# Diffusion-dynamics laws in stochastic reaction networks

Tan Van Vu[*] and Yoshihiko Hasegawa[†]

*Department of Information and Communication Engineering, Graduate School of Information Science and Technology,
The University of Tokyo, Tokyo 113-8656, Japan*

Many biological activities are induced by cellular chemical reactions of diffusing reactants. The dynamics of such systems can be captured by stochastic reaction networks. A recent numerical study has shown that diffusion can significantly enhance the fluctuations in gene regulatory networks. However, the universal relation between diffusion and stochastic system dynamics remains veiled. Within the approximation of reaction-diffusion master equation (RDME), we find general relation that the steady-state distribution in complex balanced networks is diffusion-independent. Here, complex balance is the nonequilibrium generalization of detailed balance. We also find that for a diffusion-included network with a Poisson-like steady-state distribution, the diffusion can be ignored at steady state. We then derive a necessary and sufficient condition for networks holding such steady-state distributions. Moreover, we show that for linear reaction networks the RDME reduces to the chemical master equation, which implies that the stochastic dynamics of networks is unaffected by diffusion at any arbitrary time. Our findings shed light on the fundamental question of when diffusion can be neglected, or (if nonnegligible) its effects on the stochastic dynamics of the reaction network.

## I. INTRODUCTION

Diverse biological phenomena, such as cellular signal transductions and gene expression systems, are commonly studied by stochastic reaction network modeling [1–3]. These systems involve a set of reactant species which react through several channels. In most of the existing studies, such systems are often assumed to be well mixed, meaning that the diffusion coefficients of the reactants are infinitely large [4–10]. However, experiments have shown that reactants in cells diffuse at considerably low rates [11], and that the smallest timescale of the system is a little larger than the timescale of molecular diffusion. In such cases, the well-mixed assumption cannot accurately obtain the stochastic dynamics of the system. For example, living cells continuously receive signals at their receptors, which are subsequently transmitted to the nucleus through biochemical reaction networks [12–15]. This process is strongly influenced by extrinsic and intrinsic noise arising from fluctuations in the input and reactions. These effects induce unavoidable fluctuations in the biomolecule concentrations, which deteriorate the fidelity of information transfer [16,17]. By accurately evaluating the fluctuations, we would better understand the mechanism underlying signal transmission in cells. In a numerical study of gene regulatory networks, Ref. [18] showed that the fluctuations are larger in the model with diffusion than in its well-mixed counterpart. Thus, how diffusion relates to the stochastic dynamics of reaction networks is a pertinent question. Recently, Ref. [19] has numerically studied the effects of diffusion on single-cell

variability in multicellular organisms, and the limits of slow and fast diffusion have been investigated.

Two commonly used models for studying stochastic reaction-diffusion systems are the reaction-diffusion master equation (RDME) [20] and the Smoluchowski model [21]. The RDME, which is a mesoscopic model, is an extension of the nonspatial chemical master equation (CME) [20] and can be interpreted as an asymptotic approximation to spatially continuous stochastic reaction-diffusion models [22]. The RDME has been successfully applied in studying many biological systems [19,23–25]. It is worth noting that the Langevin equation, which can be derived from an equivalent Fokker–Planck equation or the Poisson representation, can handle continuum-limit diffusion in reaction networks [20,26]. However, the Langevin equation is applicable to biochemical reactions occurring in infinite space with no physical boundary, which is unrealistic in biological cells.

In the present work, we investigate the relations between diffusion and the stochastic dynamics of reaction networks within a physical reflecting boundary. In this system, reactants diffuse within a closed three-dimensional space without escaping. With the aid of the RDME, we find an intriguing law stating that diffusion does not affect the steady-state distribution of complex balanced networks, which have a Poisson-like distribution. Our proof reveals that if the network presents a steady-state distribution of product-of-Poissonians form, diffusion can be neglected. We then calculate the necessary and sufficient conditions for such steady-state network distributions. We also find another result, wherein steady state is not a requirement. Specifically, we prove that for linear reaction networks, one can derive the CME from the RDME, which indicates that diffusion can be ignored in this case. This result can be restated as follows: the stochastic dynamics of linear networks are diffusion-independent, which is consistent with

the Smoluchowski model. In addition, we perform stochastic simulations on both linear and nonlinear networks to verify our results.

## II. MODELS

We consider a general reaction network consisting of $N$ reactant species $X_1, \ldots, X_N$ and $K$ reactions $R_1, \ldots, R_K$. Assume that all reactions occur inside a cell with fixed volume $\Omega$, and that reaction $R_j$ $(1 \leqslant j \leqslant K)$ is of the form

$$s_{1j}X_1 + \cdots + s_{Nj}X_N \xrightarrow{k_j} r_{1j}X_1 + \cdots + r_{Nj}X_N, \quad (1)$$
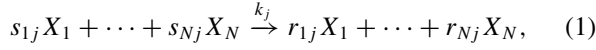
where $s_{ij}, r_{ij} \in \mathbb{N}_{\geqslant 0}$ are the stoichiometric coefficients and $k_j \in \mathbb{R}_{>0}$ is the macroscopic reaction rate. Here, $\mathbb{N}_{\geqslant 0}$ denotes the set of nonnegative integers. $\mathbb{R}_{>0}$ and $\mathbb{R}_{\geqslant 0}$ are defined analogously. If $\sum_{i=1}^{N} s_{ij} \leqslant 1$ for all $j = 1, \ldots, K$, then the reaction network is linear; otherwise, it is nonlinear. The state of the system is fully determined by the molecule-number vector of all reactant species in the system, $\boldsymbol{n} = [n_1, \ldots, n_N]^\top$, where $n_i \in \mathbb{N}_{\geqslant 0}$ is the number of molecules of species $X_i$. Assuming mass-action kinetics, the time evolution of a well-mixed system can be described by the following chemical master equation (CME):
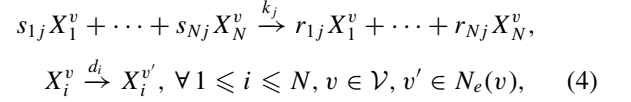
$$\partial_t P(\boldsymbol{n}, t) = \sum_{j=1}^{K} (\mathbb{E}^{-\boldsymbol{V}_j} - 1) f_j(\boldsymbol{n}, \Omega) P(\boldsymbol{n}, t), \quad (2)$$

where $\boldsymbol{V} = [r_{ij} - s_{ij}] \in \mathbb{Z}^{N \times K}$ is a stoichiometric matrix, $\boldsymbol{V}_j$ denotes the $j$th column of matrix $\boldsymbol{V}$, and $\mathbb{E}^{\boldsymbol{x}}$ is an operator that replaces $\boldsymbol{n}$ with $\boldsymbol{n} + \boldsymbol{x}$. $P(\boldsymbol{n}, t)$ is the probability of the system being in state $\boldsymbol{n}$ at time $t$, and the propensity function $f_j(\boldsymbol{n}, \Omega)$ of reaction $R_j$ is given by

$$f_j(\boldsymbol{n}, \Omega) = k_j \Omega^{1 - \sum_{i=1}^{N} s_{ij}} \prod_{i=1}^{N} \frac{n_i!}{(n_i - s_{ij})!}. \quad (3)$$

To include diffusion in stochastic spatial dynamics, many researchers apply the RDME, in which space is partitioned discretely into many voxels. It is known that the RDME is accurate if an appropriate combination of the time- and length-scale is chosen [22,27–30]. We assume from now on that the volume of the system is optimally divided into small voxels and as such, the RDME yields a good description of the time evolution of the probability distribution. Diffusion then occurs among the voxels, and the reaction can occur within the same voxel considered to be a well-mixed system. Assume that the volume $\Omega$ is divided into a set $\mathcal{V}$ of voxels labeled by integers $v = 1, 2, \ldots, |\mathcal{V}|$. Each voxel $v$ occupies a constant volume $\omega$ and contains $n_{vi}$ molecules of reactant species $X_i$. The state vector of voxel $v$ is denoted as $\boldsymbol{n}_v = [n_{v1}, \ldots, n_{vN}]^\top$. The state of the whole system is then described as the molecule-number vector $\boldsymbol{n}$ of each species in each voxel, namely, $\boldsymbol{n} = [\boldsymbol{n}_1^\top, \ldots, \boldsymbol{n}_{|\mathcal{V}|}^\top]^\top$. We also define a vector $\mathbf{1}_{vi} \in \mathbb{Z}^{|\mathcal{V}|N}$, in which the number of molecules of all species in all voxels is zero except for species $X_i$ in voxel $v$ (which is one), and a vector $\widetilde{\boldsymbol{V}}_{vj} \in \mathbb{Z}^{|\mathcal{V}|N}$, in which all elements are zero except in voxel $v$ (which holds $\boldsymbol{V}_j$). As the diffusion of each species into neighboring voxels can be modeled as a first-order reaction, the diffusion-included reaction network can be described in

the following form:

$$s_{1j}X_1^v + \cdots + s_{Nj}X_N^v \xrightarrow{k_j} r_{1j}X_1^v + \cdots + r_{Nj}X_N^v,$$

$$X_i^v \xrightarrow{d_i} X_i^{v'}, \; \forall 1 \leqslant i \leqslant N, v \in \mathcal{V}, v' \in N_e(v), \quad (4)$$

where $X_i^v$ refers to species $X_i$ in voxel $v$, $d_i$ is the diffusion rate of species $X_i$, and $N_e(v)$ is the set of voxels neighboring $v$. The stochastic dynamics of the system can then be described by the following RDME:

$$\partial_t P(\boldsymbol{n}, t) = \sum_{v \in \mathcal{V}} \sum_{v' \in N_e(v)} \sum_{i=1}^{N} (\mathbb{E}^{\mathbf{1}_{vi} - \mathbf{1}_{v'i}} - 1) d_i n_{vi} P(\boldsymbol{n}, t)$$

$$+ \sum_{v \in \mathcal{V}} \sum_{j=1}^{K} (\mathbb{E}^{-\widetilde{\boldsymbol{V}}_{vj}} - 1) f_{vj}(\boldsymbol{n}, \omega) P(\boldsymbol{n}, t), \quad (5)$$

where $f_{vj}(\boldsymbol{n}, \omega)$ is the propensity function, given by

$$f_{vj}(\boldsymbol{n}, \omega) = k_j \omega^{1 - \sum_{i=1}^{N} s_{ij}} \prod_{i=1}^{N} \frac{n_{vi}!}{(n_{vi} - s_{ij})!}. \quad (6)$$

In the large-diffusion limit, the RDME converges to the CME [31].

