

Spatial self-organization of enzymes in complex reaction networks

Vincent Ouazan-Reboul,^{1,2} Ramin Golestanian,^{1,3} and Jaime Agudo-Canalejo^{1,4}

¹*Max Planck Institute for Dynamics and Self-Organization, Am Fassberg 17, D-37077, Göttingen, Germany*

²*LPTMS, CNRS-Université Paris-Sud, 91400, Orsay, France*

³*Rudolf Peierls Centre for Theoretical Physics, University of Oxford, OX1 3PU, Oxford, UK*

⁴*Department of Physics and Astronomy, University College London, WC1E 6BT, London, UK*

Living systems contain intricate biochemical networks whose structure is closely related to their function and allows them to exhibit robust behavior in the presence of external stimuli. Such networks typically involve catalytic enzymes, which can have non-trivial transport properties, in particular chemotaxis-like directed motion along gradients of substrates and products. Here, we find that taking into account enzyme chemotaxis in models of catalyzed reaction networks can lead to their spatial self-organization in a process similar to biomolecular condensate formation. We develop a general theory for arbitrary reaction networks, and systematically study all closed unimolecular reaction networks involving up to six chemicals. Importantly, we find that network-wide propagation of concentration perturbations can be key to enabling self-organization. The ability to self-organize is highly dependent on the relative signs of the chemotactic mobilities of the enzymes to their substrate and product and on the global network structure. We find that spontaneous self-organization through chemotaxis can provide an avenue for the self-regulation of metabolic activity in complex catalyzed reaction networks. The network-induced interaction mechanism we uncover operates in the regime where the substrate molecules are diffusion-limited, suggesting that signaling molecules could take advantage of this scenario towards their functionality.

A universal feature of biological systems is their ability to robustly perform complex tasks as a result of elementary stochastic processes, from performing the right biochemical reactions when and where needed to maintaining homeostasis [1]. Many biochemical processes are performed or regulated by intricate biochemical networks whose structure is intrinsically linked to their function [2, 3]. Examples include circadian clocks [4], genetic regulation networks [5], and metabolic reaction networks [6]. Biochemical reaction networks typically involve kinetically unfavorable reactions that would not occur spontaneously in any reasonable time [7], and only take place due to the action of catalytic enzymes that therefore drive the system out of (quasi-) equilibrium. Structural analysis tools have previously been leveraged in order to understand the structure-function relationship of these networks [8], to help characterize their structure from experimental data [9–11], and to design networks which can be incorporated in synthetic biological systems [12].

However, these and other studies of metabolic networks have so far relied on the so-called well-stirred assumption, i.e. they assume that the concentrations of all the chemicals and enzymes in the system have spatially homogeneous concentrations. In contrast, biological cells are spatially heterogeneous environments, as highlighted in particular by the existence of biomolecular condensates: many proteins, including enzymes, aggregate into dense droplets at the right place and at the right time [13–15]. The precise mechanisms behind the formation of enzyme-rich biomolecular condensates, also known as metabolons [16, 17], are not well understood. In particular, it is unclear to what extent they can be fully described within equilibrium physics (e.g. weak, multi-

valent enzyme-enzyme or enzyme-scaffold interactions) or whether nonequilibrium effects due to enzyme catalytic activity play a role. Intriguingly, the composition of metabolons is often finely tuned to the metabolic pathway at hand: all the enzyme species in a particular pathway are present, in the right proportions so that the product of one enzyme species can be “channeled” as the substrate of the next enzyme species in the pathway [18–20]. Moreover, metabolons form and dissolve dynamically in response to the changing metabolic needs of the cell [16, 17]. These observations, as well as in vitro experiments [21] and theoretical studies of model systems [22], strongly suggest that nonequilibrium effects due to catalysis play a role in the formation of metabolically active enzyme complexes, enabling a link between network topology and spatial structure.

A minimal yet powerful hypothesis that could provide this link is that chemotactic-like motion of enzymes in the presence of self-generated substrate and product gradients causes effective enzyme-enzyme interactions [23], in analogy with chemotactic microbial aggregation [24] and with the self-organization of chemically-active colloids [25–28]. Indeed, many in vitro experiments have shown that biological enzymes [29–32] can chemotax in response to gradients of their substrate. For the underlying mechanism, phoretic-hydrodynamic effects and binding-induced conformational changes of the enzyme have been proposed as possible candidates [23, 33]. Minimal theoretical models and simulations have shown that enzyme-enzyme aggregation through chemotaxis is indeed possible, and can lead to metabolically optimized stoichiometries of the enzyme aggregate [34–40]. One of the key nonequilibrium features of such

interactions is non-reciprocity [41], which can lead to a wealth of complex behavior in multicomponent condensates [42, 43]. Importantly, these effective enzyme-enzyme interactions due to chemotaxis can be made thermodynamically consistent and arise generically in standard models for the nonequilibrium (thermo)dynamics of multicomponent mixtures [44]. However, all studies so far have focused on very simple metabolic networks, considering only the interconversion between two metabolites by one [39, 44] or several enzymes [34, 35, 40], or simple metabolic cycles without branches [36–38].

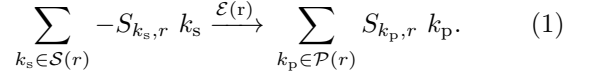
Motivated by these considerations, here we study the spatial self-organization of enzymes participating in arbitrarily complex chemical reaction networks. We generalize the mathematical framework typically used for the structural analysis of biochemical networks to account for spatial dynamics, including the chemotaxis of the enzymes in response to gradients of metabolites (Fig. 1), and analyze the stability of homogeneous (well-stirred) states to spatial perturbations. We arrive to a general, concise, and readily applicable expression describing the spatial stability of any metabolic network. We then focus our study on the set of all 1,052,145 closed unimolecular reaction networks involving up to six chemicals. For each of these networks, we systematically compare results obtained assuming concentration-dependent rates (saturated regime in the Michaelis-Menten kinetics) to results obtained with concentration-independent reaction rates (linear regime in the Michaelis-Menten kinetics). We find that, while in the former case the emergent enzyme-enzyme interactions conform to our expectations, with enzymes interacting with one another only if they chemotax in response to chemicals that the other produces or consumes, in the latter case network-wide effects reflecting a far-ranging propagation of concentration perturbations throughout the catalytic network can have a strong effect on the enzyme-enzyme interactions. These network-wide effects in turn affect the self-organization of the system, and thus imply a sensitive dependence upon the full structure of the reaction network, leading to a wide variation in the ability of different reaction networks to self-organize.

RESULTS

Catalyzed reaction network formalism

We consider a minimal model for a generic catalyzed reaction network involving K chemicals, with concentrations given by the space-and-time-dependent vector $\mathbf{c}(\mathbf{r}, t) = (c_1(\mathbf{r}, t), c_2(\mathbf{r}, t), \dots, c_K(\mathbf{r}, t))^T$, which act as the substrates and products of M enzymes with concentrations $\boldsymbol{\rho}(\mathbf{r}, t) = (\rho_1(\mathbf{r}, t), \rho_2(\mathbf{r}, t), \dots, \rho_M(\mathbf{r}, t))^T$ according to a set of R chemical reactions. Each reaction r corresponds to the conversion of a set of substrate chemicals $\mathcal{S}(r)$ into a set of product chemicals $\mathcal{P}(r)$ assisted

by a set of enzymes $\mathcal{E}(r)$, summarized as the chemical equation



Equation (1) features a set of stoichiometries $\mathbf{S}_r = (S_{1,r}, S_{2,r}, \dots, S_{K,r})^T$, which give the number of molecules of each chemical k that participates in reaction r : $S_{k,r}$ is negative if $k \in \mathcal{S}(r)$, positive if $k \in \mathcal{P}(r)$, and null otherwise (Fig. 1A). The structure of the network formed by the set of R reactions is then defined by the stoichiometry matrix $\mathbf{S} \in \mathbb{Z}^{K \times R}$, whose r^{th} column is the stoichiometry vector of reaction r (Fig. 1B).

In order to fully define the reaction network, we also need to specify the kinetics of its constitutive reactions, which are given by a vector of reaction speeds $\mathbf{v}(\mathbf{c}, \boldsymbol{\rho}) = (v_1(\mathbf{c}, \boldsymbol{\rho}), v_2(\mathbf{c}, \boldsymbol{\rho}), \dots, v_R(\mathbf{c}, \boldsymbol{\rho}))^T$ whose r^{th} component is the rate at which reaction r takes place. Taking into account the diffusion of each chemical k with a coefficient $D_k^{(c)}$, the spatio-temporal dynamics of the chemical concentration vector can then be written as

$$\partial_t \mathbf{c}(\mathbf{r}, t) = \mathbf{S} \mathbf{v}(\mathbf{c}, \boldsymbol{\rho}) + \mathcal{D}^{(c)} \nabla^2 \mathbf{c}(\mathbf{r}, t). \quad (2)$$

Here, $\mathcal{D}_{k,l}^{(c)} = D_k^{(c)} \delta_{kl}$ is the diagonal matrix of the diffusion coefficients of the chemicals (where δ represents the Kronecker symbol) and ∇ is the spatial gradient operator. Note that Einstein's summation convention is not used throughout this work. Equation (2) adds a space-dependence to the reaction dynamics, which is commonly written in the context of well-stirred chemical reaction networks [45–47], directly through the diffusion term it contains and indirectly through the reaction speeds being a function of the space-dependent concentrations.

