Stationary distributions of systems with discreteness-induced transitions

(Article begins on next page)

Stationary distributions of systems with Discreteness Induced Transitions

Enrico Bibbona[*]
Dipartimento di Scienze Matematiche G.L. Lagrange
Politecnico di Torino, Turin, Italy

Jinsu Kim
Department of Mathematics University of California, Irvine, USA.

Carsten Wiuf
Department of Mathematical Sciences
University of Copenhagen, Copenhagen, Denmark
(Dated: March 31, 2020)

We provide a theoretical analysis of some autocatalytic reaction networks exhibiting the phenomenon of discreteness induced transitions. The family of networks that we address includes the celebrated Togashi and Kaneko model. We prove positive recurrence, finiteness of all moments, and geometric ergodicity of the models in the family. For some parameter values, we find the analytic expression for the stationary distribution, and discuss the effect of volume scaling on the stationary behavior of the chain. We find the exact critical value of the volume for which discreteness induced transitions disappear.

I. INTRODUCTION

In 2001, Togashi and Kaneko described a cycle of stochastic autocatalytic reactions that displays a highly peculiar dynamics in some regions of the parameter space [1].

It is characterized by switches between patterns where one or more reactants are present in small or vanishing molecule number while other reactants are abundant. The switching is triggered by a single molecule of a previously extinct species that drives the system to a different pattern through a sequence of quick reactions. The switches were named Discreteness Induced Transitions (DIT) since deterministic ODE models are not able to reproduce them [1].

The paper raised much interest and similar effects have been observed in more complicated and realistic models, e.g., large scale networks [2], particle systems with finite interaction radius [3], reaction-diffusion systems [4], models of ant foraging [5], chiral autocatalysis [6], tumor growth [7], spatial models [8], and viral replication [9].

Several attempts have been made to underpin the phenomenon theoretically, at least in simplified toy models, through derivation of analytic expressions, without resorting to simulation or approximation. Examples in this direction are [10–13], though many questions are still unsolved.

Despite simulation of the Togashi-Kaneko (TK) model indicates a stationary behavior after a short transient time, positive recurrence (existence of a unique stationary distribution) of the corresponding continuous-time Markov chain (CTMC) has not been proved. For the original 4-dim TK model, no general result from Chem-ical Reaction Network theory is applicable. Even if the system is reduced to dimension two, the problem of finding a stationary distribution remains non-trivial, and the curious switching behavior persists. In dimensions four and two, the switching behavior causes the *seemingly* stationary distribution emerging from simulation to be multimodal for certain parameters values. When the rates are scaled in the volume V of the container and V is considered large, the multimodality disappears and a distribution with a concentrated peak emerges. In this case, the scaled stochastic model converges to the classical deterministic model (fluid limit).

In this paper, we prove that a family of autocatalytic networks, including the TK model, is positive recurrent in arbitrary dimension (Theorem 2). For some parameter values, an explicit expression for the stationary distribution is derived. In 2-dim (cf. Theorem 1) the parameter region for which the stationary distribution is known covers the 2-dim TK model. In higher dimension (cf. Theorem 3) the parameter region for which the stationary distribution is known, does not include the general TK model. However, it includes a large family of TK-like models exhibiting DITs..

The analytic form of the stationary distribution provides a clear theoretical demonstration of the effect of volume scaling on the stationary behavior of the system. It also allows us to find the exact critical value of the volume from which the DIT stops to appear. This value we also find for the TK model.
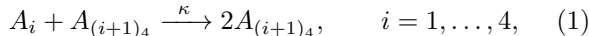
II. BACKGROUND MATERIAL

A. The original 4-dim TK model

Let $\mathbb{N} = \{0, 1, 2, \ldots\}$ denote the integers including zero. For any two integers i and n, let $(i)_n$ be the remain-

[*] enrico.bibbona@polito.it

der after integer division of i by n (elsewhere denoted by $i \bmod n$). The network proposed by Togashi and Kaneko [1] consists of the following cycle of autocatalytic reactions

$$A_i + A_{(i+1)_4} \xrightarrow{\kappa} 2A_{(i+1)_4}, \qquad i = 1, \ldots, 4, \quad (1)$$

together with inflow and outflow reactions

$$A_i \underset{\lambda}{\overset{\delta}{\rightleftharpoons}} \emptyset, \qquad i = 1, \ldots, 4.$$

The state of the system is a tuple of four non-negative integers $\mathbf{a} = (a_1, a_2, a_3, a_4)'$. Denote by $\mathbf{e}_j$ the j-th unit vector, $j = 1, \ldots, 4$. The transitions rates generated by the autocatalytic reactions are

$$q_{\mathbf{a}, \mathbf{a} - \mathbf{e}_i + \mathbf{e}_{(i+1)_4}} = \kappa a_i a_{(i+1)_4},$$

while those corresponding to inflow and outflow reactions are

$$q_{\mathbf{a}, \mathbf{a} + \mathbf{e}_i} = \lambda \quad \text{and} \quad q_{\mathbf{a}, \mathbf{a} - \mathbf{e}_i} = \delta a_i.$$

The qualitative behavior of the system depends on the parameter values. The classical volume scaling (cf. [14, Chapter 11] or [15]) is adopted in [1], where the initial molecule counts of the species are proportional to the scaling parameter V. It implies the rate constants are given by

$$\kappa = \frac{\kappa'}{V} \qquad \delta = \delta' \qquad \lambda = \lambda' V$$

One parameter can always be set to one by linear scaling of time. In [1], $\kappa' = 1$, and further $\lambda' = \delta' = D$ for simplification. According to [16] or [14, Chapter 11], when $V \to \infty$, the density process, which is the CTMC rescaled by dividing the molecule numbers by V, converges to the solution of a system of ordinary differential equations with stable equilibrium $(1, 1, 1, 1)$. Indeed when $VD \gg 1$, the reaction rates are large and the trajectories of the density process only display small fluctuations around the deterministic equilibrium.

For $VD \ll 1$, a completely different behavior appears, triggered by the slow rate of inflow and DIT appears. If the system is initialized at a state where all species counts are large, one of the species at random (say, species 3) is quickly driven to extinction by the fast autocatalytic dynamics. At this point, several molecules of species 2 are produced and not consumed and they catalyze the consumption of all molecules of species 1. We end up with a configuration where the species 1 and 3 are both consumed, the count of species 2 is very high, and that of species 4 is quite low. We call this pattern 2H4L. In this configuration only slow inflows and outflows are active, and one needs to wait until a molecule of species 3 or 1 flows in before the autocatalytic dynamics starts again leading to another pattern with two non-contiguous species extinct. The dynamics of the system then proceeds by switching between such patterns in a way that a 2H4L configuration is much more often followed by a 2L4H pattern and only rarely switches to either a 1H3L or 1L3H configuration (cf. FIG 1 in [1]).