Before stating our results, we describe several existing concepts and results of deterministic reaction networks. For each reaction $R_j$ $(1 \leqslant j \leqslant K)$, the linear combinations $\sum_{i=1}^{N} s_{ij} X_i$ and $\sum_{i=1}^{N} r_{ij} X_i$ of the species in Eq. (1) are called the complexes of the reaction. Defining $\mathcal{C} = \{C_1, C_2, \ldots, C_M\}$ as the set of complexes, with $M = |\mathcal{C}|$, each reaction can be expressed as $C_i \xrightarrow{a_{ii'}} C_{i'}$, where $a_{ii'}$ denotes the reaction rate. For each $1 \leqslant i, i' \leqslant M$, $a_{ii'} = 0$ if $C_i \rightarrow C_{i'}$ is not present in the reaction network; otherwise, $a_{ii'} = k_j$ for some $j$ $(1 \leqslant j \leqslant K)$. The matrix $\boldsymbol{A} \in \mathbb{R}^{M \times M}$, called the *Kirchhoff matrix* of the reaction network, is defined as follows:

$$[\boldsymbol{A}]_{ii'} = \begin{cases} -\sum_{j=1}^{M} a_{ij}, & \text{if } i = i' \\ a_{i'i}, & \text{if } i \neq i' \end{cases}. \quad (7)$$

Let $\mathcal{X} = \{X_1, \ldots, X_N\}$ be the set of species and $\mathcal{R} = \bigcup_{i,i':a_{ii'}>0}\{C_i \rightarrow C_{i'}\}$ be the set of reactions in the network. The triple $\{\mathcal{X}, \mathcal{C}, \mathcal{R}\}$ then defines a reaction network. A reaction network $\{\mathcal{X}, \mathcal{C}, \mathcal{R}\}$ is called *weakly reversible* if for any reaction $C_i \rightarrow C_{i'} \in \mathcal{R}$, there exists a sequence of complexes $C_{i_1}, \ldots, C_{i_p} \in \mathcal{C}$ such that $C_{i'} \rightarrow C_{i_1}, C_{i_1} \rightarrow C_{i_2}, \ldots, C_{i_p} \rightarrow C_i \in \mathcal{R}$. One can construct a directed graph $\mathcal{G}$ corresponding to a reaction network in the following manner. For each $1 \leqslant i, i' \leqslant M$, draw a directed edge from $C_i$ to $C_{i'}$ if and only if $C_i \rightarrow C_{i'} \in \mathcal{R}$. We denote by $\ell$ the number of connected components of the underlying undirected graph of $\mathcal{G}$. The *deficiency* of a reaction network is an integer defined as $\delta = |\mathcal{C}| - \ell - \text{rank}(\boldsymbol{V})$. According to Ref. [32], $\delta$ is always nonnegative.

In a deterministic system, the vector of species concentrations, $\boldsymbol{c} = [c_1, c_2, \ldots, c_N]^\top \in \mathbb{R}_{\geqslant 0}^N$, temporally evolves as described by the following differential equations, which express the different form of rate equations:

$$\partial_t \boldsymbol{c} = \boldsymbol{Y} \cdot \boldsymbol{A} \cdot \Psi(\boldsymbol{c}). \quad (8)$$

Here, $\boldsymbol{Y} = [y_{ij}] \in \mathbb{N}_{\geqslant 0}^{N \times M}$ is the matrix of stoichiometric compositions of the complexes, i.e., $y_{ij}$ is the stoichiometric

coefficient of $C_j$ corresponding to species $X_i$, and $\Psi : \mathbb{R}^N \mapsto \mathbb{R}^M$ is a mapping given by

$$\Psi_j(\boldsymbol{c}) = \prod_{i=1}^{N} c_i^{y_{ij}}, \; j = 1, \ldots, M. \tag{9}$$

A reaction network is called *complex balanced* at $\boldsymbol{c} \in \mathbb{R}_{>0}^N$ if $\boldsymbol{A} \cdot \Psi(\boldsymbol{c}) = 0$. This condition means that for each complex $C_i \in \mathcal{C}$, $\sum_{C_i \to C_{i'}} a_{ii'} \Psi_i(\boldsymbol{c}) = \sum_{C_{i'} \to C_i} a_{i'i} \Psi_{i'}(\boldsymbol{c})$. In this case, $\boldsymbol{c}$ is a positive equilibrium of the network.

## III. RESULTS

The following states our first result.

*Theorem 1.* If a reaction network is complex balanced, its steady-state distribution is unaffected by diffusion.

*Proof.* As the network is complex balanced, there exists a positive equilibrium $\boldsymbol{c} = [c_1, \ldots, c_N]^\top \in \mathbb{R}_{>0}^N$ such that $\boldsymbol{A} \cdot \Psi(\boldsymbol{c}) = 0$. We note that the only requirement in our proof is the existence of some $\boldsymbol{c}$ such that $\boldsymbol{A} \cdot \Psi(\boldsymbol{c}) = 0$. Here, $\boldsymbol{c}$ is not the steady-state concentration in the presence of diffusion. Let $\Gamma \subseteq \mathbb{N}_{\geqslant 0}^N$ be the state space of the network, which may depend on the initialization. First, we prove the following ansatz: that the steady-state distribution of the RDME is given by a product $P_\Gamma(\boldsymbol{n}, t)$ of Poisson distributions:

$$P_\Gamma(\boldsymbol{n}, t) = \begin{cases} \mathcal{N}_\Gamma \prod_{v \in \mathcal{V}} \prod_{i=1}^{N} \frac{(\omega c_i)^{n_{vi}}}{n_{vi}!}, & \text{if } \sum_{v \in \mathcal{V}} \boldsymbol{n}_v \in \Gamma, \\ 0, & \text{if } \sum_{v \in \mathcal{V}} \boldsymbol{n}_v \notin \Gamma \end{cases}, \tag{10}$$

where $\mathcal{N}_\Gamma$ is the normalizing constant. For each $\boldsymbol{n}_v \in \mathbb{N}_{\geqslant 0}^N$, we define $P_\Gamma^*(\boldsymbol{n}_v, t) = \prod_{i=1}^{N} (\omega c_i)^{n_{vi}} / n_{vi}!$. $P_\Gamma(\boldsymbol{n}, t)$ can then be expressed as $P_\Gamma(\boldsymbol{n}, t) = \mathcal{N}_\Gamma \prod_{v \in \mathcal{V}} P_\Gamma^*(\boldsymbol{n}_v, t)$. Now, we need to show that $\partial_t P_\Gamma(\boldsymbol{n}, t) = 0$. Substituting $P_\Gamma(\boldsymbol{n}, t)$ in Eq. (10) into Eq. (5), the first term of the right-hand side becomes

$$\sum_{v \in \mathcal{V}} \sum_{v' \in N_e(v)} \sum_{i=1}^{N} (\mathbb{E}^{\mathbf{1}_{vi} - \mathbf{1}_{v'i}} - 1) d_i n_{vi} P_\Gamma(\boldsymbol{n}, t) = 0. \tag{11}$$

The second term on the right-hand side becomes the sum of the following values over all voxels $v \in \mathcal{V}$:

$$\sum_{j=1}^{K} (\mathbb{E}^{-\widetilde{V}_{vj}} - 1) f_{vj}(\boldsymbol{n}, \omega) P_\Gamma(\boldsymbol{n}, t)$$

$$= \mathcal{N}_\Gamma \prod_{v' \neq v} P_\Gamma^*(\boldsymbol{n}_{v'}, t) \sum_{j=1}^{K} (\mathbb{E}^{-V_j} - 1) f_j(\boldsymbol{n}_v, \omega) P_\Gamma^*(\boldsymbol{n}_v, t).$$

Exploiting the condition $\boldsymbol{A} \cdot \Psi(\boldsymbol{c}) = 0$, one can prove that [33]

$$\sum_{j=1}^{K} (\mathbb{E}^{-V_j} - 1) f_j(\boldsymbol{n}_v, \omega) P_\Gamma^*(\boldsymbol{n}_v, t) = 0. \tag{12}$$

Therefore, the second term also disappears and we obtain the desired result $\partial_t P_\Gamma(\boldsymbol{n}, t) = 0$. Let $\widehat{\boldsymbol{n}} = \sum_{v \in \mathcal{V}} \boldsymbol{n}_v$ represent the number of molecules of all species, i.e., $\widehat{n}_i$ is the total number of molecules of species $X_i$ in the system. To complete our theorem, we compute the steady-state distribution $P_\Gamma(\widehat{\boldsymbol{n}})$, and show its diffusion-independence. For $\widehat{\boldsymbol{n}} \notin \Gamma$, obviously
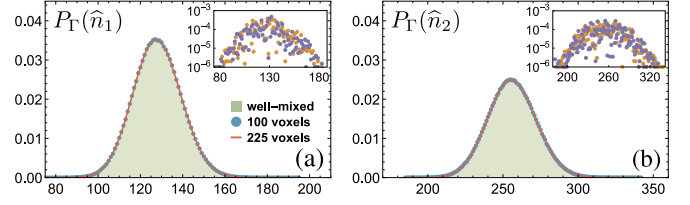


FIG. 1. Steady-state distributions (a) $P_\Gamma(\widehat{n}_1)$ of species $X_1$ and (b) $P_\Gamma(\widehat{n}_2)$ of species $X_2$ of nonlinear reaction network. Each panel shows the distributions of the 1-voxel system (green region), 100-voxel system (blue dots), and 225-voxel system (red line). The parameters are $k_1 = 4$, $k_2 = 1$, $k_3 = 2$, $\Omega = 128$. The diffusion rates of species $X_1$, $X_2$ are $d_1 = 1$, $d_2 = 2$ (100 voxels), and $d_1 = 2$, $d_2 = 1$ (225 voxels). Insets show the absolute probability differences $|P_{100} - P_1|$ (orange dots) and $|P_{225} - P_1|$ (violet dots), where $P_1$, $P_{100}$, and $P_{225}$ denote the probabilities in the 1-, 100-, and 225-voxel systems, respectively.

