ELSEVIER

A stochastic analysis of first-order reaction networks

Chetan Gadgil[a,1], Chang Hyeong Lee[a,1], Hans G. Othmer[b,*,1,2]

[a]*School of Mathematics, University of Minnesota, Minneapolis, MN 55455, United States*
[b]*Max Planck Institute for Mathematics in the Sciences, Inselstrasse 22, Leipzig, Germany*

Abstract

A stochastic model for a general system of first-order reactions in which each reaction may be either a conversion reaction or a catalytic reaction is derived. The governing master equation is formulated in a manner that explicitly separates the effects of network topology from other aspects, and the evolution equations for the first two moments are derived. We find the surprising, and apparently unknown, result that the time evolution of the second moments can be represented explicitly in terms of the eigenvalues and projections of the matrix that governs the evolution of the means. The model is used to analyze the effects of network topology and the reaction type on the moments of the probability distribution. In particular, it is shown that for an open system of first-order conversion reactions, the distribution of all the system components is a Poisson distribution at steady state. Two different measures of the noise have been used previously, and it is shown that different qualitative and quantitative conclusions can result, depending on which measure is used. The effect of catalytic reactions on the variance of the system components is also analyzed, and the master equation for a coupled system of first-order reactions and diffusion is derived.
© 2004 Society for Mathematical Biology. Published by Elsevier Ltd. All rights reserved.

1. Introduction

Understanding the time-dependent behavior of a system of interacting species is necessary for analyzing numerous problems, including the dynamics of chemical reactions,

---

* Corresponding author.
  E-mail address: othmer@math.umn.edu (H.G. Othmer).

[1] All authors contributed equally to this work.

[2] Permanent address: School of Mathematics, University of Minnesota, Minneapolis, MN 55455, United States.

gene expression profiles, signal transduction, and other biochemical processes. Many of these systems are characterized by low numbers of interacting species: for example, gene transcription involves interactions between 1–3 promoter elements, 10–20 polymerase holoenzyme units, 10–20 molecules of repressor proteins, 3000 RNA polymerase molecules, and ca. 1000 ribosomes (Kuthan, 2001). Since interactions at the molecular level are inherently stochastic there is an inherent "irreproducibility" in these dynamics, which has been demonstrated experimentally for single cell gene expression events (Ozbudak et al., 2002; Spudich and Koshland, 1976; Levsky and Singer, 2003). A major unsolved problem is to understand how the interplay between the nature of the individual steps and the connectivity or topology of the entire network affects the dynamics of the system, irrespective of whether a deterministic or a stochastic description is the most appropriate. In this paper we formulate and analyze the master equation that governs the time evolution of the number density of species that participate in a network of first-order reactions. The network may comprise both conversion reactions of the form $A \rightarrow B$, in which one component is converted to another, and catalytic reactions of the form $\phi \xrightarrow{B} A$, in which the rate of formation of a particular component depends on the concentration of the other (the catalyst), but the concentration of the catalyst is unchanged by the reaction. This is the first step in the analysis of higher-order reaction networks.

There are numerous examples of first-order reaction networks that involve a small number of molecules, for which this analysis is directly applicable. Transcription and translation have been modeled as first-order catalytic reactions (Thattai and van Oudenaarden, 2001). The evolution of the surface morphology during epitaxial growth involves the nucleation and growth of atomic islands, and these processes may be described by first-order adsorption and desorption reactions coupled with diffusion along the surface. Proteins exist in various conformational states, and the reversible transitions between states may be described as a first-order conversion processes (Mayor et al., 2003). Fluctuating protein conformations are important in the movement of small molecules through proteins such as myoglobin; hence it is important to understand the distribution of these states (Iorio et al., 1991; Austin et al., 1975). RNA also exists in several conformations, and the transitions between various folding states follow first-order kinetics (Bokinsky et al., 2003).

One of the earliest investigations of stochastic effects in reactions is by Delbruck (1940), who studied the distribution of the number of molecules for a single reacting species in an auto-catalytic one-component system, and derived an expression for the variance as a function of the mean and initial values of the mean and variance. Siegert (1949) derived the probability distribution for the momentum of a gas as a function of time, and formulated the first stochastic model of a system of first-order conversion reactions, using a matrix formulation to derive the master equation for first and second-order reaction networks. He also outlined the generating function approach for characterizing the distribution of the network components. The system he studied is equivalent to a closed system (i.e., the total mass is conserved) of conversion reactions. He proved that one eigenvalue of the matrix of reaction rate constants is zero and the rest are real and negative. Unfortunately, this important work was largely overlooked in the field of

stochastic chemical reaction kinetics for more than a decade, until Krieger and Gans (1960) re-derived these results formulating the problem as a chemical reaction network problem.

Klein (1956) used Siegert's results to analyze the Ehrenfest Urn problem, in which balls are transferred between two urns with given probabilities. Klein treated the number of balls in an urn as a measure of the occupancy of an energy state, and calculated the probability of the number of balls in an urn as a function of the transition probability and the initial distribution. This can be interpreted as a closed system with one first-order reversible reaction, with the urns characterizing the reactant and product of the reversible reaction. He showed that the stationary distribution is independent of the initial distribution, but assumed that transitions occur at fixed intervals of time. Kendall (1948) formulated a master equation for a birth–death process starting with one ancestor and computed the extinction time of the population. He also discussed the case of time-dependent rate constants. Birth processes may be modeled as auto-catalytic production reactions and death as a first-order degradation reaction. Bartholomay (1958) was apparently the first to derive the master equation for a unimolecular reaction allowing steps at random times, and he used the generating function approach to calculate the mean and standard deviation of the number of reactant molecules. He also showed that the stochastic process is "consistent in the mean" with the deterministic description, and later showed how to calculate the observed first-order rate constant (Bartholomay, 1959).

At about the same time Montroll and Shuler (1958) modeled chemical decomposition as a random walk between reflecting and absorbing barriers, and Kim (1958) computed mean first passage times for general first-order stochastic processes. Shuler (1960) analyzed the relaxation kinetics of a multi-state system, which is equivalent to a closed first-order system of reversible conversion reactions, and pointed out that there was no single characteristic "relaxation time". He re-derived the property that such a system cannot exhibit any form of periodic behavior, and showed that the relaxation of any one of the moments of the distribution does not convey any information about the relaxation of the distribution.

A systematic analysis of a closed system of first-order conversion reactions was done by Krieger and Gans (1960), who showed that a multinomial distribution characterizes the temporal evolution of the system. This generalized results of Montroll and Shuler, who had proved that the end states were characterized by a multinomial distribution. Gans extended this and previous analyses of closed systems to an open system of first-order conversion reactions (Gans, 1960). Following Krieger and Gans (1960), he derived a condition for the eigenvalues associated with the evolution of the mean to be negative ($\alpha_{ii} \leq -\sum_j \alpha_{ij}$), but did not analyze the evolution of the higher moments or their relation to the mean. McQuarrie (1963) derived solutions for the mean and variance of closed systems with reactions of the type as $A \rightarrow B$, $A \rightleftharpoons B$, and $A \rightarrow B$, $A \rightarrow C$. He also discussed the use of a cumulant generating function as a method of generating lower-order moments. Gani (1965) formulated a birth–death model for bacteriophage kinetics that was similar to an open system model for one species. Fredrickson (1966) computed the stochastic mean and variance for the concentrations in a closed cyclic ternary system, and again showed that these moments do not oscillate in time. Darvey and Staff (1966) presented the first derivation for the time-dependent mean and variance of all the species present in a closed system with first-order conversion reactions. They derived an expression for the moment

generating function for a case when only one species is present initially, and showed that the first moment is the same as the solution of the corresponding deterministic system, but did not analyze the evolution of the variance.

Other processes such as the waiting times in a queue have also been analyzed, and some of these results can be applied to reaction networks. Kelly (1979) considered reversible queuing processes in which "customers" enter a queue either with a defined distribution from a source or defined transition probabilities from other queues. Every queue has a finite number of "servers", with a characteristic waiting time associated with the service. The entry into a queue from another queue can be regarded as a conversion process, entry from the source is equivalent to production from the source, and service can be thought of as a degradation process that removes customers from the queue. A pure conversion process can be considered as an infinite server queue (see Appendix), but catalytic reactions have no apparent analog in queuing theory. Kelly considered several aspects of the process, such as whether it was open or closed, and the nature of the connections between the queues (linear vs. looped), and derived the important result that the equilibrium distribution for a closed system tends to the equilibrium distribution for an open system when the number of individuals is large. We show later that this result is also true for open and closed systems where all the reactions are first-order conversion reactions. For a particular class of open migration processes in which the transition probabilities between queues (or colonies) is proportional to the number of individuals in the colony, Kelly proved that the number of individuals in each queue has a Poisson distribution, assuming that the inflow of individuals to queues (or colonies) from outside the system (source) are Poisson processes. If one considers each colony to be a distinct species, the open migration process is equivalent to an open conversion reaction system, and the proof for the stationary distribution of the number of individuals in each colony stated by Kelly (1979) may be considered as another proof for the distribution of the number of each species in an open conversion network that we derive later. Branching Markov processes also give rise to problems with a similar mathematical structure to that in kinetics (Harris, 1963; Athreya and Ney, 1972). Athreya and Ney (1972) considered continuous-time, multi-type branching processes initiated by one particle of each type and derived first and second moments of the distribution of the number of particles of each type. Catalytic reactions can be interpreted as the death of a particle with two offspring, one identical to the original and the other possibly different, but a systematic analysis of the effect of catalytic reactions on the resulting probability distribution has not been reported earlier to our knowledge.

Thattai and van Oudenaarden (2001) presented the first analysis of a system of first-order catalytic reactions. They formulated a procedure for deriving the master equation for such systems, similar to the general procedure given in Gardiner (1983). They incorporated a negative feedback regulation of some reactions (production from source), and derived the steady-state means and covariances for a system of catalytic transformations with one source term, and first-order degradation of all the species. Recently, Brown (2003) derived the probability of the number of forward and reverse steps in a reversible first-order conversion reaction in which the transition probabilities are time-dependent.

A major objective of many of the analyses treating biological systems is prediction of the stochastic variations or noise of the concentrations. Two measures of the noise have been used in the past. Until recently the standard measure was the coefficient of

variation (CV), defined as the standard deviation divided by the mean (Delbruck, 1940; Singer, 1953; Kepler and Elston, 2001), or its square. The CV is used as a measure of noise in McQuarrie et al. (1964), Darvey et al. (1966), Laurenzi (2000), Elowitz et al. (2002) and Swain et al. (2002), while the Fano factor $\mathcal{F}$, defined as the variance divided by the mean, was introduced by Thattai and van Oudenaarden (2001), and used in Blake et al. (2003) and Ozbudak et al. (2002). It has been shown that the use of different measures of noise may lead to different conclusions concerning the importance of noise in the underlying process (Swain et al., 2002).

Our objectives here are (i) to introduce a derivation of the master equation that clarifies the separate roles of reaction rates and network topology in the master equation, (ii) to develop a unified treatment of first-order networks, including the evolution of both the mean and the variance of any species and (iii) to understand the effect of network topology on the stochastic fluctuations in specified components. We compare the noise in the amount of a species as measured by the Fano factor or the coefficient of variation for both conversion and catalytic reactions, as well as for species in open and closed systems. We also demonstrate through simple examples the utility of this framework in the analysis of the effect of reaction network topology on the variation of the number of molecules of each network component. All of the preceding analyses can be treated as special cases of the general framework that we develop here. No previous analysis of first-order reaction systems has to our knowledge considered a system consisting of both catalytic and conversion reactions, nor has a systematic comparison of the stochastic behavior of conversion and catalytic systems been made.

2. Formulation of the master equation

We first derive the master equation for a general system of reactions to provide a framework for the analysis of reactions of arbitrary order. We then focus on first-order reactions and specialize the general result for a very large class of first-order processes.

2.1. The master equation for a general system of reactions

We begin with some background on a general deterministic description of reacting systems, and then derive the master equation for an arbitrary network of reacting species. The abstract formulation is presented in brief here and follows that given elsewhere (Othmer, 1979, 1981).

Suppose that the reacting mixture contains the set $\mathcal{M}$ of s chemical species $\mathcal{M}_i$ that participate in a total of r reactions. Let ν_{ij} be the stoichiometric coefficient of the ith species in the jth reaction. The ν_{ij} are non-negative integers that represent the normalized molar proportions of the species in a reaction. Each reaction is written in the form

$$\sum_{i}^{reac.} \nu_{ij}^{reac} \mathcal{M}_i = \sum_{i}^{prod} \nu_{ij}^{prod} \mathcal{M}_i \qquad j = 1, \ldots, r, \tag{1}$$

where the sums are over reactants and products, respectively in the jth reaction. In this formulation, the forward and reverse reaction of a reversible pair are considered separately as two irreversible reactions.

For each reaction, once the reactants and products are specified, the significant entities so far as the network topology is concerned are not the species themselves, but rather the linear combinations of species that appear as reactants or products in the various elementary steps. Following Horn and Jackson (1972), these linear combinations of species will be called complexes. A species may also be a complex (as is the case for first-order reactions). We assume that changes in temperature, pressure and volume V of the mixture during reaction are negligible. Thus the state of the system is specified by the concentration vector $c = (c_1, \ldots, c_s)^T$, where c_i is the non-negative concentration of species $\mathcal{M}_i$ measured in moles/liter.