B. Lumpability

In the next Section we exploit the notion of lumpability to find the stationary distribution in some cases. We summarize here the meaning of this property.
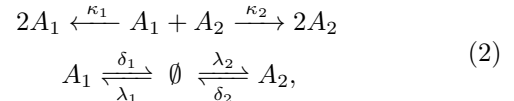
Let $\{S_I\}_{I \in \mathcal{I}}$ be a partition of a denumerable state space S of a CTMC $X(t)$, $t \geq 0$, with rates q_{ij}, $i, j \in S$. Let moreover ι be the function that maps $x \in S$ to the index of the element of the partition to which x belong, (i.e $\iota(x) = K$ if and only if $x \in S_K$). The process $X(t)$, $t \geq 0$, is (strongly) lumpable if the *lumped* process $\overline{X}(t) = \iota(X(t))$, $t \geq 0$, is a CTMC on $\mathcal{I}$ for any choice of initial distribution. Sufficient conditions (cf. [17]) that guarantees lumpability of a regular, irreducible, positive recurrent CTMC $X(t)$, $t \geq 0$, on the partition $\{S_I\}_{I \in \mathcal{I}}$ are that every subset S_I is finite, and that for any $I, J \in \mathcal{I}$, and any $i, i' \in S_I$,

$$\sum_{j \in S_J} q_{ij} = \sum_{j \in S_J} q_{i'j} = \overline{q}_{IJ}$$

The rates of the lumped chain $\overline{X}(t)$, $t \geq 0$, are $\overline{q}_{IJ}$, $I, J \in \mathcal{I}$, and for any $s < t$, the lumped variable $\overline{X}(t)$ is independent of X_s given $\overline{X}_s$

III. THE 2-DIM TK MODEL

If the number of species in the TK model is reduced to two, the reaction network becomes

$$2A_1 \xleftarrow{\kappa_1} A_1 + A_2 \xrightarrow{\kappa_2} 2A_2$$
$$A_1 \underset{\lambda_1}{\overset{\delta_1}{\rightleftharpoons}} \emptyset \underset{\delta_2}{\overset{\lambda_2}{\rightleftharpoons}} A_2, \qquad (2)$$

where we allow κ_1 and κ_2 to be different. The state of the network is denoted by $\mathbf{a} = (a_1, a_2)' \in \mathbb{N}^2$, the molecule counts of each species. The transitions rates due to the autocatalytic reactions are

$$q_{\mathbf{a}, \mathbf{a} - \mathbf{e}_1 + \mathbf{e}_2} = \kappa_2 a_1 a_2, \qquad q_{\mathbf{a}, \mathbf{a} - \mathbf{e}_2 + \mathbf{e}_1} = \kappa_1 a_1 a_2, \quad (3)$$

while those corresponding to inflow and outflow reactions are

$$q_{\mathbf{a}, \mathbf{a} + \mathbf{e}_i} = \lambda_i, \qquad q_{\mathbf{a}, \mathbf{a} - \mathbf{e}_i} = \delta_i a_i, \qquad i = 1, 2. \quad (4)$$

The dynamics is simplified, but not too much. When the inflows occur at a much slower rate than the autocatalytic reactions, the system switches between two patterns in a similar way to the original 4-dimensional TK system, where one or the other compound is mostly absent. A plot of the two simulated trajectories in this parameter range is shown in FIG. 1.

A. Positive recurrence and stationary distribution

A proof of positive recurrence in a more general setting is given in Section IV A. In this section, we show the
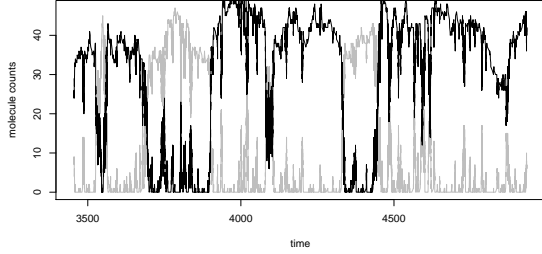
FIG. 1. Molecule counts of the two species of network (2) along time. Patterns where the grey species is mostly absent alternates with patterns where the mostly absent species is the black one. Parameters are $\lambda_i = 0.2$, $\delta_i = 0.01$, $\kappa_i = 0.05$ for $i = 1, 2$.

sketch of how to derive the stationary distribution of (2) by using its lumpability. The subsets

$$E_n = \{\mathbf{a} \in \mathbb{N}^2 \colon a_1 + a_2 = n\},$$

form a partition $\{E_n\}_{n \in \mathbb{N}}$ of the state space. The CTMC model $X(t)$, $t \geq 0$, of (2) under stochastic mass-action kinetics is lumpable with respect to this partition if $\delta = \delta_1 = \delta_2$. With this choice, the rate at which the total molecule count n is increased by one, is equal to the sum of the rates of the inflows

$$q_{n,n+1} = \sum_{i=1}^{2} q_{\mathbf{a},\mathbf{a}+\mathbf{e}_i} = \lambda_1 + \lambda_2,$$

independently of $\mathbf{a}$. The rate at which n is decreased by one, is the sum of the rates of the outflows

$$q_{n,n-1} = \sum_{i=1}^{2} q_{\mathbf{a},\mathbf{a}-\mathbf{e}_i} = \delta(a_1 + a_2) = \delta n$$

and therefore it does not depend on $\mathbf{a}$ as long as $\mathbf{a} \in E_n$.

The lumped process $\overline{X}(t)$, $t \geq 0$, is described by the following reaction network where a single species B aggregates all molecules of A_1 and A_2

$$B \underset{\delta}{\overset{\lambda_1 + \lambda_2}{\rightleftharpoons}} \emptyset. \tag{5}$$

Network (5) is weakly reversible and has *deficiency* zero [18]. By [18, Theorem 3.6 and 3.7], it admits a unique stationary distribution with Poisson law

$$\nu(n) = \frac{\mu^n}{n!} \exp(-\mu), \quad \mu = \frac{\lambda_1 + \lambda_2}{\delta}. \tag{6}$$

where n is the state of the lumped process (i.e., $\overline{X}(t) = n$ if and only if $X(t) \in E_n$). We now aim at factorizing the stationary distribution $\Pi(\mathbf{a})$ of the process $X(t)$, $t \geq 0$, of (2) by conditioning on the stationary probability $\nu(n)$ of the lumped process being in state $n = a_1 + a_2$. We write

$$\Pi(\mathbf{a}) = \pi(a_1|n)\nu(n). \tag{7}$$

A careful rewriting of the master equation for the stationary distribution $\Pi(\mathbf{a})$ shows that $\Pi(\mathbf{a})$ is stationary if and only if $\pi(\mathbf{a}|n)$ fulfils

$$R_n = L_{n-1} + L_n + L_{n+1} \tag{8}$$

with

$$R_n = (\lambda_1 + \lambda_2 + n\delta + (\kappa_1 + \kappa_2)a(n-a))\pi(a|n)$$

$$L_{n-1} = \frac{n\delta\lambda_1}{\lambda_1 + \lambda_2}\pi(a-1|n-1) + \frac{n\delta\lambda_2}{\lambda_1 + \lambda_2}\pi(a|n-1)$$

$$L_n = \kappa_1(a+1)(n-a-1)\pi(a+1|n) +$$
$$+ \kappa_2(a-1)(n-a+1)\pi(a-1|n)$$

$$L_{n+1} = \frac{\lambda_1 + \lambda_2}{n+1}(a+1)\pi(a+1|n+1)$$
$$+ \frac{\lambda_1 + \lambda_2}{n+1}(n-a+1)\pi(a|n+1),$$

for $n \geq 0$ and $a = 0, \dots, n$.