$P_\Gamma(\widehat{\boldsymbol{n}}) = 0$. For $\widehat{\boldsymbol{n}} \in \Gamma$, the explicit form of $P_\Gamma(\widehat{\boldsymbol{n}})$ is obtained as follows:

$$P_\Gamma(\widehat{\boldsymbol{n}}) = \sum_{\boldsymbol{n} : \sum_v \boldsymbol{n}_v = \widehat{\boldsymbol{n}}} P_\Gamma(\boldsymbol{n}, t) = \mathcal{N}_\Gamma \prod_{i=1}^{N} \frac{(\Omega c_i)^{\widehat{n}_i}}{\widehat{n}_i!}. \tag{13}$$

As $\mathcal{N}_\Gamma = (\sum_{\widehat{\boldsymbol{n}} \in \Gamma} \prod_{i=1}^{N} \frac{(\Omega c_i)^{\widehat{n}_i}}{\widehat{n}_i!})^{-1}$ does not depend on diffusion, the distribution $P_\Gamma(\widehat{\boldsymbol{n}})$ is also independent of diffusion. The details of these derivations can be seen in Appendix A.

Our theoretical result is empirically verified in simulations of the following complex balanced network:

$$\varnothing \xrightarrow{k_1} X_1 + 2X_2 \xrightarrow{k_2} X_2 \xrightarrow{k_3} \varnothing. \tag{14}$$
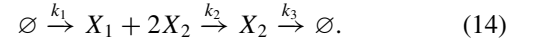
We consider three cases with different numbers of voxels in the system volume: 1 voxel (a well-mixed system), 100 voxels, and 225 voxels. The diffusion coefficients of the species in the 100-voxel system differ from those in the 225-voxel system. The steady-state distributions of species $X_1$ and $X_2$ are plotted in Fig. 1. As can be seen, the distributions of both species are consistent in all three cases. From these result, it is pertinent to ask which conditions define a complex balanced network. Reference [32] proved that a weakly reversible reaction network with zero deficiency is a complex balanced network. This implies that in some cases, a complex balanced network can be identified by its network topology. In Ref. [34], complex balanced realizations of a given kinetic polynomial system were computed by a linear programming algorithm.

Thus far, we show that the steady-state distribution of a complex balanced network is a product of Poisson distributions. A network with such a distribution implies that the system is diffusion-independent at steady state. Therefore, we desire to know the condition under which the system establishes a Poisson-like steady-state distribution. This condition is embodied in the following theorem.

*Theorem 2.* The network possesses the steady-state distribution defined in Eq. (10) in all state spaces $\Gamma \subseteq \mathbb{N}_{\geqslant 0}^N$ if and only if it is complex balanced.

*Proof.* We use the Fock space representation [35] to describe the molecule-number changes of each species inside each voxel. A state vector $|\boldsymbol{n}\rangle$ with configuration $\boldsymbol{n}$ means

that $n_{vi}$ molecules of species $X_i$ exist in voxel $v$. Using the annihilation and creation operators $a_{vi}, a_{vi}^\dagger$, i.e., $a_{vi}|n_{vi}\rangle = n_{vi}|n_{vi}-1\rangle$, $a_{vi}^\dagger|n_{vi}\rangle = |n_{vi}+1\rangle$, we can map the probability distribution $P_\Gamma(\boldsymbol{n}, t)$ to a state vector $|\psi(t)\rangle_\Gamma$, defined by

$$|\psi(t)\rangle_\Gamma = \sum_{\boldsymbol{n}} P_\Gamma(\boldsymbol{n}, t)|\boldsymbol{n}\rangle = \sum_{\boldsymbol{n}} P_\Gamma(\boldsymbol{n}, t)(\boldsymbol{a}^\dagger)^{\boldsymbol{n}}|0\rangle. \quad (15)$$

This expression sums over all possible configurations $\boldsymbol{n}$ weighted by their occurrence probabilities at time $t$. To establish the time evolution of this state vector, we apply the master equation to obtain the following Schrödinger equation:

$$\partial_t|\psi(t)\rangle_\Gamma = -\mathcal{H}(\boldsymbol{a}^\dagger, \boldsymbol{a})|\psi(t)\rangle_\Gamma, \quad (16)$$

where $\mathcal{H}(\boldsymbol{a}^\dagger, \boldsymbol{a})$ represents the Hamiltonian action on the Fock space, expressed as shown in Appendix B. In general, $\mathcal{H}(\boldsymbol{a}^\dagger, \boldsymbol{a})$ is the sum of several sub-actions created by each reaction of the system, e.g., a reaction of the form $\sum_{i=1}^N s_{ij} X_i^v \xrightarrow{k_j} \sum_{i=1}^N r_{ij} X_i^v$ yields a sub-action $k_j \omega^{1-\sum_{i=1}^N s_{ij}} (\prod_{i=1}^N (a_{vi}^\dagger)^{r_{ij}} - \prod_{i=1}^N (a_{vi}^\dagger)^{s_{ij}}) \prod_{i=1}^N (a_{vi})^{s_{ij}}$. The action $\mathcal{H}(\boldsymbol{a}^\dagger, \boldsymbol{a})$ is considered to be normally ordered, i.e., $a_{vi}^\dagger$ is always to the left of $a_{vi}$. In a steady-state system, $\mathcal{H}(\boldsymbol{a}^\dagger, \boldsymbol{a})|\psi(t)\rangle_\Gamma = 0$. Consequently, $\mathcal{H}(\boldsymbol{a}^\dagger, \boldsymbol{a})|\psi(t)\rangle$ is also 0, where the state $|\psi(t)\rangle$ is defined as follows:

$$|\psi(t)\rangle = \sum_\Gamma \frac{|\psi(t)\rangle_\Gamma}{\mathcal{N}_\Gamma} = \prod_{i=1}^N \prod_{v=1}^{|\mathcal{V}|} \sum_{n_{vi} \geqslant 0} \frac{(\omega c_i a_{vi}^\dagger)^{n_{vi}}}{n_{vi}!}|0\rangle.$$
$$= e^{\sum_{v,i} \omega c_i a_{vi}^\dagger}|0\rangle.$$

In other words, the condition

$$\mathcal{H}(\boldsymbol{a}^\dagger, \boldsymbol{a})e^{\sum_{v,i} \omega c_i a_{vi}^\dagger}|0\rangle = 0 \quad (17)$$

must hold. To derive a further condition with no involvement of $\boldsymbol{a}^\dagger$ and $\boldsymbol{a}$, we consider the coherent states $|\phi_{vi}\rangle$ and $\langle\phi_{vi}|$, defining the right and left eigenstates of $a_{vi}$ and $a_{vi}^\dagger$, respectively. Specifically, $a_{vi}|\phi_{vi}\rangle = \phi_{vi}|\phi_{vi}\rangle$ and $\langle\phi_{vi}|a_{vi}^\dagger = \langle\phi_{vi}|\phi_{vi}^*$, with complex eigenvalue $\phi_{vi} \in \mathbb{C}$. Multiplying both sides of Eq. (17) by the left coherent state $\langle\boldsymbol{\phi}|$, we obtain

$$0 = \langle\boldsymbol{\phi}|\mathcal{H}(\boldsymbol{a}^\dagger, \boldsymbol{a})e^{\sum_{v,i} \omega c_i a_{vi}^\dagger}|0\rangle \Leftrightarrow 0 = \mathcal{H}(\boldsymbol{\phi}^*, \omega\widetilde{\boldsymbol{c}}), \quad (18)$$

where $\widetilde{\boldsymbol{c}} \in \mathbb{R}^{|\mathcal{V}|N}$ is defined as $\widetilde{c}_{vi} = c_i$. As $\mathcal{H}(\boldsymbol{\phi}^*, \omega\widetilde{\boldsymbol{c}})$ is a polynomial of $\boldsymbol{\phi}^*$, this result is possible only when the coefficients of all monomials are zero. Each reaction of the form $C_i \xrightarrow{a_{ii'}} C_{i'}$ in voxel $v$ contributes to $\mathcal{H}(\boldsymbol{\phi}^*, \omega\widetilde{\boldsymbol{c}})$ a quantity $\omega a_{ii'}(\Psi_{i'}(\boldsymbol{\phi}_v^*) - \Psi_i(\boldsymbol{\phi}_v^*))\Psi_i(\boldsymbol{c})$. Therefore, by collecting the coefficients of $\Psi_i(\boldsymbol{\phi}_v^*)$ for each $i = 1, \dots, M$ and $v \in \mathcal{V}$, we obtain the following relation:

$$\mathcal{H}(\boldsymbol{\phi}^*, \omega\widetilde{\boldsymbol{c}}) = 0, \ \forall \boldsymbol{\phi} \in \mathbb{C}^{|\mathcal{V}|N} \Leftrightarrow \boldsymbol{A} \cdot \Psi(\boldsymbol{c}) = 0, \quad (19)$$

meaning that the network is complex balanced at $\boldsymbol{c}$. The details of these calculations are shown in Appendix B. From these results, we conclude that the necessary and sufficient condition for a steady-state distribution [Eq. (10)] is that the network is complex balanced.

