




On Non-ideal Chemical-Reaction Networks and Phase Separation

Ander Movilla Miangolarra^{1,2,3}  · Michele Castellana^{1,2}

Received: 5 May 2022 / Accepted: 14 November 2022 / Published online: 29 November 2022
© The Author(s) 2022

Abstract

Much of the theory on chemical-reaction networks (CRNs) has been developed in the ideal-solution limit, where interactions between the solutes are negligible. However, there is a large variety of phenomena in biological cells and soft-matter physics which appear to deviate from the ideal-solution behaviour. Particularly striking is the case of liquid-liquid phase separation, which is typically caused by inter-particle interactions. Here, we revisit a number of known results in the domain of ideal CRNs, and we generalise and adapt them to arbitrary interactions between the solutes which stem from a given free energy. We start by reviewing the theory of chemical reaction networks, linking it to concepts in statistical physics. Then we obtain a number of new results for non-ideal complex-balanced networks, where the creation and annihilation rates are equal for all chemical complexes which appear as reactants or products in the CRN. Among these is the form of the steady-state probability distribution and Lyapunov functions for such networks. Finally, this allows us to draw a phase diagram for complex-balanced reaction-diffusion systems based on the minimisation of such Lyapunov function, with a rationale similar to that of equilibrium thermodynamics but for systems that may sustain non-equilibrium chemical currents at steady state. In addition, we show that for complex-balanced networks at steady-state, there are no diffusion currents.

Keywords Statistical physics · Nonequilibrium systems · Stochastic systems

Communicated by Udo Seifert.

✉ Ander Movilla Miangolarra
ander.movilla-miangolarra@jic.ac.uk
Michele Castellana
michele.castellana@curie.fr

¹ Laboratoire Physico-Chimie Curie, Institut Curie, PSL Research University, 11 rue Pierre et Marie Curie, 75005 Paris, France

² Sorbonne Universités, UPMC Univ. Paris 06, Paris, France

³ Present Address: Computational and Systems Biology Department, John Innes Centre, Norwich, UK

1 Introduction

The cytoplasm of a cell does not behave like an ideal solution [1], since, in many cases, interactions among the solutes cannot be neglected. Indeed, in the cytoplasm there is a plethora of interactions among proteins, other macromolecules, and ions. Some of the most common interactions that are relevant in the cellular cytoplasm are steric and crowding effects [2, 3], as well as electrostatic interactions [4, 5]. Arguably, the most striking phenomenon caused by these interactions is the emergence of phase-separated condensates, also known as membraneless organelles in the cell-biology literature, which are now widely studied [6–8]. The composition of these membraneless organelles is different from the one of the cytoplasm, because they are typically enriched in a specific type of molecules while they exclude others [9]. Moreover, it has been hypothesised that such organelles spatially control biochemical reactions, by modulating their rates and specificity within the condensate [10–13].

Given their important role in the internal spatial organisation of cells, the regulation of phase-separation phenomena is crucial for many cellular functions. One of the ways in which cells can dynamically control the onset, composition and function of membraneless organelles is through chemical reactions, notably post-translational modifications like phosphorylation [14, 15] or methylation [16]. However, the rates of these chemical reactions are also affected by the fact that the solution is non-ideal, i.e., by the interactions between solutes, requiring a paradigm beyond that of ideal chemical reactions networks (CRNs). Moreover, in this type of system, diffusion plays a crucial role (since it organises the molecules in areas of high and low concentration), suggesting that an effort should be made to understand these non-ideal reaction-diffusion systems from the theoretical standpoint.

Most of these theoretical efforts [17–20] have been based on effective reaction-diffusion models that can describe patterning and non-equilibrium phenomena in a simple way, but lack thermodynamic consistency. More precisely, in these approaches the reaction dynamics is modelled with mass-action kinetics (MAK), which implicitly assumes that the solution is ideal (while the interaction-influenced diffusion that drives phase-separation is not), which leads to the aforementioned lack of consistency. Early progress in reconciling the spatial patterns predicted by these models with a thermodynamically consistent description was limited to a linear-stability analysis of binary systems [21]. More recently, some works aimed at establishing a deterministic theory for non-ideal CRNs [22], the relation between phase coexistence and chemical kinetics [23], and exploring minimal examples for pattern formation with non-ideal CRNs [24–26]. Nevertheless, the link between non-equilibrium CRNs and phase separation has not yet been elucidated in full generality.

Here, we aim at building a thermodynamically consistent framework for interacting reaction-diffusion systems which may exhibit phase-separation at steady state. Therefore, in this framework, in the same way diffusion is governed by a free energy (that takes into account the interactions), the dynamics of the chemical reactions must also reflect this free-energetic dependency. Here, previous efforts are complemented by analysing the behaviour of non-ideal CRNs in the stochastic limit, in an effort to build a more complete theory. This explicit description of non-ideality in the CRN allows us to naturally adapt and generalise the results from the well-established theory of ideal CRNs. We do so by first constructing a framework such that, in the absence of explicit non-equilibrium driving, the system relaxes to thermodynamical equilibrium. Unlike much of the previous literature, which deals with particular reaction motifs (with a few exceptions, e.g. Ref. [27]), we focused on a broad class of networks, named complex-balanced networks, for which the steady-state creation and annihilation rate of each chemical complex are equal. For this type of non-dilute CRNs,

we derive the steady-state probability distributions and Lyapunov functionals, which allows us to obtain the steady-state concentration profiles.

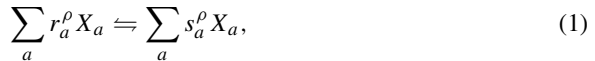
The paper is organised as follows: The first half of the paper consists of a review of the theory of CRNs and its connection to thermodynamical quantities. In Sect. 2 we describe the dynamics of spatially homogeneous CRNs in the stochastic and deterministic limit, introduce the concept of complex balance, and recall the main features of MAK. In Sect. 3, we impose thermodynamical constraints on the reaction rates for CRNs at equilibrium, by consistently relating these rates to the free energy, and discuss how they can be modified in non-equilibrium settings.

Sections 4 and 5 contain the main novel results of the paper. In Sect. 4 we generalise to non-ideal CRNs the known result for the steady-state distribution of complex-balanced networks. Building on this result, in Sect. 5 we propose a candidate Lyapunov function of complex-balanced systems. In the same Section, we generalise the previous Lyapunov function to systems with spatial inhomogeneities, and derive the resulting phase diagram for a non-equilibrium, complex-balanced, chemically reactive mixture.

Finally, in Sect. 6 we discuss the interpretations and implications of our results.

2 Chemical-Reaction Networks

A chemical-reaction network (CRN) is composed of N chemical species and M reaction pathways; which we assume reversible, for a better alignment with thermodynamic principles. A reaction within the CRN, denoted by the label ρ , is specified as follows:



where X_a , $a = 1, \dots, N$ is one the N species in the network. In the rest of this paper, the indexes a and b will be used for chemical species only.

The matrices r_a^ρ and s_a^ρ denote the number of particles of each species participating in the forward and backward reaction, respectively, i.e., r_a^ρ specifies the number of reactants of type a in the forward reaction ρ , and s_a^ρ , that of the products of type a in the backward reaction. Note that, given that the reactions are taken to be reversible, the distinction between reactants and products is arbitrary.

The amount of particles of species a created along the forward reaction is denoted by

$$v_a^\rho = s_a^\rho - r_a^\rho. \tag{2}$$

We also introduce the vectors $\mathbf{v}^\rho = (v_1^\rho, \dots, v_N^\rho)$, $\mathbf{r}^\rho = (r_1^\rho \dots r_N^\rho)$ and $\mathbf{s}^\rho = (s_1^\rho, \dots, s_N^\rho)$, and the matrices

$$\mathbf{V} = \begin{pmatrix} \vdots \\ \mathbf{v}^\rho \\ \vdots \end{pmatrix}, \mathbf{R} = \begin{pmatrix} \vdots \\ \mathbf{r}^\rho \\ \vdots \end{pmatrix}, \mathbf{S} = \begin{pmatrix} \vdots \\ \mathbf{s}^\rho \\ \vdots \end{pmatrix}. \tag{3}$$

Finally, we define a complex \mathbf{z} as the number and type of particles that participate in a chemical reaction as either reactants ($\mathbf{z} = \mathbf{r}^\rho$) or products ($\mathbf{z} = \mathbf{s}^\rho$). A single complex \mathbf{z} may appear in more than one reaction within the network. In order to clarify the definitions above, we illustrate them for the following network.

2.1 Example

For the CRN



the vectors of reactants are

$$\mathbf{r}^1 = (1, 1, 0, 0), \mathbf{r}^2 = (0, 1, 0, 0), \quad (6)$$

and the product vectors are

$$\mathbf{s}^1 = (0, 0, 1, 0), \mathbf{s}^2 = (0, 0, 0, 1), \quad (7)$$

where each of the entries in the vector correspond to different species and the reactions are labelled by the superindices. Finally, we have the vectors

$$\mathbf{v}^1 = (-1, -1, 1, 0), \mathbf{v}^2 = (0, -1, 0, 1), \quad (8)$$

that specify the net amount of particles of each species created by each of the reactions occurring once in the forward direction.

Building on the definitions above, in what follows we introduce the stochastic and deterministic description of a CRN, the starting point of the rest of this work.

2.2 Stochastic Description

If the chemical species in the solution diffuse fast (with respect to the typical timescale of chemical reactions) and is stirred regularly, the system may be considered to be well mixed and it can be described in terms of a single homogeneous concentration of each of the species across space. Then, a state of the system—the number of particles of each type—is determined by the vector

$$\mathbf{n} = (n_1, \dots, n_N). \quad (9)$$

Each state \mathbf{n} of the system has a probability measure $P(\mathbf{n}, t)$ at any instant of time t . The dynamics for the probability of states of homogeneous CRNs is given by the Chemical Master Equation (CME), which reads [28]:

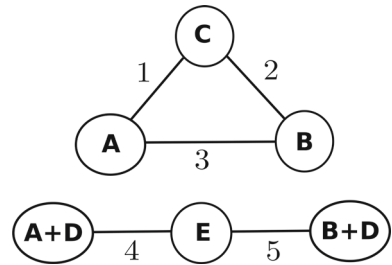
$$\begin{aligned} \frac{\partial P(\mathbf{n}, t)}{\partial t} = & \sum_{\rho} f_{+\rho}(\mathbf{n} - \mathbf{v}^{\rho}) P(\mathbf{n} - \mathbf{v}^{\rho}) + \sum_{\rho} f_{-\rho}(\mathbf{n} + \mathbf{v}^{\rho}) P(\mathbf{n} + \mathbf{v}^{\rho}) \\ & - \sum_{\rho} [f_{+\rho}(\mathbf{n}) + f_{-\rho}(\mathbf{n})] P(\mathbf{n}), \end{aligned} \quad (10)$$

where the summations over ρ run over all reactions in the CRN, \mathbf{v} is given by Eq. (2), and the rate of the transitions in the network is given by the *propensity functions* $f_{\pm\rho}$, $+\rho$ corresponding to the forward direction of the reaction and $-\rho$ corresponding to the backward direction. Eq. (10) is different to other statements of the CME because we made explicit the fact that every reaction is reversible.

2.3 Deterministic Description

For large particle numbers, by averaging both sides of Eq. (10) and assuming vanishing correlations according to the mean-field picture—assumptions that are supposed to be accurate

Fig. 1 Graphical representation of a chemical-reaction network (CRN). The CRN has six complexes, where each is represented as a node in the graph: A, B, C, A + D, E and B + D. The five reactions present in the CRN are numbered



in the large-particle number limit—one can work out a set of equations for the concentrations $c_a = n_a/V$ in the macroscopic limit, where both n_a and V are large. In this limit, the state of the system is specified by the concentrations c_1, \dots, c_N , and one obtains the following classical set of equations for the dynamics of the concentrations in a CRN [29]:

$$\frac{\partial c_a}{\partial t} = \sum_{\rho} v_a^{\rho} (J_{+\rho} - J_{-\rho}), \tag{11}$$

where the currents J still need to be determined. While both the deterministic and stochastic descriptions refer to the same system, the former one is only accurate for large particle numbers, also known as the thermodynamic limit, where fluctuations are negligible.

2.4 Complex Balance

In a CRN, a *complex* is a set of chemical species and their respective particle numbers, which take part in a reaction, as either reactants or products. Its most general expression is the vector

$$\mathbf{z} = (z_1, \dots, z_a, \dots), \tag{12}$$

where the index a runs over all chemical species and the integer z_a is the number of molecules of the species X_a in the complex \mathbf{z} . Any CRN can be represented as a graph whose nodes denote the complexes that take part in the reactions, where there is an edge between two complexes if and only if there is a reaction $\mathbf{z}^m \rightleftharpoons \mathbf{z}^n$ in the CRN, and \mathbf{z}^m and \mathbf{z}^n are two different complexes—see for example Fig 1.

In a deterministic CRN, whose kinetics are given by Eq. (11), the *creation rate of the complex \mathbf{z}* , $J_{+\mathbf{z}}$, is defined as

$$J_{+\mathbf{z}} = \sum_{\rho | \mathbf{s}^{\rho} = \mathbf{z}} J_{+\rho} + \sum_{\rho | \mathbf{r}^{\rho} = \mathbf{z}} J_{-\rho} \tag{13}$$

where both addends in Eq. (13) are source terms for the complex \mathbf{z} : The subscript $\rho | \mathbf{s}^{\rho} = \mathbf{z}$ indicates that the sum is taken over the reactions ρ whose product complex equals the complex \mathbf{z} , and, similarly, $\rho | \mathbf{r}^{\rho} = \mathbf{z}$ indicates that that the sum is taken over those reactions whose reactant complex equals \mathbf{z} . Proceeding along the same lines, we define the *rate of annihilation of the complex \mathbf{z}* as

$$J_{-\mathbf{z}} = \sum_{\rho | \mathbf{s}^{\rho} = \mathbf{z}} J_{-\rho} + \sum_{\rho | \mathbf{r}^{\rho} = \mathbf{z}} J_{+\rho}. \tag{14}$$

A deterministic network is said to have a *complex-balanced steady state* if its steady state satisfies the condition that the creation rate and the annihilation rate of each complex are

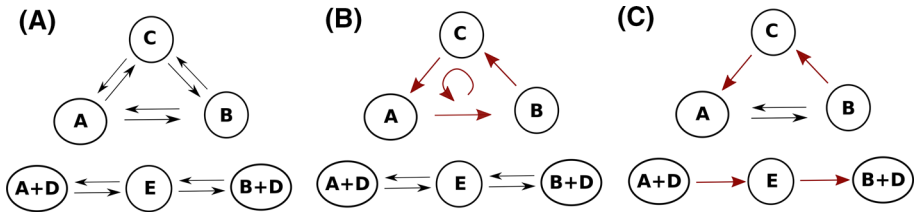


Fig. 2 Graphical depiction of the different types of steady states, taking the network of Fig. 1 as example. A) Detailed-balanced steady state. Here, the backward and forward rates of each reaction are equal. B) Complex-balanced steady state. There exist non-vanishing net currents (red arrows) but the steady state still satisfies the complex-balanced requirement, Eq. (15). It only allows for cycles at steady state that can be visualised from the network in terms of complexes (red cycle). C) General steady state. There exist non-vanishing net currents (red arrows) and the complex-balanced requirement does not hold

equal [30]:

$$J_{+z} = J_{-z} \quad \forall z. \tag{15}$$

While there exist some topological conditions in the CRN which ensure that the steady state is complex balanced [31], not every CRN possesses a complex-balanced non-equilibrium steady-state. We refer the interested reader to Refs. [32–34] for more detailed discussions on the topological constraints that determine complex-balancing and its consequences for MAK networks.