Let $\hat{\mathcal{M}}$ be the set of linear combinations with integral coefficients of the species, and let $\mathcal{C} = \{C(1), \ldots, C(p)\}$ be a set of complexes. A *reaction network* consists of the triple $\{\mathcal{M}, \hat{\mathcal{M}}, \mathcal{C}\}$, together with a stoichiometric function $\hat{v} : \hat{\mathcal{M}} \to \mathcal{C}$ and a binary relation $R \subset \mathcal{C} \times \mathcal{C}$. The function $\hat{v}$, which identifies a linear combination of species as a complex is onto, and the relation R has the properties (i) $(C(i), C(j)) \in R$ if and only if there exists one and only one reaction of the form $C(i) \to C(j)$, (ii) for every i there is a $j \neq i$ such that $(C(i), C(j)) \in R$, (iii) $(C(i), C(i)) \notin R$. Thus every complex is related to at least one other complex and the trivial reaction $C(i) \to C(i)$ that produces no change is not admitted. Therefore R is never reflexive and in general it is neither symmetric nor transitive.

The relation on $\mathcal{C}$ gives rise to a directed graph $\mathcal{G}$ in the following way. Each complex is identified with a vertex V_k in $\mathcal{G}$ and a directed edge E_ℓ is introduced into $\mathcal{G}$ for each reaction. Each edge carries a non-negative weight $\hat{\mathcal{R}}_\ell(c)$ given by the intrinsic rate of the corresponding reaction. $\mathcal{G}$ provides a concise representation of the reaction network.

The topology of $\mathcal{G}$ is in turn represented in its vertex–edge incidence matrix $\mathcal{E}$, which is defined as follows.

$$\mathcal{E}_{i\ell} = \begin{cases} +1 & \text{if } E_\ell \text{ is incident at } V_i \text{ and is directed toward it} \\ -1 & \text{if } E_\ell \text{ is incident at } V_i \text{ and is directed away from it} \\ 0 & \text{otherwise.} \end{cases} \qquad (2)$$

If there are r reactions on $\mathcal{C}$, then $\mathcal{E}$ has p rows and r columns and every column has exactly one $+1$ and one -1. The rate $\hat{\mathcal{R}}_\ell(C)$ of an elementary reaction $C(j) \to C(k)$ is generally not a function of $C(j)$, but of the concentration or activity of the individual species in the complex. Once the complexes and reactions are fixed, the stoichiometry of the complexes is specified unambiguously, and we let v denote the $s \times p$ matrix whose jth column encodes the stoichiometric amounts of the species in the jth complex. Then the temporal evolution of the composition of a reacting mixture is governed by

$$\frac{dc}{dt} = v\mathcal{E}\hat{\mathcal{R}}(c), \qquad (3)$$

where the columns of v are given by the columns of v^{reac} and v^{prod}, and the initial condition is $c(0) = c_0$. It follows from (2) that the columns of the product $v\mathcal{E}$ are the stoichiometric vectors of reactions written according to the standard convention. When the reactions are first-order this deterministic equation also governs the evolution of the mean in the Markov process description discussed later.

A special but important class of rate functions is that in which the rate of the ℓth reaction can be written as

$$\hat{\mathcal{R}}_\ell(c) = k_\ell R_j(c) \tag{4}$$

for every reaction that involves the jth complex as the reactant. This includes ideal mass action rate laws, in which the rate is proportional to the product of the concentrations of the species in the reactant complex, each concentration raised to a power equal to the stoichiometric coefficient of the corresponding species in the complex. In that case

$$R_j = \prod_{i=1}^{s} (c_i)^{v_{ij}}. \tag{5}$$

For mass-action kinetics (4) implies that

$$\hat{\mathcal{R}}(c) = K R(c) \tag{6}$$

where K is an $r \times p$ matrix with $k_{\ell j} > 0$ if and only if the ℓth edge leaves the jth vertex, and $k_{\ell j} = 0$ otherwise. The topology of the underlying graph $\mathcal{G}$ enters into K as follows. Define the exit matrix $\mathcal{E}_o$ of $\mathcal{G}$ by replacing all 1's in $\mathcal{E}$ by zeros, and changing the sign of the resulting matrix. Let $\hat{K}$ be the $r \times r$ diagonal matrix with the k_ℓ's, $\ell = 1, \ldots, r$, along the diagonal. Then it is easy to see that $K = \hat{K} \mathcal{E}_0^T$ and therefore

$$\frac{dc}{dt} = v \mathcal{E} K R(c) = v \mathcal{E} \hat{K} \mathcal{E}_o^T R(c). \tag{7}$$

It follows from the definitions that (i) the (p, q)th entry, $p \neq q$, of $\mathcal{E} \hat{K} \mathcal{E}_0^T$ is nonzero (and positive) if and only if there is a directed edge $(q, p) \in \mathcal{G}$, (ii) each diagonal entry of $\mathcal{E} \hat{K} \mathcal{E}_0^T$ is minus the sum of the k's for all edges that leave the jth vertex, and (iii) the columns of $\mathcal{E} \hat{K}_0^T$ all sum to zero, and so the rank of $\mathcal{E} \hat{K} \mathcal{E}_0^T$ is $\leq p - 1$. When all complexes are species and all reactions are first-order, $v = $ I for a closed system and $v = [$I $\mid$ **0**$]$ for an open system, where I is the $s \times s$ identity matrix and **0** is the zero vector, and the right-hand side reduces to the usual form $K c$ for a suitably-defined matrix K. In the following section we will treat the stochastic analysis of first-order systems in detail.

As it stands, (5) includes all reacting species, but those whose concentration is constant on the time scale of interest can be deleted from each of the complexes in which it appears and its concentration or mole fraction can be absorbed into the rate constant of that reaction in which it participates as reactant.[3] As a result of these deletions, it will appear that reactions which involve constant species do not necessarily conserve mass. Furthermore, some complexes may not comprise any time-dependent species; these will be called zero or null complexes. Each null complex gives rise to a column of zeros in v and the rate of any reaction in which the reactant complex is a null complex is usually constant. For instance, any transport reaction of the form $\mathcal{M}^0 \rightarrow \mathcal{M}_i$ introduces a null complex and the corresponding flux of $\mathcal{M}_i$ represents a constant input to the reaction network, provided that the rate of the transport step does not depend on the concentration of a time-dependent species. Of course, a constant species that appears in a complex which also contains a

---

[3] Hereafter s will denote the number of species whose concentration may be time-dependent.

variable species likewise represents an input to the network, and to distinguish these from inputs due to null complexes, the former are called *implicit inputs* and the latter are called *explicit inputs*.

An alternate description of the deterministic dynamics is obtained by introducing an extent for each reaction and expressing composition changes in terms of extents. It follows from (3) that the composition changes due to reaction lie in a coset by c_0 of the range of $v\mathcal{E}$, and this coset is called the reaction simplex (Othmer, 1979, 1981). Therefore, by choosing coordinates in the simplex, the composition changes can be expressed in terms of an extent for each reaction as follows

$$c = c_0 + v\mathcal{E}\xi. \tag{8}$$

If the reactions are all independent, i.e. if the rank of $v\mathcal{E}$ is r, it follows from (3) and (8) that

$$\frac{d\xi}{dt} = \hat{\mathcal{R}}(c_0 + v\mathcal{E}\xi), \tag{9}$$

but in general we can only conclude that

$$\frac{d\xi}{dt} = \hat{\mathcal{R}}(c_0 + v\mathcal{E}\xi) + \sum_k \gamma_k \mathcal{N}_k, \tag{10}$$

where $\{\mathcal{N}_k\}$ is a basis for the null space of $v\mathcal{E}$. The γ_k can be chosen so as to remove all dependent steps, in particular, those that arise from cycles in the graph.

We can also describe the evolution in terms of the number of molecules present for each species. Let $n = (n_1, n_2, \ldots, n_s)$ denote the discrete composition vector whose ith component n_i is the number of molecules of species $\mathcal{M}_i$ present in the volume V. This is the discrete version of the composition vector c, and they are related by $n = \mathcal{N}_A V c$, where $\mathcal{N}_A$ is Avogadro's number. From (3) we obtain the deterministic evolution for n as

$$\frac{dn}{dt} = v\mathcal{E}\tilde{\mathcal{R}}(n) \tag{11}$$

where $\tilde{\mathcal{R}}(n) \equiv \mathcal{N}_A V \hat{\mathcal{R}}(n/\mathcal{N}_A V)$. In particular, for mass-action kinetics

$$\tilde{\mathcal{R}}_\ell(n) = \mathcal{N}_A V k_\ell \mathcal{R}_j(n/\mathcal{N}_A V) = \mathcal{N}_A V k_\ell \prod_{i=1}^{s} \left(\frac{n_i}{\mathcal{N}_A V} \right)^{v_{ij}}$$

$$= \frac{k_\ell}{(\mathcal{N}_A V)^{\sum_i v_{ij} - 1}} \prod_{i=1}^{s} (n_i)^{v_{i\ell}} = \hat{k}_\ell \prod_{i=1}^{s} (n_i)^{v_{ij}}. \tag{12}$$

The number of molecules can be expressed in terms of the integer extents of each reaction as

$$n = n_0 + v\mathcal{E}\eta, \tag{13}$$

and it follows from (11) that

$$\frac{d\eta}{dt} = \tilde{\mathcal{R}}(n_0 + v\mathcal{E}\eta) + \sum_k \gamma_k \mathcal{N}_k. \tag{14}$$

The description in terms of the number of molecules present assumes that there are sufficient numbers present so that we can assume they vary continuously in time, but the same assumption is needed for (4).

2.2. The stochastic description

The first level of stochastic description is to consider an ensemble of deterministic systems that differ in the initial condition. Let $P(c, t)$ be the probability that the state of the system is c; then the evolution of P is governed by

$$\frac{\partial P}{\partial t} + \sum_{i=1}^{s} \frac{\partial}{\partial c_i} \left((\nu \mathcal{E} \hat{R}(c))_i P \right) = 0 \tag{15}$$

subject to the initial condition $P(c, 0) = P_0$. The characteristic equations for this hyperbolic equation are precisely the evolution equations given at (3).

At the next level of description the numbers of the individual components are followed in time and the reactions are modeled as a continuous-time Markov jump process. Let $N_i(t)$ be a random variable that represents the number of molecules of species $\mathcal{M}_i$ at time t, and let N denote the vector of N_is. Further, let $P(n, t)$ be the joint probability that $N(t) = n$, i.e., $N_1 = n_1, N_2 = n_2, \ldots, N_s = n_s$. Clearly the state of the system at any time is now a point in $\mathcal{Z}_0^s$, where $\mathcal{Z}_0$ is the set of non-negative integers. Formally the master equation that governs the evolution of P is

$$\frac{\mathrm{d}}{\mathrm{d}t} P(n, t) = \sum_{m \in \mathcal{S}(n)} \mathcal{R}(m, n) \cdot P(m, t) - \sum_{m \in \mathcal{T}(n)} R(n, m) \cdot P(n, t) \tag{16}$$

where $\mathcal{R}(m, n)$ is the probability per unit time of a transition from state m to state n, $\mathcal{R}(n, m)$ is the probability per unit time of a transition from state n to state m, $\mathcal{S}(n)$ is the set of all states that can terminate at n after one reaction step, and $\mathcal{T}(n)$ is the set of all states reachable from n in one step of the feasible reactions. The notation is meant to suggest the 'source' and 'target' states at n; one could also call $\mathcal{S}(n)$ the predecessors of state n and $\mathcal{T}(n)$ the successors of state n. The predecessor states must be non-negative for production reactions and positive for conversion, degradation and catalytic reactions. Similar bounds on the target states are naturally enforced by zero rates of reaction when the reactants are absent.