We note that the sufficient condition of Theorem 2 has been studied in Ref. [36] [i.e., if the network is complex balanced, then the steady-state distribution has a form as in Eq. (10)]. Above we investigate the relation between diffusion

and the distributions of the reactant species in steady state. We now present another result that holds under non-steady-state conditions.

*Theorem 3.* When the reaction network is linear, the RDME can be reduced to the CME. Equivalently, the diffusion can be ignored in such case.

*Proof.* The system volume $\Omega$ is related to the voxel volume $\omega$ as $\Omega = |\mathcal{V}|\omega$. Now, for each state vector $\widehat{\boldsymbol{n}} = [\widehat{n}_1, \widehat{n}_2, \dots, \widehat{n}_N]^\top \in \mathbb{N}_{\geqslant 0}^N$ representing the number of molecules of the reactant species, i.e., $\widehat{n}_i$ is the total number of molecules of species $X_i$ in the system, we define the set $\mathcal{S}(\widehat{\boldsymbol{n}}) = \{\boldsymbol{n} \in \mathbb{N}_{\geqslant 0}^{|\mathcal{V}|N} \mid \sum_{v\in\mathcal{V}} \boldsymbol{n}_v = \widehat{\boldsymbol{n}}\}$. Let $P(\widehat{\boldsymbol{n}}, t)$ be the probability of the system being in state $\widehat{\boldsymbol{n}}$ at time $t$. In terms of $P(\boldsymbol{n}, t)$, this probability becomes $P(\widehat{\boldsymbol{n}}, t) = \sum_{\boldsymbol{n}\in\mathcal{S}(\widehat{\boldsymbol{n}})} P(\boldsymbol{n}, t)$. As $\sum_{\widehat{\boldsymbol{n}}} P(\widehat{\boldsymbol{n}}, t) = \sum_{\boldsymbol{n}} P(\boldsymbol{n}, t) = 1$, $P(\widehat{\boldsymbol{n}}, t)$ is a probability distribution. To show that this probability distribution satisfies the CME given by Eq. (2), we calculate the time derivative of $P(\widehat{\boldsymbol{n}}, t)$ as follows:

$$\partial_t P(\widehat{\boldsymbol{n}}, t) = \sum_{\boldsymbol{n}\in\mathcal{S}(\widehat{\boldsymbol{n}})} \partial_t P(\boldsymbol{n}, t). \quad (20)$$

Substituting Eq. (5) into the right-hand side of Eq. (20), we obtain an equation with both diffusion and reaction terms on the right. After some algebraic transformations, the diffusion term disappears and only the reaction term remains (see Appendix C). As the reaction network is linear, i.e., $\sum_i s_{ij} \leqslant 1 \forall j = 1, \dots, K$, the propensity function $f_{vj}(\boldsymbol{n}, \omega)$ must be one of two forms: $f_{vj}(\boldsymbol{n}, \omega) = k_j n_{vi}$ or $k_j \omega$. Substituting the exact form of each propensity function into Eq. (20), we finally obtain the following master equation for $P(\widehat{\boldsymbol{n}}, t)$:

$$\partial_t P(\widehat{\boldsymbol{n}}, t) = \sum_{j=1}^K (\mathbb{E}^{-V_j} - 1) f_j(\widehat{\boldsymbol{n}}, \Omega) P(\widehat{\boldsymbol{n}}, t). \quad (21)$$

Obviously, this differential equation is identical to the CME stated in Eq. (2), and contains no diffusion factors. Therefore, it can be concluded that diffusion can be ignored in linear reaction networks.

In Theorem 3, we demonstrate that the RDME reduces to the CME in the case of linear reaction networks, which implies that diffusion does not affect the stochastic dynamics of the system at an arbitrary time. From the view of the Smoluchowski model, this statement appears to be obvious. By regarding the network as interacting many-particle system and introducing diffusion and reaction operators [37], the same result can be derived. However, it is not evident from the view of the RDME. The agreement of results in these different models serves as validation for the RDME.

We numerically verify the result of Theorem 3 on a simple linear reaction network, namely, a coarse-grained model of enzymatic reactions and gene expressions. The network consists of two reactant species $X_1$ and $X_2$ and four reactions [6,38]:

$$\varnothing \underset{k_2}{\overset{k_1}{\rightleftharpoons}} X_1, \ X_1 \xrightarrow{k_3} X_1 + X_2, \ X_2 \xrightarrow{k_4} \varnothing. \quad (22)$$
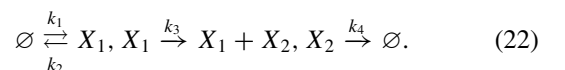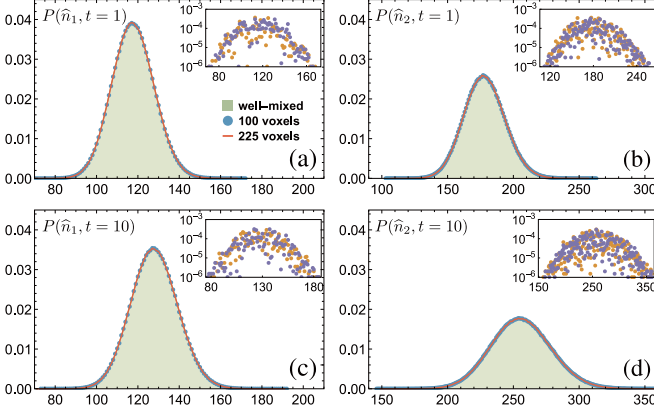
FIG. 2. Probability distributions (a) $P(\widehat{n}_1, t = 1)$, (b) $P(\widehat{n}_2, t = 1)$, (c) $P(\widehat{n}_1, t = 10)$, and (d) $P(\widehat{n}_2, t = 10)$ of two species $X_1$, $X_2$ at times $t = 1$ (upper panels) and $t = 10$ (lower panels) of linear reaction network. Each panel shows the distributions of the 1-voxel system (green region), 100-voxel system (blue dots), and 225-voxel system (red line). The parameters are $k_1 = 1, k_2 = 1, k_3 = 2, k_4 = 1, \Omega = 128$. The diffusion rates of species $X_1$, $X_2$ are $d_1 = 1, d_2 = 2$ (100 voxels) and $d_1 = 2, d_2 = 1$ (225 voxels). Insets show the absolute probability differences $|P_{100} - P_1|$ (orange dots) and $|P_{225} - P_1|$ (violet dots), where $P_1$, $P_{100}$ and $P_{225}$ indicate the probability in the 1-, 100-, and 225-voxel systems, respectively.

Again, we divide the cell volume into 1, 100, and 225 voxels with different diffusion coefficients of $X_1$ and $X_2$. The result is displayed in Fig. 2. As before, the distributions of each species at times $t = 1$ and $t = 10$ are identical in all three cases. These numerical results empirically validate Theorem 3.

## IV. CONCLUSIONS

In summary, within the approximation of the RDME, we proved that diffusion in complex-balanced networks does not affect the steady-state distribution of the system. We also showed that a diffusion-included reaction network has a Poisson-like steady-state distribution if and only if it is complex balanced, analogously to the well-mixed case described in [39]. Moreover, we demonstrated that the RDME can be reduced to the CME in the case of linear reaction networks. These results help to clarify the conditions under which diffusion is negligible. Under such conditions, the system can be described by the CME instead of the intractable RDME. In nonlinear networks that are not complex-balanced, how diffusion affects the stochastic system dynamics, or whether it can be ignored, requires further investigation.

It appears that functional biological networks satisfying the complex balanced condition are not widespread in real-world systems. Nevertheless, weakly reversible networks have been successfully applied in modeling signal transduction pathways [40] and asymmetric stem-cell division [41]. Besides that, complex balanced networks whose Fano factor is equal to one can be used in analyzing and approximating cascade networks or metabolic pathways, wherein noise is not propagated from upstream to downstream [42]. Although our results are obtained with the approximation of the RDME, it is expected that the derived diffusion-dynamics laws provide suggestive results for real physical systems.