We describe the dynamics of the enzymes with a continuity equation that enforces the conservation of the total number of enzymes. It involves a chemotactic drift term, which characterizes how a particle of enzyme species m in a concentration gradient of chemical k develops a velocity $\mathbf{u}_{m,k} = -\mu_{m,k} \nabla c_k$ (Fig. 1C). $\mu_{m,k}$ is the chemotactic mobility of enzyme species m in a concentration gradient of chemical k . The set of all mobilities defines the mobility matrix $\boldsymbol{\mu} \in \mathbb{R}^{M \times K}$. This drift term is motivated by the observation of chemotactic motion in biological enzymes [23], as discussed above. The evolution equation for the enzyme concentration vector is therefore

$$\partial_t \boldsymbol{\rho}(\mathbf{r}, t) = \nabla \cdot \left[\mathcal{D}^{(e)} \nabla \boldsymbol{\rho}(\mathbf{r}, t) + \hat{\boldsymbol{\rho}}(\mathbf{r}, t) \boldsymbol{\mu} \nabla \mathbf{c}(\mathbf{r}, t) \right], \quad (3)$$

where $\hat{\boldsymbol{\rho}}_{m,n} = \rho_m \delta_{mn}$ and $\mathcal{D}_{m,n}^{(e)} = D_m^{(e)} \delta_{mn}$ are diagonal matrices representing the enzyme concentrations and diffusion coefficients, which are introduced to help with the brevity of the notation.

Equations (2) and (3) describe a complex dynamics that

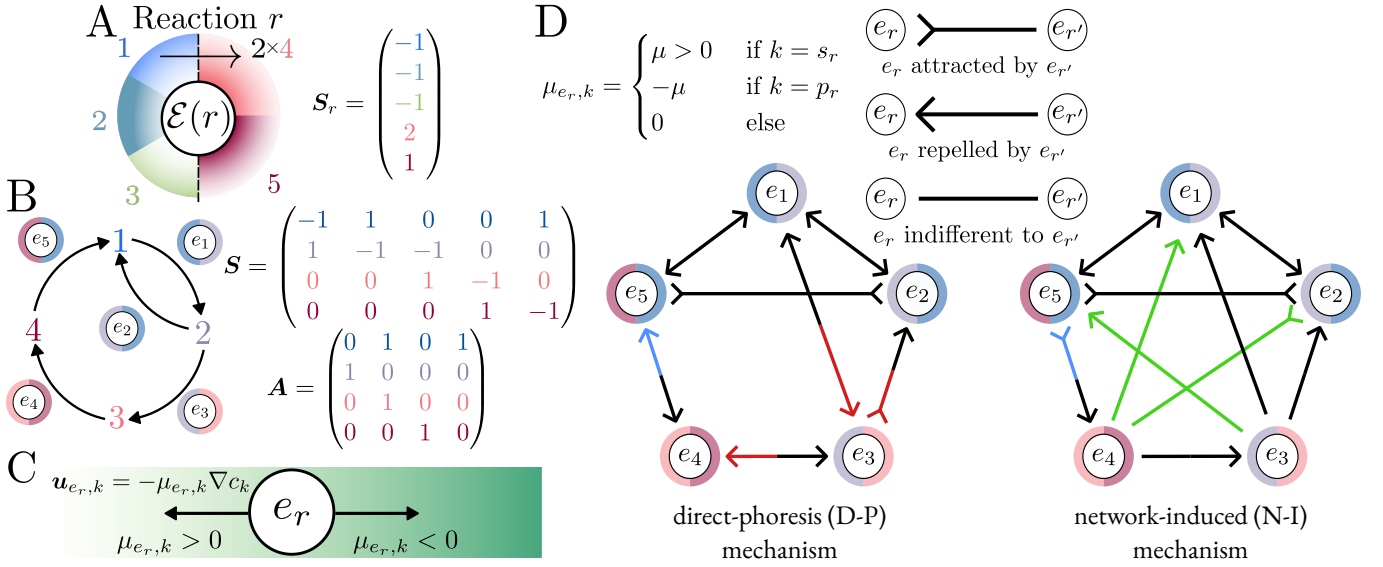


FIG. 1. Definition of a catalyzed reaction network. (A) Reaction formalism: a given reaction r transforms a set of substrates $\mathcal{S}(r) = \{1, 2, 3\}$ into a set of products $\mathcal{P}(r) = \{4, 5\}$, catalyzed by a set of enzymes $\mathcal{E}(r)$. The reaction is associated with a stoichiometry vector $\mathbf{S}_r$ describing the proportions of the production or consumption of the involved chemicals. (B) A set of reactions defines a chemical reaction network, whose structure is given by a stoichiometry matrix $\mathbf{S}$. The r th column of $\mathbf{S}$ is the stoichiometry vector $\mathbf{S}_r$. In this example, all reactions are unimolecular and each reaction r is catalyzed by a different single enzyme e_r . The reaction network is thus equivalently described by an adjacency matrix $\mathbf{A}$. (C) The enzymes (e_r) involved in the reaction network are chemotactic, moving directionally in concentration gradients of the chemicals (k) which can take them towards high or low concentrations respectively for negative or positive chemotactic mobility $\mu_{e_r,k}$. (D) The combination of chemical activity and chemotaxis creates effective interactions between the enzymes, given by Eq. (10) and described by an interaction network. In this example, enzymes move towards higher (lower) concentrations of product (substrate). Saturated reaction kinetics (independent of substrate concentration) lead to interactions originating from *direct-phoresis* (D-P), where an enzyme only interacts with another if one has a chemotactic response to the substrates or products of the other. Non-saturated reaction kinetics (dependent on substrate concentration) create *network-wide* effects that can cause *network-induced* (N-I) interactions between enzymes that might even not interact in the D-P sense (green arrows), and can switch the direction of (blue half-arrows), or suppress (red half-arrows) the D-P mechanism.

couples the concentration fields of enzymes and chemicals, and result in effective interactions between the M enzyme species, mediated by the K chemical concentration fields (Fig. 1D)

[34–36, 38, 48, 49]. The enzymes $\mathcal{E}(r)$ associated with the reaction r locally modify the concentrations of substrates $\mathcal{S}(r)$ and products $\mathcal{P}(r)$ of r , thus creating concentration gradients to which other enzymes can respond through their chemotactic mobility. Throughout the rest of this work, we aim to determine if, and under which conditions, the dynamics just described can lead to the spatial self-organization of the enzymes and chemicals participating in a reaction network.

Homogeneous steady-state and linear stability analysis

In order to determine whether a catalyzed reaction network undergoes self-organization, we perform a linear stability analysis on Eqs. (2) and (3). To do so, we perturb a homogeneous steady-state of the network. Because the enzyme concentrations are conserved, any ar-

bitrary homogeneous enzyme concentration vector $\boldsymbol{\rho}_0$ is a valid steady-state. On the other hand, the homogeneous chemical concentration vector $\mathbf{c}_0$ is defined by the requirement that the chemical reactions are balanced, that is, must be a solution of the equation

$$\mathbf{S}\mathbf{v}(\mathbf{c}_0, \boldsymbol{\rho}_0) = 0. \quad (4)$$

In other words, the reaction rates associated with the homogeneous steady-state must be in the kernel $\ker(\mathbf{S})$ of the stoichiometry matrix. The corresponding chemical concentration vector can then be obtained by deriving a vector of rates $\mathbf{v}_0 \in \ker(\mathbf{S})$ and solving $\mathbf{v}(\mathbf{c}_0, \boldsymbol{\rho}_0) = \mathbf{v}_0$ for $\mathbf{c}_0$. The details of this calculation will depend on the network structure and choice of rate functions. In particular, depending on whether or not the chemical reaction network satisfies local detailed balance, the homogeneous steady-state will represent an equilibrium state or a nonequilibrium steady-state. In the equilibrium case, there will be a single steady-state, and the concentrations $\mathbf{c}_0$ will be independent of the choice of enzyme concentrations $\boldsymbol{\rho}_0$. In the nonequilibrium case, there may be more than one steady-state, and $\mathbf{c}_0$ will generally depend on $\boldsymbol{\rho}_0$. The homogeneous concentrations obtained

by solving Eq. (4) are perturbed with space- and time-dependent perturbations $\rho(\mathbf{r}, t) = \rho_0 + \delta\rho(\mathbf{r}, t)$ and $\mathbf{c}(\mathbf{r}, t) = \mathbf{c}_0 + \delta\mathbf{c}(\mathbf{r}, t)$. Expanding Eqs. (2) and (3) to first order in the perturbations, and representing the perturbations as $\delta\mathbf{c}(\mathbf{r}, t) = \delta\tilde{\mathbf{c}}(\mathbf{q}) e^{i\mathbf{q}\cdot\mathbf{r}} e^{\lambda(q^2)t}$ and $\delta\rho(\mathbf{r}, t) = \delta\tilde{\rho}(\mathbf{q}) e^{i\mathbf{q}\cdot\mathbf{r}} e^{\lambda(q^2)t}$, yields the linearized equation

$$\lambda(q^2) \begin{pmatrix} \delta\tilde{\rho} \\ \delta\tilde{\mathbf{c}} \end{pmatrix} = \begin{pmatrix} -\mathcal{D}^{(e)} q^2 & -\hat{\rho}_0 \mu q^2 \\ \mathbf{S} \partial_{\rho} \mathbf{v}|_0 & \mathbf{S} \partial_{\mathbf{c}} \mathbf{v}|_0 - \mathcal{D}^{(c)} q^2 \end{pmatrix} \begin{pmatrix} \delta\tilde{\rho} \\ \delta\tilde{\mathbf{c}} \end{pmatrix}, \quad (5)$$

where $(\partial_{\rho} \mathbf{v}|_0)_{r,m} = \partial_{\rho_m} v_r|_{\mathbf{c}_0, \rho_0}$ and $(\partial_{\mathbf{c}} \mathbf{v}|_0)_{r,k} = \partial_{c_k} v_r|_{\mathbf{c}_0, \rho_0}$ are the Jacobians of the reaction rate vector. Equation (5) is an eigenvalue problem that involves a matrix which couples the perturbation concentrations of the enzymes and chemicals, whose eigenvalues $\lambda(q^2)$ are the growth rates of perturbations associated with wavevector $\mathbf{q}$. It follows that if the matrix shown in Eq. (5) has at least one eigenvalue with positive real part over some range of wavevectors, the corresponding system undergoes a self-organizing instability.