In general, a complex-balanced steady state is one of many steady states which a CRN may have. We can order these types of steady states in terms of their generality as follows: *Equilibrium steady state*. The most restrictive condition that we can impose to a steady state is detailed balance:

$$J_{+\rho} = J_{-\rho} \quad \forall \rho. \tag{16}$$

Equation (16) corresponds to a system at thermodynamic equilibrium, and implies that the rate of the forward reaction equals the rate of the backward reaction for every reaction ρ in the CRN (example depicted in Fig. 2A).

Complex-balanced steady state. More general than detailed-balanced steady states are complex-balanced steady states, which satisfy

$$\sum_{\rho|s^\rho=z} J_{+\rho} + \sum_{\rho|r^\rho=z} J_{-\rho} = \sum_{\rho|s^\rho=z} J_{-\rho} + \sum_{\rho|r^\rho=z} J_{+\rho} \quad \forall z. \tag{17}$$

Equation (17) implies that the creation and the annihilation rate of each complex z are equal. In the example of Fig. 1, the only way to have a complex-balanced steady state which is not at equilibrium (i.e. is not entirely detailed-balanced) is by taking the net rate $J_{+\rho} - J_{-\rho}$ in reaction 1 equal to the net rate of those of reactions 2 and 3, and thus reactions 4 and 5 must be detailed balanced (since the system has to be at steady state). This steady state is depicted in Fig. 2 B.

General steady state. The most general class of steady states is defined by the vanishing time derivatives of the dynamical equation (11). By splitting the contributions of each complex z , this condition can be rewritten as

$$\sum_m z_a^m \left(\sum_{\rho|s^\rho=z_m} J_{+\rho} + \sum_{\rho|r^\rho=z_m} J_{-\rho} \right) = \sum_m z_a^m \left(\sum_{\rho|r^\rho=z_m} J_{+\rho} + \sum_{\rho|s^\rho=z_m} J_{-\rho} \right) \quad \forall a, \tag{18}$$

where m is an index that labels each of the complexes in the network and the integer z_a^m represents its components. As a result, there are no constraints between the net rates of each reaction other than those imposed by the stationarity condition of Eq. (11). In the example of Fig. 1, this implies that, at steady state, there can be current cycles where, for example, species A is created by reaction 1, but annihilated by reaction 4 through the complex $A + D$, thus breaking complex balance (see Fig. 2 C).

From this hierarchical classification, it can be clearly seen that detailed balance, Eq. (16), implies complex balance, Eq. (17), which, in turn, implies the steady-state condition, Eq. (18). However, the converse is not true: A general steady state is not necessarily complex balanced, and a complex-balanced steady state is not necessarily detail balanced. Therefore, complex balance is less restrictive of a constraint than detailed balance, but it is more restrictive than a generic steady state.

2.5 Complex Balance in Networks with Mass-Action Kinetics

As a particular instance of special importance, in what follows we will discuss complex balance in ideal CRNs with MAK.

In short, MAK consists of the hypothesis that the rate of the chemical reaction is proportional to the product of the concentrations of the reactants: As a result, in the deterministic description, the MAK expressions for the currents read

$$J_{+\rho} = k_{+\rho} \prod_a c_a^{r_a^\rho}, \quad J_{-\rho} = k_{-\rho} \prod_a c_a^{s_a^\rho}, \tag{19}$$

where $k_{\pm\rho}$ are the rate constants. In what follows, we will denote by c_a^* the steady-state concentration of species a in the deterministic description. Then, for a deterministic system with MAK, the complex-balance condition is given by

$$\begin{aligned} & \sum_{\rho|s^\rho=z} k_{+\rho} \prod_a (c_a^*)^{r_a^\rho} + \sum_{\rho|r^\rho=z} k_{-\rho} \prod_a (c_a^*)^{s_a^\rho} \\ &= \sum_{\rho|s^\rho=z} k_{-\rho} \prod_a (c_a^*)^{s_a^\rho} + \sum_{\rho|r^\rho=z} k_{+\rho} \prod_a (c_a^*)^{r_a^\rho} \quad \forall z. \end{aligned} \tag{20}$$

Conversely, in the stochastic description [with dynamics is given by Eq. (10)] the MAK expressions for $f_{\pm\rho}$ become

$$f_{+\rho}(\mathbf{n}) = k_{+\rho} \prod_a \frac{n_a!}{(n_a - r_a^\rho)!}, \tag{21}$$

$$f_{-\rho}(\mathbf{n}) = k_{-\rho} \prod_a \frac{n_a!}{(n_a - s_a^\rho)!}. \tag{22}$$

Previous studies [33] have shown that the steady state of complex-balanced CRNs with MAK is known to have a product-form expression in terms of independent Poisson distributions, and it reads

$$\pi(\mathbf{n}) = \prod_{a=1}^M \frac{(c_a^*)^{n_a}}{n_a!} e^{-c_a^*}, \tag{23}$$

linking the deterministic steady state (c_a^*) to the stochastic steady state π . Furthermore, in Ref. [33] it is shown that, if the propensity functions take the more general form

$$f_{+\rho}(\mathbf{n}) = k_{+\rho} \frac{\theta(\mathbf{n})}{\theta(\mathbf{n} - \mathbf{r}^\rho)}, \quad (24)$$

$$f_{-\rho}(\mathbf{n}) = k_{-\rho} \frac{\theta(\mathbf{n})}{\theta(\mathbf{n} - \mathbf{s}^\rho)}, \quad (25)$$

then the steady-state distribution reads

$$\pi(\mathbf{n}) = \frac{M}{\theta(\mathbf{n})} \prod_{a=1}^M (c_a^*)^{n_a}, \quad (26)$$

where M is a normalisation constant and θ a function which maps the vector of integer numbers \mathbf{n} into a real-valued positive number.

In what follows, we will demonstrate that the result (23) can be generalised to the non-ideal case, i.e., to a class of propensity functions f which take into account the physical interactions between molecules.

3 Reaction Rates for Non-ideal Chemical-Reaction Networks

In the previous Section we introduced the general description of CRNs, both on a stochastic and deterministic level: in either cases, a choice for the propensity functions, or currents, must be made to set the network dynamics. For ideal solutions, the most common choice is MAK, as outlined above. However, in what follows we consider solutes which mutually interact and which are, therefore, not ideal, and specify the propensity functions.

3.1 Equilibrium Systems

Here, we consider CRNs at thermodynamic equilibrium, i.e., systems which are not subject to external, non-equilibrium driving. Given that the system is an equilibrium one, at steady state the principle of detailed balance must hold for every reaction ρ : The probability flux across a reaction ρ in the forward direction must equal the one in the backward direction. In this Section, we will impose the detailed-balance condition on the propensity functions at thermal equilibrium and suggest a generalisation for systems out of equilibrium.

3.1.1 Stochastic Description

In the stochastic description, the detailed-balance condition at steady state reads

$$P^{\text{eq}}(\mathbf{n}) f_{+\rho}(\mathbf{n}) = f_{-\rho}(\mathbf{n} + \mathbf{v}^\rho) P^{\text{eq}}(\mathbf{n} + \mathbf{v}^\rho), \quad (27)$$

where the equilibrium probability distribution $P^{\text{eq}}(\mathbf{n})$ for closed stochastic systems—total number of particles fixed—is given by the canonical Boltzmann distribution:

$$P^{\text{eq}}(\mathbf{n}) = \frac{1}{Z} e^{-\beta F(\mathbf{n})}, \quad (28)$$

with $\beta = 1/(k_B T)$, k_B is the Boltzmann constant, T the temperature, $F(\mathbf{n})$ the Helmholtz free energy of the system in state \mathbf{n} , and Z a normalisation factor—the partition function in statistical physics. For systems that exchange mass with a single particle reservoir, the

equilibrium distribution (28) is replaced by the distribution for the grand-canonical ensemble [35].

Combined with Eq. (28), the detailed-balance condition in Eq. (27) yields the following constraint for the propensity functions:

$$\frac{f_{+\rho}(\mathbf{n})}{f_{-\rho}(\mathbf{n} + \mathbf{v}^\rho)} = \frac{P^{\text{eq}}(\mathbf{n} + \mathbf{v}^\rho)}{P^{\text{eq}}(\mathbf{n})} = e^{-\beta[F(\mathbf{n} + \mathbf{v}^\rho) - F(\mathbf{n})]}. \tag{29}$$

Then, we choose the following functional form for the propensity functions:

$$f_{+\rho}(\mathbf{n}) = k_\rho e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - \mathbf{r}^\rho)]}, \tag{30}$$

$$f_{-\rho}(\mathbf{n}) = k_\rho e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - \mathbf{s}^\rho)]}, \tag{31}$$

where k_ρ is the reaction constant, which needs to be equal in both the forward and the backward reaction for Eq. (29) to be satisfied. Given that the free energy F may, in general, depend on the inter-particle interactions—such as steric, electrostatic, or other interactions—Eq. (29) implies that the chemical-reaction rates may depend on these inter-particle interactions.

The choices (30) and (31) for the propensity functions are not unique, but it is particularly appealing because it reduces to MAK for ideal systems. In fact, consider an ideal lattice-model solution with \mathcal{N} particles including both solvent and solute—see Appendix A for details. The free energy is

$$F_{\text{id}} = \sum_a n_a \mu_a^0 + \frac{1}{\beta} \left[\sum_a \log(n_a!) - \log(\mathcal{N}!) \right], \tag{32}$$

where $\mathcal{N} = \sum_a n_a$ (including solvent particles in the sum) and μ_a^0 is the standard-state chemical potential of species a (taken with respect a given reference state noted as ‘0’), which may depend on parameters like temperature or nature of the solvent and the solute a . Then, the rates take the following form:

$$\begin{aligned} f_{+\rho}(\mathbf{n}) &= k_\rho e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - \mathbf{r}^\rho)]} \\ &= k_\rho e^{\beta \sum_a r_a^\rho \mu_a^0} \frac{(\mathcal{N} - \sum_a r_a^\rho)!}{\mathcal{N}!} \prod_a \frac{n_a!}{(n_a - r_a^\rho)!}, \end{aligned} \tag{33}$$

where $(\mathcal{N} - \sum_a r_a^\rho)!/\mathcal{N}!$ can be approximated by $\mathcal{N}^{-\sum_a r_a^\rho}$. Setting

$$k_{+\rho} = k_\rho \frac{\exp(\beta \sum_a r_a^\rho \mu_a^0)}{\mathcal{N}^{\sum_a r_a^\rho}}, \tag{34}$$

we obtain that $f_{+\rho}$ coincides with the MAK propensity function (21), and similarly for $f_{-\rho}$ and Eq. (22).

We conclude this Section with a remark on the reaction constant, k_ρ : In Eqs. (30) and (31) we have assumed that k_ρ is a constant of the reaction, independent on the state \mathbf{n} of the system. However, in general k_ρ may depend on \mathbf{n} , because the system itself is part of the environment where the chemical reactions take place. These effects can be disregarded for most cases in ideal solutions (since they are usually dilute), but they may not be negligible in non-ideal systems. For instance, in the case of phase separation, the multiple phases of the system may constitute very different environments for the chemical reactions, accelerating them or slowing them down.

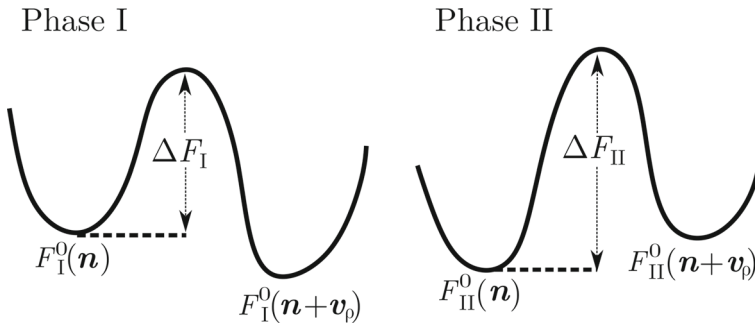


Fig. 3 Free-energy landscape for a chemical reaction. The horizontal dimension is the reaction coordinate and the vertical one specifies the height of the reaction free energy F^0 —the free energy F with the entropic term $\sum_a \log n_a!$ removed [37]. The quantities $F^0(\mathbf{n})$ and $F^0(\mathbf{n} + \mathbf{v}_\rho)$ denote the reaction free energy of the system before and after the reaction, subscripts I and II specify the system phase, and ΔF the height of the free-energy barrier. For each phase, there are two minima in the free-energy landscape, corresponding to whether the reaction has occurred or not, see the left- and right-hand minimum, respectively. In this example, ΔF depends on the phase the reaction takes place in, and, thus, the reaction constant k_ρ would also depend on the environment in which the reaction occurs

Independently of whether k_ρ in Eqs. (30) and (31) depends on the system state or not, detailed balance, Eq. (29), must still hold. This means that the forward reaction constant for a state \mathbf{n} must be equal to the backward reaction constant for a state $\mathbf{n} + \mathbf{v}^\rho$. One way to ensure this equality while keeping the state-dependency of the reaction constants, is to make k_ρ a function of the state \mathbf{n} deprived of the reactant complex, i.e., $\mathbf{n} - \mathbf{r}^\rho$, for the forward case, and of $\mathbf{n} + \mathbf{v}^\rho - \mathbf{s}^\rho$ for the backward one:

$$f_{+\rho}(\mathbf{n}) = k_\rho(\mathbf{n} - \mathbf{r}^\rho) e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - \mathbf{r}^\rho)]}, \tag{35}$$

$$f_{-\rho}(\mathbf{n} + \mathbf{v}^\rho) = k_\rho(\mathbf{n} + \mathbf{v}^\rho - \mathbf{s}^\rho) e^{\beta[F(\mathbf{n} + \mathbf{v}^\rho) - F(\mathbf{n} + \mathbf{v}^\rho - \mathbf{s}^\rho)]}, \tag{36}$$

where Eqs. (35) and (36) satisfy Eq. (29) because

$$\mathbf{n} - \mathbf{r}^\rho = \mathbf{n} + \mathbf{v}^\rho - \mathbf{s}^\rho, \tag{37}$$

see Eq. (2).