## APPENDIX A: DETAILED CALCULATIONS IN THEOREM 1

### 1. Detailed calculations of Eq. (11)

The detailed calculation of Eq. (11) is given below:

$$\sum_{v \in \mathcal{V}} \sum_{v' \in N_e(v)} \sum_{i=1}^{N} (\mathbf{E}^{\mathbf{1}_{vi} - \mathbf{1}_{v'i}} - 1) d_i n_{vi} P_\Gamma(\boldsymbol{n}, t)$$

$$= \sum_{v \in \mathcal{V}} \sum_{v' \in N_e(v)} \sum_{i=1}^{N} [d_i (n_{vi} + 1) P_\Gamma(\boldsymbol{n} + \mathbf{1}_{vi} - \mathbf{1}_{v'i}, t) - d_i n_{vi} P_\Gamma(\boldsymbol{n}, t)]$$

$$= \sum_{i=1}^{N} d_i \sum_{v \in \mathcal{V}} \sum_{v' \in N_e(v)} \left[ \mathcal{N}_\Gamma (n_{vi} + 1) \frac{(\omega c_i)^{n_{vi} + 1}}{(n_{vi} + 1)!} \frac{(\omega c_i)^{n_{v'i} - 1}}{(n_{v'i} - 1)!} \prod_{i' \neq i} \frac{(\omega c_{i'})^{n_{vi'}} (\omega c_{i'})^{n_{v'i'}}}{n_{vi'}! n_{v'i'}!} \prod_{\widetilde{v} \neq v, v'} P_\Gamma^*(\boldsymbol{n}_{\widetilde{v}}, t) - n_{vi} P_\Gamma(\boldsymbol{n}, t) \right]$$

$$= \sum_{i=1}^{N} d_i \sum_{v \in \mathcal{V}} \sum_{v' \in N_e(v)} \left[ \mathcal{N}_\Gamma \frac{(\omega c_i)^{n_{vi}}}{n_{vi}!} \frac{(\omega c_i)^{n_{v'i}}}{(n_{v'i} - 1)!} \prod_{i' \neq i} \frac{(\omega c_{i'})^{n_{vi'}} (\omega c_{i'})^{n_{v'i'}}}{n_{vi'}! n_{v'i'}!} \prod_{\widetilde{v} \neq v, v'} P_\Gamma^*(\boldsymbol{n}_{\widetilde{v}}, t) - n_{vi} P_\Gamma(\boldsymbol{n}, t) \right]$$

$$= \sum_{i=1}^{N} d_i \sum_{v \in \mathcal{V}} \sum_{v' \in N_e(v)} [n_{v'i} P_\Gamma(\boldsymbol{n}, t) - n_{vi} P_\Gamma(\boldsymbol{n}, t)]$$

$$= 0.$$

### 2. Detailed calculations of Eq. (12)

We here reveal the details of Eq. (12). The equation $\boldsymbol{A} \cdot \Psi(\boldsymbol{c}) = 0$ means that $\sum_{C_i \to C_{i'}} a_{ii'} \Psi_i(\boldsymbol{c}) - \sum_{C_{i'} \to C_i} a_{i'i} \Psi_{i'}(\boldsymbol{c}) = 0$ for each complex $C_{i'} \in \mathcal{C}$. The left side of Eq. (12) can be transformed as

$$\sum_{j=1}^{K} (\mathbb{E}^{-\boldsymbol{V}_j} - 1) f_j(\boldsymbol{n}_v, \omega) P_\Gamma^*(\boldsymbol{n}_v, t)$$

$$= \sum_{j=1}^{K} [f_j(\boldsymbol{n}_v - \boldsymbol{V}_j, \omega) P_\Gamma^*(\boldsymbol{n}_v - \boldsymbol{V}_j, t) - f_j(\boldsymbol{n}_v, \omega) P_\Gamma^*(\boldsymbol{n}_v, t)]$$

$$= \sum_{C_i \to C_{i'}} a_{ii'} \omega^{1 - \sum_{k=1}^{N} y_{ki}} \left[ \prod_{k=1}^{N} \frac{(n_{vk} + y_{ki} - y_{ki'})!}{(n_{vk} - y_{ki'})!} \frac{(\omega c_k)^{n_{vk} + y_{ki} - y_{ki'}}}{(n_{vk} + y_{ki} - y_{ki'})!} - \prod_{k=1}^{N} \frac{n_{vk}!}{(n_{vk} - y_{ki})!} \frac{(\omega c_k)^{n_{vk}}}{n_{vk}!} \right]$$

$$= \sum_{C_i \to C_{i'}} a_{ii'} \omega^{1 - \sum_{k=1}^{N} y_{ki}} \left[ \prod_{k=1}^{N} \frac{(\omega c_k)^{n_{vk} + y_{ki} - y_{ki'}}}{(n_{vk} - y_{ki'})!} - \prod_{k=1}^{N} \frac{(\omega c_k)^{n_{vk}}}{(n_{vk} - y_{ki})!} \right]$$

$$= \omega \sum_{C_i \to C_{i'}} \left[ a_{ii'} \frac{\Psi_i(\boldsymbol{c})}{\Psi_{i'}(\boldsymbol{c})} \prod_{k=1}^{N} \frac{\omega^{n_{vk} - y_{ki'}} c_k^{n_{vk}}}{(n_{vk} - y_{ki'})!} - a_{ii'} \prod_{k=1}^{N} \frac{\omega^{n_{vk} - y_{ki}} c_k^{n_{vk}}}{(n_{vk} - y_{ki})!} \right]$$

$$= \omega \sum_{C_{i'} \in \mathcal{C}} \left[ \sum_{C_i \to C_{i'}} a_{ii'} \frac{\Psi_i(\boldsymbol{c})}{\Psi_{i'}(\boldsymbol{c})} \prod_{k=1}^{N} \frac{\omega^{n_{vk} - y_{ki'}} c_k^{n_{vk}}}{(n_{vk} - y_{ki'})!} - \sum_{C_{i'} \to C_i} a_{i'i} \prod_{k=1}^{N} \frac{\omega^{n_{vk} - y_{ki'}} c_k^{n_{vk}}}{(n_{vk} - y_{ki'})!} \right]$$

$$= \omega \sum_{C_{i'} \in \mathcal{C}} \frac{1}{\Psi_{i'}(\boldsymbol{c})} \prod_{k=1}^{N} \frac{\omega^{n_{vk} - y_{ki'}} c_k^{n_{vk}}}{(n_{vk} - y_{ki'})!} \left[ \sum_{C_i \to C_{i'}} a_{ii'} \Psi_i(\boldsymbol{c}) - \sum_{C_{i'} \to C_i} a_{i'i} \Psi_{i'}(\boldsymbol{c}) \right]$$

$$= 0.$$

The same result (but omitting the details) is given in Ref. [33].

### 3. Detailed calculations of Eq. (13)

Finally, we compute the explicit form of the distribution $P_\Gamma(\widehat{\boldsymbol{n}})$ in Eq. (13). We have

$$P_\Gamma(\widehat{\boldsymbol{n}}) = \sum_{\boldsymbol{n}: \sum_v \boldsymbol{n}_v = \widehat{\boldsymbol{n}}} P_\Gamma(\boldsymbol{n}, t) = \mathcal{N}_\Gamma \sum_{\substack{\boldsymbol{n}_1, \ldots, \boldsymbol{n}_{|\mathcal{V}|} \in \mathbb{N}_{\geqslant 0}^N \\ \sum_v \boldsymbol{n}_v = \widehat{\boldsymbol{n}}}} \prod_{v \in \mathcal{V}} \prod_{i=1}^{N} \frac{(\omega c_i)^{n_{vi}}}{n_{vi}!} = \mathcal{N}_\Gamma \prod_{i=1}^{N} \left\{ \sum_{\substack{n_{1i}, \ldots, n_{|\mathcal{V}|i} \geqslant 0 \\ \sum_{v \in \mathcal{V}} n_{vi} = \widehat{n}_i}} \prod_{v \in \mathcal{V}} \frac{(\omega c_i)^{n_{vi}}}{n_{vi}!} \right\}$$

$$= \mathcal{N}_\Gamma \prod_{i=1}^{N} \left\{ \frac{(\omega c_i)^{\widehat{n}_i}}{\widehat{n}_i!} \sum_{\substack{n_{1i}, \ldots, n_{|\mathcal{V}|i} \geqslant 0 \\ \sum_{v \in \mathcal{V}} n_{vi} = \widehat{n}_i}} \frac{(\sum_{v \in \mathcal{V}} n_{vi})!}{\prod_{v \in \mathcal{V}} n_{vi}!} \right\} = \mathcal{N}_\Gamma \prod_{i=1}^{N} \frac{(\omega c_i)^{\widehat{n}_i}}{\widehat{n}_i!} |\mathcal{V}|^{\widehat{n}_i} = \mathcal{N}_\Gamma \prod_{i=1}^{N} \frac{(\Omega c_i)^{\widehat{n}_i}}{\widehat{n}_i!}.$$

In transforming the fourth equation to the fifth one, we exploited the following equality:

$$\sum_{\substack{x_1, \ldots, x_m \geqslant 0 \\ \sum_{i=1}^{m} x_i = n}} \frac{(x_1 + \cdots + x_m)!}{x_1! \ldots x_m!} = m^n, \forall m \in \mathbb{N}_{>0}, n \in \mathbb{N}_{\geqslant 0}.$$

### APPENDIX B: DETAILED CALCULATIONS IN THEOREM 2

Before presenting the calculations, we state several properties of the bosonic operators $a_{vi}^\dagger$ and $a_{vi}$.

$$|n_{vi}\rangle = (a_{vi}{}^\dagger)^{n_{vi}} |0\rangle,$$

$$(a_{vi})^l (a_{vi}{}^\dagger)^k |n_{vi}\rangle = \prod_{j=0}^{l-1} (n_{vi} + k - j) |n_{vi} + k - l\rangle,$$

$$(a_{vi}^\dagger)^k (a_{vi})^l |n_{vi}\rangle = \prod_{j=0}^{l-1}(n_{vi}-j)|n_{vi}+k-l\rangle,$$

$$[a_{vi}, a_{v'i'}^\dagger] = a_{vi} a_{v'i}'^\dagger - a_{v'i'}^\dagger a_{vi} = \delta_{vv'}\delta_{ii'},$$

$$[a_{vi}^\dagger, a_{v'i'}^\dagger] = [a_{vi}, a_{v'i'}] = 0.$$

For a general configuration $\boldsymbol{n}$, we define the corresponding state vector $|\boldsymbol{n}\rangle$ as

$$|\boldsymbol{n}\rangle = (\boldsymbol{a}^\dagger)^{\boldsymbol{n}}|0\rangle = \prod_{v\in\mathcal{V}}\prod_{i=1}^{N}(a_{vi}^\dagger)^{n_{vi}}|0\rangle. \tag{B1}$$

For convenience, we note that

$$e^{ca_{vi}}f(a_{vi}^\dagger) = f(a_{vi}^\dagger + c)e^{ca_{vi}}, \tag{B2}$$

$$e^{ca_{vi}^\dagger}f(a_{vi}) = f(a_{vi}-c)e^{ca_{vi}^\dagger}, \tag{B3}$$

where $c \in \mathbb{C}$ is a complex number and $f$ is an arbitrary function.