The sets $\mathcal{S}(n)$ and $\mathcal{T}(n)$ are easily determined using the underlying graph structure. It follows from the definition of ν and $\mathcal{E}$ that the ℓth reaction $C(j) \rightarrow C(k)$ induces a change $\Delta n^{(\ell)} = \nu \mathcal{E}_{(\ell)}$ in the number of molecules of all species after one step of the reaction, where subscript ℓ denotes the ℓth column. Therefore the state $m = n - \nu \mathcal{E}_{(\ell)}$ is a source or predecessor to n under one step of the ℓth reaction. Similarly, states of the form $m = n + \nu \mathcal{E}_{(\ell)}$ are reachable from n in one step of the ℓth reaction.[4] Once the graph of the network and the stoichiometry are fixed, we can sum over reactions rather than sources

---

[4] A slightly more abstract way of stating this is that each complex defines an equivalence class $E \subset \mathcal{Z}_0^s$, and the change in number of molecules due to one step of the kth reaction lies in the direction $\nu_{(j)} - \nu_{(i)}$ in $E_i \cap E_j$.

and targets, and consequently the master equation takes the form

$$\frac{\mathrm{d}}{\mathrm{d}t} P(n, t) = \sum_\ell \mathcal{R}_\ell(n - \nu\mathcal{E}_{(\ell)}) \cdot P(n - \nu\mathcal{E}_{(\ell)}, t) - \sum_\ell \mathcal{R}_\ell(n) \cdot P(n, t). \qquad (17)$$

However, the transition probabilities $\mathcal{R}_\ell(n)$ are not simply the macroscopic rates $\hat{\mathcal{P}}$ if the reactions are second-order (or higher), because as Gillespie (1976) and others have noted, combinatorial effects may play a significant role when the number of molecules is small. Hereafter we restrict attention to mass-action kinetics, and we suppose that the ℓth reaction involves conversion of the jth to the kth complex: $C(j) \to C(k)$. Then using the notation of Gillespie (1976), we can write,

$$\mathcal{R}_\ell = c_\ell h_{j(\ell)}(n) \qquad (18)$$

where c_ℓ is the probability per unit time that the molecular species in the jth complex react, $j(\ell)$ denotes the reactant complex for the ℓth reaction, and $h_{j(\ell)}(n)$ is the number of independent combinations of the molecular components in this complex. Thus

$$c_\ell = \frac{k_\ell}{(\mathcal{N}_A V)^{\sum_i \nu_{ij(\ell)} - 1}} = \hat{k}_\ell \qquad (19)$$

and

$$h_{j(\ell)} = \prod_i \binom{n_i}{\nu_{ij(\ell)}}. \qquad (20)$$

In the definition of h we use the standard convention that $\binom{n}{0} = 1$.

We can write the master equation in terms of integer extents in the form

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} \tilde{\mathcal{P}}(\eta, t) &= \sum_\ell \mathcal{R}_\ell(n_0 + \nu\mathcal{E}\eta - \nu\mathcal{E}_{(\ell)}) \cdot P(n_0 + \nu\mathcal{E}\eta - \nu\mathcal{E}_{(\ell)}, t) \\
&\quad - \sum_\ell \mathcal{R}_\ell(n_0 + \nu\mathcal{E}\eta) \cdot P(n_0 + \nu\mathcal{E}\eta, t) \\
&= \sum_\ell \tilde{\mathcal{R}}_\ell(\eta - \nu\mathcal{E}_{(\ell)}) \cdot \tilde{P}(\eta - \nu\mathcal{E}_{(\ell)}, t) - \sum_\ell \tilde{\mathcal{R}}_\ell(\eta) \cdot \tilde{P}(\eta, t). \qquad (21)
\end{aligned}$$

Moments of this equation or of (16) can be used to obtain the evolution equations for average extents and from this, the equations for the average change in the numbers. Only in the linear case is the right-hand side of the equation for the first moment the deterministic rate, as is shown in the following subsection. Others have derived a similar master equation for reacting systems, without the explicit inclusion of the underlying graph-theoretic structure (Gardiner, 1983; Rao and Arkin, 2003).

2.3. The master equation for general system of first-order reactions

The stochastic analysis of first-order reaction networks can be done in essentially complete generality, and in this section we analyze all cases in which every reactant and product complex is a species. We thereby exclude only those first-order splitting reactions of the type $\mathcal{M}_i \to \mathcal{M}_j + \mathcal{M}_k$. Our aim is to separate the effects of various types of

Table 1
The four classes of first-order reactions considered in the stochastic model

Label	Type of reaction	Reaction	Rate
I	Production from a source	$\phi \to \mathcal{M}_i$	k_i^s
II	Degradation	$\mathcal{M}_i \to \phi$	$k_i^d n_i$
III	Conversion	$\mathcal{M}_j \to \mathcal{M}_i$	$k_{ij}^{con} n_j$
IV	Catalytic production from source	$\phi \overset{\mathcal{M}_j}{\longrightarrow} \mathcal{M}_i$	$k_{ij}^{cat} n_j$

reactions (catalytic, conversion) on the distribution of the chemical species, and to this end we divide the set of all reactions, represented by the directed edges $E_\ell, \ell = 1, 2, \ldots, r$ into four subsets corresponding to the following reactions: production from a constant source (which in fact is a zero-order step), degradation, conversion to another species, and production catalyzed by another species. These four types are summarized in Table 1. The first type represents an explicit input to the system, whereas the last type represents an implicit input.

Every species $\mathcal{M}_i$ can be produced from a source at a specific rate k_i^s, and every species can be removed by degradation at a rate proportional to its concentration, with rate constant given by k_i^d. Each species may participate in two other types of first-order reactions: conversion reactions, in which species $\mathcal{M}_i$ is converted to species $\mathcal{M}_j$ at a rate proportional to its concentration, and catalytic reactions, in which species $\mathcal{M}_i$ catalyzes the formation of species $\mathcal{M}_j$ from a source, but is itself unchanged during the process. The first-order rate constant for the conversion reaction of species $\mathcal{M}_i$ to species $\mathcal{M}_j$ is denoted by k_{ji}^{con}, and the first-order rate constant for the catalytic production of species $\mathcal{M}_j$, with species $\mathcal{M}_i$ catalyzing the reaction, is denoted by k_{ji}^{cat}. Thus there are s uncatalyzed source reactions, s first-order decay reactions, $s(s - 1)$ conversion and s^2 catalytic reactions, for a total of up to $2s^2 + s$ reactions.

Since all reactant and product complexes are species, the stoichiometric matrix is

$$\nu = [\ \mathrm{I} \mid \mathbf{0}\ ]$$

if at least one reaction of type I, II, or IV is present, and

$$\nu = [\ \mathrm{I}\ ]$$

if the system is closed. The corresponding incidence matrices for the different types are equally simple, and if we order the types as in Table 1, then $\mathcal{E}$ can be written as follows.

$$\mathcal{E} = \left[\ \frac{\mathrm{I}}{-\mathbf{1}^T}\ \middle|\ \frac{-\mathrm{I}}{\mathbf{1}^T}\ \middle|\ \mathcal{E}^{con}\ \middle|\ \frac{\mathrm{I}}{-\mathbf{1}^T}\ \right] \tag{22}$$

where $\mathbf{1}^T = (1, 1, \ldots, 1)$, I is the identity matrix of the appropriate dimension, and $\mathcal{E}^{con}$ is the incidence matrix for the conversion network. Thus the stoichiometry of the reactions and the topology of the network are easily encoded in ν and $\mathcal{E}$, respectively.

It follows easily that the deterministic equations for the first-order reaction network can be written as

$$\frac{\mathrm{d}c}{\mathrm{d}t} = K^s \mathbf{1} - K^d c + K^{cat} c + K^{con} c \tag{23}$$

where $K^s = diag\{k_i^s\}$, $K^d = diag\{k_i^d\}$, $K_{ij}^{cat} = k_{ij}^{cat}$ and K^{con} is defined as follows.

$$K_{ij}^{con} = \begin{cases} k_{ij}^{con} & \text{if } i \neq j \\ -\sum_k{}' k_{kj}^{con} & \text{if } i = j. \end{cases}$$

It is clear that type I and IV reactions induce an increase of 1 in the number of species i without other changes, type IV induces a decrease of one in i alone, and type III induces a decrease of one in j and an increase of 1 in i. Therefore, for reactions of type I the predecessor state to state n is $\mathcal{S}_i^{-1} n \equiv (n_1, n_2, \ldots, n_i - 1, \ldots, n_s)$, and the successor state is $\mathcal{S}_i^{+1} n \equiv (n_1, n_2, \ldots, n_i + 1, \ldots, n_s)$, where $\mathcal{S}_i^k$ is the shift operator that increases the ith component of n by an integer amount k. For degradation of $\mathcal{M}_i$, the predecessor state is $\mathcal{S}_i^{+1} n$ and the successor state is $\mathcal{S}_i^{-1} n$. In type III reactions the predecessor state is $\mathcal{S}_i^{-1} \mathcal{S}_j^{+1} n$ and the successor state is $\mathcal{S}_i^{+1} \mathcal{S}_j^{-1} n$. Finally, for the catalytic reaction the predecessor and successor states are $\mathcal{S}_i^{-1} n$ and $\mathcal{S}_i^{+1} n$, respectively. Using these, the master equation for the first-order reaction network can be written as follows (here and hereafter we drop the explicit time dependence in $P(.,.)$).

$$\begin{aligned} \frac{\mathrm{d}P(n)}{\mathrm{d}t} = \sum_{i=1}^{s} \Bigg[& K_{ii}^s (\mathcal{S}_i^{-1} - 1) P(n) + \sum_{j=1}^{s} \Big(K_{ij}^{con} (\mathcal{S}_i^{-1} \mathcal{S}_j^{+1}) \\ & + K_{ij}^{cat} (\mathcal{S}_i^{-1} - 1) - K_{ij}^d (\mathcal{S}_i^{+1} - 1) \Big) (n_j P(n)) \Bigg] \end{aligned} \tag{24}$$

where $S_i^k (n_i P(n)) = S_i^k n_i \cdot P(S_i^k n)$.

3. Evolution equations for the mean and variance

The master equation derived in the previous section cannot be solved analytically except for a small number of specific simple systems. Usually the objective of a stochastic analysis is to calculate the moments of the distribution of the number of reactant molecules. There are several ways in which evolution equations for the moments of the distribution function can be obtained. One is to multiply both sides of the master equation by $n_i, n_i n_j, n_i n_j n_k$ etc and sum over all possible values of n. Alternatively, one can use the moment generating function (MGF) approach, which is what we use here to calculate the mean and variance for all the reactants in an arbitrary network. The advantage of the MGF approach is that it allows us to get an analytical solution for the MGF of purely conversion systems, thus enabling the calculation of the probability distribution function for the distribution of each reactant in such systems.

Let $z \equiv (z_1, \ldots, z_s)$ where $z_i \in [0, 1]$; then the MGF is defined as

$$G(z, t) = \sum_{n_i=0}^{\infty} z_1^{n_1} \cdots z_s^{n_s} P(n).$$

It follows that $G(z, t)|_{z=1} = 1$. The first and second moments can be obtained through successive derivatives of the MGF evaluated after setting all z_i to one.

$$M_k(t) = \left.\frac{\partial G(z, t)}{\partial z_k}\right|_{z=1} \equiv G_k(z, t)|_{z=1} = E[N_k]$$

$$V_{lk}(t) = G_{lk}(z, t)|_{z=1} = \begin{cases} E[N_l N_k] & \text{if } l \neq k \\ (E[N_k^2] - E[N_k]) & \text{if } l = k \end{cases}$$

where $E[\ ]$ denotes the expectation of the quantity in the square brackets, and for any combination of indices we define

$$G_{ijk\ldots,t} = \frac{\partial}{\partial z_i} \frac{\partial}{\partial z_j} \frac{\partial}{\partial z_k} \cdots \frac{\partial}{\partial t} G.$$

The probability distribution of the number of molecules of the ith species (N_i) at steady state can be derived from the MGF. Differentiating the MGF k times w.r.t. z_i, we get

$$\frac{\partial^k G(z, t)}{\partial z_i^k} = \sum_{n_i=k}^{\infty} \frac{n_i!}{(n_i - k)!} z_i^{n_i-k} \sum_{n_j=0, j \neq i}^{\infty} z_1^{n_1} \cdots z_{i-1}^{n_{i-1}} z_{i+1}^{n_{i+1}} \cdots z_s^{n_s} P(n, t),$$

and therefore

$$\left.\frac{\partial^k G(z, t)}{\partial z_i^k}\right|_{z_i=0, z_j=1, j \neq i} = k! \sum_{n_j, j \neq i} P_{n_1, \ldots, n_{i-1}, k, n_{i+1}, \ldots, n_s}(t) = k! P_{N_i}(k, t),$$

where $P_{N_i}(k, t)$ represents the marginal probability density function of N_i. Therefore

$$P_{N_i}(k, t) = \frac{1}{k!} \left.\frac{\partial^k G(z, t)}{\partial z_i^k}\right|_{z_i=0, z_j=1, j \neq i}.$$

These relationships are valid for the MGF corresponding to any reaction network. In the case of a system of reactions with first-order kinetics we obtain the partial differential equation for the MGF as

$$G_t = \sum_{i=1}^{s} (z_i - 1) \left(K_{ii}^s G + \sum_{j=1}^{s} (K_{ij}^{con} + K_{ij}^{cat} z_j - K_{ij}^d) G_j \right). \tag{25}$$

Alternatively, and somewhat more directly, one can use a backward equation to obtain Eq. (25) (see the Appendix).