We observe that instabilities at the scale of the system size (corresponding to $q \rightarrow 0$ in Eq. (5)) occur for two different categories of modes: (i) $\delta\tilde{\rho} = 0$ and $\lambda(0)\delta\tilde{\mathbf{c}} = \mathbf{S} \partial_{\mathbf{c}} \mathbf{v}|_0 \delta\tilde{\mathbf{c}}$, or (ii) $\delta\tilde{\rho} \neq 0$ and $\lambda(0) = 0$. Case (i) corresponds to perturbations associated only with the chemicals, which are not number conserving and originate from the well-stirred reaction dynamics. As we are interested in the spatial self-organization of the full enzyme-chemical mixture, we assume that these chemical-associated modes are stable, i.e. $\mathbf{S} \partial_{\mathbf{c}} \mathbf{v}|_0$ has only negative eigenvalues, which is equivalent to assuming that the homogeneous steady-state $(\mathbf{c}_0, \rho_0)$ is a stable fixed-point of the well-stirred reaction dynamics. Finding such stable states is a well-studied problem in the theory of chemical reaction networks [46, 50], which we will not cover here. Similarly, we will assume that these do not become positive at a finite band of $q^2 > 0$, as this would correspond to the well-known Turing instabilities of the chemicals. Instead, in the following, we will focus on the modes of case (ii), which are associated with the enzymes and respect number conservation.

Small wavevector expansion and effective enzyme-enzyme interactions

For these enzyme-associated modes, the stability of the system is determined by the slope $\lambda'(0) \equiv \frac{d\lambda}{dq^2}|_{q^2=0}$. If $\text{Re}(\lambda'(0)) > 0$, the enzymes will display a large-scale (but number-conserving) spatial instability. In order to move forward, we make a quasi-static approximation $\partial_t \mathbf{c} \simeq 0$, equivalent to considering that the chemical concentration fields equilibrate very fast relative to the timescale of variation of the enzyme concentration fields. The chemical part of the system in Eq. (5), i.e. its last K rows,

then reduce to

$$(\mathbf{S} \partial_{\mathbf{c}} \mathbf{v}|_0) \delta\tilde{\mathbf{c}} \simeq -(\mathbf{S} \partial_{\rho} \mathbf{v}|_0) \delta\tilde{\rho} + O(\mathcal{D}^{(c)} \delta\tilde{\mathbf{c}} q^2), \quad (6)$$

an equation relating the perturbations of the enzyme and chemical. To calculate the slope of the enzyme-related eigenvalues at the origin, we must solve for $\delta\tilde{\mathbf{c}}$ in Eq. (6) and introduce it into the enzyme part of the system in Eq. (5), i.e. its first M rows. Note that the $O(q^2)$ term in Eq. (6) will only lead to $O(q^4)$ terms in the eigenvalues, and thus can be ignored in the following.

Importantly, if the reaction network contains conserved moieties, i.e. if at least one linear combination of the chemical concentrations is conserved under the reaction dynamics, the matrix $\mathbf{S}$ is not of full rank and inverting Eq. (6) is not possible [51–53]. An invertible equation can be derived by Gaussian elimination on $\mathbf{S}$. This process yields a new stoichiometry matrix $\mathbf{S}_{\text{ind}} \in \mathbb{Z}^{K_{\text{ind}} \times R}$ associated with a set of K_{ind} independent chemicals. As a byproduct of this process, a set of $K - K_{\text{ind}}$ moiety conservation laws are derived, which can be written as $\mathbf{C} \delta\tilde{\mathbf{c}} = \mathbf{0}$ with $\mathbf{C} \in \mathbb{Z}^{(K-K_{\text{ind}}) \times K}$ a matrix of coefficients. We then build the Jacobian matrices of the independent system, which have expressions

$$(\mathbf{S} \partial_{\mathbf{c}} \mathbf{v}|_0)_{\text{ind}} = \begin{pmatrix} \mathbf{S}_{\text{ind}} \partial_{\mathbf{c}} \mathbf{v}|_0 \\ \mathbf{C} \end{pmatrix}, \quad (7)$$

and

$$(\mathbf{S} \partial_{\rho} \mathbf{v}|_0)_{\text{ind}} = \begin{pmatrix} \mathbf{S}_{\text{ind}} \partial_{\rho} \mathbf{v}|_0 \\ \mathbf{0} \end{pmatrix}. \quad (8)$$

The system $(\mathbf{S} \partial_{\mathbf{c}} \mathbf{v}|_0)_{\text{ind}} \delta\tilde{\mathbf{c}} = -(\mathbf{S} \partial_{\rho} \mathbf{v}|_0)_{\text{ind}} \delta\tilde{\rho}$ is in general invertible [52, 53] and gives solutions to Eq. (6) with the appropriate moiety conservation laws.

Plugging in $\delta\tilde{\mathbf{c}} = -(\mathbf{S} \partial_{\mathbf{c}} \mathbf{v}|_0)_{\text{ind}}^{-1} (\mathbf{S} \partial_{\rho} \mathbf{v}|_0)_{\text{ind}} \delta\tilde{\rho}$ into the enzyme part of Eq. (5), we finally obtain an eigenvalue equation for the slope $\lambda'(0)$, given by

$$\lambda'(0) \delta\tilde{\rho} = -(\hat{\rho}_0 \boldsymbol{\eta} + \mathcal{D}^{(e)}) \delta\tilde{\rho}, \quad (9)$$

where we have defined the *interaction matrix*

$$\boldsymbol{\eta} \equiv -\boldsymbol{\mu} (\mathbf{S} \partial_{\mathbf{c}} \mathbf{v}|_0)_{\text{ind}}^{-1} (\mathbf{S} \partial_{\rho} \mathbf{v}|_0)_{\text{ind}} \quad (10)$$

with $\boldsymbol{\eta} \in \mathbb{R}^{M \times M}$. Equations (9) and (10) constitute a central result of this work. Whenever $\text{Re}(\lambda'(0)) > 0$, i.e. if the matrix $\hat{\rho}_0 \boldsymbol{\eta} + \mathcal{D}^{(e)}$ has at least one eigenvalue with negative real part, the system will display spontaneous spatial self-organization of the enzymes. The interaction matrix has coefficients $\eta_{m,n}$ representing the effective, chemical-field-mediated response of enzyme species m to enzyme species n . Note that $\boldsymbol{\eta}$ has the same physical dimensions as the chemotactic mobility $\boldsymbol{\mu}$: effectively, $\boldsymbol{\eta}$ behaves as a chemotactic mobility in response to gradients of enzymes, rather than gradients of chemicals. Neg-

ative (respectively positive) coefficients correspond to enzyme species m being attracted to (respectively repelled by) species n .

We note that the emergence of enzyme-enzyme interactions in this system is a nonequilibrium effect, caused by the lack of detailed balance in the catalyzed chemical reaction network. Indeed, in an equilibrium state, the homogeneous concentrations $\mathbf{c}_0$ are independent of the enzyme concentrations, implying $\mathbf{S} \partial_{\rho} \mathbf{v}|_0 = 0$ and therefore $\boldsymbol{\eta} = 0$. Moreover, the interaction matrix $\boldsymbol{\eta}$ is generally not symmetric. That is, the chemical-field-mediated interactions are generally *nonreciprocal*, so that the velocity of enzyme m in response to the presence of enzyme n is not necessarily of equal magnitude and opposite direction to that of n in the presence of m , if n and m are not of the same type. In particular, n may be attracted to m while m is repelled from n . Such “chasing” interactions [41, 54], which would be impossible in an equilibrium system, are a manifestation of the nonequilibrium origin of the effective enzyme-enzyme interactions, and are known to lead to collective phenomena including travelling clusters and waves [34, 55–57].

There are two fundamentally different mechanisms with which enzymes can interact with each within this framework (see Fig. 1D). Naturally, a given enzyme species develops an effective interaction with another species if the latter produces or consumes chemicals that induce a chemotactic response in the former. We shall call this mechanism *direct-phoresis* (D-P). However, this is not the only way that interactions can emerge in complex chemical networks, as the dependence of $\mathbf{v}$ on $\mathbf{c}$ (as well as the existence of conserved moieties) can induce a more complex structure in Eq. (10) through the screening factor $(\mathbf{S} \partial_{\mathbf{c}} \mathbf{v}|_0)^{-1}_{\text{ind}}$. This leads to the emergence of a *network-induced* (N-I) mechanism, which can lead to non-intuitive behavior: for example, an enzyme species m can interact with another enzyme species n despite having null chemotactic mobilities for the substrates and products of n (see Fig. 1D).

It is interesting to note that the D-P form of the interactions can be recovered in the special limit in which the catalytic reactions take place in the saturated regime $\mathbf{v}(\boldsymbol{\rho})$, i.e. when they depend on the enzyme concentrations but not on the chemical concentrations. In this special limit, an analogous calculation shows that the stability of the system (away from the point of discontinuity $q^2 = 0$ for which $\lambda(0) = 0$; see Refs. [34, 35]) is governed by $\lambda(q^2) \delta \tilde{\boldsymbol{\rho}} = - \left(\tilde{\boldsymbol{\rho}}_0 \boldsymbol{\eta}^{\text{sat}} + \mathcal{D}^{(e)} q^2 \right) \delta \tilde{\boldsymbol{\rho}}$ with an interaction matrix $\boldsymbol{\eta}^{\text{sat}} \equiv \boldsymbol{\mu} \mathcal{D}^{(c)}{}^{-1} \mathbf{S} \partial_{\rho} \mathbf{v}|_0$ which, noting that $\mathcal{D}^{(c)}$ is a diagonal matrix, conforms to the picture of D-P interactions.