### 1. Detailed calculations of Eq. (16)

We first derive the explicit form of the Hamiltonian action $\mathcal{H}(\boldsymbol{a}^\dagger, \boldsymbol{a})$ in Eq. (16). Suppose that the network contains a set $\mathcal{R}$ of reactions $R_j$ of the general form $\sum_{v\in\mathcal{V}}\sum_{i=1}^{N} p_{vi}^j X_i^v \xrightarrow{k_j} \sum_{v\in\mathcal{V}}\sum_{i=1}^{N} q_{vi}^j X_i^v$, where $p_{vi}^j$ and $q_{vi}^j$ are the stoichiometric coefficients. For each reaction $R_j$, we define a stoichiometric vector $\boldsymbol{V}^j \in \mathbb{Z}^{|\mathcal{V}|N}$ as $V_{vi}^j = q_{vi}^j - p_{vi}^j$. Starting from the master equation, we have

$$\partial_t |\psi(t)\rangle_\Gamma = \sum_{\boldsymbol{n}} \partial_t P_\Gamma(\boldsymbol{n}, t)(\boldsymbol{a}^\dagger)^{\boldsymbol{n}}|0\rangle = \sum_{\boldsymbol{n}} \sum_{R_j\in\mathcal{R}} k_j \omega^{1-\sum_{v,i} p_{vi}^j}\left[\prod_{v,i}\frac{(n_{vi}+p_{vi}^j-q_{vi}^j)!}{(n_{vi}-q_{vi}^j)!}P_\Gamma(\boldsymbol{n}-\boldsymbol{V}^j, t)\right.$$
$$\left. - \prod_{v,i}\frac{n_{vi}!}{(n_{vi}-p_{vi}^j)!}P_\Gamma(\boldsymbol{n}, t)\right](\boldsymbol{a}^\dagger)^{\boldsymbol{n}}|0\rangle.$$

Note that the two terms inside the bracket can be obtained using operators as follows:

$$\prod_{v,i}\frac{(n_{vi}+p_{vi}^j-q_{vi}^j)!}{(n_{vi}-q_{vi}^j)!}P_\Gamma(\boldsymbol{n}-\boldsymbol{V}^j, t)(\boldsymbol{a}^\dagger)^{\boldsymbol{n}}|0\rangle = \prod_{v,i}(a_{vi}^\dagger)^{q_{vi}^j}(a_{vi})^{p_{vi}^j}P_\Gamma(\boldsymbol{n}-\boldsymbol{V}^j, t)(\boldsymbol{a}^\dagger)^{\boldsymbol{n}-\boldsymbol{V}^j}|0\rangle,$$

$$\prod_{v,i}\frac{n_{vi}!}{(n_{vi}-p_{vi}^j)!}P_\Gamma(\boldsymbol{n}, t)(\boldsymbol{a}^\dagger)^{\boldsymbol{n}}|0\rangle = \prod_{v,i}(a_{vi}^\dagger)^{p_{vi}^j}(a_{vi})^{p_{vi}^j}P_\Gamma(\boldsymbol{n}, t)(\boldsymbol{a}^\dagger)^{\boldsymbol{n}}|0\rangle.$$
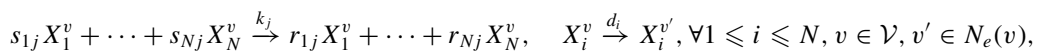
Using these equalities, $\partial_t|\psi(t)\rangle_\Gamma$ is calculated as follows:

$$\partial_t |\psi(t)\rangle_\Gamma = \sum_{\boldsymbol{n}} \sum_{R_j\in\mathcal{R}} k_j \omega^{1-\sum_{v,i} p_{vi}^j}\left[\prod_{v,i}(a_{vi}^\dagger)^{q_{vi}^j}(a_{vi})^{p_{vi}^j}P_\Gamma(\boldsymbol{n}-\boldsymbol{V}^j, t)(\boldsymbol{a}^\dagger)^{\boldsymbol{n}-\boldsymbol{V}^j}|0\rangle - \prod_{v,i}(a_{vi}^\dagger)^{p_{vi}^j}(a_{vi})^{p_{vi}^j}P_\Gamma(\boldsymbol{n}, t)(\boldsymbol{a}^\dagger)^{\boldsymbol{n}}|0\rangle\right]$$

$$= \sum_{R_j\in\mathcal{R}} k_j \omega^{1-\sum_{v,i} p_{vi}^j}\left[\prod_{v,i}(a_{vi}^\dagger)^{q_{vi}^j}(a_{vi})^{p_{vi}^j}\sum_{\boldsymbol{n}}P_\Gamma(\boldsymbol{n}-\boldsymbol{V}^j, t)(\boldsymbol{a}^\dagger)^{\boldsymbol{n}-\boldsymbol{V}^j}|0\rangle - \prod_{v,i}(a_{vi}^\dagger)^{p_{vi}^j}(a_{vi})^{p_{vi}^j}\sum_{\boldsymbol{n}}P_\Gamma(\boldsymbol{n}, t)(\boldsymbol{a}^\dagger)^{\boldsymbol{n}}|0\rangle\right]$$

$$= \sum_{R_j\in\mathcal{R}} k_j \omega^{1-\sum_{v,i} p_{vi}^j}\left[\prod_{v,i}(a_{vi}^\dagger)^{q_{vi}^j}(a_{vi})^{p_{vi}^j} - \prod_{v,i}(a_{vi}^\dagger)^{p_{vi}^j}(a_{vi})^{p_{vi}^j}\right]|\psi(t)\rangle_\Gamma.$$

Thus, the general form of $\mathcal{H}$ is obtained as

$$\mathcal{H}(\boldsymbol{a}^\dagger, \boldsymbol{a}) = \sum_{R_j\in\mathcal{R}} k_j \omega^{1-\sum_{v,i} p_{vi}^j}\left[\prod_{v,i}(a_{vi}^\dagger)^{q_{vi}^j} - \prod_{v,i}(a_{vi}^\dagger)^{p_{vi}^j}\right]\prod_{v,i}(a_{vi})^{p_{vi}^j}. \tag{B4}$$

For a diffusion-included reaction network involving the following reactions

$$s_{1j}X_1^v + \cdots + s_{Nj}X_N^v \xrightarrow{k_j} r_{1j}X_1^v + \cdots + r_{Nj}X_N^v, \quad X_i^v \xrightarrow{d_i} X_i^{v'}, \forall 1 \leqslant i \leqslant N, v \in \mathcal{V}, v' \in N_e(v),$$

the Hamiltonian action $\mathcal{H}(\boldsymbol{a}^\dagger, \boldsymbol{a})$ in Eq. (16) takes the following form

$$\mathcal{H}(\boldsymbol{a}^\dagger, \boldsymbol{a}) = \sum_{j=1}^{K} \sum_{v \in \mathcal{V}} k_j \omega^{1-\sum_{i=1}^{N} s_{ij}} \left[ \prod_{i=1}^{N} (a_{vi}^\dagger)^{r_{ij}} - \prod_{i=1}^{N} (a_{vi}^\dagger)^{s_{ij}} \right] \prod_{i=1}^{N} (a_{vi})^{s_{ij}} + \sum_{i=1}^{N} \sum_{v \in \mathcal{V}} \sum_{v' \in N_e(v)} d_i (a_{v'i}^\dagger - a_{vi}^\dagger) a_{vi}. \tag{B5}$$

We note that this form of $\mathcal{H}$ is already normal-ordered.

### 2. Detailed calculations of Eq. (18)

Equation (18) is derived through the following steps:

$$0 = \langle \boldsymbol{\phi} | \mathcal{H}(\boldsymbol{a}^\dagger, \boldsymbol{a}) e^{\sum_{v,i} \omega c_i a_{vi}^\dagger} |0\rangle \tag{B6}$$

$$\Leftrightarrow 0 = \langle \boldsymbol{\phi} | e^{\sum_{v,i} \omega c_i a_{vi}^\dagger} \mathcal{H}(\boldsymbol{a}^\dagger, \boldsymbol{a} + \omega \widetilde{\boldsymbol{c}}) |0\rangle \tag{B7}$$

$$\Leftrightarrow 0 = e^{\sum_{v,i} \omega c_i \phi_{vi}^*} \langle \boldsymbol{\phi} | \mathcal{H}(\boldsymbol{a}^\dagger, \boldsymbol{a} + \omega \widetilde{\boldsymbol{c}}) |0\rangle \tag{B8}$$

$$\Leftrightarrow 0 = e^{\sum_{v,i} \omega c_i \phi_{vi}^*} \langle \boldsymbol{\phi} | \mathcal{H}(\boldsymbol{\phi}^*, \omega \widetilde{\boldsymbol{c}}) |0\rangle \tag{B9}$$

$$\Leftrightarrow 0 = e^{\sum_{v,i} \omega c_i \phi_{vi}^*} \mathcal{H}(\boldsymbol{\phi}^*, \omega \widetilde{\boldsymbol{c}}) \langle \boldsymbol{\phi} |0\rangle \tag{B10}$$

$$\Leftrightarrow 0 = \mathcal{H}(\boldsymbol{\phi}^*, \omega \widetilde{\boldsymbol{c}}). \tag{B11}$$