For simple network topologies with a small number of nodes, (25) can be solved analytically to get the complete characterization of the evolution of the probability distribution function $P(n)$. We shall later derive such expressions for systems where only conversion reactions occur. However, this is not feasible for an arbitrary network structure, and therefore we first outline the procedure for obtaining the evolution equation for the moments,

and then we focus on the first two moments. Differentiating Eq. (25) with respect to z_k, we obtain

$$G_{k,t} = \sum_{i=1}^{s}(z_i - 1)\left(K_{ii}^s G_k + \sum_{j=1}^{s}(K_{ij}^{con} + K_{ij}^{cat} z_j - K_{ij}^d)G_{jk} + K_{ik}^{cat} G_k\right)$$
$$+ \left(K_{kk}^s G + \sum_{j=1}^{s}(K_{kj}^{con} + K_{kj}^{cat} z_j - K_{kj}^d)G_j\right). \tag{26}$$

Therefore the evolution equation for the mean of the kth component is

$$E[N_k]' = \sum_{j=1}^{s}(K_{kj}^{con} + K_{kj}^{cat} - K_{kj}^d)E[N_j] + K_{kk}^s$$

or in matrix form

$$M(t)' = (K^{con} + K^{cat} - K^d)M(t) + K^s \mathbf{1} \tag{27}$$
$$= \mathcal{K}M(t) + K^s \mathbf{1}, \tag{28}$$

where $M(t) = [E[N_1(t)], \ldots, E[N_s(t)]]^T$ and $\mathcal{K}$ is defined by the second equality. From this one sees that the explicit inputs contained in the last term serve as a nonhomogeneous forcing term for the evolution of the mean. This equation is identical to Eq. (23) given earlier for the evolution of the deterministic first-order system. Therefore a general stochastic system of first-order interactions is "consistent in the mean" with the corresponding system for systems comprised of catalytic and conversion reactions, as is well known (Darvey and Staff, 1966).

Higher moments of the distribution are obtained by successive differentiation of (25). For the second moment we obtain

$$G_{lk,t} = \sum_{i=1}^{s}(z_i - 1)\left(K_{ii}^s G_{kl} + \sum_{j=1}^{s}(K_{ij}^{con} + K_{ij}^{cat} z_j - K_{ij}^d)G_{jkl} + K_{ik}^{cat} G_{kl}\right.$$
$$\left. + K_{il}^{cat} G_{lk}\right) + \left(K_{ll}^s G_k + \sum_{j=1}^{s}(K_{lj}^{con} + K_{lj}^{cat} z_j - K_{lj}^d)G_{jk} + K_{lk}^{cat} G_k\right)$$
$$+ \left(K_{kk}^s G_l + \sum_{j=1}^{s}(K_{kj}^{con} + K_{kj}^{cat} z_j - K_{kj}^d)G_{jl} + K_{kl}^{cat} G_l\right).$$

Thus the matrix containing the second-order moments can be written as

$$V(t)' = \mathcal{K}V(t) + (\mathcal{K}V(t))^T + \Gamma(t) + \Gamma^T(t) \tag{29}$$

where

$$\Gamma_{ij}(t) \equiv (K_{ij}^{cat} + K_{ii}^s)M_j(t).$$

It is noteworthy that in this equation both the implicit and explicit inputs, as well as the mean, appear in the forcing term for the evolution of the second moments. In the equation for the mean the forcing is time-independent, but here the forcing is time-dependent via the appearance of the mean in this term. Later we will see how this time-dependence is filtered

via the action of the kinetic matrix. The somewhat unusual structure on the right-hand side arises from the fact that V is a symmetric matrix, and thus the evolution equation for it must be symmetric as well.

The equations for the mean and the second moments can be integrated to get the first two moments of the distribution. Thus we now have an algorithm to compute both the steady-state and time-dependent behavior of the mean and variance of every species in a general first-order network. Next we analyze this behavior in detail.

3.1. The steady-state and time-dependent solution for the mean

The steady-state solution M^s for the mean is the solution of

$$\mathcal{K}M^s = -K^s \mathbf{1} = -k^s \tag{30}$$

where $k^s = (k_1^s, \ldots, k_s^s)$. The nature of the reaction types and rates, as reflected in the spectral structure of $\mathcal{K}$, dictates the steady-state mean. In general, if $\mathcal{K}$ is singular then $K^s \mathbf{1}$ must lie in the range of $\mathcal{K}$, and in particular, if there are no explicit inputs ($K^s = 0$) then M_1^s is given by an eigenvector corresponding to a zero eigenvalue of $\mathcal{K}$. On the other hand, if degradation reactions are the only type present, and all species react, then $\mathcal{K}$ is nonsingular and $M_1^s = 0$ is the only solution. This case is of little interest and will be excluded in what follows. Results concerning the localization of the spectrum of $\mathcal{K}$ are summarized in the following theorem.

Theorem 1. *The eigenvalues of $\mathcal{K}$ have non-positive real parts if either of the following conditions hold.*

(a) *The sum of the specific rates of formation for each species $\mathcal{M}_i$ by conversion and catalytic reactions does not exceed the sum of the specific rates of loss of $\mathcal{M}_i$ by conversion reactions and degradation (this implies that the column sums of $\mathcal{K}$ are non-positive).*

(b) *The sum of the specific rates of formation catalyzed by each species $\mathcal{M}_i$ is less than or equal to the sum of the specific rates of degradation of that species.*

Proof. The statement in (a) can be translated into the inequality

$$\sum_{j \neq i} K_{ij}^{con} + \sum_{j} K_{ij}^{cat} \leq \sum_{j \neq i} K_{ji}^{con} + K_{ii}^{d}. \tag{31}$$

Since all the terms are non-negative, each of the terms on the left-hand side of the inequality is less than the right-hand side. In particular,

$$K_{ii}^{cat} \leq \sum_{j \neq i} K_{ji}^{con} + K_{ii}^{d}, \tag{32}$$

and therefore

$$\mathcal{K}_{ii} = K_{ii}^{con} + K_{ii}^{cat} - K_{ii}^{d} \tag{33}$$

$$= -\sum_{j \neq i} K_{ji}^{con} + K_{ii}^{cat} - K_{ii}^{d} \tag{34}$$

$$\leq 0. \tag{35}$$

Using the definition of K_{ii}^{con} and $\mathcal{K}_{ii}$, we can rewrite (31)

$$\sum_{j \neq i} [K_{ij}^{con} + K_{ij}^{cat}] \leq -[K_{ii}^{con} + K_{ii}^{cat} - K_{ii}^{d}] \tag{36}$$

$$\sum_{j \neq i} \mathcal{K}_{ij} \leq -\mathcal{K}_{ii}. \tag{37}$$

Now, bounds on the eigenvalues of $\mathcal{K}$ are given by the Levy–Hadamard theorem (Bodewig, 1959), which states that for every eigenvalue λ of $\mathcal{K}$,

$$|\lambda - \mathcal{K}_{ii}| \leq \sum_{j \neq i} |\mathcal{K}_{ij}|.$$

Since the off-diagonal elements of $\mathcal{K}$ are non-negative, this can be written as

$$|\lambda - \mathcal{K}_{ii}| \leq \sum_{j \neq i} \mathcal{K}_{ij} \leq -\mathcal{K}_{ii} \tag{38}$$

where the second inequality follows from (37). Since $\mathcal{K}_{ii} < 0$, it follows that $Re(\lambda) \leq 0$, which proves (a).

To prove (b) we do a similar analysis, using the Levy–Hadamard theorem expressed in terms of a sum over the columns of $\mathcal{K}$. The constraint (b) can be written as

$$\sum_{j} K_{ji}^{cat} \leq K_{ii}^{d}.$$

By reasoning similar to that used in the proof of (a), this constraint leads to the relations

$$\mathcal{K}_{ii} \leq 0$$
$$\sum_{j \neq i} \mathcal{K}_{ji} \leq -\mathcal{K}_{ii}.$$

The Levy–Hadamard theorem applied to the columns of $\mathcal{K}$ states that

$$|\lambda - \mathcal{K}_{ii}| \leq \sum_{j \neq i} \mathcal{K}_{ji},$$

and the proof of (b) now follows as before. $\square$

Remark 2. (i) A special case of (a) shows that in a closed system the eigenvalues have non-negative real parts, for in that case $K_{ii}^{d} = 0$, $K_{ii}^{cat} = 0$ and $K_{jj}^{con} = -\sum_{i \neq j} K_{ij}^{con}$. In fact in that case it is known that the eigenvalues are all real as well if the system satisfies detailed balance (Gans, 1960; Wei and Prater, 1962). If the underlying graph is strongly connected, then there is exactly one zero eigenvalue (Othmer, 1979).

The theorem gives conditions for the stability of an arbitrary first-order system of reactions in terms of the specific rates of the reactions. The first sufficient condition for stability is easily understood, as it is expected that for stability of a system the specific rates of production for all species should be less than the specific rates of degradation for every component of the system. The second criterion, which requires that the rate at which any component catalyzes the formation of other species is less than its degradation rate, is less immediately obvious. This is however an important relationship that can be used

to guarantee stability of the mean of artificial transcriptional networks. In particular, once the inequality between the catalytic and degradation rates is satisfied, the system will be stable irrespective of the conversion reactions in the system, and may be used in the design of such networks. In the theory of branching processes condition b is equivalent to the assumption that the offspring distribution has mean less than or equal to 1.

Further information about the structure of the solution can be gotten from a spectral representation of $\mathcal{K}$. Hereafter we assume that $\mathcal{K}$ is semisimple (i.e., it has a complete set of eigenvectors), which is the generic case, and then it has the spectral representation

$$\mathcal{K} = \sum_i \lambda_i P_i \tag{39}$$

where λ_i is the ith eigenvalue of $\mathcal{K}$ and P_i is the associated projection onto the span of the eigenvectors associated with λ_i. Since $\mathcal{K}$ is assumed to be semisimple, they have the property that $\sum_i P_i = I$.

The projections have the representation

$$P_i = \sum_{p=1}^{m_i} \phi_p * \overline{\phi_p^*} \tag{40}$$

where m_i is the algebraic multiplicity of the ith eigenvalue, '$*$' represents the dyad product, and the ϕ's and ϕ^*'s are the corresponding eigenvectors and adjoint eigenvectors, defined via

$$\mathcal{K}\phi_p = \lambda_i \phi_p \qquad \text{for } p = 1, m_i \tag{41}$$

$$\mathcal{K}^T \phi_p^* = \overline{\lambda_i} \phi_p^* \qquad \text{for } p = 1, m_i. \tag{42}$$

They can be chosen to satisfy the orthogonality relations

$$\langle \phi_p^*, \phi_q \rangle = \sum_{i=1}^{m_i} (\overline{\phi_p^*})_i (\phi_q)_i = \delta_{pq}$$

and the projection of any vector is defined as

$$P_i u \equiv \left(\sum_{p=1}^{m_i} \phi_p * \overline{\phi_p^*} \right) u = \sum_{p=1}^{m_i} \langle \phi_p^*, u \rangle \phi_p.$$

When $\mathcal{K}$ is invertible the steady-state mean can be written as

$$M_1^s = -\left(\sum_i \frac{P_i}{\lambda_i} \right) k^s = -\sum_i \sum_{p=1}^{m_i} \frac{\langle \phi_p^*, k^s \rangle}{\lambda_i} \phi_p. \tag{43}$$

Thus in the nonsingular case the steady-state mean is the weighted sum of projections onto the ith eigenspace of $\mathcal{K}$, weighted by the corresponding eigenvalue. In particular, projections corresponding to eigenvalues of large modulus contribute less to the sum than those of small modulus. If $\mathcal{K}$ has a d-dimensional null space there are vectors $\{\eta_1, \eta_2, \ldots, \eta_d\}$ with the property that the functionals $\langle \eta, M \rangle$ are time-invariant, and this restricts the dynamics and steady-state solution to a lower-dimensional set. We leave the details of this case to the reader.

The transient solution of (28) is given by

$$
\begin{aligned}
M(t) &= e^{\mathcal{K}t} M(0) + \left(\int_0^t e^{\mathcal{K}(t-\tau)} d\tau \right) K^s \mathbf{1} \\
&= \left(\sum_i e^{\lambda_i t} P_i \right) M(0) + \left(\int_0^t \sum_i e^{\lambda_i (t-\tau)} P_i d\tau \right) K^s \mathbf{1} \\
&= \left(\sum_i e^{\lambda_i t} P_i \right) M(0) - \sum_i (1 - e^{\lambda_i t}) \frac{P_i}{\lambda_i} K^s \mathbf{1}.
\end{aligned}
\tag{44}
$$

Since $e^{\mathcal{K}t}$ is non-negative, i.e. all its entries are non-negative, the solution remains non-negative if $M(0)$ is non-negative. Therefore if there are no eigenvalues with a positive real part the solution converges to a non-negative steady state. If the real part of all eigenvalues is negative, the solution is globally asymptotically stable. However these conditions do not guarantee that the solution has strictly positive components, i.e. that it does lie on the boundary of the positive 'orthant' of R^s. Under stronger hypotheses one can guarantee that the solution lies in the interior of the orthant, as shown in the following theorem.