Unfortunately, there is not a simple way to find a closed form expression of $\pi(\cdot|n)$ satisfying equation (8). However, simulation of the process for different rate constants, corresponding to different regimes of the volume V (cf. Section III B for more details), indicates that the conditional stationary distribution may be unimodal, flat or concentrated at the boundaries (cf. FIG 2). Statistical practice suggests the beta-binomial as a natural candidate for a discrete distribution on the integers $\{0, \cdots, n\}$ that may display these behaviors. The next theorem confirms this, and FIG 2 provides a graphical comparison between simulations and theoretical values in different parameter settings.

Theorem 1. *Network (2), assuming that $\kappa = \kappa_1 = \kappa_2 > 0$ and $\delta = \delta_1 = \delta_2 > 0$, has a unique stationary distribution $\Pi(\mathbf{a})$ that factorizes as (7), where $\nu(n)$ is given by (6), and $\pi(\cdot|n)$ is given by the beta-binomial distribution*

$$\pi(i|n) = \binom{n}{i}\frac{B(i + \alpha, n - i + \beta)}{B(\alpha, \beta)}, \quad i = 0, \dots, n, \tag{9}$$

where

$$\alpha = \frac{\delta\lambda_1}{\kappa(\lambda_1 + \lambda_2)}, \quad \beta = \frac{\delta\lambda_2}{\kappa(\lambda_1 + \lambda_2)}, \tag{10}$$

and

$$B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x + y)}, \quad x, y > 0.$$

Proof. The proof is by direct verification, substituting expression (9) into equation (8). Calculations are displayed in Appendix B in a more general context. $\square$

B. Volume scaling

Molecule counts and mass-action rates can be scaled with the volume V in such a way that the scaled stochastic system $X(t)/V$ converges for large V to the solution
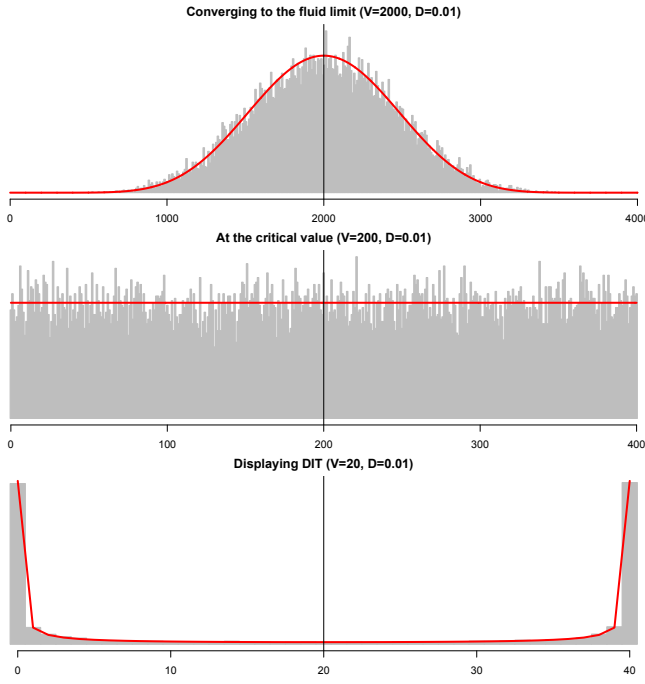
FIG. 2. The effect of scaling and DIT. The conditional stationary distribution $\pi(a_1|n)$, cf. (7) from simulation (grey histograms) and from (9) in Theorem 1 (red lines). Parameters are chosen according to (11) with $\kappa'_i = 1$, $\delta'_i = \lambda'_i = D = 0.01$ for $i = 1, 2$. The volume parameter V differs in the three panels to illustrate the effect of scaling, and n is chosen as the mean of $\nu(n)$, which is 4000, 400, and 40, respectively; implying that the mean of the scaled process X/V is $(1, 1)$ in all three cases. Simulation set-up: $2.75 \cdot 10^6$, 10^6, and 10^5 simulations (from upper to lower panel) were conducted with fixed time $T = 250, 50, 50$, respectively (the stationary regime already applies). Only values of $\mathbf{a}(T)$ with $a_1(T) + a_2(T) = n$ were kept, and the histogram of $a_1(T)$ was plotted.

of the deterministic system on any finite time horizon, cf. [14][Ch. 11, Theorem 2.1].

This is achieved for (3) and (4), under the hypothesis of Theorem 1, by setting the constants to

$$\kappa_i = \frac{\kappa'}{V} \quad \delta_i = \delta', \quad \lambda_i = \lambda'_i V, \qquad (11)$$

for $i = 1, 2$. When V is not sufficiently large the stochastic model differs significantly from the deterministic limit [1] and starts to display the switching behavior (DIT) illustrated in FIG 1.

In [1], the authors set

$$\kappa'_i = 1, \qquad \delta'_i = \lambda'_i = D, \qquad i = 1, 2. \qquad (12)$$

With this choice of the rate constants, by Theorem 1, we know the explicit form of the stationary density, and we can investigate the behavior of the system at every V without resorting to simulations. The stationary conditional density $\pi(\cdot|n)$ is beta-binomial with parameters $\alpha = \beta = DV/2$. The beta-binomial density is unimodal
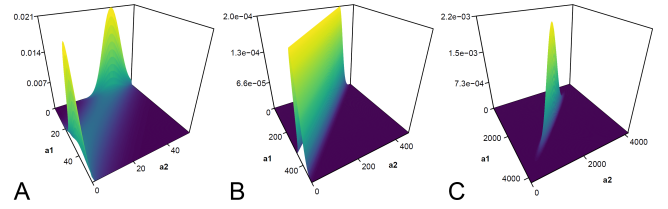

FIG. 3. Smoothed representation of the stationary density (7). In all panels $D = 0.01$. The volume parameter V is different in the three panels to illustrate the effect of the scaling. In panel A, $V = 20$ (same range as in FIG. 1). The density is bimodal and concentrated at the boundaries, since DIT are present. In panel B, $V = 200$ and the conditional density (9) is uniform. In panel C, $V = 2000$ and the density is concentrated around the deterministic equilibrium.

when α and β are both larger than one (that is, when DV is larger than two) with the mass concentrated at the equilibrium of the corresponding deterministic model. When both α and β are smaller than one (that is, when DV is small) the density becomes bimodal with most of the mass at the boundaries. The intermediate case is when $\alpha = \beta = 1$ and the conditional distribution reduces to the discrete uniform distribution on $\{0, \ldots, n\}$. In other words, at the critical value $DV = 2$, the conditional density flattens to $\pi(a|n) = \frac{1}{n+1}$ for every a. A pictorial representation of the density (7), at different values of V with D fixed to the value 0.01 is given in FIG. 3. The effect of the scaling is apparent. For graphical convenience, the discrete density has been smoothed to a continuous one.

To make this effect quantitatively apparent, in the general setting where (11) holds but not necessarily (12), we prove that for $V \to 0$ the stationary distribution concentrates at the boundaries by showing that the conditional probability $\pi(0|n) + \pi(n|n)$ tends to one, for any n. Indeed, inserting (11) into (10), we get $\alpha = \alpha' V$ and $\beta = \beta' V$ with

$$\alpha' = \frac{\delta\lambda'_1}{\kappa(\lambda'_1 + \lambda'_2)} \quad \text{and} \quad \beta' = \frac{\delta\lambda'_2}{\kappa(\lambda'_1 + \lambda'_2)}.$$

The sum of the two conditional probabilities reduces to

$$\pi(0|n) + \pi(n|n) = \qquad\qquad\qquad (13)$$
$$= \left[\frac{\Gamma(n + \alpha'V)}{\Gamma(\alpha'V)} + \frac{\Gamma(n + \beta'V)}{\Gamma(\beta'V)}\right] \frac{\Gamma[(\alpha' + \beta')V]}{\Gamma[n + (\alpha' + \beta')V]}.$$

Whatever n is, since $\Gamma(z) \sim 1/z$ for $z \to 0$, it is easily seen that the sum tends to one as $V \to 0$.

