechhoefer, and C. Van den Broeck, Brownian duet: A novel tale of thermodynamic efficiency, Phys. Rev. X **6**, 041010 (2016).

[32] T. R. Gingrich, G. M. Rotskoff, G. E. Crooks, and P. L. Geissler, Near-optimal protocols in complex nonequilibrium transformations, Proc. Natl. Acad. Sci. USA **113**, 10263 (2016).

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

[34] S. J. Large and D. A. Sivak, Optimal discrete control: Minimizing dissipation in discretely driven nonequilibrium systems, J. Stat. Mech. (2019) 083212.

[35] B. Remlein and U. Seifert, Optimality of nonconservative driving for finite-time processes with discrete states, Phys. Rev. E **103**, L050105 (2021).

[36] E. Ilker, O. Güngör, B. Kuznets-Speck, J. Chiel, S. Deffner, and M. Hinczewski, Shortcuts in stochastic systems and control of biophysical processes, Phys. Rev. X **12**, 021048 (2022).

[37] S. Blaber and D. A. Sivak, Optimal control in stochastic thermodynamics, J. Phys. Commun. **7**, 033001 (2023).

[38] M. C. Engel, J. A. Smith, and M. P. Brenner, Optimal control of nonequilibrium systems through automatic differentiation, Phys. Rev. X **13**, 041032 (2023).

[39] E. Aurell, C. Mejía-Monasterio, and P. Muratore-Ginanneschi, Optimal protocols and optimal transport in stochastic thermodynamics, Phys. Rev. Lett. **106**, 250601 (2011).

[40] E. Aurell, K. Gawędzki, C. Mejía-Monasterio, R. Mohayaee, and P. Muratore-Ginanneschi, Refined second law of thermodynamics for fast random processes, J. Stat. Phys. **147**, 487 (2012).

[41] M. Nakazato and S. Ito, Geometrical aspects of entropy production in stochastic thermodynamics based on Wasserstein distance, Phys. Rev. Res. **3**, 043093 (2021).

[42] T. Van Vu and Y. Hasegawa, Geometrical bounds of the irreversibility in Markovian systems, Phys. Rev. Lett. **126**, 010601 (2021).

[43] A. Dechant, Minimum entropy production, detailed balance and Wasserstein distance for continuous-time Markov processes, J. Phys. A: Math. Theor. **55**, 094001 (2022).

[44] 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).

[45] K. Yoshimura, A. Kolchinsky, A. Dechant, and S. Ito, Housekeeping and excess entropy production for general nonlinear dynamics, Phys. Rev. Res. **5**, 013017 (2023).

[46] A. Zhong and M. R. DeWeese, Beyond linear response: Equivalence between thermodynamic geometry and optimal transport, Phys. Rev. Lett. **133**, 057102 (2024).

[47] G. Falasco, R. Rao, and M. Esposito, Information thermodynamics of turing patterns, Phys. Rev. Lett. **121**, 108301 (2018).

[48] R. Rao and M. Esposito, Nonequilibrium thermodynamics of chemical reaction networks: Wisdom from stochastic thermodynamics, Phys. Rev. X **6**, 041064 (2016).

[49] H. Ge and H. Qian, Mesoscopic kinetic basis of macroscopic chemical thermodynamics: A mathematical theory, Phys. Rev. E **94**, 052150 (2016).

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

[51] A. M. Miangolarra and M. Castellana, On non-ideal chemical-reaction networks and phase separation, J. Stat. Phys. **190**, 23 (2023).

[52] S. Liang, P. D. L. Rios, and D. M. Busiello, Thermodynamic bounds on symmetry breaking in biochemical systems, Phys. Rev. Lett. **132**, 228402 (2022).

[53] A. Mielke, A gradient structure for reaction–diffusion systems and for energy-drift-diffusion systems, Nonlinearity **24**, 1329 (2011).

[54] S. Riaz, S. Banarjee, S. Kar, and D. Ray, Pattern formation in reaction-diffusion system in crossed electric and magnetic fields, Eur. Phys. J. B **53**, 509 (2006).

[55] J. W. Cahn and J. E. Hilliard, Free energy of a nonuniform system. I. Interfacial free energy, J. Chem. Phys. **28**, 258 (1958).

[56] F. Avanzini, E. Penocchio, G. Falasco, and M. Esposito, Nonequilibrium thermodynamics of non-ideal chemical reaction networks, J. Chem. Phys. **154**, 094114 (2021).