The dependency above of k_ρ on the system state can be pictured as follows. In analogy with the classical transition-state theory, we can think of the microscopic mechanism of a reaction as a random walk in a free-energy landscape [36, 37], see Fig. 3. Then, the value of the rate constant k_ρ depends on the height of the free-energy barrier ΔF of the reaction. The stable local minima in the reaction landscape have reaction free energies F_I^0, F_{II}^0 , which are defined as the free energy of the system once the entropic term $\sum_a \log(n_a)!$, has been removed [37]. Thus, F_I^0 and F_{II}^0 need not be equal since they include a contribution due to the interactions of reactants and products with its environment, i.e., the phase they are in. Hence, F_I^0 and F_{II}^0 are determined by F but the barrier height ΔF is not. The dependency of the height of the barrier—and thus of k_ρ —on the system state is precisely the one discussed in Eqs. (35) and (36), and it may strongly affect the CRN dynamics. In summary, we are connecting the chemical reaction rates to the free energy of the system F , but also to ΔF which sets the value of the reaction constants k_ρ .

3.1.2 Deterministic Description

When the particle numbers \mathbf{n} are large compared to the number of reactants and products, \mathbf{r} and \mathbf{s} , respectively, the free-energy differences which appear in the rates (30) and (31) can be rewritten as

$$\begin{aligned}
 F(\mathbf{n}) - F(\mathbf{n} - \mathbf{r}^\rho) &\approx F(\mathbf{n}) - \left[F(\mathbf{n}) - \sum_a \frac{\partial F}{\partial n_a} r_a^\rho \right] \\
 &= \sum_a r_a^\rho \mu_a,
 \end{aligned}
 \tag{38}$$

where in the first line we expanded F to first order in \mathbf{r} , in the second line we used the definition of the chemical potential of species a :

$$\begin{aligned}
 \mu_a &= \frac{\partial F}{\partial n_a} \\
 &= \frac{\partial \mathfrak{f}}{\partial c_a},
 \end{aligned}
 \tag{39}$$

and $\mathfrak{f}(\mathbf{c})$ is the free energy per unit volume in the deterministic notation. The second line in Eq. (38) is only an approximation, however, in the large particle number or deterministic limit, the error due to retaining only the first term in the Taylor expansion becomes vanishing. In this limit, Eq. (38) matches the affinity of the reaction, which is the driving force of the chemical reaction, according to irreversible thermodynamics [37], see Appendix B for details.

Therefore, the currents in a deterministic, non-ideal CRN can be written as

$$\begin{aligned}
 J_{+\rho} &= k_\rho e^{\beta \sum_a r_a^\rho \mu_a}, \\
 J_{-\rho} &= k_\rho e^{\beta \sum_a s_a^\rho \mu_a},
 \end{aligned}
 \tag{40}$$

which is an expression conceptually similar to that given by other approaches to construct thermodynamically consistent dynamics for deterministic CRNs [22, 24]. Once again, the currents (40) match their ideal MAK counterpart (19) if the chemical potentials used in the rates are those of an ideal solution, i.e., $\mu_a = 1/\beta \log c_a + \mu_a^0$. Here and in the rest of the text, dimensional arguments of the logarithms remain due to the fact that we are absorbing the effect of the total concentration in μ^0 , i.e., the original chemical potential was $\mu_a = 1/\beta \log(c_a/c_{\text{tot}}) + \mu_a^0$, where $c_{\text{tot}} = \sum_a c_a$ (the sum includes the solvent), but since variations in c_{tot} can be neglected – $\log(c_{\text{tot}})$ is just a constant and is absorbed into μ_a^0 (and, thus, into k_ρ).

As in the previous section, if we assume the rate constant is state-dependent then the currents are given by

$$\begin{aligned}
 J_{+\rho}(\mathbf{c}) &= \tilde{k}_\rho g_\rho(\mathbf{c}) e^{\beta \sum_a r_a^\rho \mu_a}, \\
 J_{-\rho}(\mathbf{c}) &= \tilde{k}_\rho g_\rho(\mathbf{c}) e^{\beta \sum_a s_a^\rho \mu_a},
 \end{aligned}
 \tag{41}$$

where \tilde{k}_ρ is still a constant and any dependency of the rate constant on the state is given by the function $g_\rho(\mathbf{c})$.

3.2 Non-equilibrium Systems

So far we considered the propensity functions of equilibrium CRNs. Given the large number of physically interesting systems which are out of equilibrium, such as living beings, in what follows we will generalise the analysis of Sect. 3.1 to a specific type of non-equilibrium systems: those in which the work is done by the chemostats they are connected to.

Let us assume that N' out of the N species in the system are connected to multiple particle reservoirs—chemostats: In the stochastic and deterministic description, each chemostat keeps constant the chemical potential of the species to which it is connected. Then, in general, the system will not relax to equilibrium, because of the work done on it by the chemostats. In the stochastic and deterministic description, the dimensions of the space of states or concentrations, respectively, is reduced to $N - N' \leq N$, since the chemostatted species are no longer dynamical variables.

3.2.1 Stochastic Description

We assume that connecting the system to several chemostats does not alter any of the mechanisms of the chemical reactions, since it only tunes the concentration of the species to which they are connected, in order to match a given value of chemical potential. Then, reactions that involve both chemostatted and non-chemostatted species are driven in one direction by the work done by the chemostats inserting and removing particles from the system (in order to keep their chemical potentials constant). Given that the mechanism of reaction remains the same, in line with the previous section the rates of these driven chemical transitions are taken to be [22, 23, 25, 26]

$$f_{+\rho}(\mathbf{n}) = k_{\rho} e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - \mathbf{r}^{\rho}) + \sum_b r_b^{\rho} \mu_b]}, \quad (42)$$

$$f_{-\rho}(\mathbf{n}) = k_{\rho} e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - \mathbf{s}^{\rho}) + \sum_b s_b^{\rho} \mu_b]}, \quad (43)$$

where now F is the free energy of the $N - N'$ non-chemostatted species, \mathbf{n} contains the particle numbers of the non-chemostatted species only, and the summation over b runs over the chemostatted species. For the sake of clarity, in what follows we will reserve the index b for the chemostatted species, and the index a for the non-chemostatted ones.

The rationale behind these relations is that the chemical reaction is still driven by free energy differences except that now the free energy differences due to the consumption of chemostatted species is just given by the chemical potential of the chemostats μ_b . The terms $\sum_b r_b^{\rho} \mu_b$ and $\sum_b s_b^{\rho} \mu_b$ in the exponential represents the chemical work done by the chemostats (with chemical potentials fixed at μ_b) when a reaction ρ occurs, which pushes the system out of equilibrium. The effect of the non-chemostatted species is still given by the free energy differences $F(\mathbf{n}) - F(\mathbf{n} - \mathbf{r}^{\rho})$ and $F(\mathbf{n}) - F(\mathbf{n} - \mathbf{s}^{\rho})$.

This implicitly assumes that the chemostatted species are abundant (so that the chemical potential does not fluctuate) and that they are ideal (negligible interactions with the non-chemostatted species). If the chemostatted species were not ideal, then the concentration of species might dynamically vary to match the chemostatted chemical potential as the particle numbers in the system change. Here, we only consider the simpler case of ideal chemostatted species and refer the interested reader to Ref. [22], where the case of non-ideal chemostatted species was analysed.

As in Sect. 3.1 [see Eqs. (35) and (36)], the rate constants in Eqs. (42) and (43) may be generalised in such a way that k_ρ depends on the system state:

$$\begin{aligned} f_{+\rho}(\mathbf{n}) &= \tilde{k}_\rho g_\rho(\mathbf{n} - \mathbf{r}^\rho) e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - \mathbf{r}^\rho) + \sum_b r_b^\rho \mu_b]}, \\ f_{-\rho}(\mathbf{n} + \mathbf{v}^\rho) &= \tilde{k}_\rho g_\rho(\mathbf{n} + \mathbf{v}^\rho - \mathbf{s}^\rho) e^{\beta[F(\mathbf{n} + \mathbf{v}^\rho) - F(\mathbf{n} + \mathbf{v}^\rho - \mathbf{s}^\rho) + \sum_b s_b^\rho \mu_b]}, \end{aligned} \tag{44}$$

where \tilde{k}_ρ is independent of \mathbf{n} . Propensity functions of this form have been suggested before in other contexts, such as in the modelling of molecular motors [38].

3.2.2 Deterministic Description

Proceeding along the lines of Sect. 3.1.2, in the deterministic limit the above propensity functions result in the currents

$$\begin{aligned} J_{+\rho}(\mathbf{c}) &= \tilde{k}_\rho g_\rho(\mathbf{c}) e^{\beta(\sum_a r_a^\rho \mu_a + \sum_b r_b^\rho \mu_b)}, \\ J_{-\rho}(\mathbf{c}) &= \tilde{k}_\rho g_\rho(\mathbf{c}) e^{\beta(\sum_a s_a^\rho \mu_a + \sum_b s_b^\rho \mu_b)}. \end{aligned} \tag{45}$$

4 Steady-State Distribution for Complex-Balanced, Non-ideal CRNs

In what follows, we will prove one of the central results of this work, i.e., that the complex-balance condition allows us to generalise to non-ideal CRNs the result (23) [33] for the steady-state distribution of the network, which is generally unique (for details see Refs. [39, 40]).

Namely, we claim that CRNs for which the complex-balance condition (15) holds, the steady state of the stochastic dynamics (10) with propensity functions (44) reads

$$\pi_{\text{neq}}(\mathbf{n}) = \frac{e^{-\beta[F(\mathbf{n}) + \sum_a \tilde{\mu}_a n_a]}}{Z}, \tag{46}$$

where the parameters $\tilde{\mu}_a$ depend on the chemostats to which the system is connected and on the reaction constants of the network, but not on F . These parameters can be obtained from the CRN in the ideal and deterministic limit, thus significantly simplifying the task of obtaining analytically the steady-state of the system. Note that we reserve μ_b for the chemical potentials of the chemostats while $\tilde{\mu}_a$ is a parameter that describes how the particle numbers at steady state of the non-chemostatted species depend on the non-equilibrium driving of the system. An additional necessary condition to prove this result is that the function g_ρ must be the same for all reactions, i.e., $g_\rho = g$; the relaxation of this hypothesis will be discussed in Sect. 6.

The proof follows closely that of Anderson, Craciun and Kurtz [33], and here we only present its main steps—see Appendix C for a full proof. We will substitute the steady-state (46) into the dynamical equations, look for solutions where the probability flux across complexes vanishes, and obtain the complex-balance condition for a network with MAK, Eq. (20). We can thus conclude that, if the network modelled deterministically with MAK is complex-balanced at steady state, i.e. Eq. (20) is satisfied, then Eq. (46) is the steady-state probability distribution of its stochastic non-ideal counterpart. Furthermore, the parameters $\tilde{\mu}_a$ in Eq. (46) can be obtained by solving Eq. (20).

By inserting the ansatz (46) in Eq. (10) with propensity functions of the form (44) and $g_\rho = g$ for all reactions, at steady state we obtain

$$\begin{aligned} & \sum_{\rho} \tilde{k}_{\rho} \left\{ g(\mathbf{n} - \mathbf{s}^{\rho}) e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - \mathbf{s}^{\rho}) + \sum_a v_a^{\rho} \tilde{\mu}_a + \sum_b r_b^{\rho} \mu_b]} \right. \\ & \quad \left. + g(\mathbf{n} - \mathbf{r}^{\rho}) e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - \mathbf{r}^{\rho}) - \sum_a v_a^{\rho} \tilde{\mu}_a + \sum_b s_b^{\rho} \mu_b]} \right\} \\ & = \sum_{\rho} \tilde{k}_{\rho} \left\{ g(\mathbf{n} - \mathbf{r}^{\rho}) e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - \mathbf{r}^{\rho}) + \sum_b r_b^{\rho} \mu_b]} \right. \\ & \quad \left. + g(\mathbf{n} - \mathbf{s}^{\rho}) e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - \mathbf{s}^{\rho}) + \sum_b s_b^{\rho} \mu_b]} \right\}. \end{aligned} \tag{47}$$

The previous equation is satisfied if, for each complex \mathbf{z} , we have

$$\begin{aligned} & \sum_{\rho | \mathbf{s}^{\rho} = \mathbf{z}} \tilde{k}_{\rho} g(\mathbf{n} - \mathbf{s}^{\rho}) e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - \mathbf{s}^{\rho}) + \sum_a v_a^{\rho} \tilde{\mu}_a + \sum_b r_b^{\rho} \mu_b]} \\ & \quad + \sum_{\rho | \mathbf{r}^{\rho} = \mathbf{z}} \tilde{k}_{\rho} g(\mathbf{n} - \mathbf{r}^{\rho}) e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - \mathbf{r}^{\rho}) - \sum_a v_a^{\rho} \tilde{\mu}_a + \sum_b s_b^{\rho} \mu_b]} \\ & = \sum_{\rho | \mathbf{r}^{\rho} = \mathbf{z}} \tilde{k}_{\rho} g(\mathbf{n} - \mathbf{r}^{\rho}) e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - \mathbf{r}^{\rho}) + \sum_b r_b^{\rho} \mu_b]} \\ & \quad + \sum_{\rho | \mathbf{s}^{\rho} = \mathbf{z}} \tilde{k}_{\rho} g(\mathbf{n} - \mathbf{s}^{\rho}) e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - \mathbf{s}^{\rho}) + \sum_b s_b^{\rho} \mu_b]}. \end{aligned} \tag{48}$$

Given that in the previous equation the complex \mathbf{z} is fixed, it can be simplified and yields

$$\begin{aligned} & \sum_{\rho | \mathbf{s}^{\rho} = \mathbf{z}} \tilde{k}_{\rho} e^{\beta[\sum_a (z_a - r_a^{\rho}) \tilde{\mu}_a + \sum_b r_b^{\rho} \mu_b]} + \sum_{\rho | \mathbf{r}^{\rho} = \mathbf{z}} \tilde{k}_{\rho} e^{\beta[-\sum_a (s_a^{\rho} - z_a) \tilde{\mu}_a + \sum_b s_b^{\rho} \mu_b]} \\ & = \sum_{\rho | \mathbf{r}^{\rho} = \mathbf{z}} \tilde{k}_{\rho} e^{\beta \sum_b r_b^{\rho} \mu_b} + \sum_{\rho | \mathbf{s}^{\rho} = \mathbf{z}} \tilde{k}_{\rho} e^{\beta \sum_b s_b^{\rho} \mu_b}. \end{aligned} \tag{49}$$

Setting

$$c_a^* = \exp[-\beta(\tilde{\mu}_a + \mu_a^0)], \tag{50}$$

$$k_{+\rho} = \tilde{k}_{\rho} \exp \left[\beta \left(\sum_a r_a^{\rho} \mu_a^0 + \sum_b r_b^{\rho} \mu_b \right) \right], \tag{51}$$

$$k_{-\rho} = \tilde{k}_{\rho} \exp \left[\beta \left(\sum_a s_a^{\rho} \mu_a^0 + \sum_b s_b^{\rho} \mu_b \right) \right], \tag{52}$$

Eq. (49) can be shown to be equivalent to the complex-balance condition for deterministic CRNs with MAK, Eq. (20), with rate constants given by Eqs. (51) and (52). These rate constants include the contribution of the standard-state chemical potentials μ_a^0 and the chemostats, as is usually the case in MAK [41] (although, without loss of generality, for the purposes of this result, all μ_a^0 can taken to be 0). Hence, a CRN for which the deterministic steady-state is complex balanced allows for a steady state of the form (46) for its stochastic and non-ideal version. Solving Eq. (20) for the steady-state concentrations with MAK and rate constants (51) and (52) yields c_a^* and, thus, the parameters $\tilde{\mu}_a$ [via Eq. (50)] which appear in the steady-state distribution (46). The exponential relationship between the concentrations c_a^* and $\tilde{\mu}_a$ reflects the logarithmic contribution of concentrations in the ideal chemical potential of solutes: $\mu_a = \mu_a^0 + 1/\beta \log c_a$.