Theorem 3. *Suppose that the graph $\mathcal{G}$ associated with the reactions is strongly connected, and that the eigenvalues of $\mathcal{K}$ have negative real parts. Then $\mathcal{K}^{-1} < 0$ and if there is at least one species produced by a source then the solution* (30) *is component-wise positive.*

Proof. Notice that $\mathcal{K}$ is irreducible since the graph $\mathcal{G}$ is strongly connected. Write $\mathcal{K} = \mathcal{K}^+ - \kappa I$, where $\mathcal{K}^+ \geq 0$ and $\kappa > 0$. Observe that $\mathcal{K}^+$ is also irreducible. Let

$$
Q = -\mathcal{K} = \kappa I - \mathcal{K}^+.
\tag{45}
$$

Since the eigenvalues of $\mathcal{K}$ have negative real parts by hypothesis, the eigenvalues $\lambda(Q)$ of Q have positive real parts and

$$
\lambda(Q) = \kappa - \lambda(\mathcal{K}^+).
$$

Let $\rho(\mathcal{K}^+)$ be the spectral radius of $\mathcal{K}^+$; then by Perron–Frobenius theorem $\rho(\mathcal{K}^+)$ is a simple positive real eigenvalue of $\mathcal{K}^+$ and $\kappa - \rho(\mathcal{K}^+)$ is an eigenvalue of Q. Thus $\kappa - \rho(\mathcal{K}^+) > 0$, i.e., $\rho(\mathcal{K}^+)/\kappa < 1$. Since $\rho(\mathcal{K}^+)/\kappa < 1$, the series

$$
I + \frac{\mathcal{K}^+}{\kappa} + \left(\frac{\mathcal{K}^+}{\kappa} \right)^2 + \cdots
$$

converges, and so

$$
\left(I - \frac{\mathcal{K}^+}{\kappa} \right)^{-1} = \sum_{n=0}^{\infty} \left(\frac{\mathcal{K}^+}{\kappa} \right)^n.
$$

Since $(I + \frac{\mathcal{K}^+}{\kappa})$ is an $s \times s$ irreducible non-negative matrix with positive diagonal elements, it can be obtained that

$$
\left(I + \frac{\mathcal{K}^+}{\kappa} \right)^{s-1} > 0,
$$

which implies that

$$I + \frac{\mathcal{K}^+}{\kappa} + \left(\frac{\mathcal{K}^+}{\kappa}\right)^2 + \cdots + \left(\frac{\mathcal{K}^+}{\kappa}\right)^{s-1} > 0.$$

Therefore

$$\begin{aligned} Q^{-1} &= \kappa^{-1}\left(I - \frac{\mathcal{K}^+}{\kappa}\right)^{-1} \\ &= \frac{1}{\kappa}\left(I + \frac{\mathcal{K}^+}{\kappa} + \left(\frac{\mathcal{K}^+}{\kappa}\right)^2 + \cdots\right) \\ &> 0. \end{aligned}$$

Thus $Q^{-1} > 0$, so $\mathcal{K}^{-1} < 0$ and finally, if there is at least one species produced by a source, i.e., $k_i^s > 0$ for at least one i, then

$$M^s = -\mathcal{K}^{-1}k^s > 0.$$

Next we analyze the evolution of the second moment, and find that the conditions (a) and (b) in Theorem 1 will also guarantee the stability of the second moments.

3.2. Evolution of the second moment

One can show (cf. Appendix) that the evolution equation for the second moment (29) can be written in the form[5]

$$\frac{\mathrm{d}v}{\mathrm{d}t} = \mathcal{V}v + \gamma \tag{46}$$

where

$$v(t) = col(V(t)) \qquad \mathcal{V} = (\mathcal{K} \otimes I_n + I_n \otimes \mathcal{K}) \quad \text{and} \quad \gamma(t) = col(\Gamma(t) + \Gamma(t)^T).$$

Here the notation $col(A)$ denotes a vector of length s^2 whose elements are the columns of A stacked in order (cf. Appendix). By the definition of the tensor product

$$\begin{aligned} \mathcal{V}(\phi_i \otimes \phi_j) &= (\mathcal{K} \otimes I_n + I_n \otimes \mathcal{K})(\phi_i \otimes \phi_j) = \mathcal{K}\phi_i \otimes I_n\phi_j + I_n\phi_i \otimes \mathcal{K}\phi_j \\ &= (\lambda_i + \lambda_j)(\phi_i \otimes \phi_j), \end{aligned} \tag{47}$$

and therefore $\phi_i \otimes \phi_j$ are the eigenvectors of $\mathcal{K} \otimes I_n + I_n \otimes \mathcal{K}$ corresponding to $\lambda_i + \lambda_j$. It follows that if P_i and P_j are the projections associated with the ith and jth eigenvalues, respectively, then

$$P_{ij} \equiv P_i \otimes P_j = \left(\sum_{p=1}^{m_i} \phi_p * \overline{\phi_p^*}\right) \otimes \left(\sum_{p=1}^{m_j} \phi_p * \overline{\phi_p^*}\right) \tag{48}$$

---

[5] A reviewer has pointed out that the equation for second moments in the form (29) is solved formally i.e. converted to an integral equation, in Athreya and Ney (1972) and estimates of the growth rate derived from this. However, an explicit solution is not given.

is the projection associated with $\lambda_i + \lambda_j$. Consequently

$$\mathcal{V} = \sum_{i,j} (\lambda_i + \lambda_j) P_{ij}.$$

After some simplification one finds (cf. Appendix) that the solution for the second moment is

$$
\begin{aligned}
v &= e^{\mathcal{V}t} v(0) + \int_0^t e^{\mathcal{V}(t-\tau)} \gamma(\tau) d\tau \\
&= \sum_{i,j} e^{(\lambda_i+\lambda_j)t} P_{ij} v(0) + \sum_{i,j} \int_0^t e^{(\lambda_i+\lambda_j)(t-\tau)} P_{ij} \gamma(\tau) d\tau \\
&= \sum_{i,j} e^{(\lambda_i+\lambda_j)t} P_{ij} v(0) + \sum_{i,j} \left(\frac{e^{(\lambda_i+\lambda_j)t} - e^{\lambda_i t}}{\lambda_j} (P_i M_0 \otimes P_j) \right. \\
&\quad + \left(\frac{1}{\lambda_i \lambda_j} (e^{(\lambda_i+\lambda_j)t} - e^{\lambda_i t}) - \frac{1}{\lambda_i(\lambda_i + \lambda_j)} (e^{(\lambda_i+\lambda_j)t} - 1) \right) (P_i S \otimes P_j) \\
&\quad + \frac{e^{(\lambda_i+\lambda_j)t} - e^{\lambda_j t}}{\lambda_i} (P_i \otimes P_j M_0) + \left(\frac{1}{\lambda_i \lambda_j} (e^{(\lambda_i+\lambda_j)t} - e^{\lambda_j t}) \right. \\
&\quad \left. \left. - \frac{1}{\lambda_j(\lambda_i + \lambda_j)} (e^{(\lambda_i+\lambda_j)t} - 1) \right) (P_i \otimes P_j S) \right) col\, K^s \\
&\quad + \sum_{i,j} \int_0^t e^{(\lambda_i+\lambda_j)(t-\tau)} (P_i K^{cat} \otimes P_j + P_i \otimes P_j K^{cat}) col\, M_d d\tau, \quad (49)
\end{aligned}
$$

where $M_0 = [M(0)|M(0)|\cdots|M(0)]$, $S = [k^s|k^s|\cdots|k^s]$ and $M_d = diag\{M_1(t), M_2(t) \cdots M_n(t)\}$. From (49) and (47) it is clear that the time-dependent behavior of the covariance is governed by the set of eigenvalues $\lambda_i^{\mathcal{K}} + \lambda_j^{\mathcal{K}}$, and if the eigenvalues of $\mathcal{K}$ have negative real parts, so do the eigenvalues of $\mathcal{V}$. Thus the sufficient conditions for the stability of the mean derived in the preceding section also guarantee the stability of the second moment. For a closed system, one of the eigenvalues is zero and hence the longest characteristic time for the evolution of M and V will be identical. For an open system, the characteristic time for the evolution of the second moment will be twice as large as that for the evolution of M. Note however that $V_{ii} = (E[N_i(t)^2] - E[N_i(t)])$, and therefore this should not be interpreted to imply that the variance of the number of molecules of a particular species evolves twice as rapidly as the mean.

From the expression for v, one obtains the variance of the lth species (cf. Appendix) explicitly as

$$
\begin{aligned}
\sigma_l^2(t) &= -\sum_{i,j} e^{(\lambda_i+\lambda_j)t} (P_i \odot P_j M(0))_l \\
&\quad + \sum_i e^{\lambda_i t} \left[(P_i M(0))_l + \frac{(P_i k^s)_l}{\lambda_i} \right] - \sum_i \frac{1}{\lambda_i} (P_i k^s)_l
\end{aligned}
$$

$$+ \sum_{i,j} \sum_k D(l, k, j) \left[\sum_r \left((P_r M(0))_k + \frac{(P_r k^s)_k}{\lambda_r} \right) f_{r,i,j}(t) \right.$$

$$\left. - \frac{(P_r k^s)_k}{\lambda_r (\lambda_i + \lambda_j)} (e^{(\lambda_i + \lambda_j)t} - 1) \right],$$

where

$$D(l, k, j) = ((P_j)_{lk} C(l, k, i) + (P_i)_{lk} C(l, k, j)),$$

$$C(l, k, i) = \sum_{m=1}^n (P_i)_{lm} K_{mk}^{cat}$$

$$f_{r,i,j}(t) = \begin{cases} \dfrac{1}{\lambda_r - \lambda_i - \lambda_j} (e^{\lambda_r t} - e^{(\lambda_i + \lambda_j)t}) & \text{if } \lambda_r \neq \lambda_i + \lambda_j \\ t e^{(\lambda_i + \lambda_j)t} & \text{if } \lambda_r = \lambda_i + \lambda_j \end{cases}$$

and $\odot$ denotes element-wise multiplication. Using the expression for the mean of the lth reactant given in (44), we obtain the relationship between the mean and variance of every species in the reaction network.

$$\sigma_l^2(t) = - \sum_{i,j} e^{(\lambda_i + \lambda_j)t} \left(P_i \odot P_j M(0) \right)_l + M_l(t)$$

$$+ \sum_{i,j} \sum_k D(l, k, j) \left[\sum_r \left((P_r M(0))_k + \frac{(P_r k^s)_k}{\lambda_r} \right) f_{r,i,j}(t) \right.$$

$$\left. - \frac{(P_r k^s)_k}{\lambda_r (\lambda_i + \lambda_j)} (e^{(\lambda_i + \lambda_j)t} - 1) \right]. \tag{50}$$

From this expression one can calculate the evolution of the Fano factor $\mathcal{F} = \sigma_l^2 / M_l$ and $CV = \sigma_l / M_l$ for every species, and thus determine the effect of various network structures and reaction types on the noise. We study several examples in the following section.

4. The effect of network structure on the dynamics

The master equation for a system of first-order chemical reactions reflects three major characteristics of the system, (i) whether it is open or closed, (ii) the topology of the network of the chemical interactions, and (iii) whether or not the reacting mixture is spatially uniform, i.e., whether or not diffusive or other transport mechanisms play an important role. The effect of each of these factors on the distribution of species undergoing first-order catalytic and conversion reactions can be studied using the general results of the preceding section.

4.1. Open and closed conversion systems

In the context of first-order reaction dynamics, catalytic systems are necessarily open because they involve production from a source catalyzed by a time-dependent species

(cf. Table 1). Thus the comparison of open and closed systems can only be made for those in which there are no catalytic reactions. Therefore we compare open conversion networks in which there is at least one Type I reaction and one Type II reaction, with closed networks in which all reactions are type III and thus the total mass is constant.

The equation for the MGF for a system of stochastic conversion reactions can be solved analytically (Gans, 1960; Darvey and Staff, 1966), and for closed conversion networks it has been shown that the distribution is multinomial when the eigenvalues are distinct (Darvey and Staff, 1966). We derive the general result via a backward equation. For open systems we prove that the distribution is Poisson, and we demonstrate how the choice of the noise measure leads to differing conclusions about the noise in open and closed conversion networks.

4.1.1. Closed conversion systems

In a closed system of linear reactions the molecules independently execute a random walk through the states, where the state of a molecule is the property of being a molecule of type i; $i = 1, s$, where s is the number of species or states. We let $X(t)$ be the random process whose value at t is the state of a given molecule. Let P_{ij} be $Pr\{X(t) = j \mid X(0) = i\}$; then the matrix P of transition probabilities satisfies the backward Kolmogorov differential equation

$$\frac{dP(t)}{dt} = K^T P(t), \tag{51}$$

where $K = K^{con}$. Since $P(0) = I$, this has the solution

$$P(t) = e^{K^T t}.$$

If we denote by $p_i(t)$ the probability that the molecule is of type i at time t, then $p(t) = (p_1(t), \ldots, p_s(t))^T$ is the solution of

$$\frac{dp}{dt} = Kp$$

and therefore

$$p(t) = P(t)^T p(0) = e^{Kt} p(0). \tag{52}$$

Assume that the graph of the underlying network is strongly-connected; then K has exactly one zero eigenvalue and the equilibrium probability distribution is given by

$$\pi_i = \frac{\phi_{1,i}}{\sum_j \phi_{1,j}} \tag{53}$$

where ϕ_1 is the unique positive eigenvector of K corresponding to the zero eigenvalue.