For large V, we show that the stationary distribution Π_V of the scaled process $X(t)/V$ concentrates around the deterministic equilibrium $(1, 1)$. The mean μ_V and variance Σ_V of Π_V might easily be computed (by conditioning on n) from the first and second moments of the Poisson distribution and the beta-binomial distribution.

4

The explicit calculation is here only reported for two components, but can be found for the others as well,

$$(\mu_V)_1 = \frac{\mu}{V}\frac{\alpha}{\alpha + \beta}$$

$$(\Sigma_V)_{11} = \frac{1}{V^2}\left[\frac{\alpha\beta}{(\alpha+\beta)^2}\frac{(\alpha+\beta)\mu + \mu^2 + \mu}{\alpha + \beta + 1} + \mu\frac{\alpha}{\alpha+\beta}\right],$$

where μ is given in (6) and α and β in (10). Scaling the parameters as in (11), it is easily observed that $(\mu_V)_1 \to 1$ and $(\Sigma_V)_{11} \to 0$ for $V \to \infty$. With a little more effort, the same result extends to the other components, that is, we have

$$\mu_V \to \left(\frac{\lambda_1'}{\delta'}, \frac{\lambda_2'}{\delta'}\right) \quad \text{and} \quad \Sigma_V \to \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

In general, the *agreement* between the stochastic and the deterministic model for large volumes only holds on a finite time horizon only. Negative examples where the two modelling paradigms differ asymptotically are discussed in [19, 20]. Our result shows that for large V, under the assumptions of Theorem 1, the stochastic and the deterministic models of (2) are in agreement asymptotically.

IV. HIGHER DIMENSIONAL MODELS

In higher dimension there exist different models whose 2-dim reduction corresponds to network (2). One of them is the 4-dim TK model (1), but also the network

$$2A_i \xleftarrow{\kappa_{ji}} A_i + A_j \xrightarrow{\kappa_{ij}} 2A_j \qquad A_i \xrightleftharpoons[\lambda_i]{\delta_i} \emptyset, \qquad (14)$$
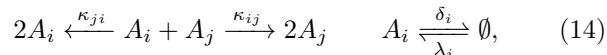
$i, j = 1, \ldots, d$, $i \neq j$, can be seen as a d-dimensional version of model (2). Network (14) includes (1) as a special case for κ_{ij} equal to κ when $j = (i+1)_d$ and zero otherwise. Reaction rates are the obvious generalizations of (3) and (4).

A. Positive recurrence

In this section, we state the positive recurrence of the Markov process underlying the general d-dimensional model (14). To do so, we show that $V(x) = e^{\|x\|_1}$, where $\|x\|_1 = \sum_{i=1}^d |x_i|$, is a Lyapunov function. Non-explosivity and positivity, then, follow by the Foster-Lyapunov criterion [21]. Additionally, as a by-product, all moments of the stationary distribution are positive and convergence to the stationary distribution is exponentially fast. The detailed proof can be found in Appendix A.

Theorem 2. *For any non-negative values of the parameters κ_{ij}, κ_{ji}, and for positive λ_i and δ_i, the CTMC associated to the system (14) is positive recurrent on $\mathbb{N}^d$ (for*

any d). Consequently, it has a unique stationary distribution supported on $\mathbb{N}^d$. Moreover, all moments are finite and the convergence to the stationary distribution is exponentially fast.

B. Stationary distribution

1.   The model

By the same argument as we used in dimension 2, under the assumption of equal outflow rates ($\delta_i = \delta$ for all $i = 1 \cdots d$), the process $X(t)$, $t \geq 0$, that counts the molecules of each species is lumpable on the partition $\{E_n\}$, where $E_n = \{\mathbf{a} \in \mathbb{N}^2 \colon \sum_i a_i = n\}$.

The lumped process $\overline{X}(t) = \sum_i X_i(t)$ represents the total molecule count. It follows a birth and death process (as in (5)) with Poisson stationary distribution with intensity

$$\mu = \frac{\sum_{i=1}^d \lambda_i}{\delta}. \qquad (15)$$

Similarly to the 2-dim case, the stationary distribution $\Pi(\mathbf{a})$ factorizes as

$$\Pi(\mathbf{a}) = \pi(\mathbf{a}|n)\nu(n). \qquad (16)$$

Theorem 3. *Assume $\kappa_{ij} = \kappa > 0$, $i, j = 1, \ldots, d$, $i \neq j$, $\delta = \delta_1 = \cdots = \delta_d > 0$, and $\lambda_i > 0$ for all i. Then, model (14) has a unique stationary distribution $\Pi(\mathbf{a})$ expressed as in (16), where $\nu(n)$ is given as in (6) and (15), and $\pi(\cdot|n)$ is given by the Dirichlet-multinomial distribution*

$$\pi(\mathbf{a}|n) = \binom{n}{\mathbf{a}}\frac{\Gamma(\sum_{i=1}^d \alpha_i)}{\Gamma(n + \sum_{i=1}^d \alpha_i)}\prod_{i=1}^d \frac{\Gamma(a_i + \alpha_i)}{\Gamma(\alpha_i)} \qquad (17)$$

where $\mathbf{a}$ is any d-dimensional integer vector with $\|\mathbf{a}\|_1 = n$, and

$$\alpha_i = \frac{\delta\lambda_i}{\kappa \sum_{i=1}^d \lambda_i}.$$

Proof. The proof is by direct verification, substituting expression (17) into equation (16) using (6). Calculations are displayed in Appendix B. $\square$

C. Volume scaling and other properties

The scaled process $X(t)/V$ in dimension d has similar properties to that of the scaled process in dimension 2. In the case where the stationary distribution is known (Theorem 3), we might proceed similarly to what was done in dimension 2 and calculate the mean vector and covariance matrix of the molecule counts, now using moment properties of the Poisson and the Dirichlet-multinomial distributions. Parameters are scaled according to

$$\kappa = \frac{\kappa'}{V} \quad \delta = \delta', \quad \lambda_i = \lambda_i'V, \qquad (18)$$

for $i = 1, \cdots, d$.

As the volume V increases towards infinity, it can be shown that the mean vector converges to $\left(\frac{\lambda_1}{\delta}, \ldots, \frac{\lambda_d}{\delta}\right)$, the equilibrium point of the deterministic process, and the covariance matrix decreases towards the zero matrix. Thus, under the hypothesis of Theorem 3, the deterministic and the stochastic models of (14) are in agreement asymptotically for large volume size in the long run (at stationarity) as well as over the finite time horizon.

At the other extreme, for $V \to 0$, the conditional probability of a corner configuration tends to one. Indeed, such probability, that generalizes (13), is equal to

$$\sum_{i=1}^{d} \pi(\mathbf{e}_i|n) = \frac{\Gamma(V \sum_j \alpha'_j)}{\Gamma(n + V \sum_j \alpha'_j)} \sum_{i=1}^{d} \frac{\Gamma(n + V\alpha'_i)}{\Gamma(V\alpha'_i)},$$

where

$$\alpha'_i = \frac{\delta \lambda'_i}{\kappa \sum_j \lambda'_j}.$$

The convergence to one can be easily shown with the same methodology used in dimension 2.