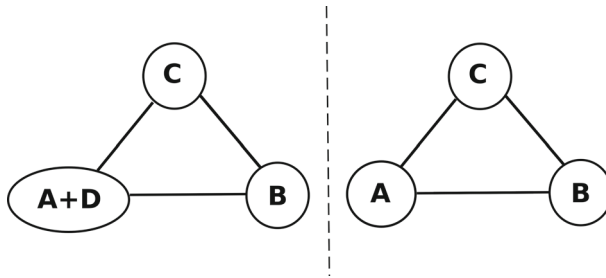


Fig. 4 Graphical representation of the illustrative CRNs of Sects. 4.1 (left) and 5.1.1 (right)

Equation (46) shows that the steady-state distribution of a non-ideal complex-balanced CRN has the form of an effective Boltzmann distribution, with the standard-state chemical potentials μ_a^0 shifted by $\tilde{\mu}_a$ (typically μ_a^0 would be included within F). From the physical standpoint, it is interesting to note that in Eq. (46) the free-energetic contribution F and the non-equilibrium term $\sum_a \tilde{\mu}_a n_a$ factor out.

This result is similar to Theorem 6.6 of Ref. [33]—here Eq. (26)—but we have generalised it slightly to include rates of the form (44), which includes the function g that could be of interest in phase-separated systems as it modulates the rates depending on the environment. Moreover, our approach relates both the rates (44) and the steady-state distribution (46) to thermodynamic quantities, like free energies and chemical potentials.

In what follows, we will illustrate the result (46) with a minimal working example of a complex-balanced CRN, and compare its predictions with numerical simulations.

4.1 Example

Let us consider the following CRN—see Fig. 4 for a graphical representation:



with a free energy taken from a regular-solution theory (where each particle, including the solvent, occupies a finite volume and thus total volume is linked to the total number of particles), see Appendix A for details.

For the sake of simplicity, we assume that the solvent particle number, n_{sol} , is conserved, and allow the total volume to vary:

$$N = N_{\text{ABC}} + n_{\text{D}} + n_{\text{sol}},
 \tag{54}$$

where the total number of particles of species A, B and C,

$$N_{\text{ABC}} = n_{\text{A}} + n_{\text{B}} + n_{\text{C}},
 \tag{55}$$

is kept constant in the CRN defined in (53).

Since the CRN (53) is complex balanced (which can be checked *a posteriori*), its steady state in the stochastic description and with propensity functions (44) can be obtained from its

deterministic dynamical equations (11). To achieve this, we write the stoichiometry matrices

$$R = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, S = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}, \tag{56}$$

which, together with the reaction constants given by Eqs. (51) and (52) and the free energy (32), completely define an ideal CRN. For simplicity, we assume that the standard-state chemical potentials μ^0 take the value 0 and that $\tilde{k}_\rho = 1$ for every reaction ρ . Finally, as an example, we take the non-equilibrium contribution of the chemostats to be present only in the reaction $C \rightleftharpoons A+D$, with $\sum_b r_b \mu_b = 0$ and $\sum_b s_b \mu_b = \log(5/2)\beta^{-1}$. These considerations, together with MAK [Eq. (19)] and the dynamics (11), yield the following set of deterministic and ideal equations for the CRN:

$$\begin{aligned} \frac{dn_A}{dt} &= n_B + \frac{5}{2} n_C - 2 n_A n_D \\ \frac{dn_B}{dt} &= n_C + n_A n_D - 2 n_B \\ \frac{dn_C}{dt} &= n_B + n_A n_D - \frac{7}{2} n_C. \end{aligned} \tag{57}$$

Note that in the system derived from the matrices (56) $d_t n_A = d_t n_D$. For the sake of concreteness, we take as initial conditions

$$N_{ABC} = 40 \tag{58}$$

and

$$n_A - n_D = 5, \tag{59}$$

the solution of the above system at steady state is $n_A^* \simeq 8.1$, $n_B^* \simeq 19.1$, $n_C^* \simeq 12.7$ and $n_D^* \simeq 3.1$. It can be checked that this steady-state solution satisfies the complex-balance requirement for MAK, Eq. (20). According to Eq. (50), we have the following identity: $n_a^* = e^{-\beta \tilde{\mu}_a}$, which enables us to obtain the values of $\tilde{\mu}_a$ and the steady-state probability (46).

Note that there are two conservation laws, Eqs. (58) and (59), and four chemical species: hence, $\pi_{\text{neq}}(\mathbf{n})$ is a distribution with only two independent variables.

In order to evaluate Eq. (46) explicitly, let us assume that the system has the following regular-solution free energy

$$F(\mathbf{n}) = \frac{1}{\beta} \log \left(\frac{n_A! n_B! n_C! n_D! n_{\text{sol}}!}{N!} \right) + (n_A \mu_A^0 + n_C \mu_C^0 + n_D \mu_D^0) + \chi \frac{n_A n_C}{N}, \tag{60}$$

where the first addend is an entropic term, the second corresponds to the internal energies of the chemical species taken with respect to that of species B, and the third to an interaction between species A and C.

Setting $\chi = 10$ and $\mu_A^0 = \mu_C^0 = \mu_D^0 = \log 2$, we obtain the bimodal steady-state probability depicted in Fig. 5, which closely matches the one obtained from a simulation of the same CRN using the Gillespie algorithm [42]. Simulations were started in parallel from random Poissonian initial conditions satisfying the constraints above, and the samples were obtained after the simulation relaxed to steady state. Note that, in order to arrive to the set of Eqs. (57), we assumed all $\mu_a^0 = 0$, while in the free energy we are giving them a different value. It would have been equivalent to insert these values of μ_a^0 into the system (57) and omit them in the free energy (60).

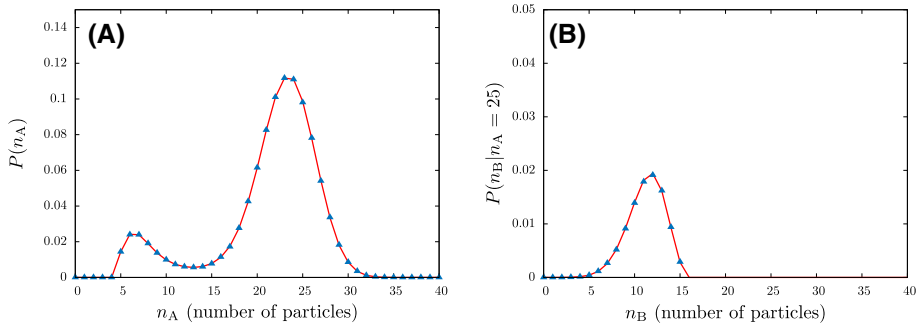


Fig. 5 Examples of marginal probabilities (A) and conditional probabilities (B) of the CRN discussed in Sect. 4.1. Blue triangles are the probabilities obtained from the Gillespie simulation, and the red lines theoretical predictions from our analytical result (46)

5 Lyapunov Function for Complex-Balanced Steady States

A Lyapunov function is a function that is minimised by the dynamics of the system and takes the value 0 at steady state. Under fairly general conditions, the logarithm of the steady-state probability distribution in the stochastic CRN is a Lyapunov function of the deterministic one [43, 44].

While the exact form of the Lyapunov function has been obtained for ideal and complex-balanced CRNs [30, 45], here we demonstrate that for non-ideal, complex-balanced CRNs the following function decreases with the dynamics

$$\begin{aligned} \mathcal{L}(\mathbf{c}) &= - \lim_{V \rightarrow \infty} \frac{1}{V} \log[\pi_{\text{neq}}(\mathbf{n})] \\ &= \beta \left[\mathfrak{f}(\mathbf{c}) + \sum_a \tilde{\mu}_a c_a \right] + \frac{\log Z}{V}, \end{aligned} \tag{61}$$

where the factor $1/V$ has been inserted to maintain the magnitude intensive while $V \rightarrow \infty$. Our approach generalises the results of Anderson and Nguyen [46] for product-form stationary states of CRNs. We will call the function (61) a Lyapunov function: This is a slight abuse of terminology, because we will only prove that \mathcal{L} decreases with the dynamics, not that its value is zero at steady state.

The function \mathcal{L} equals zero only if

$$\frac{\log Z}{V} = -\beta \left[\mathfrak{f}(\mathbf{c}^*) + \sum_a \tilde{\mu}_a c_a^* \right], \tag{62}$$

where the asterisk denotes values at steady state. Given that Z is a normalisation factor for the stochastic complex-balanced CRN at steady state, see Eq. (46), it reads

$$Z = \sum_{\mathbf{n}} e^{-\beta[F(\mathbf{n}) + \sum_a \tilde{\mu}_a n_a]}, \tag{63}$$

which, for a deterministic CRN involving a large number of particles, can be evaluated using the saddle-point approximation, where the sum is evaluated at the minimum value of the argument of the exponential. If the deterministic system is monostable, then the argument of the exponential has a single local minimum. Therefore, for monostable CRNs, this approximation will yield the correct value and the Lyapunov function Eq. (61) will take the value 0

at steady state. However, care must be taken when handling multistable CRNs in this way, which is why, in order to avoid this complexities, we will not prove that Eq. (61) takes the value 0 at steady state in general. Nevertheless, the fact that this function decreases with the dynamics is sufficient for our purposes.

In what follows we sketch the proof that, for complex-balanced non-ideal CRNs, the Lyapunov function (61) is a decreasing function of time—for a full step-by-step proof, see Appendix D.

Given that the normalisation factor Z does not depend on time but only on the non-equilibrium steady state, the time derivative of \mathcal{L} is

$$\begin{aligned} \frac{d\mathcal{L}}{dt} &= \sum_k \frac{\partial \mathcal{L}}{\partial c_k} \frac{\partial c_k}{\partial t} \\ &= \beta \sum_k (\mu_k + \tilde{\mu}_k) \left\{ \sum_{\rho} v_k^{\rho} k_{\rho} g(\mathbf{c}) \left[e^{\beta(\sum_a r_a^{\rho} \mu_a + \sum_b r_b^{\rho} \mu_b)} - e^{\beta(\sum_a s_a^{\rho} \mu_a + \sum_b s_b^{\rho} \mu_b)} \right] \right\}, \end{aligned} \tag{64}$$

where in the second line we used Eqs. (11) and (45), together with the assumption $g_{\rho} = g$. After adding and subtracting terms of the form $\sum_a r_a^{\rho} \tilde{\mu}_a$ in the exponentials (of the form $\sum_a s_a^{\rho} \tilde{\mu}_a$ for the second exponential), we repeatedly apply the inequality $e^x (y - x) \leq e^y - e^x$ to the sums of chemical potentials, and obtain

$$\begin{aligned} \frac{d\mathcal{L}}{dt} &\leq \sum_{\rho} k_{\rho} g(\mathbf{c}) e^{\beta(\sum_b r_b^{\rho} \mu_b - \sum_a r_a^{\rho} \tilde{\mu}_a)} \left[e^{\beta \sum_a (\mu_a + \tilde{\mu}_a) s_a^{\rho}} - e^{\beta \sum_a r_a^{\rho} (\mu_a + \tilde{\mu}_a)} \right] \\ &\quad + \sum_{\rho} k_{\rho} g(\mathbf{c}) e^{\beta(\sum_b s_b^{\rho} \mu_b - \sum_a s_a^{\rho} \tilde{\mu}_a)} \left[e^{\beta \sum_a (\mu_a + \tilde{\mu}_a) r_a^{\rho}} - e^{\beta \sum_a s_a^{\rho} (\mu_a + \tilde{\mu}_a)} \right]. \end{aligned} \tag{65}$$

The expression in the right-hand side (RHS) above can be split in terms of the different complexes in the system:

$$\begin{aligned} \frac{d\mathcal{L}}{dt} &\leq \sum_z g(\mathbf{c}) \left\{ \sum_{\rho | s^{\rho} = z} k_{\rho} e^{\beta[\sum_b r_b^{\rho} \mu_b - \sum_a r_a^{\rho} \tilde{\mu}_a + \sum_a (\mu_a + \tilde{\mu}_a) s_a^{\rho}]} \right. \\ &\quad - \sum_{\rho | r^{\rho} = z} k_{\rho} e^{\beta[\sum_b r_b^{\rho} \mu_b - \sum_a r_a^{\rho} \tilde{\mu}_a + \sum_a r_a^{\rho} (\mu_a + \tilde{\mu}_a)]} \\ &\quad + \sum_{\rho | r^{\rho} = z} k_{\rho} e^{\beta[\sum_b s_b^{\rho} \mu_b - \sum_a s_a^{\rho} \tilde{\mu}_a + \sum_a (\mu_a + \tilde{\mu}_a) r_a^{\rho}]} \\ &\quad \left. - \sum_{\rho | s^{\rho} = z} k_{\rho} e^{\beta[\sum_b s_b^{\rho} \mu_b - \sum_a s_a^{\rho} \tilde{\mu}_a + \sum_a s_a^{\rho} (\mu_a + \tilde{\mu}_a)]} \right\}. \end{aligned} \tag{66}$$

For a complex-balanced system, it can be shown that the expression in curly brackets in Eq. (66) vanishes for each complex z independently, as a consequence of the complex-balance condition for MAK systems, Eq. (20). Then

$$\frac{d\mathcal{L}}{dt} \leq 0, \tag{67}$$

and \mathcal{L} decreases, or remains unchanged, along a trajectory.

We conclude that, unlike in classical equilibrium systems, here it is not the F that is minimised by the dynamics, but a free energy (61) where the standard chemical potentials μ_a^0 are shifted by $\tilde{\mu}_a$. This shift, which is entirely due to the non-equilibrium contribution

of the chemostats, enables the system to present non-vanishing chemical-reaction net flows between species at steady-state, which, in the following, we will call chemical currents.