Now if there are a total of N molecules in the system, the joint equilibrium distribution is the multinomial distribution

$$\pi(n_1, n_2, \ldots, n_s) = \frac{N!}{n_1! n_2! \cdots n_s!} \prod_{i=1}^{s} \pi_i^{n_i}, \tag{54}$$

since there are s classes to put the N molecules into, and the probability of the ith class is π_i. In particular, the number of individuals of the ith species is binomially distributed according to

$$\pi(n_i) = \binom{N}{n_i} \pi_i^{n_i} (1 - \pi_i)^{N-n_i}. \tag{55}$$

If in addition the joint distribution of molecular numbers of two species is multinomial initially, then the joint distribution at any time t is also multinomial. First suppose there are only two molecular species and N molecules initially. Then for $0 \leq m \leq N$, one can show that (see the Appendix for details)

$$
\begin{aligned}
Pr(N_1(t) &= m, N_2(t) = N - m) \\
&= \sum_{i=0}^{N} Pr(N_1(t) = m, N_2(t) = N - m | N_1(0) = i, N_2(0) = N - i) \\
&\qquad \cdot Pr(N_1(0) = i, N_2(0) = N - i) \\
&= \sum_{k=0}^{m} \sum_{i=k}^{N-m+k} \binom{i}{k} \left(p_t^k(1, 1) p_t^{i-k}(1, 2) \right) \binom{N-i}{m-k} \\
&\qquad \times \left(p_t^{m-k}(2, 1) p_t^{N-i-m+k}(2, 2) \right) \binom{N}{i} p_1(0)^i p_2(0)^{N-i} \\
&= \binom{N}{m} (p_1(t))^m (p_2(t))^{N-m}
\end{aligned}
$$

where the last step follows from (52). Thus $\{N_1(t), N_2(t)\}$ has a binomial distribution if initially it has a binomial distribution. By induction, it follows that the joint distribution of $\{N_1(t), N_2(t), \ldots, N_s(t)\}$ is multinomial if the initial distribution is multinomial. That is to say, if the initial joint distribution of molecular numbers of species is multinomial, i.e.,

$$P(n, 0) \equiv Pr(N_1(0) = n_1, \ldots, N_s(0) = n_s) = \frac{N!}{n_1! \cdots n_s!} (p_1(0))^{n_1} \cdots (p_s(0))^{n_s},$$

then the joint distribution at any time t is also multinomial and the density function is given by

$$P(N_1(t) = n_1, \ldots, N_s(t) = n_s) = \frac{N!}{n_1! \cdots n_s!} (p_1(t))^{n_1} \cdots (p_s(t))^{n_s},$$

where the probabilities $p_i(t)$, $1 \leq i \leq s$ are given by (52).

It follows from (55) that the mean and variance for the mth species at the steady state are given by

$$
\begin{aligned}
M_m &= N\pi_m = E[N_m] \\
\sigma^2(N_m) &= N\pi_m(1 - \pi_m) = E[N_m]\left(1 - \frac{E[N_m]}{N}\right).
\end{aligned} \tag{56}
$$

Notice that π_m is the steady-state fraction of the mth molecular species in a deterministic description, and since this is fixed by the reaction rates, the variance $\sigma^2(N_m)$ does not approach the mean even as $N \to \infty$. Thus the distribution is never Poisson in a closed

conversion network. The maximum variance is attained for that species for which π_m is closest to $1/2$. The Fano factor $\mathcal{F}_m = 1 - \pi_m$ is always less than 1 for all m, it is independent of N, and it is fixed entirely by the network topology and reaction rates. Thus it is an inappropriate measure of stochastic fluctuations in these networks.

In contrast,

$$CV_m = \frac{1}{\sqrt{N}}\sqrt{\frac{1-\pi_m}{\pi_m}} = \frac{\sqrt{1-\pi_m}}{\sqrt{M_m}} \tag{57}$$

varies as $1/\sqrt{N}$, while for fixed N it is monotone decreasing with π_m. In both cases the measures are smallest for the most-abundant species in the system.

4.1.2. Open conversion systems

Next we derive the steady-state distribution of species in an open conversion network. Here $\mathcal{K} = K^{con} - K^d$, and (25) can be written as

$$\frac{\partial G}{\partial t} - \sum_{j=1}^{s}\sum_{i=1}^{s}\mathcal{K}_{ij}(z_i-1)\frac{\partial G}{\partial z_j} = \sum_{i=1}^{s}K_i^s(z_i-1)G.$$

We find the solution of this PDE using the method of characteristics, analogous to the procedure outlined by Gans (1960). The result is that

$$G = \exp\left(\frac{1}{\det\Phi}\sum_{i}^{s}K_i^s\sum_{k}^{s}\frac{1}{\lambda_k}\Phi^{ki}(e^{\lambda_k t}-1)\sum_{j}^{s}\phi_{kj}(z_j-1)\right)$$
$$\cdot\prod_{i}\left(1+\frac{1}{\det\Phi}\sum_{k}\Phi^{ki}\sum_{j}\phi_{kj}(z_j-1)e^{\lambda_k t}\right)^{m_i},$$

where Φ is the matrix whose rows are the eigenvectors ϕ_k, Φ^{kj} is the cofactor of ϕ_{kj}, and m_i is the number of molecules of species i present initially. Therefore at steady state

$$G(z,\infty) = \exp\left(-\frac{1}{\det\Phi}\sum_{i}K_i^s\sum_{k}\frac{1}{\lambda_k}\Phi^{ki}\sum_{j}\phi_{kj}(z_j-1)\right),$$

and the steady-state value of the mean is given by

$$M_m = E(N_m) = \frac{\partial G(z,\infty)}{\partial z_m} = -\frac{1}{\det\Phi}\sum_{i}^{s}K_i^s\sum_{k}\frac{1}{\lambda_k}\Phi^{ki}\phi_{ki}.$$

To obtain the probability distribution, we differentiate k times to get

$$\frac{\partial^k G(z,\infty)}{\partial z_m^k} = \left(-\frac{1}{\det\Phi}\sum_{i}K_i^s\sum_{k}\frac{1}{\lambda_k}\Phi^{ki}\phi_{km}\right)^k$$
$$\times\exp\left(-\frac{1}{\det\Phi}\sum_{i}K_i^s\sum_{k}\frac{1}{\lambda_k}\Phi^{ki}\sum_{j}\phi_{kj}(z_j-1)\right)$$

and then

$$P_{N_m}(k, \infty) = \frac{1}{k!} \left. \frac{\partial^k G(z, \infty)}{\partial z_m^k} \right|_{z_m=0, z_j=1, j\neq m} = \frac{1}{k!} M_m^k e^{-M_m}.$$

This defines the density for a Poisson distribution, and as in any Poisson distribution, the variance is equal to the mean. This shows that the steady-state distribution of any species in an open first-order conversion network is a Poisson distribution, but this does not extend to the transient dynamics unless one assumes that the initial distribution of molecules is a Poisson distribution, rather than the Dirac distribution generally used.[6]

4.1.3. The noise during transients in conversion systems

The time-dependent variance in an open conversion system is given by (50) wherein $K^{cat} = 0$. In that case $C(l, k, i) = 0$ and (50) reduces to

$$\sigma_m^2(t) = -\sum_{i,j} e^{(\lambda_i + \lambda_j)t} (P_i \odot P_j M(0))_m + M_m(t).$$

Evidently $\lim_{t\to\infty} \sigma_m^2(t) = \lim_{t\to\infty} M_m(t)$ and if $M(0) = 0$, then $\sigma_m^2(t) = M_m(t)$. Thus $\mathcal{F}_m$ will always be 1 for open conversion systems when no species is present initially ($M(0) = 0$, which is by definition a Poisson initial distribution), and for more general initial conditions, $\mathcal{F}_m = 1$ for all species at the steady state. On the other hand, it is easy to see that the CV_m is always inversely proportional to the square root of the mean for all these situations where $\mathcal{F}_m$ is a constant. There is thus a clear difference in the qualitative estimation of the noise predicted by the two factors during transients as well as at steady state.

For closed systems, $\mathcal{F}_m = 1 - p_m$ at steady state, while CV_m is given by (57). When $M_m \sim N$ for some m, $p_m \sim 1$ and $\mathcal{F}_m$ and CV_m are both small. However when $N \gg M_n$, $\mathcal{F}_m \to 1$, while CV_m is inversely proportional to $\sqrt{M_m}$. We illustrate this and the transient behavior of the noise measures with an example of a closed three-component system shown in the following figure, where arrows indicate conversion reactions, and the symbols on the arrows indicate the specific rate constant associated with that reaction.
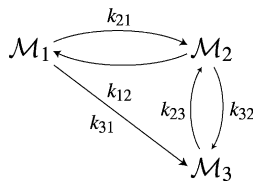
Using the procedure given earlier, one can find the means and covariances by solving

$$M(t)' = \mathcal{K}M(t) \tag{58}$$
$$V(t)' = \mathcal{K}V(t) + (\mathcal{K}V(t))^T, \tag{59}$$

---

[6] A reviewer has remarked that the distribution in an open system is Poisson at time t if the initial distribution is Poisson.

where

$$\mathcal{K} = \begin{bmatrix} -k_{21} - k_{31} & k_{12} & 0 \\ k_{21} & -k_{12} - k_{32} & k_{23} \\ k_{31} & k_{32} & -k_{23} \end{bmatrix}.$$

The eigenvalues of $\mathcal{K}$ are given by

$$\lambda_1 = 0$$

$$\lambda_{2,3} = \frac{1}{2} \Bigg[-k_{12} - k_{21} - k_{23} - k_{31} - k_{32}$$

$$\pm \sqrt{(k_{12} + k_{21} + k_{23} + k_{31} + k_{32})^2 - 4(k_{12}k_{23} + k_{21}k_{23} + k_{12}k_{31} + k_{23}k_{31} + k_{21}k_{32} + k_{31}k_{32})} \Bigg]$$

and because

$$(k_{12} + k_{21} + k_{23} + k_{31} + k_{32})^2$$
$$- 4(k_{12}k_{23} + k_{21}k_{23} + k_{12}k_{31} + k_{23}k_{31} + k_{21}k_{32} + k_{31}k_{32}) > 0,$$

the eigenvalues are distinct and $\mathcal{K}$ is semisimple. Thus,

$$M(t) = e^{\mathcal{K}t} M(0)$$
$$= \sum_{i=1}^{3} e^{\lambda_i t} (P_i M(0))_k$$

and furthermore

$$\sigma_k^2(t) = \sum_{i,j} e^{(\lambda_i + \lambda_j)t} \left(-P_i \odot P_j M(0) \right)_k + \sum_{i=1}^{3} e^{\lambda_i t} (P_i M(0))_k.$$

Assuming that only one species $\mathcal{M}_l$ for $l \in \{1, 2, 3\}$ is present initially with N molecules, we find that for $k = 1, 2, 3$,

$$M_k(t) = N \sum_{i=1}^{3} e^{\lambda_i t} (P_i)_{kl}$$

and so

$$\sigma_k^2(t) = \sum_{i,j=1}^{3} e^{(\lambda_i + \lambda_j)t} \left(-P_i \odot P_j M(0) \right)_k + M_k(t)$$

$$= M_k(t) - \sum_{i,j=1}^{3} e^{(\lambda_i + \lambda_j)t} (P_i)_{kl} (P_j)_{kl} N$$

$$= M_k(t) \left(1 - \frac{M_k(t)}{N} \right).$$

Fig. 1 shows the evolution of the Fano factor and CV for two components of the network. Since the sum of the number of molecules of all components at any instant is equal to the sum of the initial number, the mean and variance of the first component can be calculated from the mean and variance of the other two. It is clear that for these values of the rate parameters, which produce a steady state in which most of the total molecules exist
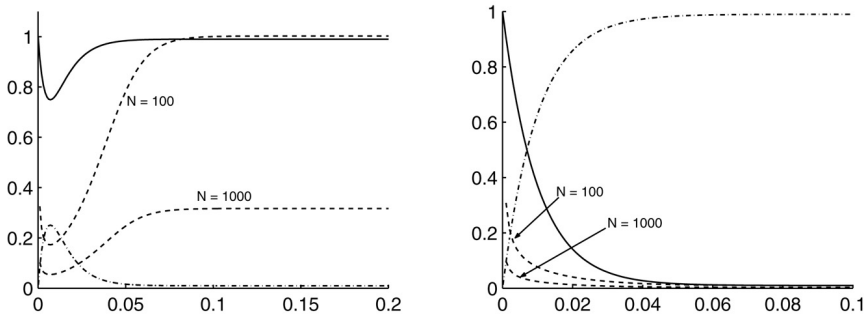
Fig. 1. Time-evolution (*x*-axis) of the mean scaled to the total number of molecules (-.-.-.), Fano factor (solid line) and CV (dashed line) for Species 2 (left) and Species 3 (right) for two values of $N = 100$ and $N = 1000$. Profiles for the fraction in each state and the Fano factor are independent of N. These plots illustrate the fact that use of the Fano factor leads to the conclusion that the noise of both species does not change on increasing N, whereas use of the CV predicts that increasing N reduces the noise in Species 2, whereas the noise in Species 3 remains essentially unchanged at steady state. The values of the rate constants are (in units of time^{-1}) $k_{21} = 100$, $k_{31} = 100$, $k_{12} = 1$, $k_{23} = 1$, $k_{32} = 100$ and all N molecules exist as $\mathcal{M}_1$ initially.

as Species 3 and the steady-state value of the mean for Species 2 is a small fraction of the total number of molecules, the Fano factor for species 2 is close to one at steady state, and does not change as the total number of molecules is increased 10-fold. For Species 3 the steady-state value of the mean is almost equal to the total number of molecules, and both the Fano factor and the CV predict that the steady-state noise will not change appreciably when the total number of molecules in the system is changed.