The peaks at the vertexes reflect the presence of DIT that causes the switch between dynamical patterns where only one of the species is present in large quantity at a time, while all the others are almost extinct. A graphical illustration of the presence of DITs in a three dimensional version of model (14) is given in FIG. 4. In dimension three it is no longer possible to plot the stationary distribution $\Pi(\mathbf{a})$. However, choosing $D = 0.01$ and $V = 20$, we can plot a set of simulated trajectories, and the values of the conditional stationary distribution $\pi(\mathbf{a}|n)$.

If parameter are further chosen as

$$\kappa'_i = 1, \qquad \delta'_i = \lambda'_i = D, \qquad i = 1, \cdots, d, \qquad (19)$$

in analogy of what was done in [1], the distribution becomes symmetric in the labels of the species and the $\alpha_i = DV/d$, $i = 1, \ldots, d$, are all equal. Moreover if $V = \frac{d}{D}$, the conditional distribution is flat, providing a transition point from the multimodal case to the unimodal case. If the α_is are not equal (i.e. the λ'_i are not), the transition will not proceed through a flat conditional distribution.

Other relevant properties of the Dirichlet-multinomial distribution, like aggregation, marginals, conditional distributions are discussed in [22, 23].

D.   Back to the d-dim TK model

Model (1) motivated our interest in autocatalytic networks. Theorem 2 guarantees that it is positive recurrence, but an explicit form of the stationary distribution cannot be derived by Theorem 3. Indeed, it is a special case of (14), where some of the κ_{ij} are set to zero (those for which $j \neq (i+1)_d$) and all others are set to the same value κ. However, it is still possible to find the explicit expression in a very special case.
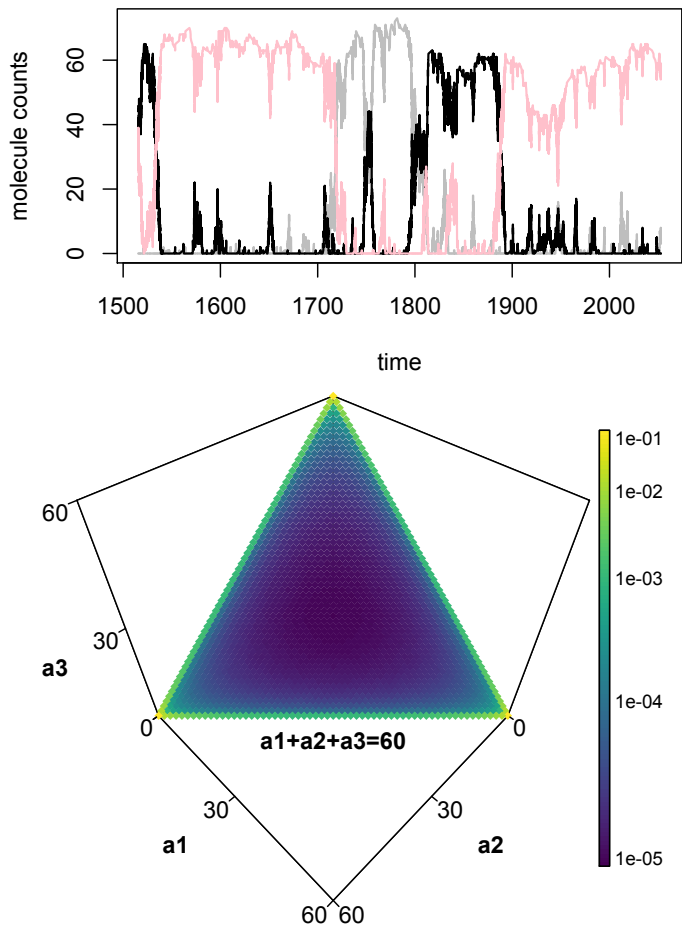

FIG. 4.   Simulated trajectories and the conditional stationary distribution $\pi(\mathbf{a}|n)$, from model (14) in dimension three. Parameters are chosen according to (19) with $D = 0.01$ and $V = 20$. The value of n is fixed to 60 in the lower plot. The presence of DITs is apparent both from the trajectories and from the conditional distribution that is concentrated at the corners of the simplex.

Theorem 4. *Assume that $\kappa = \kappa_1 = \cdots = \kappa_d \geq 0$ and $\delta = \delta_1 = \cdots = \delta_d = \frac{d}{d-1}\kappa$ and $\lambda = \lambda_1 = \cdots = \lambda_d > 0$. Then, model (1) has a unique stationary distribution $\Pi(\mathbf{a})$ whose expression is (16) with $\nu(n)$ given by (6) and (15) and with $\pi(\cdot|n)$ given by a uniform distribution*

$$\pi(\mathbf{a}|n) = \frac{n!(d-1)!}{(n+d-1)!} \qquad (20)$$

on the simplex $\{\mathbf{a} \in \{0, \ldots n\}^d : \|\mathbf{a}\|_1 = n\}$.

The proof is by direct verification, substituting expression (20) into equation (16) with κ_{ij} set to zero for all $j \neq (i+1)_d$ and to the same value κ otherwise. Calculations are displayed in Appendix C. If the rate constants are scaled in the volume as in equation (18) and further set to (19), the critical value of the volume that makes the distribution flat is $V = \frac{d}{(d-1)D}$, in agreement with the result for $d = 2$. In 4-dim, in [1], it was noticed from

simulation that the order of the magnitude of this critical value should be around $V \sim 1/D$. However, determining the exact value was pursued. Our result allows us to ensure that the exact value is $V = 4/3D$.

APPENDIX

In this Appendix, we report the detailed proofs of the three main theorems (Theorem 2, 3, 4) that are stated in Section IV. Theorem 1 is not proved separately since it is a special case of Theorem 3. The most general high dimensional model that we consider is that in (14), that we repeat here

$$2A_i \xleftarrow{\kappa_{ji}} A_i + A_j \xrightarrow{\kappa_{ij}} 2A_j \qquad A_i \underset{\lambda_i}{\overset{\delta_i}{\rightleftarrows}} \emptyset.$$

Appendix A

As we stated in the main text, the associated CTMC for the general dimensional model (14) is positive recurrent and admits a unique stationary distribution. We prove this in the following theorem. We further show that the CTMC is *exponentially ergodic* meaning that the associated distribution P^t at time t converges to the unique stationary distribution exponentially fast. The proof relies on the Foster-Lyapunov criterion [21]. We begin with a formal statement and necessary concepts for the Foster-Lyapunov criterion.

Definition 1. *For a CTMC $X(t)$, $t \geq 0$, defined on a countable state space χ, the infinitesimal generator $\mathcal{L}$ is the operator*

$$\mathcal{L}V(x) = \sum_\eta \lambda_\eta(x)(V(x+\eta) - V(x)),$$

where η is a transition of $X(t)$, λ_η is the transition rate associated with η, and V is any real function defined on the state space.

For a CTMC $X(t)$, $t \geq 0$, we define a truncated process X_M such that $X_M(t) = X(t)$ if $|X(t)| < M$ and $X_M(t) = x_M$ otherwise for some fixed state x_M with $|x_M| \geq M$. We denote by $\mathcal{L}_M$ the infinitesimal generator of X_M.