[57] T. Aslyamov, F. Avanzini, É. Fodor, and M. Esposito, Non-ideal reaction-diffusion systems: Multiple routes to instability, Phys. Rev. Lett. **131**, 138301 (2023).

[58] Y. Oono and M. Paniconi, Steady state thermodynamics, Prog. Theor. Phys. Suppl. **130**, 29 (1998).

[59] T. Hatano and S.-I. Sasa, Steady-state thermodynamics of Langevin systems, Phys. Rev. Lett. **86**, 3463 (2001).

[60] M. Esposito and C. Van den Broeck, Three detailed fluctuation theorems, Phys. Rev. Lett. **104**, 090601 (2010).

[61] C. Maes and K. Netočný, A nonequilibrium extension of the Clausius heat theorem, J. Stat. Phys. **154**, 188 (2014).

[62] A. Dechant, S.-I. Sasa, and S. Ito, Geometric decomposition of entropy production into excess, housekeeping, and coupling parts, Phys. Rev. E **106**, 024125 (2022).

[63] A. Dechant, S.-I. Sasa, and S. Ito, Geometric decomposition of entropy production in out-of-equilibrium systems, Phys. Rev. Res. **4**, L012034 (2022).

[64] N. Shiraishi, K. Funo, and K. Saito, Speed limit for classical stochastic processes, Phys. Rev. Lett. **121**, 070601 (2018).

[65] A. Dechant and Y. Sakurai, Thermodynamic interpretation of Wasserstein distance, arXiv:1912.08405.

[66] A. Kolchinsky, A. Dechant, K. Yoshimura, and S. Ito, Information geometry of excess and housekeeping entropy production, arXiv:2206.14599.

[67] T. Van Vu and K. Saito, Topological speed limit, Phys. Rev. Lett. **130**, 010402 (2023).

[68] O. M. Miangolarra, A. Taghvaei, and T. T. Georgiou, Minimal entropy production in the presence of anisotropic fluctuations, arXiv:2302.04401.

[69] S. Chennakesavalu and G. M. Rotskoff, Unified, geometric framework for nonequilibrium protocol optimization, Phys. Rev. Lett. **130**, 107101 (2023).

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

[71] J. M. Horowitz and T. R. Gingrich, Thermodynamic uncertainty relations constrain non-equilibrium fluctuations, Nat. Phys. **16**, 15 (2020).

[72] C. Villani, *Optimal Transport: Old and New* (Springer, New York, 2009), Vol. 338.

[73] F. Santambrogio, *Optimal Transport for Applied Mathematicians, Calculus of Variations*, PDEs, and Modeling (Birkäuser, New York, 2015).

[74] G. Schiebinger, J. Shu, M. Tabaka, B. Cleary, V. Subramanian, A. Solomon, J. Gould, S. Liu, S. Lin, P. Berube *et al.*, Optimal-transport analysis of single-cell gene expression identifies developmental trajectories in reprogramming, Cell **176**, 928 (2019).

[75] M. Arjovsky, S. Chintala, and L. Bottou, Wasserstein generative adversarial networks, in *Proceedings of the 34th International Conference on Machine Learning*, edited by D. Precup and Y. W. Teh, Proceedings of Machine Learning Research Vol. 70 (PMLR, 2017), pp. 214–223.

[76] G. Peyré, M. Cuturi *et al.*, Computational optimal transport: With applications to data science, Found. Trends Mach. Learn. **11**, 355 (2019).

[77] J.-D. Benamou and Y. Brenier, A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem, Numer. Math. **84**, 375 (2000).

[78] R. Fu, A. Taghvaei, Y. Chen, and T. T. Georgiou, Maximal power output of a stochastic thermodynamic engine, Automatica **123**, 109366 (2021).

[79] O. M. Miangolarra, A. Taghvaei, Y. Chen, and T. T. Georgiou, Geometry of finite-time thermodynamic cycles with anisotropic thermal fluctuations, IEEE Control Syst. Lett. **6**, 3409 (2022).

[80] Y. Fujimoto and S. Ito, Game-theoretical approach to minimum entropy productions in information thermodynamics, arXiv:2112.14035.

[81] R. Nagase and T. Sagawa, Thermodynamically optimal information gain in finite time measurement, arXiv:2307.16757.