In Eq. (B7), we use the property stated in Eq. (B3). In Eq. (B9), the operators $a_{vi}$ ($v \in \mathcal{V}, i = 1, \dots, N$) are absorbed into $|0\rangle$ ($\because a_{vi}|0\rangle = 0$), and the operators $a_{vi}^\dagger$ are replaced by $\phi_{vi}^*$ ($\because \langle \phi_{vi} | a_{vi}^\dagger = \langle \phi_{vi} | \phi_{vi}^*$). The result Eq. (B11) is obtained by noting that

$$e^{\sum_{v,i} \omega c_i \phi_{vi}^*} \neq 0,$$

$$\langle \boldsymbol{\phi} |0\rangle = \prod_{v,i} \langle \phi_{vi} |0\rangle = \prod_{v,i} e^{-\frac{1}{2}|\phi_{vi}|^2} \neq 0.$$

### 3. Detailed calculations of Eq. (19)

Equation (19) is given by

$$\mathcal{H}(\boldsymbol{\phi}^*, \omega \widetilde{\boldsymbol{c}}) = 0, \forall \boldsymbol{\phi} \in \mathbb{C}^{|\mathcal{V}|N} \Leftrightarrow \boldsymbol{A} \cdot \Psi(\boldsymbol{c}) = 0.$$

The Hamiltonian action $\mathcal{H}(\boldsymbol{a}^\dagger, \boldsymbol{a})$ is described by

$$\mathcal{H}(\boldsymbol{a}^\dagger, \boldsymbol{a}) = \sum_{j=1}^{K} \sum_{v \in \mathcal{V}} k_j \omega^{1-\sum_{i=1}^{N} s_{ij}} \left[ \prod_{i=1}^{N} (a_{vi}^\dagger)^{r_{ij}} - \prod_{i=1}^{N} (a_{vi}^\dagger)^{s_{ij}} \right] \prod_{i=1}^{N} (a_{vi})^{s_{ij}} + \sum_{i=1}^{N} \sum_{v \in \mathcal{V}} \sum_{v' \in N_e(v)} d_i (a_{v'i}^\dagger - a_{vi}^\dagger) a_{vi}.$$

The case $\mathcal{H}(\boldsymbol{\phi}^*, \omega \widetilde{\boldsymbol{c}}) = 0$ is equivalent to

$$\sum_{j=1}^{K} \sum_{v \in \mathcal{V}} k_j \omega^{1-\sum_{i=1}^{N} s_{ij}} \left[ \prod_{i=1}^{N} (\phi_{vi}^*)^{r_{ij}} - \prod_{i=1}^{N} (\phi_{vi}^*)^{s_{ij}} \right] \prod_{i=1}^{N} (\omega \widetilde{c}_{vi})^{s_{ij}} + \sum_{i=1}^{N} \sum_{v \in \mathcal{V}} \sum_{v' \in N_e(v)} d_i (\phi_{v'i}^* - \phi_{vi}^*) \omega \widetilde{c}_{vi} = 0$$

$$\Leftrightarrow \sum_{C_i \to C_{i'}} \sum_{v \in \mathcal{V}} \omega a_{ii'} [\Psi_{i'}(\boldsymbol{\phi}_v^*) - \Psi_i(\boldsymbol{\phi}_v^*)] \Psi_i(\boldsymbol{c}) + \sum_{i=1}^{N} d_i \sum_{\substack{v, v' \in \mathcal{V} \\ v' \in N_e(v) \\ v \in N_e(v')}} [(\phi_{v'i}^* - \phi_{vi}^*) \omega \widetilde{c}_{vi} + (\phi_{vi}^* - \phi_{v'i}^*) \omega \widetilde{c}_{v'i}] = 0$$

$$\Leftrightarrow \omega \sum_{v \in \mathcal{V}} \sum_{C_i \in \mathcal{C}} \Psi_i(\boldsymbol{\phi}_v^*) \left[ \sum_{C_{i'} \to C_i} a_{i'i} \Psi_{i'}(\boldsymbol{c}) - \sum_{C_i \to C_{i'}} a_{ii'} \Psi_i(\boldsymbol{c}) \right] = 0, \forall \boldsymbol{\phi} \in \mathbb{C}^{|\mathcal{V}|N}$$

$$\Leftrightarrow \sum_{C_{i'} \to C_i} a_{i'i} \Psi_{i'}(\boldsymbol{c}) - \sum_{C_i \to C_{i'}} a_{ii'} \Psi_i(\boldsymbol{c}) = 0, \forall C_i \in \mathcal{C} \Leftrightarrow \boldsymbol{A} \cdot \Psi(\boldsymbol{c}) = 0.$$

**APPENDIX C: DETAILED CALCULATIONS IN THEOREM 3**

The master equation of $P(\widehat{\boldsymbol{n}}, t)$ is derived as follows:

$$\partial_t P(\widehat{\boldsymbol{n}}, t) = \sum_{\boldsymbol{n} \in \mathcal{S}(\widehat{\boldsymbol{n}})} \partial_t P(\boldsymbol{n}, t) = \sum_{\boldsymbol{n} \in \mathcal{S}(\widehat{\boldsymbol{n}})} \left\{ \sum_{v \in \mathcal{V}} \sum_{v' \in N_e(v)} \sum_{i=1}^{N} [d_i (n_{vi} + 1) P(\boldsymbol{n} + \boldsymbol{1}_{vi} - \boldsymbol{1}_{v'i}, t) - d_i n_{vi} P(\boldsymbol{n}, t)] \right.$$

$$\left. + \sum_{v \in \mathcal{V}} \sum_{j=1}^{K} [f_{vj}(\boldsymbol{n} - \widetilde{\boldsymbol{V}}_{vj}, \omega) P(\boldsymbol{n} - \widetilde{\boldsymbol{V}}_{vj}, t) - f_{vj}(\boldsymbol{n}, \omega) P(\boldsymbol{n}, t)] \right\}. \tag{C1}$$

As $\boldsymbol{n} \in \mathcal{S}(\widehat{\boldsymbol{n}}) \to \boldsymbol{n} + \boldsymbol{1}_{vi} - \boldsymbol{1}_{v'i} = \widetilde{\boldsymbol{n}} \in \mathcal{S}(\widehat{\boldsymbol{n}})$, the first term of the right-hand side in Eq. (C1) becomes

$$\sum_{\boldsymbol{n} \in \mathcal{S}(\widehat{\boldsymbol{n}})} \sum_{v \in \mathcal{V}} \sum_{v' \in N_e(v)} \sum_{i=1}^{N} [d_i (n_{vi} + 1) P(\boldsymbol{n} + \boldsymbol{1}_{vi} - \boldsymbol{1}_{v'i}, t) - d_i n_{vi} P(\boldsymbol{n}, t)]$$

$$= \sum_{\boldsymbol{n} \in \mathcal{S}(\widehat{\boldsymbol{n}})} \sum_{v \in \mathcal{V}} \sum_{v' \in N_e(v)} \sum_{i=1}^{N} d_i (n_{vi} + 1) P(\boldsymbol{n} + \boldsymbol{1}_{vi} - \boldsymbol{1}_{v'i}, t) - \sum_{\boldsymbol{n} \in \mathcal{S}(\widehat{\boldsymbol{n}})} \sum_{v \in \mathcal{V}} \sum_{v' \in N_e(v)} \sum_{i=1}^{N} d_i n_{vi} P(\boldsymbol{n}, t)$$

$$= \sum_{\widetilde{\boldsymbol{n}} \in \mathcal{S}(\widehat{\boldsymbol{n}})} \sum_{v \in \mathcal{V}} \sum_{v' \in N_e(v)} \sum_{i=1}^{N} d_i \widetilde{n}_{vi} P(\widetilde{\boldsymbol{n}}, t) - \sum_{\boldsymbol{n} \in \mathcal{S}(\widehat{\boldsymbol{n}})} \sum_{v \in \mathcal{V}} \sum_{v' \in N_e(v)} \sum_{i=1}^{N} d_i n_{vi} P(\boldsymbol{n}, t)$$

$$= 0.$$

As the reaction network is linear, the propensity function $f_{vj}(\boldsymbol{n}, \omega)$ takes one of two forms: $f_{vj}(\boldsymbol{n}, \omega) = k_j n_{vi}$ or $f_{vj}(\boldsymbol{n}, \omega) = k_j \omega$, where $k_j$ is the reaction rate and $i$ is the index of some species. When $f_{vj}(\boldsymbol{n}, \omega) = k_j n_{vi}$, the second term can be transformed as follows:

$$\sum_{\boldsymbol{n} \in \mathcal{S}(\widehat{\boldsymbol{n}})} \sum_{v \in \mathcal{V}} [f_{vj}(\boldsymbol{n} - \widetilde{\boldsymbol{V}}_{vj}, \omega) P(\boldsymbol{n} - \widetilde{\boldsymbol{V}}_{vj}, t) - f_{vj}(\boldsymbol{n}, \omega) P(\boldsymbol{n}, t)]$$

$$= k_j \sum_{\boldsymbol{n} \in \mathcal{S}(\widehat{\boldsymbol{n}})} \sum_{v \in \mathcal{V}} [(n_{vi} - V_{ij}) P(\boldsymbol{n} - \widetilde{\boldsymbol{V}}_{vj}, t) - n_{vi} P(\boldsymbol{n}, t)]$$

$$= k_j \left[ \sum_{\boldsymbol{n} \in \mathcal{S}(\widehat{\boldsymbol{n}})} \sum_{v \in \mathcal{V}} (n_{vi} - V_{ij}) P(\boldsymbol{n} - \widetilde{\boldsymbol{V}}_{vj}, t) - \sum_{\boldsymbol{n} \in \mathcal{S}(\widehat{\boldsymbol{n}})} \sum_{v \in \mathcal{V}} n_{vi} P(\boldsymbol{n}, t) \right]$$