We further call $V(x)$ a norm-like function if $V(x)$ is a positive function such that $|V(x)| \to \infty$, as $|x| \to \infty$.

The following theorem is Theorem 6.1 in [21], in the case of a countable state space. It is one version of the Foster-Lyapunov criterion for exponential ergodicity.

Theorem A (Foster-Lyapunov criterion [21]). *Let $X(t)$, $t \geq 0$, be a CTMC defined on a countable state space χ. Then $X(t)$, $t \geq 0$, is non-explosive and positive recurrent if there exist a norm-like function V on χ, positive constants C and D such that for any $M > 0$*

$$\mathcal{L}_M V(x) \leq -CV(x) + D \quad \text{for all } x \in \chi.$$

Furthermore, $X(t)$, $t \geq 0$, admits a unique stationary distribution π on each irreducible component, and there exist $B > 0$ and $\beta \in (0,1)$ such that

$$\sup_A |P^t(x,A) - \pi(A)| \leq BV(x)\beta^t \text{ for all } x \in \chi.$$

To show positive recurrence and exponential ergodicity of the CTMC associated with the general model (14), it is therefore sufficient to prove that there exist a norm-like function V and positive constants C, D such that

$$\mathcal{L}V(x) \leq -CV(x) + D \quad \text{for all } x. \tag{21}$$

In the proof of the following theorem, we prove (21) for an exponential function V. Using this specific function, we also show that all moments of the unique stationary distribution of $X(t)$, $t \geq 0$, are finite.

Proof of Theorem 2. Let $X(t)$, $t \geq 0$, be the CTMC associated with the system (14). Let $V(x) = e^{\|x\|_1}$, where $\|x\|_1 = \sum_{i=1}^d |x_i|$. Then we show that (21) holds for some positive constants C and D.

Let $e_i \in \mathbb{N}^d$ be the vector with i-th component 1 and zero otherwise. We have

$$\begin{aligned}
\mathcal{L}V(x) = &\sum_{i,j} \kappa_{ij} x_i x_j (V(x+e_i-e_j) - V(x)) \\
&+ \sum_{i,j} \kappa_{ij} x_i x_j (V(x+e_i-e_j) - V(x)) \\
&+ \sum_{i=1}^d \delta_i x_i (V(x-e_i) - V(x)) \\
&+ \sum_{i=1}^d \lambda_i (V(x+e_i) - V(x))
\end{aligned}$$

$$\mathcal{L}V(x) = \sum_{i=1}^{d} \delta_i x_i (V(x - e_i) - V(x))$$
$$+ \sum_{i=1}^{d} \lambda_i x_i (V(x + e_i) - V(x)).$$

Let $K_n = \{x \in \mathbb{N}^d : x_i \geq n \text{ for each } i\}$. Then note that for $x \in K_n$,

$$\mathcal{L}V(x) = V(x) \left(\sum_{i=1}^{d} \delta_i x_i (e^{-1} - 1) + \sum_{i=1}^{d} \lambda_i (e - 1) \right)$$
$$\leq \left(\min_i \delta_i (e^{-1} - 1) dn + \sum_{i=1}^{d} \lambda_i (e - 1) \right) V(x)$$

Hence, by choosing sufficiently large N such that

$$C = -\left(\min_i \delta_i (e^{-1} - 1) dN + \sum_{i=1}^{d} \lambda_i (e - 1) \right) > 0,$$

we conclude that (21) holds with $D = 2C \max_{x \in K_N^c} V(x)$. This implies that $X(t)$, $t \geq 0$, is non-explosive, positive recurrent and exponential ergodicity by Theorem A. This implies existence of a unique stationary distribution π.

To show that π has finite mth moment for any $m \in \mathbb{N}^d$, we use (22) below combined with the ergodic theorem [24]. Let $\tau_M = \inf\{t > 0 : |X(t)| \geq M\}$. Then by using Dynkin's formula [25, 26] and (21), we have

$$\mathbb{E}_x(V(X(t))) = V(x) + \mathbb{E}_x \left(\int_0^{t \wedge \tau_M} \mathcal{L}V(X(s)) ds \right)$$
$$\leq V(x) - C\mathbb{E}_x \left(\int_0^{t \wedge \tau_M} V(X(s)) ds \right) + Dt$$
$$= V(x) - C\mathbb{E}_x \left(\int_0^{t} V(X(s)) \mathbb{1}_{\{|X(s)| < M\}} ds \right) + Dt,$$
$$(22)$$

where $\mathbb{E}_x$ denotes the expectation of $X(t)$ with $X(0) = x$, and $t \wedge \tau_M = \min\{t, \tau_M\}$. By rearranging terms in (22) and dividing by t, C, it follows that

$$\mathbb{E}_x \left(\frac{1}{t} \int_0^t V(X(s)) \mathbb{1}_{\{|X(s)| < M\}} ds \right) \leq \frac{V(x)}{Ct} + \frac{D}{C}. \quad (23)$$

Then by the dominant convergence theorem, taking lim for $t \to \infty$ on both sides in (23) gives that

$$\lim_{t \to \infty} \mathbb{E}_x \left(\frac{1}{t} \int_0^t V(X(s)) \mathbb{1}_{\{|X(s)| < M\}} ds \right)$$
$$= \sum_{x \in \mathbb{N}^d} V(x) \mathbb{1}_{\{|x| < M\}} \pi(x) \leq \frac{D}{C}.$$

Then the monotone convergence theorem applies for $M \to \infty$ to conclude that $\sum_{x \in \mathbb{N}^d} V(x) \pi(x) \leq \frac{D}{C}$. Since $V(x) = e^{\|x\|_1}$, any moment of π is finite. $\square$

Appendix B. Stationary distribution

Proof of Theorem 3. Under the assumption of equal outflow rates, the process $X(t)$ that counts the molecules of each species is lumpable on the partition $\{E_n\}_{n \in \mathbb{N}}$, where $E_n = \{\mathbf{a} \in \mathbb{N}^2 : \sum_{i=1}^d a_i = n\}$.

The lumped process $\overline{X}(t) = \sum_{i=1}^d X_i(t)$ has Poisson stationary distribution $\nu(n)$ with intensity (15). As stated earlier, the stationary distribution $\Pi(\mathbf{a})$ factorizes as $\Pi(\mathbf{a}) = \pi(\mathbf{a}|n)\nu(n)$. Under the given assumptions on the parameters, $\pi(\mathbf{a}|n)$ solves the equation, similar to (8),

$$R_n = L_{n-1} + L_n + L_{n+1}, \quad (24)$$

where

$$R_n = \pi(\mathbf{a}|n) \left[\sum_{i=1}^{d} \lambda_i + \delta n + \sum_{i=1}^{d} \sum_{j \neq i} \kappa a_i a_j \right]$$

$$L_{n-1} = \frac{\delta n}{\sum_{i=1}^{d} \lambda_i} \sum_{i=1}^{n} \pi(\mathbf{a} - \mathbf{e}_i | n - 1) \lambda_i$$

$$L_n = \sum_{i,j=1}^{n} \pi(\mathbf{a} - \mathbf{e}_i + \mathbf{e}_j | n) \kappa (a_i - 1)(a_j + 1)$$

$$L_{n+1} = \frac{\sum_{i=1}^{n} \lambda_i}{(n+1)} \sum_{i=1}^{n} \pi(\mathbf{a} + \mathbf{e}_i | n + 1)(a_i + 1),$$