[82] S. Ito, Geometric thermodynamics for the Fokker–Planck equation: Stochastic thermodynamic links between information geometry and optimal transport, Info. Geo. **7**, 441 (2023).

[83] K. Yoshimura and S. Ito, Two applications of stochastic thermodynamics to hydrodynamics, Phys. Rev. Res. **6**, L022057 (2024).

[84] K. Yoshimura and S. Ito, Thermodynamic uncertainty relation and thermodynamic speed limit in deterministic chemical reaction networks, Phys. Rev. Lett. **127**, 160601 (2021).

[85] M. Liero and A. Mielke, Gradient structures and geodesic convexity for reaction–diffusion systems, Philos. Trans. R. Soc. A **371**, 20120346 (2013).

[86] M. Liero, A. Mielke, and G. Savaré, Optimal transport in competition with reaction: The Hellinger–Kantorovich distance and geodesic curves, SIAM J. Math. Anal. **48**, 2869 (2016).

[87] A. Graps, An introduction to wavelets, IEEE Comput. Sci. Eng. **2**, 50 (1995).

[88] C. Torrence and G. P. Compo, A practical guide to wavelet analysis, Bull. Am. Meteorol. Soc. **79**, 61 (1998).

[89] C. Nardini, É. Fodor, E. Tjhung, F. Van Wijland, J. Tailleur, and M. E. Cates, Entropy production in field theories without time-reversal symmetry: Quantifying the non-equilibrium character of active matter, Phys. Rev. X **7**, 021007 (2017).

[90] T. Harada and S.-I. Sasa, Equality connecting energy dissipation with a violation of the fluctuation-response relation, Phys. Rev. Lett. **95**, 130602 (2005).

[91] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient Flows: In Metric Spaces and in the Space of Probability Measures*, Lectures in Mathematics: ETH Zürich (Birkhäuser, Basel, 2005).

[92] L. Brasco, A survey on dynamical transport distances, J. Math. Sci. **181**, 755 (2012).

[93] L. C. Evans and W. Gangbo, *Differential Equations Methods for the Monge-Kantorovich Mass Transfer Problem* (American Mathematical Soc., Providence, RI, 1999).

[94] M. Beckmann, A continuous model of transportation, Econometrica **20**, 643 (1952).

[95] Y. Chen, T. T. Georgiou, L. Ning, and A. Tannenbaum, Matricial Wasserstein-1 distance, IEEE Control Syst. Lett. **1**, 14 (2017).

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

[97] K. Liu, Z. Gong, and M. Ueda, Thermodynamic uncertainty relation for arbitrary initial states, Phys. Rev. Lett. **125**, 140602 (2020).

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

[99] C. R. Rao, Information and the accuracy attainable in the estimation of statistical parameters, in *Breakthroughs in Statistics: Foundations and Basic Theory* (Springer, New York, 1992), pp. 235–247.

[100] S. Ito and A. Dechant, Stochastic time evolution, information geometry, and the Cramér-Rao bound, Phys. Rev. X **10**, 021056 (2020).

[101] S. B. Nicholson, L. P. García-Pintos, A. del Campo, and J. R. Green, Time–information uncertainty relations in thermodynamics, Nat. Phys. **16**, 1211 (2020).

[102] J. A. Griepentrog, On the unique solvability of a nonlocal phase separation problem for multicomponent systems, Banach Center Publications **66**, 153 (2004).

[103] L. Onsager, Reciprocal relations in irreversible processes. I., Phys. Rev. **37**, 405 (1931).

[104] C. Maes and K. Netočný, Canonical structure of dynamical fluctuations in mesoscopic nonequilibrium steady states, Europhys. Lett. **82**, 30003 (2008).

[105] C. Maes, *Non-Dissipative Effects in Nonequilibrium Systems* (Springer International Publishing, New York, 2018).

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

[107] J. D. Murray, *Mathematical Biology: I. An Introduction*, Interdisciplinary Applied Mathematics (Springer, New York, 2002).

[108] J. Maas, Gradient flows of the entropy for finite Markov chains, J. Funct. Anal. **261**, 2250 (2011).

[109] 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).

[110] M. Feinberg, Foundations of chemical reaction network theory, in *Applied Mathematical Sciences* (Springer, New York, 2019).