5.1 Spatially Heterogeneous Systems

In order to describe phase-separating systems, in what follows we will incorporate in our framework spatial inhomogeneities. In the deterministic description, concentrations are now a function of space, $c_a(\mathbf{x})$ within a volume Ω , and the free energy is a functional of these concentrations, $F[\mathbf{c}]$.

The time derivative of the concentrations is given by the following reaction-diffusion (RD) equation

$$\frac{\partial c_a(\mathbf{x})}{\partial t} = -\nabla \cdot \mathbf{J}_a + \sum_{\rho} v_a^{\rho} [J_{+\rho}(\mathbf{x}) - J_{-\rho}(\mathbf{x})], \tag{68}$$

where the time dependence of c is omitted, the first term in the RHS of the equation represents diffusion, and the second one the chemical reactions. As in the linear irreversible thermodynamics framework [47], the driving force of the diffusion current \mathbf{J}_a is the gradient of chemical potentials, $\nabla \mu_a$: the diffusive currents then read

$$\mathbf{J}_a = - \sum_k \mathcal{M}_{ak} \nabla \mu_k, \tag{69}$$

where \mathcal{M}_{ak} is the mobility matrix. We assume no-flux boundary conditions

$$\mathbf{J}_a|_{\mathbf{x} \in \partial\Omega} = 0 \tag{70}$$

for the non-chemostatted species, where $\partial\Omega$ denotes the boundaries of the volume Ω .

We now consider a generalisation of the Lyapunov function (61) to inhomogeneous systems. In the following, we will show that the dynamics (68) for complex-balanced networks minimise the Lyapunov functional

$$L = \beta \int_{\Omega} d\mathbf{x} \mathcal{L}(c(\mathbf{x})) = \beta F[\mathbf{c}] + \beta \int_{\Omega} d\mathbf{x} \sum_a \tilde{\mu}_a c_a(\mathbf{x}) \tag{71}$$

where $F[\mathbf{c}]$ is the free energy of the system, which depends on the concentration profile through

$$F[\mathbf{c}] = \int_{\Omega} d\mathbf{x} f(\mathbf{c}). \tag{72}$$

The time derivative of the Lyapunov functional (71) yields

$$\begin{aligned} \frac{dL}{dt} &= \int_{\Omega} \frac{\partial \mathcal{L}}{\partial t} d\mathbf{x} \\ &= \sum_a \int_{\Omega} d\mathbf{x} \frac{\partial \mathcal{L}}{\partial c_a(\mathbf{x})} \frac{\partial c_a(\mathbf{x})}{\partial t} \\ &= \sum_a \int_{\Omega} d\mathbf{x} \beta [\mu_a(\mathbf{x}) + \tilde{\mu}_a] \left[\nabla \cdot \mathbf{J}_a + \sum_{\rho} v_a^{\rho} (J_{+\rho} - J_{-\rho}) \right], \end{aligned} \tag{73}$$

where $\mu_a(\mathbf{x}) = \delta F[\mathbf{c}(\mathbf{x})] / \delta c_a(\mathbf{x})$ is the local chemical potential.

By applying the results of Sect. 5, Eq. (67), at every spatial point \mathbf{x} , we obtain that the second term in the square brackets of the RHS of Eq. (73) is negative or zero. Therefore, to

prove that $dL/dt \leq 0$ it is sufficient to show that the first term in the square brackets of the RHS is also negative or zero. In this regard, we note that

$$\begin{aligned} & \sum_a \int_{\Omega} d\mathbf{x} [\mu_a(\mathbf{x}) + \tilde{\mu}_a] \nabla \cdot \mathbf{J}_a \\ &= \sum_a \int_{\Omega} d\mathbf{x} \{ \nabla \cdot [(\mu_a(\mathbf{x}) + \tilde{\mu}_a) \mathbf{J}_a] - [\nabla(\mu_a(\mathbf{x}) + \tilde{\mu}_a)] \cdot \mathbf{J}_a \}. \end{aligned} \tag{74}$$

The first term in the RHS of the last equality vanishes due to the divergence theorem and Neumann boundary conditions (70). By observing that $\tilde{\mu}$ does not depend on space, the addend containing $\tilde{\mu}$ in the second term vanishes ($\nabla \tilde{\mu}_a = 0$). Finally, if the Onsager reciprocal relations for the mobility matrix \mathcal{M}_{ak} hold [48], then the addend containing $\mu_a(\mathbf{x})$ in the second term is necessarily positive, because it represents the entropy production of a diffusion process [37, 47]. The Onsager reciprocity relations ensure that a system relaxes to equilibrium in the absence of external work. Thus, the condition that the Onsager relations hold is not a limitation of the result but a consequence of thermodynamical consistency.

Combining the results above, we obtain that

$$\frac{dL}{dt} \leq 0. \tag{75}$$

It follows that, for a non-ideal, complex-balanced system, L decreases, which we can now use to obtain useful information about the steady state, along the lines of the free energy minimisation for systems at thermodynamic equilibrium.

Therefore, for a complex-balanced system, we can minimise L (subject to constraints in particle numbers) in order to obtain the concentration profile at steady state. This minimisation results in constraints for the steady-state profile of the form

$$\frac{\delta L}{\delta c_a(\mathbf{x})} = \beta(\mu_a(\mathbf{x}) + \tilde{\mu}_a) - \lambda = 0, \tag{76}$$

where λ is a Lagrange multiplier that enforces the particle-conservation constraint—for further details see Sect. 5.2. Equation (76) implies that in a complex-balanced solution at steady state there cannot be any diffusive currents, since the chemical potential is constant throughout space and the force driving diffusion currents is $\nabla \mu_a(\mathbf{x})$. Nevertheless, chemical currents can exist at steady state, as noted in the previous section, and the concentration profile may not be homogeneous. This is a major consequence of the present work and it stems directly from the form of the chemical reaction rates and the fact that the network is complex-balanced.

5.1.1 Example

Let us consider the following CRN (see Fig. 4 for a graphical representation):



with a free energy taken from a regular solution theory, as before. We will first assume that the system is homogeneous and later we will analyse the full reaction-diffusion system.

For simplicity, we assume that the system is driven out of equilibrium solely by imposing a non-equilibrium chemical potential difference in the transition from $C \rightleftharpoons A$, with $\sum_b s_b \mu_b = -\Delta\beta^{-1}$ and $\sum_b r_b \mu_b = 0$.

We take all the reaction constants k equal to each other, and note that the network is necessarily complex-balanced, as all chemical reactions are unimolecular (in unimolecular networks each of the species is a complex, hence the steady-state condition is equivalent to the complex-balance condition, if $g_\rho = g$ for every reaction ρ).

Proceeding along the lines of Sect. 4.1, in the stochastic description the steady-state of the CRN (77) with propensity functions (44) can be obtained from the following ideal and deterministic rate equations:

$$\begin{aligned} \frac{dn_A}{dt} &= n_B + e^{-\Delta}n_C - 2n_A \\ \frac{dn_B}{dt} &= n_C + n_A - 2n_B \\ \frac{dn_C}{dt} &= n_B + n_A - (1 + e^{-\Delta})n_C. \end{aligned}$$

The solution of the above system at steady-state is

$$\begin{aligned} n_A &= \frac{2n_C}{3} \left(\frac{1}{2} + e^{-\Delta} \right), \\ n_B &= n_C \left[\left(\frac{1}{2} + e^{-\Delta} \right) \frac{1}{3} + \frac{1}{2} \right]. \end{aligned}$$

We obtain the values of $\tilde{\mu}_a$ by identifying n_a with $e^{-\beta\tilde{\mu}_a}$ [as in Eq. (50)]. Noting that we can express such potentials with respect to that of species C, we obtain the Lyapunov function of the system

$$\mathcal{L}(\mathbf{c}) = \beta f(\mathbf{c}) - c_A \log \frac{1 + 2e^{-\Delta}}{3} - c_B \log \left(\frac{1}{2} + \frac{\frac{1}{2} + e^{-\Delta}}{3} \right) - \frac{\log Z}{V}, \tag{78}$$

where a detailed expression of the normalisation constant Z is not essential here, because Z is constant along the dynamics, and it does not alter the location of the minima of \mathcal{L} in the space of concentrations \mathbf{c} .

Unlike above, we will now describe the amount of species with reference to the fraction of volume they occupy at each point of space $\phi(\mathbf{x})$. Then,

$$\sum_a \phi_a(\mathbf{x}) + \phi_{\text{solv}}(\mathbf{x}) = 1, \tag{79}$$

where $\phi_{\text{solv}}(\mathbf{x})$ is the volume fraction of the solvent and the sum runs over solutes only. Equation (79) states that the solution is incompressible and, thus, $\phi_{\text{solv}}(\mathbf{x}) = 1 - \sum_a \phi_a(\mathbf{x})$. The reason for using volume fractions instead of concentrations is threefold: It is the convention normally used in phase separation studies and regular solution models, it enforces incompressibility (which is the case in most liquids) and is dimensionless. For simplicity we will assume that the molecular volumes of every species is the same, so that $\phi_a(\mathbf{x})$ is proportional to $c_a(\mathbf{x})$.

Hence, the following regular solution free-energy density can describe spatial inhomogeneities in an incompressible solution:

$$\beta f(\mathbf{c}) = \sum_a \phi_a \log \phi_a + \left(1 - \sum_a \phi_a \right) \log \left(1 - \sum_a \phi_a \right)$$

$$+ \chi_{AA}\phi_A^2 + \chi_{AB}\phi_A\phi_B + \sum_{a,k} \kappa_{ak}(\nabla\phi_a) \cdot \nabla\phi_k, \tag{80}$$

where the first two terms in the RHS are entropic terms and the following two represent the interactions between the solutes. The last addend represents the free-energetic cost of spatial inhomogeneities in the concentration profiles, and is known as Cahn-Hilliard term [49].

Assuming the system is one-dimensional, the resulting Lyapunov functional for the RD system is

$$L = \beta \int dx \left[f(\phi) + \sum_a \tilde{\mu}_a \phi_a(x) \right], \tag{81}$$

where $\tilde{\mu}_a$ are the ones obtained for the homogeneous system and do not depend on the coordinate x .

We set $\chi_{AA} = -2$, $\chi_{AB} = -7$, $\Delta = -2$, $\kappa_{AA} = \kappa_{BB} = 5$, $\kappa_{AB} = 1$ and any other Cahn-Hilliard coefficient equal to 0. With this parameter set, the reaction-diffusion system exhibits phase separation at steady state (see Fig. 6). By entering this free energy in the RD equations (68) and assuming no state dependency of the reaction constants k_ρ , we obtain a set of equations which describes the dynamics of the system. Figure 6 shows that the Lyapunov functional (81) is minimised by the dynamics, and that the non-equilibrium steady state is characterised by phase coexistence.

Finally, in Fig 6 C, the net reaction flux $J_{+\rho} - J_{-\rho}$ at steady state as a function of the spatial coordinate is depicted. This net reaction flux is constant in space and, given the topology of the CRN (77), is equal for all reactions ρ . The fact that the net reaction flux is independent of the spatial coordinate despite the varying concentrations (see Fig. 6 B) is a result of chemical reaction fluxes being driven by the chemical potential, which, as argued above, is constant—see Eq. (76). Note that this is also a consequence of having dropped the dependency of the reaction constants on the environment via a function $g(c)$. If all reaction constants were subject to this modulation (which has to be the same for every reaction for our results to hold), then the reaction rates at steady state could be space-dependent but the chemical potential would still be constant.

5.2 Phase Diagram of a Chemically Reactive Mixture

Since the Lyapunov functional for complex-balanced systems discussed in Sect. 5.1 is minimised by the dynamics, it carries plenty of information on the steady state.

Along the lines of phase separation for equilibrium systems, the steady state can be obtained by minimising $L[c]$ with respect to c , subject to certain constraints, e.g., particle conservation. The concentration profiles which realise the absolute minimum of $L[c]$ may be either spatially uniform, or depend on space, according to the system parameters. On a qualitative level, the phenomenology of a complex-balanced system does not change much with respect to that of a non-ideal solution at equilibrium, but the non-equilibrium terms may alter the phase diagram, thus tweaking the onset of phase separation.

To illustrate this point, in this Section we consider a non-ideal solution with the CRN (77) in the deterministic description, and obtain its phase diagram. Therefore, we minimise the Lyapunov functional (81) of Sect. 5.1.1, with the particle-conservation constraint

$$\phi_N = \frac{1}{|\Omega|} \int_{\Omega} dx [\phi_A(x) + \phi_B(x) + \phi_C(x)], \tag{82}$$

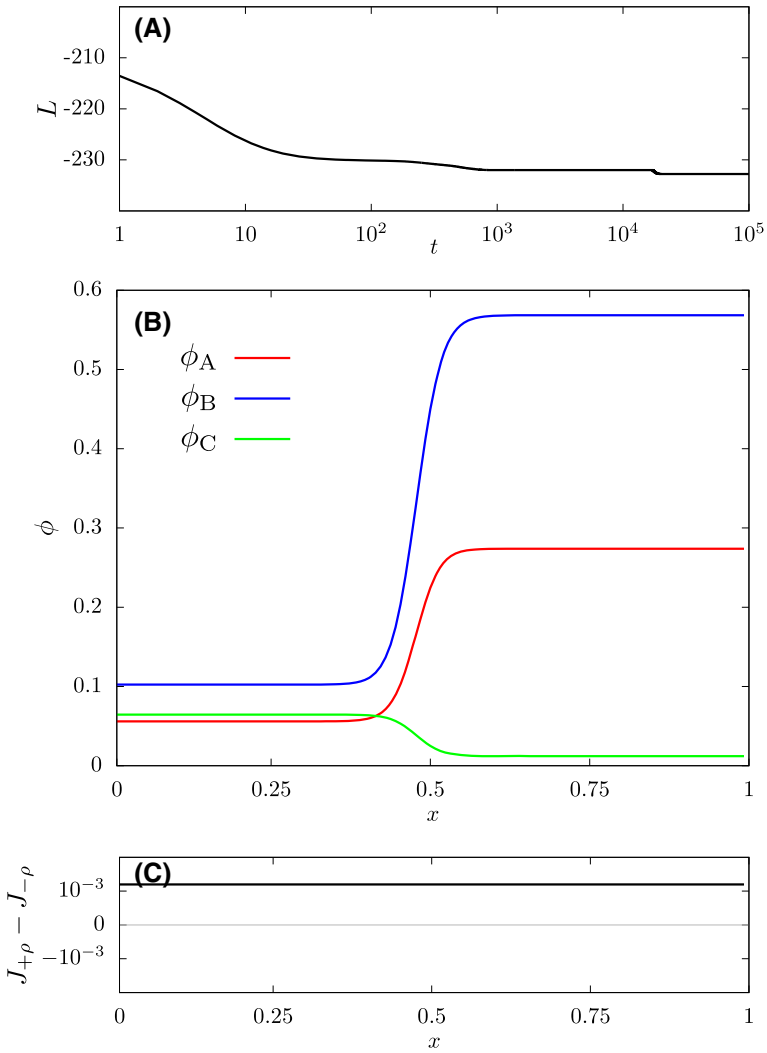


Fig. 6 Numerical results for the CRN (77) obtained by perturbing a homogeneous solution and integrating forward in time the reaction-diffusion equations (68) until a steady state is reached. The system is considered one-dimensional, and its normalised spatial coordinate x lies between zero and one. A) Evolution of the Lyapunov functional as a function of time in logarithmic scale. Here, the Lyapunov functional does not include the constant term $\log Z$ in Eq. (61). B) Steady-state profile of the system, displaying phase coexistence. The volume fractions ϕ of species A, B and C are plotted as functions of x . C) Net reaction flux $J_{+\rho} - J_{-\rho}$ (black line, arbitrary units) at steady state, as a function of x

where $\phi_a(\mathbf{x})$ is the volume fraction of species a , and the constant ϕ_N fixes the total volume fraction of the solutes. Then, the function that needs to be minimised is the Lagrangian

$$\Lambda = \int_{\Omega} d\mathbf{x} \mathcal{L}(\boldsymbol{\phi}) - \lambda \left\{ |\Omega| \phi_N - \int_{\Omega} d\mathbf{x} [\phi_A(\mathbf{x}) + \phi_B(\mathbf{x}) + \phi_C(\mathbf{x})] \right\}, \quad (83)$$

where λ is the Lagrange multiplier associated with the conservation of solutes, and $|\Omega|$ the volume of Ω .