The proof now proceeds by showing that the ansatz $\pi(\cdot|n)$ specified by equation (17) solves equation (24). First we note that if the ansatz is true, then the following recurrence relations hold

$$\pi(\mathbf{a}|n) = \frac{1}{n+1} \sum_{i=1}^{d} (a_i + 1) \, \pi(\mathbf{a} + \mathbf{e}_i | n + 1),$$

$$\pi(\mathbf{a} - \mathbf{e}_i | n - 1) = \frac{a_i (n - 1 + \sum_{i=1}^{d} \alpha_i)}{n(a_i - 1 + \alpha_i)} \pi(\mathbf{a}|n), \quad (25)$$

$$\pi(\mathbf{a} - \mathbf{e}_i + \mathbf{e}_j | n) = \frac{a_i (a_j + \alpha_j)}{(a_j + 1)(a_i - 1 + \alpha_i)} \pi(\mathbf{a}|n).$$

Applying (25) and dividing by $\pi(\mathbf{a}|n)$ in (24) we get

$$\delta n + \sum_{i=1}^{d} \sum_{j \neq i} \kappa a_i a_j = \frac{\delta (n - 1 + \sum_{i=1}^{d} \alpha_i)}{\sum_{i=1}^{d} \lambda_i} \sum_{i=1}^{d} \frac{\lambda_i a_i}{a_i - 1 + \alpha_i}$$

$$+ \sum_{i=1}^{d} \sum_{j \neq i} \frac{\kappa a_i (a_i - 1)(a_j + \alpha_j)}{a_i - 1 + \alpha_i} \quad (26)$$

By fixing $a_i = n$, the following condition is necessary

$$\delta(n - 1 + \alpha_i)$$
$$= \delta \left(n - 1 + \sum_{i=1}^{d} \alpha_i \right) \frac{\lambda_i}{\sum_{i=1}^{d} \lambda_i} + \kappa(n-1) \sum_{j \neq i} \alpha_j. \quad (27)$$

If we further set $n = 1$ we get

$$\frac{\lambda_i}{\sum_{i=1}^d \lambda_i} = \frac{\alpha_i}{\sum_{i=1}^d \alpha_i}. \tag{28}$$

Moreover if we take equation (27) and sum over all $i = 1, \cdots, d$, we get

$$(d-1)\delta(n-1) + \delta \sum_{i=1}^d \alpha_i = \delta \sum_{i=1}^d \alpha_i + \kappa(n-1)(d-1) \sum_{i=1}^d \alpha_i,$$

which further implies

$$\sum_{i=1}^d \alpha_i = \frac{\delta}{\kappa}. \tag{29}$$

Together with equation (28), this implies

$$\alpha_i = \frac{\delta \lambda_i}{\kappa \sum_{i=1}^d \lambda_i}. \tag{30}$$

Taking again equation (26), we can further recast it into the following form

$$\delta n = -\sum_{i=1}^d \sum_{j \neq i} \kappa a_i a_j + \frac{\delta(n - 1 + \sum_{i=1}^d \alpha_i)}{\sum_{i=1}^d \lambda_i} \sum_{i=1}^d \frac{\lambda_i a_i}{a_i - 1 + \alpha_i}$$

$$+ \sum_{i=1}^d \sum_{j \neq i} \frac{\kappa a_i(a_i - 1 + \alpha_i)(a_j + \alpha_j)}{a_i - 1 + \alpha_i}$$

$$- \sum_{i=1}^d \sum_{j \neq i} \frac{\kappa a_i \alpha_i (a_j + \alpha_j)}{a_i - 1 + \alpha_i}$$

$$= \frac{\delta(n - 1 + \sum_{i=1}^d \alpha_i)}{\sum_{i=1}^d \lambda_i} \sum_{i=1}^d \frac{\lambda_i a_i}{a_i - 1 + \alpha_i} +$$

$$+ \sum_{i=1}^d \sum_{j \neq i} \kappa a_i \alpha_j - \sum_{i=1}^d \sum_{j \neq i} \frac{\kappa a_i \alpha_i (a_j + \alpha_j)}{a_i - 1 + \alpha_i}$$

$$= \frac{\delta(n - 1 + \sum_{i=1}^d \alpha_i)}{\sum_{i=1}^d \lambda_i} \sum_{i=1}^d \frac{\lambda_i a_i}{a_i - 1 + \alpha_i} +$$

$$+ \sum_{i=1}^d \sum_{j=1}^d \kappa a_i \alpha_j - \sum_{i=1}^d \kappa a_i \alpha_i$$

$$- \sum_{i=1}^d \sum_{j=1}^d \frac{\kappa a_i \alpha_i (a_j + \alpha_j)}{a_i - 1 + \alpha_i} + \sum_{i=1}^d \sum_{i=1}^d \frac{\kappa a_i \alpha_i (a_i + \alpha_i)}{a_i - 1 + \alpha_i},$$

$$\delta n = \frac{\delta(n - 1 + \sum_{i=1}^d \alpha_i)}{\sum_{i=1}^d \lambda_i} \sum_{i=1}^d \frac{\lambda_i a_i}{a_i - 1 + \alpha_i} + \sum_{i=1}^d \sum_{j=1}^d \kappa a_i \alpha_j$$

$$- \sum_{i=1}^d \kappa a_i \alpha_i - \left(n + \sum_{i=1}^d \alpha_i\right) \sum_{i=1}^d \frac{\kappa a_i \alpha_i}{a_i - 1 + \alpha_i}$$

$$+ \sum_{i=1}^d \sum_{i=1}^d \frac{\kappa a_i \alpha_i (a_i + \alpha_i)}{a_i - 1 + \alpha_i}$$

$$= \frac{\delta(n - 1 + \sum_{i=1}^d \alpha_i)}{\sum_{i=1}^d \lambda_i} \sum_{i=1}^d \frac{\lambda_i a_i}{a_i - 1 + \alpha_i} + \sum_{i=1}^d \sum_{j=1}^d \kappa a_i \alpha_j$$

$$- \sum_{i=1}^d \kappa a_i \alpha_i - \left(n + \sum_{i=1}^d \alpha_i\right) \sum_{i=1}^d \frac{\kappa a_i \alpha_i}{a_i - 1 + \alpha_i}$$

$$+ \sum_{i=1}^d \sum_{i=1}^d \frac{\kappa a_i \alpha_i (a_i + \alpha_i - 1)}{a_i - 1 + \alpha_i}$$

$$- \sum_{i=1}^d \sum_{i=1}^d \frac{\kappa a_i \alpha_i}{a_i - 1 + \alpha_i}$$

and, finally,

$$\delta n = \frac{\delta(n - 1 + \sum_{i=1}^d \alpha_i)}{\sum_{i=1}^d \lambda_i} \sum_{i=1}^d \frac{\lambda_i a_i}{a_i - 1 + \alpha_i} \tag{31}$$

$$+ \sum_{i=1}^d \sum_{j=1}^d \kappa a_i \alpha_j - \left(n + \sum_{i=1}^d \alpha_i - 1\right) \sum_{i=1}^d \frac{\kappa a_i \alpha_i}{a_i - 1 + \alpha_i}$$