[111] A. Chambolle and T. Pock, A first-order primal-dual algorithm for convex problems with applications to imaging, J. Math. Imaging Vision **40**, 120 (2011).

[112] W. Li, E. K. Ryu, S. Osher, W. Yin, and W. Gangbo, A parallel method for earth mover's distance, J. Sci. Comput. **75**, 182 (2018).

[113] I. Borg and P. J. Groenen, *Modern Multidimensional Scaling: Theory and Applications*, Springer Series in Statistics (Springer, New York, NY, 2007).

[114] O. Niggemann and U. Seifert, Field-theoretic thermodynamic uncertainty relation: General formulation exemplified with the Kardar–Parisi–Zhang equation, J. Stat. Phys. **178**, 1142 (2020).

[115] S. Rana and A. C. Barato, Precision and dissipation of a stochastic Turing pattern, Phys. Rev. E **102**, 032135 (2020).

[116] Y. Song and C. Hyeon, Cost-precision trade-off relation determines the optimal morphogen gradient for accurate biological pattern formation, Elife **10**, e70034 (2021).

[117] M. Nguyen, Y. Qiu, and S. Vaikuntanathan, Organization and self-assembly away from equilibrium: Toward thermodynamic design principles, Annu. Rev. Condens. Matter Phys. **12**, 273 (2021).

[118] E. Zimmermann and U. Seifert, Efficiencies of a molecular motor: A generic hybrid model applied to the F1-APTase, New J. Phys. **14**, 103023 (2012).

[119] S.-I. Amari and H. Nagaoka, *Methods of Information Geometry*, Vol. 191 (American Mathematical Soc., Providence, RI, 2000).

[120] S.-I. Amari, *Information Geometry and Its Applications* (Springer Japan, Tokyo, 2016).

[121] H. Shima, *The Geometry of Hessian Structures* (World Scientific, Singapore, 2007).

[122] C. A. Weber, D. Zwicker, F. Jülicher, and C. F. Lee, Physics of active emulsions, Rep. Prog. Phys. **82**, 064601 (2019).

[123] J. Kirschbaum and D. Zwicker, Controlling biomolecular condensates via chemical reactions, J. R. Soc. Interface. **18**, 20210255 (2021).

[124] J. Bauermann, C. A. Weber, and F. Jülicher, Energy and matter supply for active droplets, Ann. Phys. (Leipzig) **534**, 2200132 (2022).

[125] D. Zwicker, The intertwined physics of active chemical reactions and phase separation, Curr. Opin. Colloid Interface Sci. **61**, 101606 (2022).

[126] S. K. Manikandan, D. Gupta, and S. Krishnamurthy, Inferring entropy production from short experiments, Phys. Rev. Lett. **124**, 120603 (2020).

[127] R. G. Plaza, F. Sanchez-Garduno, P. Padilla, R. A. Barrio, and P. K. Maini, The effect of growth and curvature on pattern formation, J. Dyn. Differ. Equations **16**, 1093 (2004).

[128] R. Nishide and S. Ishihara, Pattern propagation driven by surface curvature, Phys. Rev. Lett. **128**, 224101 (2022).

[129] H. G. Othmer and L. Scriven, Instability and dynamic pattern in cellular networks, J. Theor. Biol. **32**, 507 (1971).

[130] H. Nakao and A. S. Mikhailov, Turing patterns in network-organized activator–inhibitor systems, Nat. Phys. **6**, 544 (2010).

[131] T. H. Tan, J. Liu, P. W. Miller, M. Tekant, J. Dunkel, and N. Fakhri, Topological turbulence in the membrane of a living cell, Nat. Phys. **16**, 657 (2020).

[132] V. Colizza, R. Pastor-Satorras, and A. Vespignani, Reaction–diffusion processes and metapopulation models in heterogeneous networks, Nat. Phys. **3**, 276 (2007).

[133] J. Howard, S. W. Grill, and J. S. Bois, Turing's next steps: The mechanochemical basis of morphogenesis, Nat. Rev. Mol. Cell Biol. **12**, 392 (2011).

[134] M. Mercker, F. Brinkmann, A. Marciniak-Czochra, and T. Richter, Beyond turing: Mechanochemical pattern formation in biological tissues, Biol. Direct **11**, 22 (2016).

[135] J. Dolbeault, B. Nazaret, and G. Savaré, A new class of transport distances between measures, Calc. Var. **34**, 193 (2009).