Unimolecular reaction networks with bilinear rates

We now apply the general framework developed in the previous sections to a particular class of networks and reaction rate functions, to gain a better understanding on how the properties of a catalytic network influence its stability.

We consider reaction networks containing only unimolecular reactions, i.e. a reaction r has a single substrate s_r and product p_r . Each column r of the stoichiometry matrix $\mathbf{S}$ then only contains one -1 at line s_r and one $+1$ at row p_r , the rest being only zeroes. Here, s_r and p_r are functions mapping reaction indexes to chemical indexes. Moreover, every reaction is taken to be catalyzed by a single enzyme, so that $M = R$, and the function e_r mapping reaction indexes to enzyme indexes is a bijection (without loss of generality, one could consider $e_r = r$). The structure of such unimolecular reaction networks can be characterized by an adjacency matrix $\mathbf{A} \in \{0, 1\}^{K \times K}$ defining a simple directed graph with K nodes and M edges, with $A_{j,k} = 1$ (corresponding to a directed edge) if a reaction converts chemical k into chemical j , and 0 otherwise (Fig. 1B).

To these networks, we associate bilinear reaction kinetics corresponding to enzymes operating in the linear (non-saturated) concentration-dependence regime

$$v_r(\mathbf{c}, \boldsymbol{\rho}) = \alpha_r \rho_{e_r} c_{s_r}, \quad (11)$$

where $\alpha_r > 0$ is an activity parameter corresponding, within a characteristic Michaelis-Menten picture, to the ratio k_{cat}/K_M between the catalytic rate and the Michaelis constant of the enzyme e_r . Note that Eq. (11) represents unidirectional chemical reactions, with a negligible reverse rate, implying a strongly out-of-equilibrium system. This could correspond e.g. to a system where each catalyzed reaction is fuelled by the consumption of a high-energy fuel such as ATP (in the biological context), or a light-fuelled photocatalytic system.

Closed chemical reaction networks defined in this way have a unique steady-state with nontrivial dynamics only if the associated adjacency matrix $\mathbf{A}$ corresponds to a strongly connected directed graph, i.e. if any chemical can be converted into any other through some series of reactions. If this condition is not satisfied, the reaction dynamics will lead to the accumulation of all the reaction material into a subset of dead-end chemical species, with all the other chemical species having zero concentration. We thus focus on these nontrivial networks, starting from the adjacency matrix of a strongly connected digraph, and associating one chemical species to each of its nodes and a reaction together with an associated enzyme to each of its directed edges (see Fig. 1B for an example of such a network). Networks generated according to this procedure always have a single homogeneous steady-state for the chemical concentrations, $\mathbf{c}_0$, which is stable under space-independent, purely reactive dynam-

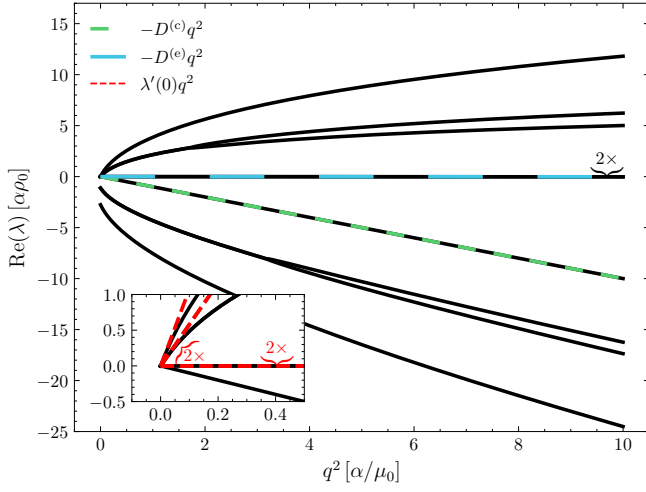


FIG. 2. Linear stability analysis for the unimolecular reaction network in Fig. 1B and the chemotactic mobilities in Fig. 1D, for linear reaction kinetics. The real parts of the eigenvalues for the full stability matrix (5) are shown as solid black lines. Closed unimolecular reaction networks exhibit one purely diffusive chemical-associated diffusive mode (dashed green) and $\max(1, M - K + 1)$ purely diffusive enzyme-associated modes (dashed cyan). Inset: Eq. (9) (dashed red) accurately captures the slope at the origin of the enzyme-associated eigenvalues of the full stability matrix. For this example, we chose $D^{(c)} = \mu_0 c_{\text{tot}}$, $D^{(e)} = D^{(c)}/500$, $\mu_p = -\mu_s = 20$.

ics according to the zero deficiency theorem [46]. As the reaction rates in Eq. (11) are linear in c_{s_r} , finding the homogeneous concentration vector $\mathbf{c}_0$ involves solving a linear system according to Eq. (4). Because all transitions are unimolecular, there is automatically a moiety conservation corresponding to the total concentration of chemicals, i.e. $c_{\text{tot}} = \sum_{k=1}^K c_{0,k}$ is conserved and is an input parameter of the model. This conservation is also reflected in the fact that each column of the stoichiometry matrix sums to zero.

Stability of networks with uniform parameters

To further reduce the parameter space and focus on the role of network topology, we make a number of simplifying assumptions in all that follows. We take $\alpha_r = \alpha$ to be equal for all enzymes, and choose all enzymes to have identical homogeneous concentrations $\rho_{0m} = \rho_0$. We further assume that each enzyme species only chemotaxes in concentration gradients of its substrate and product, and that the mobilities towards substrate and product are equal for all enzyme species, implying that the only non-null coefficients of the mobility matrix are of the form $\mu_{e_r, s_r} = \mu_s \mu_0$ and $\mu_{e_r, p_r} = \mu_p \mu_0$, where μ_s and μ_p are dimensionless parameters and $\mu_0 > 0$ is a mobility scale. Finally, we take the same diffusion coefficients $D^{(c)}$ and $D^{(e)}$ equal for all chemical and enzyme species, respec-

tively.

With these simplifying assumptions, the eigenvalue equation for $\lambda'(0)$ in (9) can be rewritten as

$$\beta \delta \tilde{\rho} = -\tilde{\eta} \delta \tilde{\rho}, \quad (12)$$

where $\lambda'(0) = \mu_0 c_{\text{tot}} \beta - D^{(e)}$, and $\tilde{\eta} \equiv (\rho_0 / c_{\text{tot}})(\eta / \mu_0)$ is a rescaled, dimensionless form of the interaction matrix that only depends on μ_s , μ_p , and the topology of the reaction network. Importantly, these parameters are sufficient to determine whether a given reaction network has the potential to display self-organization or not. Indeed, let us focus on the eigenvalue with largest real part, and define $\beta_{\text{max}} \equiv \max(\text{Re}(\beta))$. If $\beta_{\text{max}} \leq 0$, the network is guaranteed to be stable (as this implies $\text{Re}(\lambda'(0)) < 0$). On the other hand, if $\beta_{\text{max}} > 0$, such a network will display a spatial instability whenever the condition

$$\frac{D^{(e)}}{\mu_0 c_{\text{tot}}} < \beta_{\text{max}} \quad (13)$$

is satisfied, as this implies $\text{Re}(\lambda'(0)) > 0$. Thus, in systems with $\beta_{\text{max}} > 0$, an instability can occur if the chemotactic mobility scale or total chemical concentration are sufficiently large, or if the enzyme diffusion coefficient is sufficiently small.

To test the validity of Eq. (12) and its general form Eq. (9), we have compared its predictions for the slope $\lambda'(0)$ of the M enzyme-associated eigenvalues, to the $M + K$ enzyme- and chemical-associated eigenvalues $\lambda(q^2)$ calculated from the full problem, (5), for a variety of networks and choices of parameters. The results for an example network are shown in Fig. 2. Because of the conservation of the total chemical concentration, we can expect one purely diffusive chemical-associated mode going as $-D^{(c)}q^2$, which we observe (Fig. 2, dashed cyan line). Similar purely diffusive modes going as $-D^{(e)}q^2$ are observed for the enzymes. Our numerical calculations show that there is one such mode if $M \leq K$, and $M - K + 1$ such modes if $M > K$ (Fig. 2, dashed green line). Eq. (12) always correctly captures the slope at the origin of the M enzyme-associated modes (Fig. 2, inset, dashed red lines). Importantly, of all $M + K$ modes, only the enzyme-associated ones can become unstable, and the instability always occurs at $q^2 \rightarrow 0$. Overall, comparison with the numerical solution of the full eigenvalue problem in (5) confirms that, in order to study the stability of a unimolecular reaction network with uniform parameters, it is sufficient to consider the reduced eigenvalue problem in (12)

Statistics for networks with uniform parameters

In order to gather statistics on the stability of closed unimolecular reaction networks, and to uncover the influence of network-wide effects, we systematically gen-

erate (see Methods) and analyze all such reaction networks up to $K = 6$ chemicals. To this end, we generated all 1,052,145 strongly connected simple directed graphs containing up to $K = 6$ nodes, which includes 1, 5, 83, 5,048, and 1,047,008 graphs with $K = 2, 3, 4, 5$, and 6 nodes, respectively [58]. For examples of all reaction networks with $K = 3$ and $K = 4$ chemicals, see Figures S1 and S2 in the Supporting Information. As described above, with the simplifying assumption of uniform parameters, the stability of a reaction network is given by (12), and depends only on the network topology as well as the value of the dimensionless mobilities μ_s and μ_p . We first study the stability of all the generated networks under eight choices of the mobilities: $(\mu_s, \mu_p) \in \{(\pm 1, 0), (0, \pm 1), (\pm 1, \pm 1), (\pm 1, \mp 1)\}$. In the following, we call these “mobility patterns”, and for conciseness write + or – instead of +1 or –1.