Now, using (29) and (30), we have

$$\sum_{i=1}^d \kappa a_i \sum_{j=1}^d \alpha_j = n\delta$$

and

$$\frac{\delta}{\sum_{i=1}^d \lambda_i} \sum_{i=1}^d \frac{\lambda_i a_i}{a_i - 1 + \alpha_i} = \sum_{i=1}^d \frac{\kappa a_i \alpha_i}{a_i - 1 + \alpha_i},$$

making equation (31) identically satisfied. $\square$

Appendix C

Proof of Theorem 4. We first rewrite equation (24), where the κ_{ij} are set to zero for all $j \neq (i+1)_d$ and to a constant value κ otherwise. We get the condition

$$R_n = L_{n-1} + L_n + L_{n+1}, \tag{32}$$

where

$$R_n = \pi(\mathbf{a}|n) \left[\sum_{i=1}^d \lambda_i + \delta n + \sum_{i=1}^d \kappa a_i a_{(i+1)_d}\right]$$

$$L_{n-1} = \frac{\delta n}{\sum_{i=1}^d \lambda_i} \sum_{i=1}^n \pi(\mathbf{a} - \mathbf{e}_i|n-1)\lambda_i$$

$$L_n = \sum_{i=1}^{n} \pi(\mathbf{a} - \mathbf{e}_i + \mathbf{e}_{(i+1)_d}|n)\kappa(a_i - 1)(a_{(i+1)_d} + 1)$$

$$L_{n+1} = \frac{\sum_{i=1}^{n}\lambda_i}{(n+1)}\sum_{i=1}^{n}\pi(\mathbf{a} + \mathbf{e}_i|n+1)(a_i + 1).$$

We now notice that if the uniform ansatz is true, the following recurrence relations also hold

$$\pi(\mathbf{a} + \mathbf{e}_i|n+1) = \frac{n+1}{n+d}\,\pi(\mathbf{a}|n)$$

$$\pi(\mathbf{a} - \mathbf{e}_i|n-1) = \frac{n+d-1}{n}\,\pi(\mathbf{a}|n)$$

Plugging the ansatz (20) and these recurrence relations into (32), we get that equation (32) holds if and only if

$$\delta n + \sum_{i=1}^{d}\kappa a_i a_{(i+1)_d} = \delta(n+d-1) + \sum_{i=1}^{d}\kappa(a_i-1)(a_{(i+1)_d}+1).$$

It simplifies to $0 = \delta(d-1) - \kappa d$. Such a condition is identically satisfied under the hypothesis of the theorem which guarantees

$$\kappa = \frac{d-1}{d}\delta.$$

$\square$

[1] Y. Togashiand K. Kaneko, Transitions induced by the discreteness of molecules in a small autocatalytic system, Physical Review Letters **86**, 2459 (2001), cited By 58.

[2] A. Awazuand K. Kaneko, Discreteness-induced transition in catalytic reaction networks, Phys. Rev. E **76**, 041915 (2007).

[3] S. Berti, C. López, D. Vergni, and A. Vulpiani, Discreteness effects in a reacting system of particles with finite interaction radius, Phys. Rev. E **76**, 031139 (2007).

[4] Y. Togashiand K. Kaneko, Molecular discreteness in reaction-diffusion systems yields steady states not seen in the continuum limit, Phys. Rev. E **70**, 020901 (2004).

[5] T. Biancalani, L. Dyson, and A. J. McKane, Noise-induced bistable states and their mean switching time in foraging colonies, Phys. Rev. Lett. **112**, 038101 (2014).

[6] G. Lente, Stochastic kinetic models of chiral autocatalysis: A general tool for the quantitative interpretation of total asymmetric synthesis, Journal of Physical Chemistry A **109**, 11058 (2005), cited By :54.

[7] J. Sardanyésand T. Alarcón, Noise-induced bistability in the fate of cancer phenotypic quasispecies: A bit-strings approach, Scientific Reports **8** (2018).

[8] P. de Anna, F. Di Patti, D. Fanelli, A. J. McKane, and T. Dauxois, Spatial model of autocatalytic reactions, Phys. Rev. E **81**, 056110 (2010).

[9] J. Sardanys, A. Arderiu, S. F. Elena, and T. Alarcn, Noise-induced bistability in the quasi-neutral coexistence of viral rnas under different replication modes, Journal of The Royal Society Interface **15**, 20180129 (2018).

[10] N. Saitoand K. Kaneko, Theoretical analysis of discreteness-induced transition in autocatalytic reaction dynamics, Phys. Rev. E **91**, 022707 (2015).

[11] M. Nakagawaand Y. Togashi, An analytical framework for studying small-number effects in catalytic reaction networks: A probability generating function approach to chemical master equations, Frontiers in Physiology **7** (2016).

[12] T. Biancalani, T. Rogers, and A. J. McKane, Noise-induced metastability in biochemical networks, Physical Review E - Statistical, Nonlinear, and Soft Matter Physics **86** (2012), cited By :18.

[13] B. Houchmandzadehand M. Vallade, Exact results for a noise-induced bistable system, Physical Review E - Statistical, Nonlinear, and Soft Matter Physics **91** (2015), cited By :10.

[14] S. N. Ethierand T. G. Kurtz, *Markov processes*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics (John Wiley & Sons, Inc., New York, 1986) pp. x+534, characterization and convergence.

[15] P. Mozgunov, M. Beccuti, A. Horvath, T. Jaki, R. Sirovich, and E. Bibbona, A review of the deterministic and diffusion approximations for stochastic chemical reaction networks, Reaction Kinetics, Mechanisms and Catalysis **123**, 289 (2018).

[16] T. G. Kurtz, Solutions of ordinary differential equations as limits of pure jump Markov processes, J. Appl. Probability **7**, 49 (1970).

[17] F. Balland G. Yeo, Lumpability and marginalisability for continuous-time markov-chains, Journal of Applied Probability **30**, 518 (1993).

[18] D. F. Andersonand T. G. Kurtz, *Stochastic analysis of biochemical systems*, Mathematical Biosciences Institute Lecture Series. Stochastics in Biological Systems, Vol. 1 (Springer, Cham; MBI Mathematical Biosciences Institute, Ohio State University, Columbus, OH, 2015) pp. x+84.

[19] D. F. Andersonand D. Cappelletti, Discrepancies between extinction events and boundary equilibria in reaction networks, Journal of Mathematical Biology 10.1007/s00285-019-01394-9 (2019).

[20] A. Agazziand J. C. Mattingly, Seemingly stable chemical kinetics can be stable, marginally stable, or unstable (2018), arXiv:1810.06547.

[21] S. P. Meynand R. L. Tweedie, Stability of Markovian Processes III : Foster-Lyapunov Criteria for Continuous-Time Processes, Advances in Applied Probability **25**, 518 (1993).

[22] J. E. Mosimann, On the compound multinomial distribution, the multivariate β-distribution, and correlations among proportions, Biometrika **49**, 65 (1962).

[23] B. Hoadley, The compound multinomial distribution and bayesian analysis of categorical data from finite populations, Journal of the American Statistical Association **64**, 216 (1969).

[24] J. Norris, *Markov Chains* (Cambridge University Press, 1997).

[25] E. Dynkin, Markov processes, vol. ii. translated with the authorization and assistance of the author by J. Fabius, V. Greenberg, A. Maitra, G. Majone, Die Grundlehren der Mathematischen Wissenschaften **122** (1965).

[26] B. Oksendal, *Stochastic differential equations: an introduction with applications* (Springer Science & Business Media, 2013).