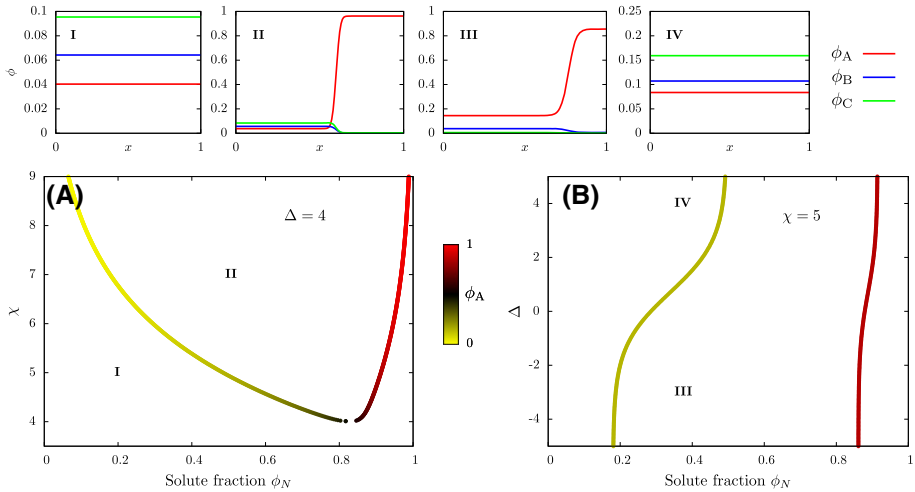


Fig. 7 Top: Steady-state concentration profiles of the example discussed in Sect. 5.2, obtained by numerical integration of the RD equations [Eq. (68)]. The concentrations of each species are plotted as functions of the one-dimensional space coordinate x , from $x = 0$ to $x = 1$. Each panel has a number (roman numerals) associated and corresponds to a different parameter configuration, which is specified by the exact position of the roman numerals in the phase diagram below. The leftmost and rightmost panel correspond to a spatially uniform steady state, while in the two middle panels phase separation occurs. Bottom: Phase diagrams as functions of the total solute fraction ϕ_N and interaction parameter χ (A) or non-equilibrium driving Δ (B), obtained by minimising the Lagrangian Λ , see Appendix E. The numerical minimisation requires higher precision as the parameters near criticality, which is why the density of data around the critical point decreases. The color code represents the amount of A particles along the phase-coexistence lines. While in the top panels the width of the inter-phase interface is finite to ease numerical integration, in the lower panels such width has been assumed to vanish

For the sake of simplicity, we take the typical lengthscale of Ω to be large with respect to inter-species interfaces: as a result, the volume fractions $\phi(x)$ can be approximated by piecewise constant functions. If the system phase separates, we assume that only two homogeneous, distinct phases, which we denote by ‘1’ and ‘2’, will appear. Within this assumption, the Lagrangian (83) reads

$$\Lambda = |\Omega_1| \mathcal{L}(\phi^1) + |\Omega_2| \mathcal{L}(\phi^2) - \lambda [|\Omega| \phi_N - |\Omega_1| (\phi_A^1 + \phi_B^1 + \phi_C^1) - |\Omega_2| (\phi_A^2 + \phi_B^2 + \phi_C^2)], \quad (84)$$

where Ω_1 and Ω_2 stand for the volumes phases 1 and 2, respectively, with $|\Omega_1| + |\Omega_2| = |\Omega|$, and we consider the following free-energy density

$$\beta f(\phi) = \sum_a \phi_a \log \phi_a + \left(1 - \sum_a \phi_a\right) \log \left(1 - \sum_a \phi_a\right) - \chi \phi_A^2, \quad (85)$$

where all species except A are considered non-interacting and A interacts with itself ($\chi > 0$ implies an effective attraction between A particles).

The minimisation of Λ yields the phase diagram in Fig. 7, see Appendix E for details. Phase separation occurs in regions II and III of the phase diagram, as shown in the concentration profiles displayed in the insets.

From the form of the free energy (85), we can see that is the species A that drives phase separation, since for $\chi > 0$ the free energy will favour segregating A from the rest of the solution. Thus, whether the steady-state displays one phase or a coexistence of phases also depends on the value of the non-equilibrium chemical potential difference Δ , which can alter

the concentration of A at steady state and, hence, modulate phase separation, as can be seen in Fig 7 B.

6 Discussion

In this work, we have shown that for a chemically reactive non-ideal solution we can obtain results for complex-balanced networks analogous to those for ideal solutions, provided that the system is modelled in a thermodynamically consistent way. This implies that the rates of the chemical reactions incorporate the interactions between the species in the system and, therefore, mass-action kinetics (MAK) no longer holds. By generalising MAK to a non-ideal solution, we obtained the steady-state probability distribution for a stochastic complex-balanced CRN and the Lyapunov function of its deterministic counterpart. Furthermore, after including diffusion in the description of the system we were also able to find a Lyapunov functional for the reaction-diffusion system, functional that determines the phase diagram of the system when minimised.

Our results are of particular importance for non-equilibrium phase-separating systems. By combining previous results from the mathematical theory of CRNs [33, 46] and concepts of non-equilibrium thermodynamics [37, 48], we found that the resulting complex-balanced RD system cannot sustain diffusion currents at steady state, see Eq. (76). Since, in many cases, diffusion currents are required for pattern formation in reaction-diffusion systems, breaking complex balance is a necessary condition to obtain such patterned steady states, at least when interactions are modelled in a thermodynamically consistent way unlike, e.g., those in Refs. [19, 20]. In this regard, complex balance can be broken in two ways: First, by choosing a suitable network topology that allows for a steady state which is not complex-balanced, as in Ref. [21]. Second, in a system where different phases coexist, by allowing the reaction rates to depend differently on local environment: For example, in Ref. [25] a patterned steady-state is produced by allowing one (and only one) of the reaction constants to depend on the concentration of an enzyme which localises in one of the phases. Mathematically, this violates one of the necessary conditions for our results to hold, namely $g_\rho = g$ (see Sect. 4), thus allowing for more general steady states.

In biological cells, phase separation has been hypothesised to perform many functions, such as, accelerating biochemical reactions within the condensate irrespective of the rate of the reaction in the dilute phase [50, 51]. The present work implies that, in order to control chemical reactions in each of the phases independently (at steady state) breaking complex balance is necessary, by virtue of Eq. (76). Indeed, in a complex-balanced system, the chemical potential of every species is constant throughout space. Then, given that the force driving the chemical reactions are the chemical potentials, the reaction rates in both phases are related, making it impossible to regulate the rates of chemical reactions in each phase in a fully independent way, and suggesting that breaking complex balance in one of the two ways outlined above is crucial for such control.

Clearly, the largest limitation of our study is the fact that we are restricted to complex-balanced networks, which, as shown above, constrains its phenomenology. However, these networks still have a broad applicability and interest in cell biology, examples of which are the *E. coli* chemotaxis system [52], kinetic proofreading [53, 54] or histone post-translational modifications [55] (the latter being particularly interesting since the model includes chemical reaction rates that are not MAK, which makes our theory particularly relevant to this case).

Overall, complex balance is known to be a key feature of CRNs which determines not only their behavior [32, 33] but also their thermodynamic properties [34]. In this analysis, we further stress the connection between the characteristics of the reaction network and the thermodynamically consistent structure of the physical system, in an effort to generalise results from ideal CRNs, and explore non-equilibrium dynamics of complex-balanced networks. However, little is known about non-complex balanced systems and, given our results, further research regarding the behaviour of this type of networks out of thermodynamic equilibrium would be of the utmost importance, both from the physical [56] and biological [51] point of view.

Acknowledgements We thank F. Brochard-Wyart, A. S. Vishen, P. Sens, U. Gerland, J.-F. Joanny, D. Lacoste, J. Palmeri, A. Šarić for valuable conversations and suggestions.

Author Contributions A.M.M. conceived the study and wrote the paper. M.C. contributed to the discussions and revisions.

Funding This study is supported by a Agence nationale de la recherche (ANR) grant ANR-17-CE11-0004.

Data Availability Statement Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Declarations

Conflict of interest/Competing interests The authors declare that they have no conflict of interest nor competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix A: Form of the Propensity Functions for a Regular-Solution Theory

We now consider a model of a solution based on a lattice where each chemical species (including the solvent) occupies one lattice site, thus neglecting differences in molecular volumes.

In a lattice with \mathcal{N} sites (note that the number of sites is proportional to the volume) occupied by N different species, with $\sum_{a=1}^N n_a = \mathcal{N}$, the configurational entropy is given by

$$S = k_B \log \frac{\mathcal{N}!}{\prod_a n_a!}, \quad (\text{A1})$$

where the argument of the logarithm is the number of microstates. The internal free energy of each species a is given by the standard-state chemical potential μ_a^0 .

We incorporate in this regular-solution model interactions among neighbouring sites, whose energy (in the mean-field approximation) reads

$$\sum_{a,k} \frac{\chi_{ak}}{2\mathcal{N}} [n_a n_k (1 - \delta_{a,k}) + n_a (n_a - 1) \delta_{a,k}] = \sum_{a,k} \frac{\chi_{ak}}{2\mathcal{N}} (n_a n_k - n_a \delta_{a,k}), \tag{A2}$$

where χ_{ak} represents the interaction energy between species a and k , and it can also be interpreted as the matrix of virial coefficients.

Taken into account the previous considerations, the free energy for a homogeneous mixture of chemical species in the regular-solution model reads

$$\begin{aligned} F &= U - TS \\ &= k_B T \left[\sum_a \log(n_a!) - \log(\mathcal{N}!) \right] + \sum_a \mu_a^0 n_a + \sum_{a,k} \frac{\chi_{ak}}{2\mathcal{N}} (n_a n_k - n_a \delta_{a,k}), \end{aligned} \tag{A3}$$

where the first two terms in the last line represent the ideal free energy (see Eq. (32) in the main text), while the last term is exclusively due to interactions between solutes.

With this expression of the free energy we can now derive an expression for the propensity functions (31) and (31). The forward (or backward) rates are a function of the free energy difference of the complex:

$$F(\mathbf{n}) - F(\mathbf{n} - \mathbf{r}^\rho) = \Delta F_{\text{id}} + \sum_a r_a^\rho \sum_k \frac{\chi_{ak}}{\mathcal{N}} n_k - \sum_{a,k} \frac{\chi_{ak}}{2\mathcal{N}} r_k^\rho r_a^\rho + \sum_a \frac{\chi_{aa}}{2\mathcal{N}} r_a^\rho \tag{A4}$$

where ΔF_{id} is the ideal part of the free energy difference of the complex, given by Eq. (33) in the main text. In the RHS of Eq. (A4), only the first two terms are non-vanishing as we approach the thermodynamic limit ($\mathcal{N} \rightarrow \infty$, while keeping n_a/\mathcal{N} fixed): hence, for large systems, the rest of the interacting terms are negligible. However, for a unimolecular reaction, since the free energy difference takes a particularly simple form, we have that

$$\begin{aligned} f_{+\rho}(\mathbf{n}) &= k_\rho e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - \mathbf{r}^\rho)]} \\ &= k_\rho \frac{n_a}{\mathcal{N}} e^{\beta(\mu_a^0 + \sum_k \frac{\chi_{ak}}{\mathcal{N}} n_k)}, \end{aligned} \tag{A5}$$

where a is the reactant of the reaction $+\rho$.

In the thermodynamic limit, Eqs. (A4), (30) and (31) imply that the deterministic rates can be written as

$$J_{+\rho} = k_{+\rho} \prod_a c_a^{r_a^\rho} \exp \left(\sum_{a,k} r_a^\rho \chi_{ak} c_k \right), \tag{A6}$$

where the particle numbers have been replaced with concentrations (an additional logarithmic factor has been absorbed into the rate constant $k_{+\rho}$, as explained in the main text, Sect. 3.1.2) and the part of the chemical potential representing the internal energy has also been absorbed in the rate constant $k_{+\rho}$. Setting

$$\mu_a = \log c_a + \mu_a^0 + \sum_k \chi_{ak} c_k, \tag{A7}$$

the rates (A6) match the general expression given in the main text, Eq. (40).

Appendix B: Chemical Affinity in the Deterministic Limit

In the deterministic limit, we describe the system in terms of intensive variables, namely, the concentration, c , and the free energy density, f . Then, the left-hand side Eq. (38) becomes

$$F(\mathbf{n}) - F(\mathbf{n} - \mathbf{r}^\rho) = V \left[f(\mathbf{c}) - f\left(\mathbf{c} - \frac{\mathbf{r}^\rho}{V}\right) \right]. \tag{B8}$$

If f is differentiable, then the derivative satisfies [57]

$$\lim_{V \rightarrow \infty} \frac{1}{\|\frac{\mathbf{r}^\rho}{V}\|} \left[f(\mathbf{c}) - f\left(\mathbf{c} - \frac{\mathbf{r}^\rho}{V}\right) + \sum_a \mu_a \frac{r_a^\rho}{V} \right] = 0, \tag{B9}$$

since μ_a has been defined as the partial derivative of f with respect to c_a , see Eq. (39) and $\|\cdot\|$ is the norm of the vector. Thus,

$$\lim_{V \rightarrow \infty} V \left[f(\mathbf{c}) - f\left(\mathbf{c} - \frac{\mathbf{r}^\rho}{V}\right) \right] = \sum_a \mu_a r_a^\rho, \tag{B10}$$

implying that the bottom line of Eq. (38) is not an approximation but is actually exact in the limit $V \rightarrow \infty$ (while keeping the concentrations constant).