A choice of reaction network and mobility pattern defines an interaction network (Fig. 1D). For an example of all interaction networks with $K = 3$ chemicals, see Figure S1 in the Supporting Information. We systematically solve the eigenvalue problem in (12) for all $8 \times 1,052,145$ interaction networks, and in each case calculate $\beta_{\max}$. This procedure yields values of $\beta_{\max}$ with magnitudes spanning nine decades. Due to this wide range, we perform the corresponding calculations symbolically, in order to distinguish small but non-null eigenvalues from null eigenvalues. Interaction networks with $\beta_{\max} > 0$ are labeled unstable.

As a point of comparison, we consider the special case of reactions in the saturated regime, $v_r(\rho) = \alpha \rho_{e_r}$, which leads to D-P interactions without network-wide effects, as previously described. In this special case, the stability of the system is fully determined by the eigenvalue problem $\beta^{\text{sat}} \delta \tilde{\rho} = -\tilde{\eta}^{\text{sat}} \delta \tilde{\rho}$ where $\tilde{\eta}^{\text{sat}} \equiv (\mu/\mu_0)\mathbf{S}$ is a dimensionless interaction matrix that depends only on μ_s , μ_p , and the topology of the reaction network. As in the case of linear kinetics, we may define $\beta_{\max}^{\text{sat}} \equiv \max(\text{Re}(\beta^{\text{sat}}))$, and classify networks with $\beta_{\max}^{\text{sat}} > 0$ as networks which are unstable in the case of saturated kinetics.

In this way we obtain the stability statistics broken down by mobility pattern shown in Table I (see Table S1 in the Supporting Information for the statistics broken down by network size). A first key observation is that slightly over 1/2 of interaction networks are unstable for linear kinetics, compared to exactly 3/8 for saturated kinetics, implying that network-wide effects facilitate self-organization. We also calculate, for each mobility pattern, the proportion of unstable networks which involve an overall self-attracting set of enzymes, i.e. whose interaction matrix $\tilde{\eta}$ has a negative trace $\sum_m \tilde{\eta}_{m,m} < 0$ (and similarly with $\tilde{\eta}^{\text{sat}}$ for saturated kinetics). Previous work has shown that, in the regime of saturated kinetics, overall self-attraction is a necessary and sufficient condition for self-organization for enzymes interacting through the production or consumption of a single chemical [34] and for enzymes participating in a metabolic cycle if they

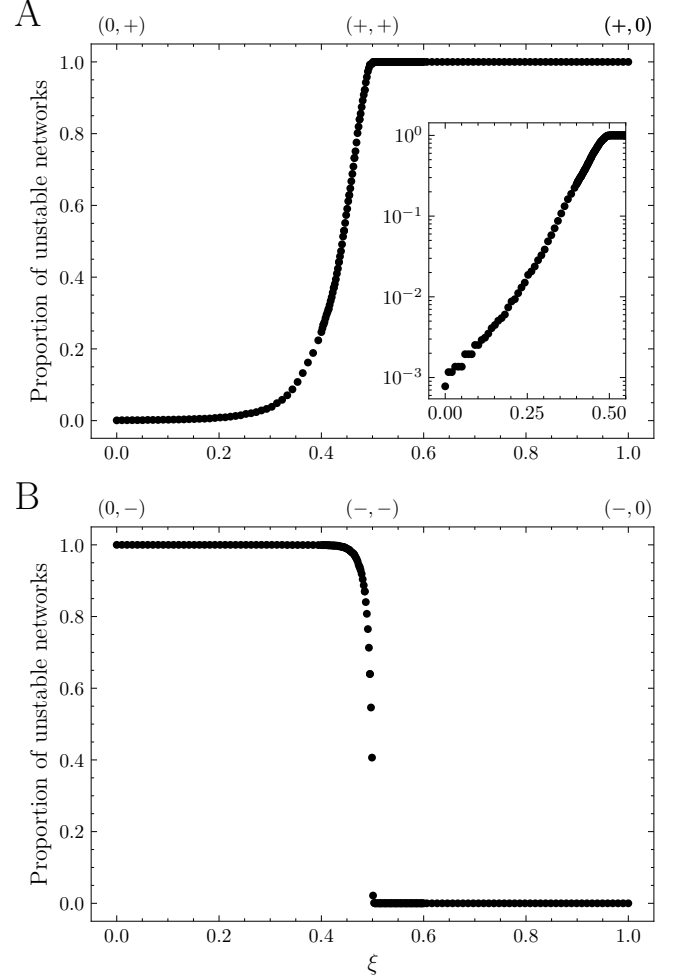


FIG. 3. Transitions in the proportion of unstable networks under continuous changes in the mobility patterns, for all 5,137 closed unimolecular reaction networks up to $K = 5$ chemicals and linear reaction kinetics. (A) $(\mu_s, \mu_p) = (\xi, 1 - \xi)$, and (B) $(\mu_s, \mu_p) = -(\xi, 1 - \xi)$. Inset in (A): magnification showing the exponential increase in the proportion of unstable networks (notice the logarithmic vertical axis).

have uniform parameters [36], but not if their parameters are not uniform [37, 38]. We find that, for saturated kinetics, an interaction network is unstable if and only if it satisfies this self-attraction condition. On the other hand, for linear kinetics, we find that self-attraction is a sufficient, but not a necessary, condition: the network-wide effects induced by the screening factor can overcome self-repulsion and lead to self-organization. Additionally, we find that network-wide effects can make a network overall self-attracting, even if it would not be so for saturated kinetics.

Focusing on individual mobility patterns, we first find that purely substrate-chemotactic enzymes [pattern $(-, 0)$], exclusively lead to stable networks for both linear and saturated kinetics. Qualitatively, this universal stability is related to the fact that such enzymes

	Linear kinetics (Full interactions)		Saturated kinetics (Naive interactions)	
Mobilities	Proportion unstable	Of which self-attr.	Proportion unstable	Of which self-attr.
$(-, 0)$	0	N/A	0	N/A
$(+, 0)$	1	1	1	1
$(0, -)$	1	1	1	1
$(0, +)$	$2.167 \cdot 10^{-4}$	0	0	N/A
$(-, -)$	$2.405 \cdot 10^{-2}$	$1.976 \cdot 10^{-4}$	0	N/A
$(+, +)$	$1 - 1.388 \cdot 10^{-4}$	$1 - 4.753 \cdot 10^{-6}$	0	N/A
$(-, +)$	0	N/A	0	N/A
$(+, -)$	1	1	1	1
Total:	0.5030	0.9939	0.3750	1

TABLE I. Statistics for the stability of closed unimolecular reaction networks up to $K = 6$ chemicals. Mobility patterns are written as (μ_s, μ_p) . For each mobility pattern, the proportion of unstable reaction networks is given out of all 1,052,145 networks, whereas the total is given out of all $8 \times 1,052,145$ combinations of mobility patterns and reaction networks, i.e. interaction networks. An interaction network is considered self-attracting if it verifies $\sum_{m=1}^M \eta_{m,m} < 0$.

are necessarily self-repelling, as they consume the chemical that they are attracted by. Conversely, purely substrate-antichemotactic enzymes [pattern $(+, 0)$] consume a chemical by which they are repelled, and are thus self-attracting and unstable for both types of reaction kinetics. Analogous qualitative arguments based on self-interaction can explain the behavior of purely product-chemotactic enzymes [pattern $(0, -)$], and the behavior of the antisymmetric patterns $(-, +)$ and $(+, -)$.

Surprisingly, in the case of product-antichemotactic enzymes [pattern $(0, +)$], a small proportion of networks, none of which are self-attracting, is unstable only under linear kinetics. The presence of self-repelling and unstable networks demonstrates that network-wide effects induced by the substrate concentration dependence of the reaction kinetics can provide an alternate route towards self-organization. This mechanism for self-organization without self-attraction is distinct from the one found in Refs. 37, 38, which was based on saturated reaction kinetics, but non-identical activities and mobilities for the enzymes. Note that, since all networks are unstable for pattern $(0, -)$, for the small subset of networks which is unstable for pattern $(0, +)$ an instability can occur both if the enzymes are attracted to or repelled from their product.

For networks with symmetric mobility patterns $(-, -)$, enzymes do not self-interact under saturated kinetics, as in this case the self-repulsion arising from substrate chemotaxis balances the self-attraction arising from product chemotaxis. For this mobility pattern, we find that network-wide effects can create attractive self-interactions, which constitutes yet another mechanism through which network-wide effects enable self-organization. This new mechanism becomes especially prevalent and impactful in the other symmetric mobility pattern, $(+, +)$, for which all but a handful of networks are unstable and self-attracting for linear kinetics, even

if all networks are stable for saturated kinetics. Again in this case, we find that there are networks that are unstable under both mobility patterns $(+, +)$ and $(-, -)$, implying that these reaction networks can show an instability both if the enzymes are attracted to or repelled from their substrate and product.