4.2. The effect of network topology on stochastic reaction networks

The structure of the graph of interactions among the network components influences the transient stochastic evolution of the network through its effect on the eigenvalues and eigenvectors of $\mathcal{K}$. However, we have shown that the steady-state distribution is always multinomial for a closed conversion system and Poisson for an open conversion system, and these conclusions are independent of the topology of the network. In both cases the distribution is completely characterized by the mean, and the effect of topology on the mean of the various species can be derived from an analysis of the structure of the vertex–edge incidence matrix $\mathcal{E}$ (Othmer, 1979). We illustrate the effects of a change in network topology on the steady-state distribution of a simple catalytic network, and on the evolution of the distribution to the steady state for general conversion or catalytic networks through a simple example.

We consider reaction networks that either form closed "loops" (Fig. 2(b)), with either a feedforward or feedback interaction, and compare the results with the evolution of the distribution in linear reaction networks of the form shown in Fig. 2(a). In general, for a conversion chain the ith species is formed from the $(i - 1)$st species and is converted to the $(i + 1)$st species; and in a catalytic chain the ith species is formed from the source with the reaction being catalyzed by the $(i - 1)$st species, and catalyzes the formation of the $(i + 1)$st species. All species undergo first-order degradation. The conversion chain is an
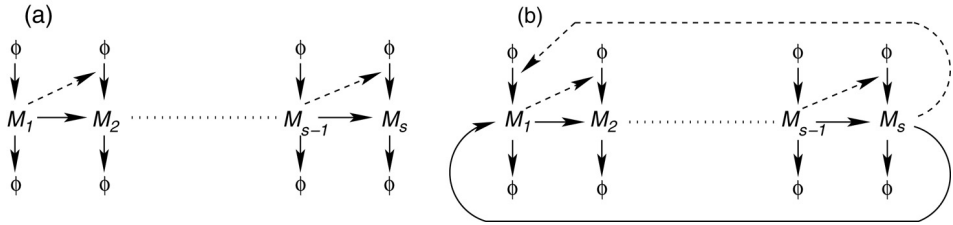
Fig. 2. Linear (a) and looped (b) reaction networks. The arrows represent the dependence of one species on the rate of formation of the species at the head of the arrow: the reactions may be conversion (solid lines) or catalytic (dashed lines).

Table 2
Mean and variance for a 2-component feedback network with either conversion or catalytic reactions

Species	$Mean_{conversion}$	$Mean_{catalytic}$	$Variance_{catalytic}$
$\mathcal{M}_1$	$\dfrac{ptk_1^s(k_{12}^{con}+k_2^d)}{(k_{12}^{con}+k_2^d)(k_{21}^{con}+k_1^d)-k_{21}^{con}k_{12}^{con}}$	$\dfrac{ptk_1^sk_2^d}{k_1^dk_2^d-k_{21}^{cat}k_{12}^{cat}}$	$\dfrac{ptk_1^dk_2^d(k_1^d+k_2^d)-k_{21}^{cat}k_{12}^{cat}(k_1^d-k_{12}^{cat})}{(k_1^dk_2^d-k_{21}^{cat}k_{12}^{cat})(k_1^d+k_2^d)}\times M_1$
$\mathcal{M}_2$	$\dfrac{ptk_1^sk_{21}^{con}}{(k_{12}^{con}+k_2^d)(k_{21}^{con}+k_1^d)-k_{21}^{con}k_{12}^{con}}$	$\dfrac{ptk_1^sk_{21}^{cat}}{k_1^dk_2^d-k_{21}^{cat}k_{12}^{cat}}$	$\dfrac{ptk_1^dk_2^d(k_1^d+k_2^d)+k_{21}^{cat}k_2^d(k_1^d-k_{12}^{cat})}{(k_1^dk_2^d-k_{21}^{cat}k_{12}^{cat})(k_1^d+k_2^d)}\times M_2$

example of an open conversion system, and we have shown that the distribution of the number of molecules of all components is a Poisson distribution. The analysis of the dynamics of the mean for feedback networks has been carried out earlier (Tyson and Othmer, 1978).

Consider an example of a feedback loop with $s = 2$, where only $\mathcal{M}_1$ has a non-negative rate of production from the source. Table 2 gives the steady-state mean of both species for conversion and catalytic reactions, and the steady-state variance for a network of catalytic reactions (the variance is equal to the mean in the conversion case).

It is clear that the presence or absence ($k_{12}^{con} = 0$) of a feedback loop in a conversion network changes the mean value of both $\mathcal{M}_1$ and $\mathcal{M}_2$, but the variance is always equal to the mean and the Fano factor is always equal to one. In contrast, for a catalytic network the presence of a feedback loop changes not only the mean, but also the steady-state fluctuations, as is most clear from the change in the Fano factor for $\mathcal{M}_1$ from a value of one in the absence of a feedback ($k_{12}^{cat} = 0$), to the value indicated in Table 2 for a network with feedback. We will further explore the differences in the variance of conversion and catalytic networks in the following subsection.

We can also compare the change in the evolution rates for these networks. For the reaction networks in Fig. 2, the matrix of reaction rates takes the form

$$\mathcal{K} = \begin{bmatrix} \mathcal{K}_{11} & 0 & 0 & \ldots & \mathcal{K}_{1s} \\ \mathcal{K}_{21} & \mathcal{K}_{22} & 0 & \ldots & 0 \\ 0 & \mathcal{K}_{32} & \mathcal{K}_{33} & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & \mathcal{K}_{ss-1} & \mathcal{K}_{ss} \end{bmatrix}.$$

For the linear network, the value of $\mathcal{K}_{1s}$ is zero and $\mathcal{K}_{ii}$ are its eigenvalues.

The characteristic equation $f(\lambda)$ of $\mathcal{K}$ is given by

$$f(\lambda) = [(\mathcal{K}_{11} - \lambda)(\mathcal{K}_{22} - \lambda) \cdots (\mathcal{K}_{ss} - \lambda)] + (-1)^{s+1} \mathcal{K}_{1s} \mathcal{K}_{21} \cdots \mathcal{K}_{ss-1}$$

$$= (-1)^s \left[\prod_{i=1}^{s} (\lambda - \mathcal{K}_{ii}) - \mathcal{K}_{1s} \mathcal{K}_{21} \cdots \mathcal{K}_{ss-1} \right] \equiv (-1)^s g(\lambda).$$

Consider a looped network that conforms to the stability condition (1) that the specific rate of production of each species is less than the specific rate of degradation. This requires that

$$\mathcal{K}_{21} < -\mathcal{K}_{11}, \quad \mathcal{K}_{32} < -\mathcal{K}_{22}, \ldots, \quad \mathcal{K}_{ss-1} < -\mathcal{K}_{s-1s-1}, \quad \mathcal{K}_{1s} < -\mathcal{K}_{ss}$$

and this implies that

$$\mathcal{K}_{1s} \mathcal{K}_{21} \mathcal{K}_{32} \cdots \mathcal{K}_{ss-1} < \prod_{i=1}^{s} (-\mathcal{K}_{ii}).$$

Hence, for all $\mathcal{K}_{kk}$, $1 \le k \le s$ we have

$$g(\mathcal{K}_{kk}) = -\mathcal{K}_{1s} \mathcal{K}_{21} \cdots \mathcal{K}_{ss-1} < 0$$

and

$$g(0) = \prod_{i=1}^{s} (-\mathcal{K}_{ii}) - \mathcal{K}_{1n} \mathcal{K}_{21} \mathcal{K}_{32} \cdots \mathcal{K}_{ss-1} > 0.$$

Therefore there is a real root λ_r of $g(\lambda) = 0$ such that $\max_i \mathcal{K}_{ii} < \lambda_r < 0$, which implies that in the loops there exists at least one real negative eigenvalue λ_r of $\mathcal{K}$ which is bigger than $\max_i \mathcal{K}_{ii}$.

Thus the presence of a loop leads to a slowing down of the evolution of the moments to the steady-state values. Notice that for a feed-forward looped network (i.e. Fig. 2(b) with the the arrow from $\mathcal{M}_1$ to $\mathcal{M}_s$ reversed), the corresponding reaction rate k_{s1} is still in the lower triangular part of $\mathcal{K}$ and the eigenvalues will be identical to those of the corresponding linear system.

4.3. Effect of the nature of the chemical reaction

Previous analyses of stochastic first-order reaction systems have been restricted to either all-conversion or all-catalytic systems, but they have not been compared directly. Let us consider reactions of the form shown in Fig. 2(a) in which the first component is produced from a source, and then is either converted to the second species (conversion chain), or catalyzes the formation of the second species (catalysis chain). As we know, the distribution of the number of molecules is a Poisson distribution for an open conversion chain, the Fano factor is one for all species at steady state, and the CV is inversely proportional to the square root of the mean. The results are quite different for a catalytic chain. In Fig. 3 we show the Fano factor and CV for a catalytic chain with rate parameters such that the means are identical 3(a) and different 3(b). It is clear that for a catalytic chain with identical means ($k_{i(i-1)}^{cat} = k_i^d$ $\forall i \ge 2$), the variance reaches a limiting value as the
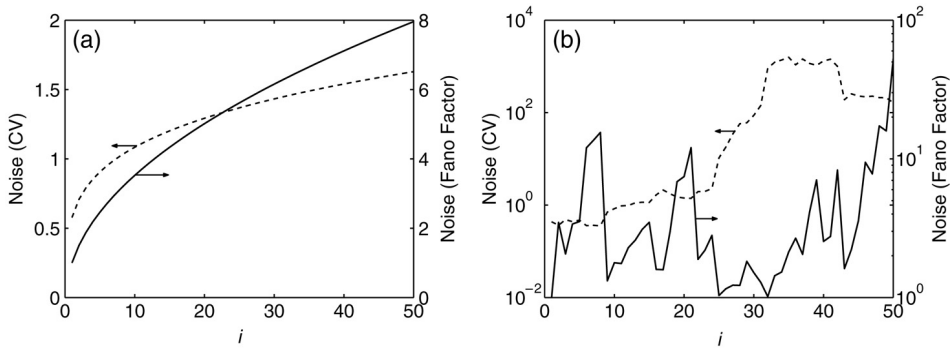
Fig. 3. Fano factor (solid line) and CV (dashed line) at steady state for the ith species in a catalytic chain with equal (a) and unequal (b) means. In (a) all species have a steady-state mean value of 3, resulting from setting all catalysis and degradation rate constants to 2 and the production rate for $\mathcal{M}_1$ to 2, random parameters for rate constants lead to mean values between 10^{-4} and 120 for species in (b). In both simulations, one molecule of each species was assumed to be present initially: but the results do not depend on the choice of the initial condition.

chain length is increased. This agrees with the results of Thattai and van Oudenaarden (2001), who studied catalytic chains with hyperbolic activation functions. However, it is important to note that variance may not saturate when the means are different, which is almost always the case in biological systems. When the means are the same, the Fano factor and CV give estimations of the noise that are qualitatively consistent, but when the means are different the use of the two measures of noise give different predictions of the effect of increasing the number of species in a catalytic chain on the level of fluctuations.

5. The effect of diffusion on stochastic reaction networks

Heretofore we have ignored the possible effects of spatial nonuniformity in the distribution of species, but when transport is solely by diffusion we can analyze a suitably-discretized spatial model within the preceding framework, because diffusion is a linear process. As the reader will appreciate after the development of the equations, the same analysis applies to compartmental models in which transport between compartments is linear in the concentration of the species. Here we restrict the discussion to a closed system containing an isotropic medium having no diffusive coupling between species, but the general formulation of the corresponding deterministic linear equations allows for bi-directional exchange with a reservoir, diffusive coupling, anisotropy in the transport (Othmer and Scriven, 1971).

For simplicity of exposition we begin with a symmetric domain that is subdivided into identical cubical compartments, and denote the number of molecules of species $\mathcal{M}_i$ present in the jth compartment as N_{ji}. Diffusive transport from compartment j to compartment k can be represented as the reaction $(N_{j,i}, N_{k,i}) \rightarrow (N_{j,i} - 1, N_{k,i} + 1)$ at a rate given by D_i/δ^2, where δ is the length scale of each compartment and D_i is the diffusion constant for species i (Nicolis and Prigogine, 1977; Stundzia and Lumsden, 1996).

Suppose that there are N_c cells and s reacting species. As before, $\mathcal{K}$ is the $s \times s$ reaction rate constant matrix $= K^{cat} + K^{con} - K^d$ defined for reactions without diffusion. Let Δ be the $N_c \times N_c$ structural matrix of the network of cells, whose elements Δ_{jk} for $j \neq k$ are 1 or 0, resp., if cell k is connected to cell j, or not, resp., and $-\Delta_{jj}$ is the number of cells connected to cell j. Δ encodes the connectivity of the network, and in the simple case treated here simply reflects the discretization of the Laplace operator on the domain in question. We define $\mathcal{D}$ as the $s \times s$ diagonal matrix of diffusion rates D_i / δ^2 for the species.