Appendix C: Proof of the Complex-Balanced Distribution

In this Section we present the full proof of the result (46).

At steady state, the CME with rates of the form (44) and $g_\rho = g$ for all reactions ρ , reads

$$\begin{aligned} & \sum_\rho \tilde{k}_\rho g(\mathbf{n} - \mathbf{s}^\rho) e^{\beta[F(\mathbf{n} - \mathbf{v}^\rho) - F(\mathbf{n} - \mathbf{s}^\rho) + \sum_b r_b^\rho \mu_b]} P(\mathbf{n} - \mathbf{v}^\rho) \\ & + \sum_\rho \tilde{k}_\rho g(\mathbf{n} - \mathbf{r}^\rho) e^{\beta[F(\mathbf{n} + \mathbf{v}^\rho) - F(\mathbf{n} - \mathbf{r}^\rho) + \sum_b s_b^\rho \mu_b]} P(\mathbf{n} + \mathbf{v}^\rho) \\ & = \sum_\rho \left\{ \tilde{k}_\rho g(\mathbf{n} - \mathbf{r}^\rho) e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - \mathbf{r}^\rho) + \sum_b r_b^\rho \mu_b]} \right. \\ & \left. + \tilde{k}_\rho g(\mathbf{n} - \mathbf{s}^\rho) e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - \mathbf{s}^\rho) + \sum_b s_b^\rho \mu_b]} \right\} P(\mathbf{n}). \end{aligned} \tag{C11}$$

By dividing the previous expression by $P(\mathbf{n})$ and substituting the ansatz (46) into it, we obtain

$$\begin{aligned} & \sum_\rho \tilde{k}_\rho \left\{ g(\mathbf{n} - \mathbf{s}^\rho) e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - \mathbf{s}^\rho) + \sum_a v_a^\rho \tilde{\mu}_a + \sum_b r_b^\rho \mu_b]} \right. \\ & \left. + g(\mathbf{n} - \mathbf{r}^\rho) e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - \mathbf{r}^\rho) - \sum_a v_a^\rho \tilde{\mu}_a + \sum_b s_b^\rho \mu_b]} \right\} \\ & = \sum_\rho \tilde{k}_\rho \left\{ g(\mathbf{n} - \mathbf{r}^\rho) e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - \mathbf{r}^\rho) + \sum_b r_b^\rho \mu_b]} \right. \\ & \left. + g(\mathbf{n} - \mathbf{s}^\rho) e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - \mathbf{s}^\rho) + \sum_b s_b^\rho \mu_b]} \right\}. \end{aligned} \tag{C12}$$

We now rewrite the relation above in terms of a summation over each of the complexes $z \in \mathcal{C}$ separately

$$\begin{aligned}
 & \sum_z \left\{ \sum_{\rho|s^\rho=z} \tilde{k}_\rho g(\mathbf{n} - \mathbf{s}^\rho) e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - \mathbf{s}^\rho) + \sum_a v_a^\rho \tilde{\mu}_a + \sum_b r_b^\rho \mu_b]} \right. \\
 & \quad \left. + \sum_{\rho|r^\rho=z} \tilde{k}_\rho g(\mathbf{n} - \mathbf{r}^\rho) e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - \mathbf{r}^\rho) - \sum_a v_a^\rho \tilde{\mu}_a + \sum_b s_b^\rho \mu_b]} \right\} \\
 & = \sum_z \left\{ \sum_{\rho|r^\rho=z} \tilde{k}_\rho g(\mathbf{n} - \mathbf{r}^\rho) e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - \mathbf{r}^\rho) + \sum_b r_b^\rho \mu_b]} \right. \\
 & \quad \left. + \sum_{\rho|s^\rho=z} \tilde{k}_\rho g(\mathbf{n} - \mathbf{s}^\rho) e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - \mathbf{s}^\rho) + \sum_b s_b^\rho \mu_b]} \right\}, \tag{C13}
 \end{aligned}$$

where the subscript ‘ $\rho|s^\rho = z$ ’ denotes that the sum runs only over reactions ρ whose product complex s^ρ is equal to z . This previous equation will be satisfied if

$$\begin{aligned}
 & \sum_{\rho|s^\rho=z} \tilde{k}_\rho g(\mathbf{n} - \mathbf{s}^\rho) e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - \mathbf{s}^\rho) + \sum_a v_a^\rho \tilde{\mu}_a + \sum_b r_b^\rho \mu_b]} \\
 & \quad + \sum_{\rho|r^\rho=z} \tilde{k}_\rho g(\mathbf{n} - \mathbf{r}^\rho) e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - \mathbf{r}^\rho) - \sum_a v_a^\rho \tilde{\mu}_a + \sum_b s_b^\rho \mu_b]} \\
 & = \sum_{\rho|r^\rho=z} \tilde{k}_\rho g(\mathbf{n} - \mathbf{r}^\rho) e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - \mathbf{r}^\rho) + \sum_b r_b^\rho \mu_b]} \\
 & \quad + \sum_{\rho|s^\rho=z} \tilde{k}_\rho g(\mathbf{n} - \mathbf{s}^\rho) e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - \mathbf{s}^\rho) + \sum_b s_b^\rho \mu_b]}, \tag{C14}
 \end{aligned}$$

for every complex z . For any given complex z , Eq. (C14) can be rewritten in the following form:

$$\begin{aligned}
 & g(\mathbf{n} - z) e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - z)]} \left\{ \sum_{\rho|s^\rho=z} \tilde{k}_\rho e^{\beta[\sum_a v_a^\rho \tilde{\mu}_a + \sum_b r_b^\rho \mu_b]} \right. \\
 & \quad \left. + \sum_{\rho|r^\rho=z} \tilde{k}_\rho e^{\beta[-\sum_a v_a^\rho \tilde{\mu}_a + \sum_b s_b^\rho \mu_b]} \right\} \\
 & = g(\mathbf{n} - z) e^{\beta[F(\mathbf{n}) - F(\mathbf{n} - z)]} \left\{ \sum_{\rho|r^\rho=z} \tilde{k}_\rho e^{\beta \sum_b r_b^\rho \mu_b} + \sum_{\rho|s^\rho=z} \tilde{k}_\rho e^{\beta \sum_b s_b^\rho \mu_b} \right\}. \tag{C15}
 \end{aligned}$$

We now divide both sides by $g(\mathbf{n} - z) \exp\{\beta[F(\mathbf{n}) - F(\mathbf{n} - z)]\}$, and obtain

$$\begin{aligned}
 & \sum_{\rho|s^\rho=z} \tilde{k}_\rho e^{\beta[\sum_a (z_a - r_a^\rho) \tilde{\mu}_a + \sum_b r_b^\rho \mu_b]} + \sum_{\rho|r^\rho=z} \tilde{k}_\rho e^{\beta[-\sum_a (s_a^\rho - z_a) \tilde{\mu}_a + \sum_b s_b^\rho \mu_b]} \\
 & = \sum_{\rho|r^\rho=z} \tilde{k}_\rho e^{\beta \sum_b r_b^\rho \mu_b} + \sum_{\rho|s^\rho=z} \tilde{k}_\rho e^{\beta \sum_b s_b^\rho \mu_b}, \tag{C16}
 \end{aligned}$$

where we have substituted $\mathbf{v}_\rho = \mathbf{s}_\rho - \mathbf{r}_\rho$ and, depending on the reactions over which the sum runs, one of this terms can be replaced by the complex z .

Finally, given that in Eq. (C16) z is fixed, we can divide both sides by $\exp(\beta \sum_a z_a \tilde{\mu}_a)$, yielding

$$\begin{aligned} & \sum_{\rho|s^\rho=z} \tilde{k}_\rho e^{\beta[-\sum_a r_a^\rho \tilde{\mu}_a + \sum_b r_b^\rho \mu_b]} + \sum_{\rho|r^\rho=z} \tilde{k}_\rho e^{\beta[-\sum_a s_a^\rho \tilde{\mu}_a + \sum_b s_b^\rho \mu_b]} \\ &= \sum_{\rho|r^\rho=z} \tilde{k}_\rho e^{\beta[\sum_b r_b^\rho \mu_b - \sum_a z_a^\rho \tilde{\mu}_a]} + \sum_{\rho|s^\rho=z} \tilde{k}_\rho e^{\beta[\sum_b s_b^\rho \mu_b - \sum_a z_a^\rho \tilde{\mu}_a]}. \end{aligned} \tag{C17}$$

By inserting Eqs. (50) to (52) in Eq. (C17), we recover the complex-balance condition (20) for an ideal and deterministic network.

Appendix D: Proof of the Minimisation of the Lyapunov Function

In this Section we provide the full proof of Eq. (67).

Since Z does not depend on time, the time derivative of the Lyapunov function can be written as

$$\begin{aligned} \frac{d\mathcal{L}}{dt} &= \sum_k \frac{\partial \mathcal{L}}{\partial c_k} \frac{\partial c_k}{\partial t} \\ &= \beta \sum_k (\mu_k + \tilde{\mu}_k) \left\{ \sum_\rho v_k^\rho k_\rho g(\mathbf{c}) \left[e^{\beta[\sum_a r_a^\rho \mu_a + \sum_b r_b^\rho \mu_b]} - e^{\beta[\sum_a s_a^\rho \mu_a + \sum_b s_b^\rho \mu_b]} \right] \right\}, \end{aligned} \tag{D18}$$

where in the third line we have used Eq. (11) with currents given by Eq. (45). By adding and subtracting terms of the form $\sum_a r_a^\rho \tilde{\mu}_a$ in the exponentials (of the form $\sum_a s_a^\rho \tilde{\mu}_a$ for the second exponential), we rewrite the previous equality as

$$\begin{aligned} \frac{d\mathcal{L}}{dt} &= \sum_\rho \sum_k k_\rho g(\mathbf{c}) (\mu_k + \tilde{\mu}_k) (s_k^\rho - r_k^\rho) e^{\beta[\sum_a r_a^\rho (\mu_a + \tilde{\mu}_a) - \sum_a r_a^\rho \tilde{\mu}_a + \sum_b r_b^\rho \mu_b]} \\ &+ \sum_\rho \sum_k k_\rho g(\mathbf{c}) (\mu_k + \tilde{\mu}_k) (r_k^\rho - s_k^\rho) e^{\beta[\sum_a s_a^\rho (\mu_a + \tilde{\mu}_a) - \sum_a s_a^\rho \tilde{\mu}_a + \sum_b s_b^\rho \mu_b]}. \end{aligned} \tag{D19}$$

We now consider the inequality $e^s(t-s) \leq e^t - e^s$ —which results from $1+x \leq e^x, \forall x \in \mathbb{R}$ with $x = t - s$ —and apply it to the sums of chemical potentials. In the first term in the RHS of Eq. (D19), we set $s = \sum_a (\mu_a + \tilde{\mu}_a) r_a^\rho$ and $t = \sum_a (\mu_a + \tilde{\mu}_a) s_a^\rho$, and conversely in the second term. We then obtain

$$\begin{aligned} \frac{d\mathcal{L}}{dt} &\leq \sum_\rho k_\rho g(\mathbf{c}) e^{\beta(\sum_b r_b^\rho \mu_b - \sum_a r_a^\rho \tilde{\mu}_a)} \left[e^{\beta \sum_a (\mu_a + \tilde{\mu}_a) s_a^\rho} - e^{\beta \sum_a r_a^\rho (\mu_a + \tilde{\mu}_a)} \right] \\ &+ \sum_\rho k_\rho g(\mathbf{c}) e^{\beta(\sum_b s_b^\rho \mu_b - \sum_a s_a^\rho \tilde{\mu}_a)} \left[e^{\beta \sum_a (\mu_a + \tilde{\mu}_a) r_a^\rho} - e^{\beta \sum_a s_a^\rho (\mu_a + \tilde{\mu}_a)} \right]. \end{aligned} \tag{D20}$$

This expression can now be separated in terms of the different complexes in the system:

$$\begin{aligned} \frac{d\mathcal{L}}{dt} &\leq \sum_{z \in \mathcal{C}} g(\mathbf{c}) \left\{ \sum_{\rho|s^\rho=z} k_\rho e^{\beta[\sum_b r_b^\rho \mu_b - \sum_a r_a^\rho \tilde{\mu}_a + \sum_a (\mu_a + \tilde{\mu}_a) s_a^\rho]} \right. \\ &\quad \left. - \sum_{\rho|r^\rho=z} k_\rho e^{\beta[\sum_b r_b^\rho \mu_b - \sum_a r_a^\rho \tilde{\mu}_a + \beta \sum_a r_a^\rho (\mu_a + \tilde{\mu}_a)]} \right\} \end{aligned}$$

$$\begin{aligned}
 & + \sum_{\rho|r^\rho=z} k_\rho e^{\beta[\sum_b s_b^\rho \mu_b - \sum_a s_a^\rho \tilde{\mu}_a + \sum_a (\mu_a + \tilde{\mu}_a)r_a^\rho]} \\
 & - \sum_{\rho|s^\rho=z} k_\rho e^{\beta[\sum_b s_b^\rho \mu_b - \sum_a s_a^\rho \tilde{\mu}_a + \sum_a s_a^\rho (\mu_a + \tilde{\mu}_a)]} \Big\}. \tag{D21}
 \end{aligned}$$

Now, for a complex-balance system with MAK, we know that

$$\begin{aligned}
 & \sum_{\rho|s^\rho=z} k_\rho e^{\beta(\sum_b r_b^\rho \mu_b - \sum_a r_a^\rho \tilde{\mu}_a)} - \sum_{\rho|r^\rho=z} k_\rho e^{\beta(\sum_b r_b^\rho \mu_b - \sum_a r_a^\rho \tilde{\mu}_a)} \\
 & + \sum_{\rho|r^\rho=z} k_\rho e^{\beta(\sum_b s_b^\rho \mu_b - \sum_a s_a^\rho \tilde{\mu}_a)} - \sum_{\rho|s^\rho=z} k_\rho e^{\beta(\sum_b s_b^\rho \mu_b - \sum_a s_a^\rho \tilde{\mu}_a)} = 0, \tag{D22}
 \end{aligned}$$

for all complex \mathbf{z} , see Eq. (C17). Multiplying Eq. (D22) by $\exp[\beta \sum_a z_a (\mu_a + \tilde{\mu}_a)]$ (since the complex \mathbf{z} is fixed), we obtain an equality for each complex \mathbf{z} :

$$\begin{aligned}
 & \sum_{\rho|s^\rho=z} k_\rho e^{\beta[\sum_b r_b^\rho \mu_b - \sum_a r_a^\rho \tilde{\mu}_a + \sum_a (\mu_a + \tilde{\mu}_a)s_a^\rho]} \\
 & - \sum_{\rho|r^\rho=z} k_\rho e^{\beta[\sum_b r_b^\rho \mu_b - \sum_a r_a^\rho \tilde{\mu}_a + \beta \sum_a r_a^\rho (\mu_a + \tilde{\mu}_a)]} \\
 & + \sum_{\rho|r^\rho=z} k_\rho e^{\beta[\sum_b s_b^\rho \mu_b - \sum_a s_a^\rho \tilde{\mu}_a + \sum_a (\mu_a + \tilde{\mu}_a)r_a^\rho]} \\
 & - \sum_{\rho|s^\rho=z} k_\rho e^{\beta[\sum_b s_b^\rho \mu_b - \sum_a s_a^\rho \tilde{\mu}_a + \sum_a s_a^\rho (\mu_a + \tilde{\mu}_a)]} = 0, \tag{D23}
 \end{aligned}$$

which is precisely the term in brackets in Eq. (D21). Summing over all complexes yields the inequality (67).