In order to characterize the transition between stability and instability in the case of linear kinetics where network-wide effects are important, we measure the proportion of unstable reaction networks under a continuous change of the mobilities between each pair of previously studied mobility patterns. The changes in mobility patterns from $(0, +)$ to $(+, 0)$ and from $(0, -)$ to $(-, 0)$ are of particular interest, as both cases involve a transition from (almost) all reaction networks being stable to all being unstable. We continuously vary the mobilities according to a parameter $\xi \in \{0, 1\}$ by imposing $(\mu_s, \mu_p) = (\xi, 1 - \xi)$ (Fig. 3A) or $(\mu_s, \mu_p) = -(\xi, 1 - \xi)$ (Fig. 3B). In both cases, $\xi = 0.5$ corresponds to a symmetric pattern, respectively $(+, +)$ and $(-, -)$. Due to the large number of calculations involved, we limited this study to the 5,137 reaction networks containing up to $K = 5$ chemical species. We find that, for the transition between $(0, +)$ and $(+, 0)$, the proportion of unstable networks undergoes an exponential (Fig. 3A, inset) increase from 0 to 1 between $\xi = 0$ and 0.5. For the transition between $(0, -)$ and $(-, 0)$, on the other hand, we observe an abrupt change from all networks being stable to all unstable around $\xi \approx 0.5$.

Distribution of growth rates for networks with uniform parameters

To further quantify the degree to which different interaction networks are unstable, we computed the probability distributions of $\beta_{\max}$ for unstable interaction networks

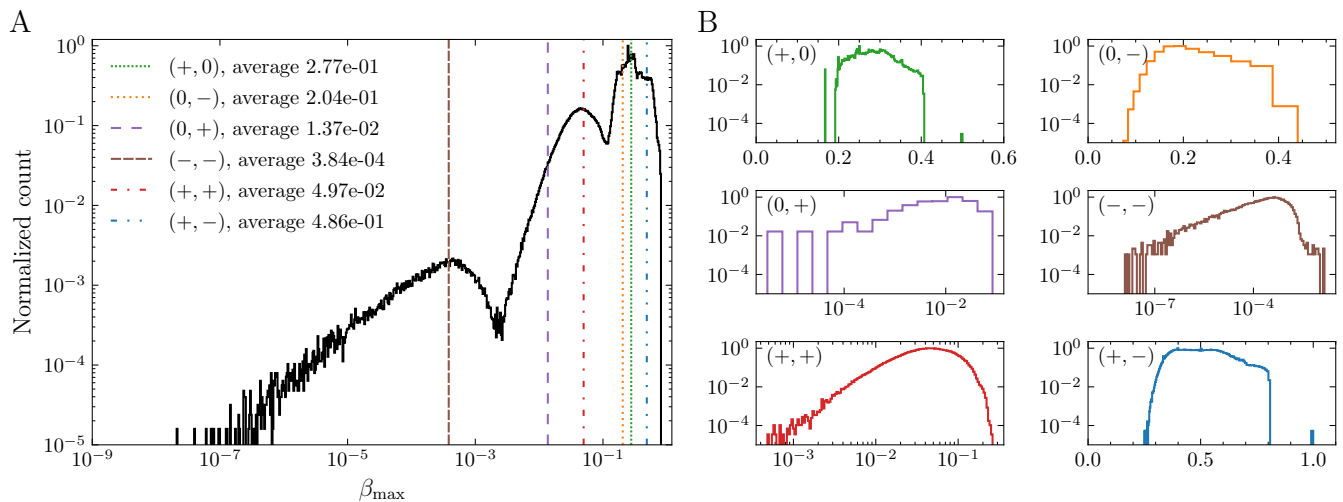


FIG. 4. Distribution of $\beta_{\max} > 0$ for all the unstable closed unimolecular reaction networks up to $K = 6$ chemicals that are unstable and all six mobility patterns that allow for instabilities. (A) Distribution over all mobility patterns, with the average $\beta_{\max}$ for each mobility pattern shown as a dashed vertical line. (B) Distribution for each of the six unstable mobility patterns. Note the different range and the use of linear and logarithmic scales for the x-axis in each case.

(i.e. confined to values $\beta_{\max} > 0$). Indeed, through the instability condition in Eq. (13), larger values of $\beta_{\max}$ correspond to wider ranges of values of $D^{(e)}$, μ_0 , and c_{tot} that can induce an instability. In this analysis, we considered all reaction networks with up to $K = 6$ chemicals, and the six mobility patterns that can become unstable (Table I). When pooling together all six mobility patterns, the resulting distribution is fat-tailed (Fig. 4A), spanning the range $\beta_{\max} \in (10^{-9}, 1)$, with an overall average of $\langle \beta_{\max} \rangle = 2.527 \cdot 10^{-1}$.

The wide range of $\beta_{\max}$ -values observed can in part be explained by considering the distributions associated with individual mobility patterns (Fig. 4B). Each pattern displays a markedly different distribution, with mean values and typical ranges that are consistent with qualitative arguments based on the presence of self-attraction in the interaction network. For the mobility patterns $(+, 0)$, $(0, -)$, and $(+, -)$, which are associated with self-attracting species for both linear and saturated reaction kinetics, the $\beta_{\max}$ -values are all on the order of 10^{-1} , and have a mean value on the order of the average of the overall distribution. In particular, patterns $(+, 0)$ and $(0, -)$, which correspond to similarly-strong degrees of self-attraction, broadly span the same range of $\beta_{\max}$ -values. The pattern $(+, -)$, on the other hand, makes the enzymes the most self-attracting, as it corresponds to species both repelled by the chemical they consume and attracted to the one they produce. As a consequence, its $\beta_{\max}$ -distribution features a markedly larger mean and contains the overall largest values.

We now focus on the mobility patterns $(-, -)$, $(0, +)$, and $(+, +)$, which lead to self-organization only when network-wide effects are present. The associated $\beta_{\max}$ -distributions have means that are several orders of mag-

nitude smaller than the previously discussed patterns, and span wider ranges (typically several decades). The pattern $(+, +)$, for which network-wide effects induce self-attraction, displays significantly larger $\beta_{\max}$ -values than $(-, -)$ and $(0, +)$, for which network-wide effects enable self-organization in spite of self-repulsion. From these statistics we deduce that, while network-wide effects can indeed favor instabilities in systems without D-P self-attraction, the instabilities in these systems will be harder to observe, because their smaller $\beta_{\max}$ -values require either significantly smaller values of the enzyme diffusion coefficient $D^{(e)}$, larger values of their chemotactic mobility μ_0 , or larger total amounts of reactant c_{tot} , as seen from the instability condition in Eq. (13). This is particularly true for interaction networks which stay self-repelling even in the presence of network-wide effects: such networks are rarely unstable, and only feature very small $\beta_{\max}$ -values. We also speculate that the wider distribution of $\beta_{\max}$ for these mobility patterns is due to an increased sensitivity to the network structure, resulting from the importance of network-wide effects for such patterns.

DISCUSSION

In this work, we have shown that the spatial dynamics of the components of a catalyzed chemical reaction network (including diffusion for the chemicals, and diffusion as well chemotaxis in response to chemical gradients for the enzymes) can lead to self-organization of the enzymes through a spatial instability. We first characterized this instability process by applying linear stability analysis on generic network structures and chemical reaction rates, obtaining a concise criterion for the observation of spa-

tial instabilities in such systems, see Eq. (9). From the structure of the effective enzyme-enzyme interaction matrix involved in this stability equation, we deduced that perturbations of the enzyme concentration fields are effectively propagated through the chemical reaction network, giving rise to network-wide effects. In particular, two enzymes can effectively interact with each other even if they do not undergo chemotaxis towards each other’s substrates or products. We then focused on unimolecular reaction networks with linear reaction rates. We systematically determined the stability of all strongly connected reaction networks involving six or fewer chemical species for various patterns of chemotactic mobilities, finding that the aforementioned network-wide effects can facilitate self-organization. By interpolating between different mobility values, we discovered that continuous changes in mobilities can either lead to an exponential increase of the number of unstable networks, or to a sharp transition. We finally quantified the extent to which different mobility patterns are unstable which, due to network-wide effects, can span nine orders of magnitude.

In order to gather gain an initial understanding of self-organization processes in catalyzed reaction networks, we limited most of our quantitative exploration to small, closed, unimolecular reaction networks. Real-life biochemical reaction networks, however, are typically larger and more complex than the ones we studied, and involve bimolecular reactions as well as external inputs and outputs of reaction material. Our general framework culminating in Eq. (9), however, can be directly applied to bimolecular reactions, and easily extended to account for non-enzymatic chemical sources or sinks such as chemostats. A possible approach for the analysis of large realistic networks could be to use small, strongly connected networks as building blocks for more complex topologies closer to metabolic pathways, akin to the network motifs formalism used in systems biology [38, 59, 60]. Another relevant extension of this work is to determine how enzyme self-organization influences the output of biochemical pathways. Concentrating enzymes into condensates has indeed been shown to result in localized metabolic factories in which pathway outputs are increased through the channeling of reaction intermediates between enzymes [61, 62], or to favor one branch of a pathway with respect to another [20]. The effect of

spatial structure on the output of biochemical pathways has so far mostly been studied by starting from a pre-defined spatial structure. However, recent works have shown that a single enzyme catalyzing a one-step reaction can undergo phase separation, lowering its activity in the process [44]. This work is a first step towards bridging these two approaches. It will be of particular interest to determine which kinds of pathway regulations are possible through phase separation in multi-enzyme systems, and how spontaneously-arising spatial structure affects metabolic output.