$$= k_j \left[ \sum_{\boldsymbol{n} \in \mathcal{S}(\widehat{\boldsymbol{n}})} \sum_{v \in \mathcal{V}} n_{vi} P(\boldsymbol{n} - \widetilde{\boldsymbol{V}}_{vj}, t) - \sum_{\boldsymbol{n} \in \mathcal{S}(\widehat{\boldsymbol{n}})} \sum_{v \in \mathcal{V}} V_{ij} P(\boldsymbol{n} - \widetilde{\boldsymbol{V}}_{vj}, t) - \sum_{\boldsymbol{n} \in \mathcal{S}(\widehat{\boldsymbol{n}})} \widehat{n}_i P(\boldsymbol{n}, t) \right]$$

$$= k_j \left[ \sum_{\widetilde{\boldsymbol{n}} \in \mathcal{S}(\widehat{\boldsymbol{n}} - \boldsymbol{V}_j)} \sum_{v \in \mathcal{V}} (\widetilde{n}_{vi} + V_{ij}) P(\widetilde{\boldsymbol{n}}, t) - V_{ij} \sum_{v \in \mathcal{V}} \sum_{\boldsymbol{n} \in \mathcal{S}(\widehat{\boldsymbol{n}})} P(\boldsymbol{n} - \widetilde{\boldsymbol{V}}_{vj}, t) - \widehat{n}_i P(\widehat{\boldsymbol{n}}, t) \right]$$

$$= k_j [(\widehat{n}_i - V_{ij}) P(\widehat{\boldsymbol{n}} - \boldsymbol{V}_j, t) + V_{ij} |\mathcal{V}| P(\widehat{\boldsymbol{n}} - \boldsymbol{V}_j, t) - V_{ij} |\mathcal{V}| P(\widehat{\boldsymbol{n}} - \boldsymbol{V}_j, t) - \widehat{n}_i P(\widehat{\boldsymbol{n}}, t)]$$

$$= k_j (\widehat{n}_i - V_{ij}) P(\widehat{\boldsymbol{n}} - \boldsymbol{V}_j, t) - k_j \widehat{n}_i P(\widehat{\boldsymbol{n}}, t).$$

When $f_{vj}(\boldsymbol{n}, \omega) = k_j \omega$, we similarly have

$$\sum_{\boldsymbol{n} \in \mathcal{S}(\widehat{\boldsymbol{n}})} \sum_{v \in \mathcal{V}} [f_{vj}(\boldsymbol{n} - \widetilde{\boldsymbol{V}}_{vj}, \omega) P(\boldsymbol{n} - \widetilde{\boldsymbol{V}}_{vj}, t) - f_{vj}(\boldsymbol{n}, \omega) P(\boldsymbol{n}, t)] = k_j \sum_{\boldsymbol{n} \in \mathcal{S}(\widehat{\boldsymbol{n}})} \sum_{v \in \mathcal{V}} [\omega P(\boldsymbol{n} - \widetilde{\boldsymbol{V}}_{vj}, t) - \omega P(\boldsymbol{n}, t)]$$

$$= k_j \left[ \sum_{\boldsymbol{n} \in \mathcal{S}(\widehat{\boldsymbol{n}})} \sum_{v \in \mathcal{V}} \omega P(\boldsymbol{n} - \widetilde{\boldsymbol{V}}_{vj}, t) - \sum_{\boldsymbol{n} \in \mathcal{S}(\widehat{\boldsymbol{n}})} \sum_{v \in \mathcal{V}} \omega P(\boldsymbol{n}, t) \right]$$

$$= k_j \left[ \sum_{v \in \mathcal{V}} \omega \sum_{\boldsymbol{n} \in \mathcal{S}(\widehat{\boldsymbol{n}})} P(\boldsymbol{n} - \widetilde{\boldsymbol{V}}_{vj}, t) - \sum_{v \in \mathcal{V}} \omega \sum_{\boldsymbol{n} \in \mathcal{S}(\widehat{\boldsymbol{n}})} P(\boldsymbol{n}, t) \right]$$

$$= k_j \Omega P(\widehat{\boldsymbol{n}} - \boldsymbol{V}_j, t) - k_j \Omega P(\widehat{\boldsymbol{n}}, t).$$

The master equation of $P(\widehat{\boldsymbol{n}}, t)$ is then obtained as

$$\partial_t P(\widehat{\boldsymbol{n}}, t) = \sum_{j=1}^{K} [f_j(\widehat{\boldsymbol{n}} - \boldsymbol{V}_j, \Omega) P(\widehat{\boldsymbol{n}} - \boldsymbol{V}_j, t) - f_j(\widehat{\boldsymbol{n}}, \Omega) P(\widehat{\boldsymbol{n}}, t)]. \tag{C2}$$

[1] W. J. Blake, M. Kærn, C. R. Cantor, and J. J. Collins, Nature **422**, 633 (2003).

[2] Y. Lan and G. A. Papoian, Phys. Rev. Lett. **98**, 228301 (2007).

[3] L. S. Tsimring, Rep. Prog. Phys. **77**, 026601 (2014).

[4] J. Elf and E. Måns, Genome Res. **13**, 2475 (2003).

[5] C. A. Gómez-Uribe and G. C. Verghese, J. Chem. Phys. **126**, 024109 (2007).

[6] F. Tostevin and P. R. ten Wolde, Phys. Rev. E **81**, 061917 (2010).

[7] K. H. Kim, H. Qian, and H. M. Sauro, J. Chem. Phys. **139**, 144108 (2013).

[8] K. R. Pilkiewicz and M. L. Mayo, Phys. Rev. E **94**, 032412 (2016).

[9] H. S. Samanta, M. Hinczewski, and D. Thirumalai, Phys. Rev. E **96**, 012406 (2017).

[10] T. E. Ouldridge, C. C. Govern, and P. R. ten Wolde, Phys. Rev. X **7**, 021004 (2017).

[11] R. Ellis, Curr. Opin. Struct. Biol. **11**, 114 (2001).

[12] B. N. Kholodenko, Nat. Rev. Mol. Cell Biol. **7**, 165 (2006).

[13] R. Cheong, A. Rhee, C. J. Wang, I. Nemenman, and A. Levchenko, Science **334**, 354 (2011).

[14] N. B. Becker, A. Mugler, and P. R. ten Wolde, Phys. Rev. Lett. **115**, 258103 (2015).

[15] Y. Hasegawa, Phys. Rev. E **97**, 022401 (2018).

[16] M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain, Science **297**, 1183 (2002).

[17] E. M. Ozbudak, M. Thattai, I. Kurtser, A. D. Grossman, and A. van Oudenaarden, Nat. Genet. **31**, 69 (2002).

[18] J. S. van Zon, M. J. Morelli, S. Tănase-Nicola, and P. R. ten Wolde, Biophys. J. **91**, 4350 (2006).

[19] S. Smith and R. Grima, Nat. Commun. **9**, 345 (2018).

[20] C. Gardiner, *Stochastic Methods* (Springer-Verlag, Berlin/ Heidelberg, 2009).

[21] M. v. Smoluchowski, Z. Phys. Chem. **92**, 129 (1917).

[22] S. A. Isaacson and D. Isaacson, Phys. Rev. E **80**, 066106 (2009).

[23] M. Howard and A. D. Rutenberg, Phys. Rev. Lett. **90**, 128102 (2003).

[24] D. Fange and J. Elf, PLoS Comput. Biol. **2**, e80 (2006).

[25] M. J. Lawson, B. Drawert, M. Khammash, L. Petzold, and T.-M. Yi, PLoS Comput. Biol. **9**, e1003139 (2013).

[26] F. Benitez, C. Duclut, H. Chaté, B. Delamotte, I. Dornic, and M. A. Muñoz, Phys. Rev. Lett. **117**, 100601 (2016).

[27] J. Elf and M. Ehrenberg, IET Syst. Biol. **1**, 230 (2004).

[28] M. Dobrzyński, J. V. Rodríguez, J. A. Kaandorp, and J. G. Blom, Bioinform. **23**, 1969 (2007).

[29] S. A. Isaacson, SIAM J. Appl. Math. **70**, 77 (2009).

[30] R. Erban and S. J. Chapman, Phys. Biol. **6**, 046001 (2009).

[31] S. Smith and R. Grima, Phys. Rev. E **93**, 052135 (2016).

[32] M. Feinberg, Arch. Ration. Mech. Anal. **132**, 311 (1995).

[33] D. F. Anderson, G. Craciun, and T. G. Kurtz, Bull. Math. Biol. **72**, 1947 (2010).

[34] G. Szederkényi and K. M. Hangos, J. Math. Chem. **49**, 1163 (2011).

[35] L. Peliti, J. Phys. France **46**, 1469 (1985).

[36] D. K. Lubensky, Phys. Rev. E **81**, 060102 (2010).

[37] M. Doi, J. Phys. A **9**, 1465 (1976).

[38] S. Tănase-Nicola, P. B. Warren, and P. R. ten Wolde, Phys. Rev. Lett. **97**, 068102 (2006).

[39] D. Cappelletti and C. Wiuf, SIAM J. Appl. Math. **76**, 411 (2016).

[40] E. Friedmann, R. Neumann, and R. Rannacher, Commun. Math. Anal. **15**, 76 (2013).

[41] B. Mayer, G. Emery, D. Berdnik, F. Wirtz-Peitz, and J. A. Knoblich, Curr. Biol. **15**, 1847 (2005).

[42] E. Levine and T. Hwa, Proc. Natl. Acad. Sci. U.S.A. **104**, 9224 (2007).