Appendix E: Minimisation of the Lagrangian to Obtain the Phase Diagram

In order to find the steady state of the system, we need to minimise the Lyapunov functional or, in the presence of particle-conservation constraints, the Lagrangian, (83). A substantial simplification can be made by neglecting the contribution of the interfaces, i.e., considering the system as composed of two homogeneous phases. In this approximation, the function which needs to be minimised is the Lagrangian (84), which depends on eight independent variables: ϕ_a^p for $a = A, B, C, p = 1, 2, \lambda$ and $|\Omega_1|$.

First, we reduce the dimensionality of the problem by equating the derivatives of the Lagrangian with respect to the concentrations of the species:

$$\frac{\partial \Lambda}{\partial \phi_a^p} = \frac{\partial \Lambda}{\partial \phi_k^p}, \tag{E24}$$

where a and k denote two chemical species, and $p = 1, 2$ refers to the phases. Equation (E24) for a system at equilibrium yields the equality of chemical potentials (with their appropriate stoichiometry). Here, however, Equation (E24) includes the shifted chemical potential term $\tilde{\mu}$ due to the out-of-equilibrium complex-balancing. For a simple free energy like Eq. (85), Eq. (E24) implies

$$\phi_B^p = \phi_A^p e^{2\chi \phi_A^p + \tilde{\mu}_A - \tilde{\mu}_B}, \tag{E25}$$

$$\phi_C^p = \phi_A^p e^{2\chi\phi_A^p + \bar{\mu}_A - \bar{\mu}_C}, \quad (\text{E26})$$

which reduces the problem to just four variables: ϕ_A^1 , ϕ_A^2 , λ and $|\Omega^1|$.

Finally, given that we are interested in the phase diagram of the mixture and not in the actual steady state of the solution (i.e., we do not need to know how much volume each of the phases occupies), we can avoid solving for $|\Omega^1|$. This can be achieved by enforcing the stationarity condition of the Lagrangian with respect to the volume:

$$\frac{\partial \Lambda}{\partial |\Omega^1|} = 0, \quad (\text{E27})$$

which, together with

$$\frac{\partial \Lambda}{\partial \phi_A^1} = 0, \quad (\text{E28})$$

$$\frac{\partial \Lambda}{\partial \phi_A^2} = 0, \quad (\text{E29})$$

yields a fully determined system for the unknowns ϕ_A^1 , ϕ_A^2 and λ (the dependency on $|\Omega^1|$ drops out). The resulting equations for such unknowns are transcendental equations which, in general, have no explicit analytical solution. Therefore, they need to be solved numerically. Even the numerical solution is involved as the parameters near criticality, which is why in Fig. 7 the density of data around the critical point decreases.

References

- Zielinski, M.W., McGann, L.E., Nychka, J.A., Elliott, J.A.W.: Nonideal solute chemical potential equation and the validity of the grouped solute approach for intracellular solution thermodynamics. *J. Phys. Chem. B* **121**(46), 10443–10456 (2017)
- Mitchison, T.J.: Colloid osmotic parameterization and measurement of subcellular crowding. *Mol. Biol. Cell* **30**(2), 173–180 (2019)
- Zhou, H.-X., Rivas, G., Minton, A.P.: Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. *Annu. Rev. Biophys.* **37**(1), 375–397 (2008)
- Fall, C.P., Keizer, J.E.: Mitochondrial modulation of intracellular Ca²⁺ signaling. *J. Theor. Biol.* **210**(2), 151–165 (2001)
- Wang, J., Choi, J.-M., Holehouse, A.S., Lee, H.O., Zhang, X., Jahnel, M., Maharana, S., Lemaitre, R., Pozniakovskiy, A., Drechsel, D., Poser, I., Pappu, R.V., Alberti, S., Hyman, A.A.: A molecular grammar governing the driving forces for phase separation of prion-like RNA binding proteins. *Cell* **174**(3), 688–699 (2018)
- Brangwynne, C.P., Eckmann, C.R., Courson, D.S., Rybarska, A., Hoege, C., Gharakhani, J., Jülicher, F., Hyman, A.A.: Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science* **324**(5935), 1729 (2009)
- Li, P., Banjade, S., Cheng, H.-C., Kim, S., Chen, B., Guo, L., Llaguno, M., Hollingsworth, J.V., King, D.S., Banani, S.F., Russo, P.S., Jiang, Q.-X., Nixon, B.T., Rosen, M.K.: Phase transitions in the assembly of multivalent signalling proteins. *Nature* **483**(7389), 336–340 (2012)
- Su, X., Ditlev, J.A., Hui, E., Xing, W., Banjade, S., Okrut, J., King, D.S., Taunton, J., Rosen, M.K., Vale, R.D.: Phase separation of signaling molecules promotes T cell receptor signal transduction. *Science* **352**(6285), 595–599 (2016)
- Ditlev, J.A., Case, L.B., Rosen, M.K.: Who's in and who's out—compositional control of biomolecular condensates. *J. Mol. Biol.* **430**(23), 4666–4684 (2018)
- Banani, S.F., Lee, H.O., Hyman, A.A., Rosen, M.K.: Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18**(5), 285–298 (2017)
- Castellana, M., Wilson, M.Z., Xu, Y., Joshi, P., Cristea, I.M., Rabinowitz, J.D., Gitai, Z., Wingreen, N.S.: Enzyme clustering accelerates processing of intermediates through metabolic channeling. *Nat. Biotechnol.* **32**(10), 1011 (2014)

12. Vagne, Q., Turner, M.S., Sens, P.: Sensing size through clustering in non-equilibrium membranes and the control of membrane-bound enzymatic reactions. *PLoS ONE* **10**(12), 0143470 (2015)
13. Buchner, A., Tostevin, F., Hinzpeter, F., Gerland, U.: Optimization of collective enzyme activity via spatial localization. *J. Chem. Phys.* **139**(13), 135101 (2013)
14. Tsang, B., Arsenault, J., Vernon, R.M., Lin, H., Sonenberg, N., Wang, L.-Y., Bah, A., Forman-Kay, J.D.: Phosphoregulated FMRP phase separation models activity-dependent translation through bidirectional control of mrna granule formation. *Proc. Natl. Acad. Sci.* **116**(10), 4218–4227 (2019)
15. Kim, T.H., Tsang, B., Vernon, R.M., Sonenberg, N., Kay, L.E., Forman-Kay, J.D.: Phospho-dependent phase separation of FMRP and CAPRIN1 recapitulates regulation of translation and deadenylation. *Science* **365**(6455), 825–829 (2019)
16. Nott, T., Petsalaki, E., Farber, P., Jervis, D., Fussner, E., Plochowietz, A., Craggs, T.D., Bazett-Jones, D., Pawson, T., Forman-Kay, J., Baldwin, A.: Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol. Cell* **57**(5), 936–947 (2015)
17. Huberman, B.A.: Striations in chemical reactions. *J. Chem. Phys.* **65**(5), 2013–2019 (1976)
18. Glotzer, S.C., Di Marzio, E.A., Muthukumar, M.: Reaction-controlled morphology of phase-separating mixtures. *Phys. Rev. Lett.* **74**, 2034–2037 (1995)
19. Li, Y.I., Cates, M.E.: Non-equilibrium phase separation with reactions: a canonical model and its behaviour. *J. Stat. Mech.: Theory Exp.* **2020**(5), 053206 (2020)
20. Wurtz, J.D., Lee, C.F.: Chemical-reaction-controlled phase separated drops: formation, size selection, and coarsening. *Phys. Rev. Lett.* **120**, 078102 (2018)
21. Carati, D., Lefever, R.: Chemical freezing of phase separation in immiscible binary mixtures. *Phys. Rev. E* **56**, 3127–3136 (1997)
22. Avanzini, F., Penocchio, E., Falasco, G., Esposito, M.: Nonequilibrium thermodynamics of non-ideal chemical reaction networks. *J. Chem. Phys.* **154**(9), 094114 (2021)
23. Bauermann, J., Laha, S., McCall, P.M., Jülicher, F., Weber, C.A.: Chemical kinetics and mass action in coexisting phases. *arXiv preprint arXiv:2112.07576* (2021)
24. Bazant, M.Z.: Theory of chemical kinetics and charge transfer based on nonequilibrium thermodynamics. *Acc. Chem. Res.* **46**(5), 1144–1160 (2013)
25. Kirschbaum, J., Zwicker, D.: Controlling biomolecular condensates via chemical reactions. *J. R. Soc. Interface* **18**(179), 20210255 (2021)
26. Bauermann, J., Weber, C.A., Jülicher, F.: Energy and matter supply for active droplets. *Annalen der Physik*, 2200132 (2022)
27. Zwicker, D.: The intertwined physics of active chemical reactions and phase separation. *Curr. Opin. Colloid Interface Sci.* **61**, 101606 (2022)
28. Gillespie, D.T.: A rigorous derivation of the chemical master equation. *Phys. A: Stat. Mech. Appl.* **188**(1), 404–425 (1992)
29. Schnoerr, D., Sanguinetti, G., Grima, R.: Approximation and inference methods for stochastic biochemical kinetics—a tutorial review. *J. Phys. A: Math. Theor.* **50**(9), 093001 (2017)
30. Horn, F., Jackson, R.: General mass action kinetics. *Arch. Ration. Mech. Anal.* **47**(2), 81–116 (1972)
31. Feinberg, M.: Complex balancing in general kinetic systems. *Arch. Ration. Mech. Anal.* **49**(3), 187–194 (1972)
32. Feinberg, M.: The existence and uniqueness of steady states for a class of chemical reaction networks. *Arch. Ration. Mech. Anal.* **132**(4), 311–370 (1995)
33. Anderson, D.F., Craciun, G., Kurtz, T.G.: Product-form stationary distributions for deficiency zero chemical reaction networks. *Bull. Math. Biol.* **72**(8), 1947–1970 (2010)
34. Polettini, M., Wachtel, A., Esposito, M.: Dissipation in noisy chemical networks: the role of deficiency. *J. Chem. Phys.* **143**(18), 184103 (2015)
35. Rao, R., Esposito, M.: Conservation laws and work fluctuation relations in chemical reaction networks. *J. Chem. Phys.* **149**(24), 245101 (2018)
36. Hänggi, P., Talkner, P., Borkovec, M.: Reaction-rate theory: fifty years after Kramers. *Rev. Mod. Phys.* **62**, 251–341 (1990)
37. Kondepudi, D., Prigogine, I.: *Modern Thermodynamics*. John Wiley & Sons Ltd, New York (2014)
38. Jülicher, F., Ajdari, A., Prost, J.: Modeling molecular motors. *Rev. Mod. Phys.* **69**(4), 1269 (1997)
39. Schnakenberg, J.: Network theory of microscopic and macroscopic behavior of master equation systems. *Rev. Mod. Phys.* **48**, 571–585 (1976)
40. Anderson, D.F., Kurtz, T.G.: *Stochastic Analysis of Biochemical Systems*, vol. 674. Springer, Zurich (2015)
41. Rao, R., Esposito, M.: Nonequilibrium thermodynamics of chemical reaction networks: wisdom from stochastic thermodynamics. *Phys. Rev. X* **6**, 041064 (2016)

42. Gillespie, D.T.: Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **81**(25), 2340–2361 (1977)
43. Gang, H.: Lyapounov function and stationary probability distributions. *Zeitschrift für Physik B Condensed Matter* **65**(1), 103–106 (1986)
44. Ge, H., Qian, H.: Mathematical formalism of nonequilibrium thermodynamics for nonlinear chemical reaction systems with general rate law. *J. Stat. Phys.* **166**(1), 190–209 (2017)
45. Anderson, D.F., Craciun, G., Gopalkrishnan, M., Wiuf, C.: Lyapunov functions, stationary distributions, and non-equilibrium potential for reaction networks. *Bull. Math. Biol.* **77**(9), 1744–1767 (2015)
46. Anderson, D.F., Nguyen, T.D.: Results on stochastic reaction networks with non-mass action kinetics. *Math. Biosci. Eng.* **16**(4), 2118 (2019)
47. De Groot, S.R., Mazur, P.: *Non-Equilibrium Thermodynamics*. Dover Books on Physics. Dover Publications, Amsterdam (2013)
48. Onsager, L.: Reciprocal relations in irreversible processes. I. *Phys. Rev.* **37**(4), 405 (1931)
49. Cahn, J.W., Hilliard, J.E.: Free energy of a nonuniform system. I. Interfacial free energy. *J. Chem. Phys.* **28**(2), 258 (1958)
50. Lyon, A.S., Peeples, W.B., Rosen, M.K.: A framework for understanding the functions of biomolecular condensates across scales. *Nat. Rev. Mol. Cell Biol.* **22**(3), 215–235 (2021)
51. Hondele, M., Heinrich, S., De Los Rios, P., Weis, K.: Membraneless organelles: phasing out of equilibrium. *Emerg. Top. Life Sci.* **4**(3), 343–354 (2020)
52. Govern, C.C., ten Wolde, P.R.: Optimal resource allocation in cellular sensing systems. *Proc. Natl. Acad. Sci.* **111**(49), 17486–17491 (2014)
53. Hopfield, J.J.: Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proc. Natl. Acad. Sci.* **71**(10), 4135–4139 (1974)
54. Owen, J.A., Gingrich, T.R., Horowitz, J.M.: Universal thermodynamic bounds on nonequilibrium response with biochemical applications. *Phys. Rev. X* **10**, 011066 (2020)
55. Berry, S., Dean, C., Howard, M.: Slow chromatin dynamics allow polycomb target genes to filter fluctuations in transcription factor activity. *Cell Syst.* **4**(4), 445–4578 (2017)
56. Mehta, P., Lang, A.H., Schwab, D.J.: Landauer in the age of synthetic biology: energy consumption and information processing in biochemical networks. *J. Stat. Phys.* **162**(5), 1153–1166 (2016)
57. Tromba, A., Marsden, J.E.: *Vector Calculus*. Freeman, London (1996)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.