MATERIALS AND METHODS

Generation of strongly connected reaction networks

Above, we systematically analyzed the stability of minimal closed networks involving unimolecular reactions. To do so, we generated the adjacency matrices corresponding to all 1,052,145 strongly connected digraphs with $K = 2$ to 6 nodes, where the nodes take the role of the chemicals involved in the reactions. This was done by first using the command line graph isomorphism program nauty [63], which can generate one component of each graph equivalence class with a certain number of nodes, yielding the set of all unique digraphs with a given size. We then applied a Julia implementation of Tarjan’s algorithm [64] to each obtained digraph, which returned the set of strongly connected components for each digraph. Only digraphs with a single strongly connected component (which are by definition strongly connected) are then kept. We verified that the number of digraphs obtained with this method is consistent with the value reported in literature, obtained from [58], and that the adjacency matrices obtained for $K = 3$ and 4 nodes correspond to the recorded examples of strongly connected digraphs.

ACKNOWLEDGEMENTS

We acknowledge support from the Max Planck School Matter to Life and the MaxSynBio Consortium which are jointly funded by the Federal Ministry of Education and Research (BMBF) of Germany and the Max Planck Society.

-
- [1] R. Phillips, J. Kondev, J. Theriot, H. G. Garcia, and N. Orme, *Physical Biology of the Cell* (Garland Science, 2012).
 - [2] M. Newman, *Networks* (Oxford University Press, 2018).
 - [3] E. Tang, J. Agudo-Canalejo, and R. Golestanian, *Physical Review X* **11**, 031015 (2021).
 - [4] J. S. van Zon, D. K. Lubensky, P. R. H. Altena, and P. R. ten Wolde, *Proceedings of the National Academy of Sciences* **104**, 7420 (2007).
 - [5] E. H. Davidson, *The Regulatory Genome: Gene Regulatory Networks in Development and Evolution* (Elsevier, 2010).
 - [6] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A.-L. Barabási, *Nature* **407**, 651 (2000).
 - [7] B. Alberts, *Molecular Biology of the Cell*, 6th ed. (W.W. Norton & Company, New York, 2017).
 - [8] J. Stelling, S. Klamt, K. Bettenbrock, S. Schuster, and E. D. Gilles, *Nature* **420**, 190 (2002).

- [9] D. A. Fell, *Metabolic Control Analysis of Exponential Growth and Product Formation* (2018).
- [10] R. Moreno-Sánchez, E. Saavedra, S. Rodríguez-Enríquez, and V. Olín-Sandoval, *Journal of Biomedicine and Biotechnology* **2008**, e597913 (2008).
- [11] D. Visser and J. J. Heijnen, *Metabolic Engineering* **4**, 114 (2002).
- [12] U. Alon, *An Introduction to Systems Biology: Design Principles of Biological Circuits*, 2nd ed. (Chapman and Hall/CRC, New York, 2019).
- [13] S. F. Banani, H. O. Lee, A. A. Hyman, and M. K. Rosen, *Nature Reviews Molecular Cell Biology* **18**, 285 (2017).
- [14] J. Söding, D. Zwicker, S. Sohrabi-Jahromi, M. Boehning, and J. Kirschbaum, *Trends in cell biology* **30**, 4 (2020).
- [15] B. G. O'Flynn and T. Mittag, *Current opinion in cell biology* **69**, 70 (2021).
- [16] L. J. Sweetlove and A. R. Fernie, *Nature communications* **9**, 2136 (2018).
- [17] A. M. Pedley, V. Pareek, and S. J. Benkovic, *Annual Review of Biochemistry* **91**, 89 (2022).
- [18] M. Castellana, M. Z. Wilson, Y. Xu, P. Joshi, I. M. Cristea, J. D. Rabinowitz, Z. Gitai, and N. S. Wingreen, *Nature biotechnology* **32**, 1011 (2014).
- [19] V. Pareek, Z. Sha, J. He, N. S. Wingreen, and S. J. Benkovic, *Molecular cell* **81**, 3775 (2021).
- [20] F. Hinzpeter, F. Tostevin, and U. Gerland, *Journal of The Royal Society Interface* **16**, 20190444 (2019).
- [21] A. Testa, M. Dindo, A. A. Rebane, B. Nasouri, R. W. Style, R. Golestanian, E. R. Dufresne, and P. Laurino, *Nat. Commun.* **12**, 6293 (2021).
- [22] R. Kokkoorakunnel, S. Saha, and R. Golestanian, *Phys. Rev. Lett.* **133**, 058401 (2024).
- [23] J. Agudo-Canalejo, T. Adeleke-Larodo, P. Illien, and R. Golestanian, *Accounts of chemical research* **51**, 2365 (2018).
- [24] E. F. Keller and L. A. Segel, *Journal of theoretical biology* **26**, 399 (1970).
- [25] R. Golestanian, *Phys. Rev. Lett.* **108**, 038303 (2012).
- [26] J. Palacci, S. Sacanna, A. P. Steinberg, D. J. Pine, and P. M. Chaikin, *Science* **339**, 936 (2013).
- [27] F. Schmidt, B. Liebchen, H. Löwen, and G. Volpe, *The Journal of chemical physics* **150** (2019).
- [28] R. Golestanian, in *Active Matter and Nonequilibrium Statistical Physics. Lecture Notes of the Les Houches Summer School* (Oxford University Press, UK, 2022) pp. 230–293.
- [29] H. Yu, K. Jo, K. L. Kounovsky, J. J. d. Pablo, and D. C. Schwartz, *Journal of the American Chemical Society* **131**, 5722 (2009).
- [30] K. K. Dey, S. Das, M. F. Poyton, S. Sengupta, P. J. Butler, P. S. Cremer, and A. Sen, *ACS nano* **8**, 11941 (2014).
- [31] X. Zhao, H. Palacci, V. Yadav, M. M. Spiering, M. K. Gilson, P. J. Butler, H. Hess, S. J. Benkovic, and A. Sen, *Nat. Chem.* **10**, 311 (2018).
- [32] A.-Y. Jee, S. Dutta, Y.-K. Cho, T. Thlusty, and S. Granick, *Proceedings of the National Academy of Sciences* **115**, 14 (2018).
- [33] J. Agudo-Canalejo, P. Illien, and R. Golestanian, *Nano letters* **18**, 2711 (2018).
- [34] J. Agudo-Canalejo and R. Golestanian, *Physical Review Letters* **123**, 018101 (2019).
- [35] V. Ouazan-Reboul, J. Agudo-Canalejo, and R. Golestanian, *European Physical Journal E* **44**, 113 (2021).
- [36] V. Ouazan-Reboul, J. Agudo-Canalejo, and R. Golestanian, *Nature Communications* **14**, 4496 (2023).
- [37] V. Ouazan-Reboul, R. Golestanian, and J. Agudo-Canalejo, *Physical Review Letters* **131**, 128301 (2023).
- [38] V. Ouazan-Reboul, R. Golestanian, and J. Agudo-Canalejo, *New Journal of Physics* **25**, 103013 (2023).
- [39] G. Giunta, H. Seyed-Allaei, and U. Gerland, *Communications Physics* **3**, 1 (2020).
- [40] C. Kocher, L. Agozzino, and K. Dill, *The Journal of Physical Chemistry B* **125**, 8781 (2021).
- [41] R. Soto and R. Golestanian, *Phys. Rev. Lett.* **112**, 068301 (2014).
- [42] L. Parkavousi, N. Rana, R. Golestanian, and S. Saha, *Phys. Rev. Lett.* **134**, 148301 (2025).
- [43] S. Osat and R. Golestanian, *Nat. Nanotechnol.* **18**, 79 (2022).
- [44] M. W. Cotton, R. Golestanian, and J. Agudo-Canalejo, *Physical Review Letters* **129**, 158101 (2022).
- [45] F. Blanchini and G. Giordano, *Automatica* **126**, 109376 (2021).
- [46] M. Feinberg, *Foundations of Chemical Reaction Network Theory* (2019).
- [47] D. Angeli, *European Journal of Control* **15**, 398 (2009).
- [48] T. Yu, P. Chuphal, S. Thakur, S. Y. Reigh, D. P. Singh, and P. Fischer, *Chemical Communications* **54**, 11933 (2018).
- [49] C. H. Meredith, P. G. Moerman, J. Groenewold, Y.-J. Chiu, W. K. Kegel, A. van Blaaderen, and L. D. Zarzar, *Nature Chemistry* **12**, 1136 (2020).
- [50] B. L. Clarke, in *Advances in Chemical Physics* (John Wiley & Sons, Ltd, 1980) pp. 1–215.
- [51] H. M. Sauro, *Biosystems* **33**, 55 (1994).
- [52] C. Reder, *Journal of Theoretical Biology* **135**, 175 (1988).
- [53] A. Cornish-bowden and J.-H. S. Hofmeyr, *Journal of Theoretical Biology* **216**, 179 (2002).
- [54] R. Soto and R. Golestanian, *Phys. Rev. E* **91**, 052304 (2015).
- [55] S. Saha, J. Agudo-Canalejo, and R. Golestanian, *Physical Review X* **10**, 041009 (2020).
- [56] Z. You, A. Baskaran, and M. C. Marchetti, *Proceedings of the National Academy of Sciences* **117**, 19767 (2020).
- [57] S. Saha and R. Golestanian, *Nat. Commun.* **16**, 7310 (2025).
- [58] Number of unlabeled strongly connected digraphs with n nodes, entry A035512 in the on-line encyclopedia of integer sequences (2023), <http://oeis.org/A035512>.
- [59] U. Alon, *Nature Reviews Genetics* **8**, 450 (2007).
- [60] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, *Science* **298**, 824 (2002).
- [61] Y. Zhang and A. R. Fernie, *Plant Communications Plant Genome Biology*, **2**, 100081 (2021).
- [62] T. S. Harmon and F. Jülicher, *Physical Review Letters* **128**, 108102 (2022).
- [63] B. D. McKay and A. Piperno, *Journal of Symbolic Computation* **60**, 94 (2014).
- [64] R. Tarjan, *SIAM Journal on Computing* **1**, 146 (1972).

SUPPLEMENTAL INFORMATION

Proportion of unstable networks with K chemicals						
Mobilities	$K = 2$	$K = 3$	$K = 4$	$K = 5$	$K = 6$	$K \leq 6$
$(-, 0)$	0	0	0	0	0	0
$(+, 0)$	1	1	1	1	1	1
$(0, -)$	1	1	1	1	1	1
$(0, +)$	0	0	0	$7.924 \cdot 10^{-4}$	$2.139 \cdot 10^{-4}$	$2.167 \cdot 10^{-4}$
$(-, -)$	0	0	0	$8.518 \cdot 10^{-2}$	$2.376 \cdot 10^{-2}$	$2.405 \cdot 10^{-2}$
$(+, +)$	0	0.6	$9.277 \cdot 10^{-1}$	$9.958 \cdot 10^{-1}$	$1 - 1.108 \cdot 10^{-4}$	$1 - 1.388 \cdot 10^{-4}$
$(-, +)$	0	0	0	0	0	0
$(+, -)$	1	1	1	1	1	1
Total:	0.375	0.45	0.491	0.5102	0.5030	0.5030

TABLE S1. Stability statistics broken down by number of chemicals.

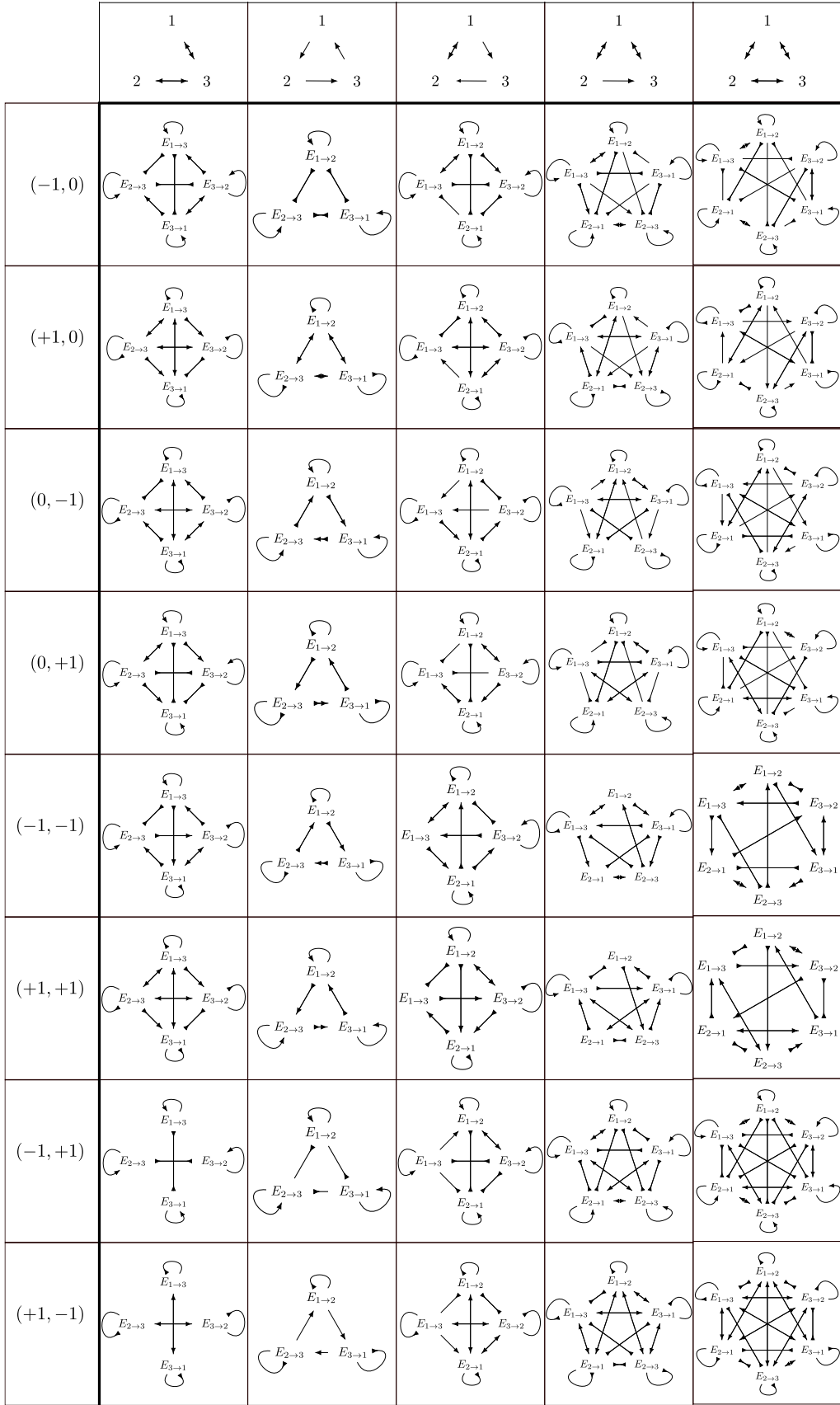


FIG. S1. All $8 \times 5 = 40$ interaction networks for $K = 3$ chemical species. Top row represents the catalyzed reaction networks between chemicals indexed 1, 2, and 3 in the form of digraphs, where a directed link between chemicals i and j represents the existence a catalyzed reaction converting i into j . Leftmost column corresponds to the relative substrate and product mobilities (μ_s, μ_p) . In the main portion of the table, catalyzed reaction networks between enzymes $E_{i \rightarrow j}$ catalyzing the conversion of chemical i into chemical j are shown. The convention for attracting and repulsing interactions is the same as in the main text.

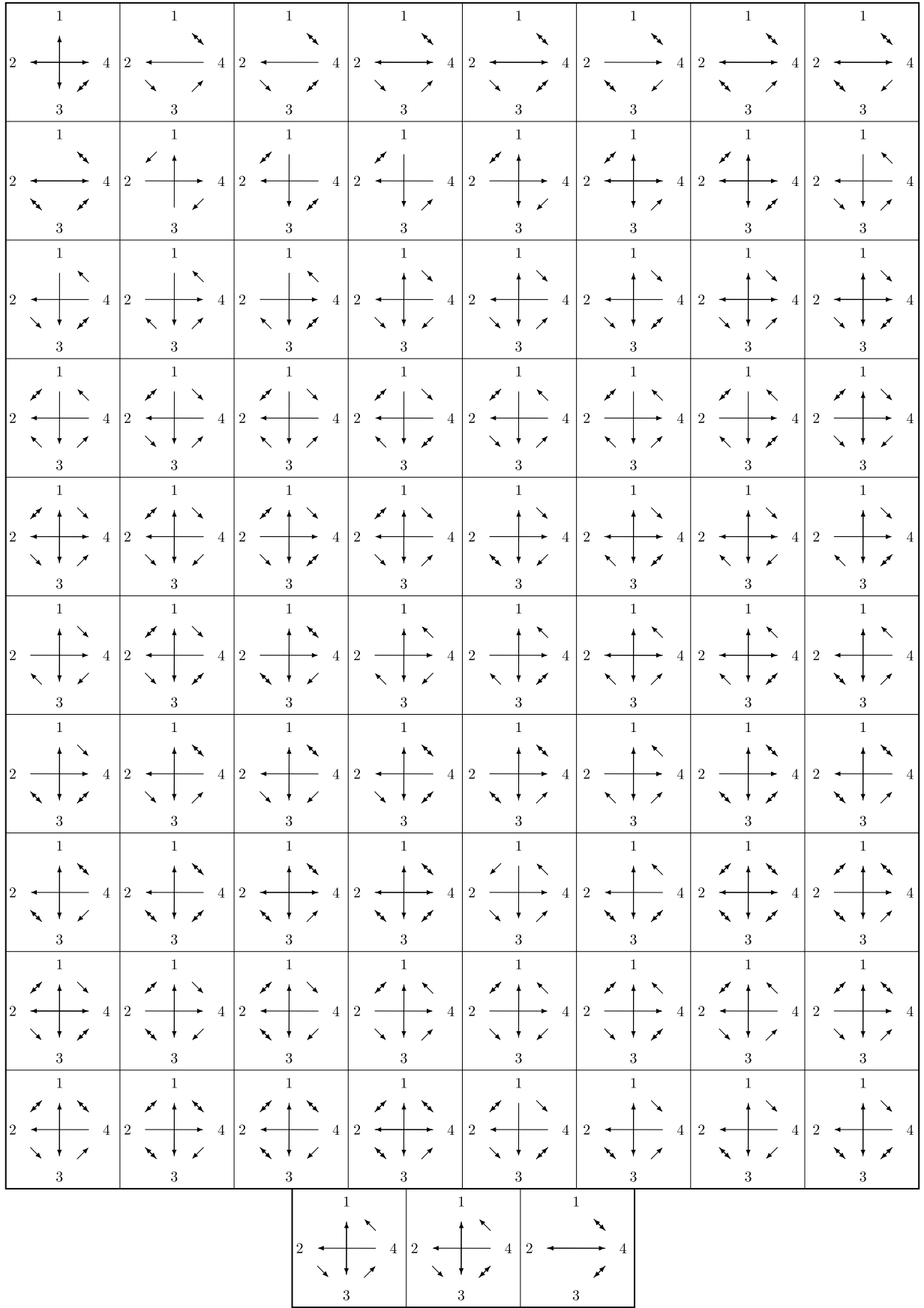


FIG. S2. All 83 reaction networks for $K = 4$ chemical species. The representation of the chemical networks is the same as in Fig. S1, top row.