The deterministic evolution of the system is governed by

$$\frac{d\mathbf{c}}{dt} = \Omega \mathbf{c},$$

where $\mathbf{c}$ is the composition vector for all cells and the $n_s \cdot N_c \times n_s \cdot N_c$ matrix $\Omega \equiv I_{N_c} \otimes K + \Delta \otimes D$. It follows immediately that the equations for the means and second moments are

$$M(t)' = \Omega M(t) + K^s \cdot \mathbf{1}$$
$$V(t)' = \Omega V(t) + [\Omega V(t)]^T + C + C^T.$$

Here the matrix of means is defined as

$$M(t) = diag[M_{11}(t), M_{12}(t), \ldots M_{1s}(t),$$
$$M_{21}(t), \ldots M_{2s}(t), \ldots M_{N_c1}(t), \ldots M_{N_cs}(t)]$$

and K^s is a matrix containing the rate constants for production of each of the $N_c \times s$ species, given by

$$K^s = diag\left[k_{11}^s, k_{12}^s, \ldots k_{1s}^s, k_{21}^s, \ldots k_{2s}^s, \ldots k_{N_cs}^s\right],$$

where k_{ij}^s is the rate of inflow of species j from sources.

The covariance matrix $V(t)$ has matrix-valued elements V_{ij} defined as

$$V_{ij}(k, m) = \begin{cases} E[N_{ik}(t)^2] - E[N_{ik}(t)] & \text{if } i = j \text{ and } k = m \\ E[N_{ik}(t)N_{jm}(t)] & \text{otherwise.} \end{cases}$$

Finally, $C = W + K^s \mathbf{1} \cdot M(t)^T$ where W is a block-diagonal matrix, with each block defined as $W_{kk}(i, j) = K_{ij}^{cat} E[N_{kj}(t)]$.

It is clear from earlier sections that the evolution of the first and second moments are formally known once the eigenvalues and eigenvectors of Ω are known, and these are known from earlier work. Let α_k be an eigenvalue of the symmetric matrix Δ; then the eigenvalues λ_{kj} of Ω are solutions of the family of N_c sth-order determinantal equations

$$|\mathcal{K} + \alpha_k \mathcal{D} - \lambda_k \mathbf{I}_s| = 0 \tag{60}$$

(Othmer and Scriven, 1971). It is known that whenever $\mathcal{K}$ is not diagonal there may be counter-intuitive effects of diffusion on the eigenvalues determined by (60), and this lies at the heart of Turing's mechanism of pattern formation (Turing, 1952; Othmer, 1969).

The foregoing has been formulated for a regular discretization of a domain, but it holds whatever the topology of the connections between the compartments, as long as transport depends only on the concentration difference between compartments. The advantage of

the present formulation is that the effects of network structure in the reaction dynamics can be separated to the maximal extent possible from the topology of the compartmental connections, and effects due to spatial variations arise from the effect of the α_k's for different spatial modes. It should be noted here that in our treatment of reaction–diffusion systems there is an assumption that individual compartments are well-mixed, and for this to be valid, the size of each compartment should be related to the diffusion coefficient (and ultimately to the mean free path) of the diffusing species. However, the basis for the choice of compartment size is not clear for a system containing species with very different diffusion coefficients. On the one hand, if the size is based on the faster-diffusing species the well-mixed assumption may not be true for species that diffuse slowly, but if the size is computed using the slower-diffusing species, any solution algorithm will become computationally inefficient. More generally, the problem of how to treat wide disparities in time scales in the full master equation remains to be solved.

6. Discussion and conclusion

We have analyzed a general system of first-order reactions amongst s species that can be produced from sources, converted to other species or degraded, and catalyze the formation of other species. All previous stochastic analyses of systems of first-order reactions can be formulated as special cases of the general model studied here. We have derived explicit evolution equations for the mean and variance of the number of molecules of each reactant, and have solved them explicitly in a number of cases when the rate matrix $\mathcal{K}$ is semisimple. We find that the evolution of the second moments is completely determined by the spectral properties of $\mathcal{K}$ and the mean itself. To our knowledge this is the first report of a method to analytically compute the first two moments for an arbitrary first-order network comprising both conversion and catalytic reactions.

We have used the general framework to explore the effect of changes in the network topology on the distribution of the number of reactant molecules, and the difference between conversion and catalytic networks with the same topology. We prove that for an open system of first-order conversion reactions, the distribution of the number of molecules of every species is always a Poisson distribution. This is not the case for closed conversion systems, since the total number of molecules is constant. This result can be directly applied to the interpretation of experimental results on protein conformational-state transitions. The folding of a protein from its unfolded state to the fully folded ("native") state occurs through a series of intermediates. The first-order rate constants governing the reversible transitions from the unfolded state to the native state are calculated experimentally (Mayor et al., 2003). The above theory suggests the distributions that the experimental data may be fitted to in order to derive accurate estimates of the first-order transition rates. If ingress and egress through flow is allowed for the measurement device, each of the protein folding states will exhibit a Poisson distribution. If the system is closed, each state will be characterized by a multinomial distribution with a mean that is lower than the variance. A protein molecule undergoing conformational-state transitions is but one example of a set of chemical reactions that may be carried out in a closed or batch process, where there is no inflow or outflow of the chemical species, or in an open or continuous process, where

one or more chemicals are introduced at a constant rate, or removed from the system. The inflows may be modeled as production reactions of the form $\phi \rightarrow \mathcal{M}_i$ where species $\mathcal{M}_i$ is produced at a constant rate from a source. The outflows are modeled as degradation reactions that result in a depletion of species at a rate proportional to their concentration, with the specific rate constant corresponding to the dilution rate for the reactor. These equations exactly describe the addition and removal of species in microfluidic devices which may be modeled as continuous stirred-tank reactors. Our analysis of open and closed systems may be used to distinguish between the effects of batch-mode operations and continuous operations on the stochastic behavior of the chemical species undergoing the same set of chemical conversion reactions.

Two measures are used to estimate the stochastic fluctuations of reactant concentrations: the Fano factor and the coefficient of variance. We have shown that the equilibrium distribution of all components is a Poisson distribution for open conversion networks, in which case use of the Fano factor as the measure of noise leads to the prediction that the fluctuations of all components in an open conversion reaction system are identical. Thus a species that has a mean of 10 molecules will exhibit the same amount of noise as a species that has a mean concentration of 1M. This is clearly not correct, and the use of the CV as the measure of the noise will correctly predict that the noise is inversely proportional to the square root of the mean value. The only instance when the use of the Fano factor has a distinct advantage is when noise is defined as a deviation from the Poisson distribution. Our work is the first instance where the two measures have been compared theoretically, and we conclude that the only instance where either measure can be used to compare the noise of two species is when the mean values are identical. We have also shown that the use of the two measures leads to contradictory conclusions about the noise when the means are not identical.

We use the example of a linear reaction chain with and without feedback to demonstrate the effect of changes in the species interconnectivity on the dynamics of the evolution of the moments of the distribution. For the same interconnectivity, the nature of the distribution changes depending on whether the reactions are conversion reactions or catalytic reactions. It has been shown for a catalytic chain whose mean values are equal at steady state that the variance of the last species in the chain increases as the number of species in the chain increases (Thattai and van Oudenaarden, 2001), but ultimately saturates for long chains. We show with a counterexample that the ratio of the variance to the mean does not show this saturation behavior when the steady-state mean values of all the components are different.

The mathematical formulation that leads to a direct solution of the moment equations for a well-stirred system can be extended to arbitrary networks of well-mixed compartments that are coupled by diffusion. We demonstrate that the eigenvalues that govern the evolution in such distributed systems are solutions of a one-parameter family of modified kinetic matrices and thus one can formally display the solution for the first two moments in this case as well. However much remains to be done for this case to develop computationally-efficient algorithms.

We anticipate that the analytical framework presented here will be extended to the stochastic analysis of nonlinear reaction networks, and our analysis of first-order reaction network will lead to insights into the local linear behavior of such networks.

Acknowledgements

Appendix

An alternate approach to the evolution equation for the MGF[7]

To obtain the generating function for the first-order network, one may use the Kolmogorov backward equation instead of the master equation (24), which is generally used in the field of chemical reaction networks.

The Kolmogorov backward equation is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}E[f(N(t))] = E[\mathcal{A}f(N(t))],$$

where

$$\mathcal{A}f(n) = \sum_i k_i^s (\mathcal{S}_i^{+1} f(n) - f(n)) + \sum_{i,j} k_{ij}^{con} n_j (\mathcal{S}_i^{+1}\mathcal{S}_j^{-1} f(n) - f(n))$$

$$+ \sum_{i,j} k_{ij}^{cat} n_j (\mathcal{S}_i^{+1} f(n) - f(n)) + \sum_i k_i^d n_i (\mathcal{S}_i^{-1} f(n) - f(n)).$$

Notice that $G(z, t) = E[\prod z_i^{N_i(t)}]$, and therefore, taking $f(N(t)) = \prod z_i^{N_i(t)}$,

$$\frac{\mathrm{d}}{\mathrm{d}t}G(z, t) = \frac{\mathrm{d}}{\mathrm{d}t}E\left[\prod z_i^{N_i(t)}\right] = \sum_i k_i^s (z_i - 1) E\left[\prod z_i^{N_i(t)}\right]$$

$$+ \sum_{i,j} k_{ij}^{con} \left(\frac{z_i}{z_j} - 1\right) E\left[N_j(t) \prod z_i^{N_i(t)}\right]$$

$$+ \sum_{i,j} k_{ij}^{cat} (z_i - 1) E\left[N_j(t) \prod z_i^{N_i(t)}\right]$$

$$+ \sum_i k_i^d \left(\frac{1}{z_i} - 1\right) E\left[N_i(t) \prod z_i^{N_i(t)}\right].$$

Using the fact that

$$E\left[N_j(t) \prod z_i^{N_i(t)}\right] = z_j \frac{\partial}{\partial z_j} E\left[\prod z_i^{N_i(t)}\right],$$

we obtain Eq. (24).

---

[7] We thank a reviewer for pointing out this approach.

Tensor products and the column operation

Here we record a few basic facts about tensor products and the *col* operation that are used throughout.

Let $x = (x_1, x_2, \ldots, x_n)^T$ and $y = (y_1, y_2, \ldots, y_m)^T$. Then we define the tensor product of x and y as (cf. Othmer and Scriven (1971) and references therein)

$$x \otimes y = (x_1 y, x_1 y, \cdot, x_1 y)$$
$$= (x_1 y_1, x_1 y_2, \ldots, x_1 y_n, x_2 y_1, x_2, y_2, \ldots, x_2 y_n, \ldots, x_n y_m)^T.$$

For any $n \times n$ matrix $B = [b_{ij}]$ and an $m \times m$ matrix C, we define

$$(\mathbf{B} \otimes \mathbf{C}) = \begin{bmatrix} b_{11}C & b_{12}C & \ldots & b_{1n}C \\ b_{21}C & b_{22}C & \ldots & b_{2n}C \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1}C & b_{n2}C & \ldots & b_{nn}C \end{bmatrix}.$$

Let $B(i)$ be the ith column of a $n \times n$ matrix B. Then define the *column* operation as

$$col\, B = \begin{pmatrix} B(1) \\ B(2) \\ \vdots \\ B(n) \end{pmatrix}.$$

Then we have

$$col(BC) = \begin{pmatrix} BC(1) \\ BC(2) \\ \vdots \\ BC(n) \end{pmatrix} = \begin{bmatrix} B & & & \\ & B & & \\ & & \ddots & \\ & & & B \end{bmatrix} \begin{pmatrix} C(1) \\ C(2) \\ \vdots \\ C(n) \end{pmatrix} = (I_n \otimes B)col\, C.$$

Similarly

$$col(CB) = (B^T \otimes I_n)col\, C.$$

We can use these results to rewrite (29) as follows. We have

$$V(t)' = \mathcal{K}V(t) + (\mathcal{K}V(t))^T + \Gamma(t) + \Gamma^T(t)$$

and therefore

$$col(V(t)') = col(\mathcal{K}V(t)) + col((\mathcal{K}V(t))^T) + col(\Gamma(t) + \Gamma^T(t)).$$

We apply the above to the first term on the right-hand side by setting $B = \mathcal{K}$ and $C = V$, and to the second term by setting $B = \mathcal{K}^T$ and $C = V$, and noting that V is symmetric. This leads to (46).

Evolution of the moments

We consider the differential equation for the second moment

$$\frac{dv}{dt} = \mathcal{V}v + \gamma$$

where

$$v(t) = col(V(t)) \quad \mathcal{V} = (\mathcal{K} \otimes I_n + I_n \otimes \mathcal{K}) \quad \text{and} \quad \gamma(t) = col(\Gamma(t) + \Gamma(t)^T).$$

Since

$$[\Gamma(t)]_{ij} = K_{ij}^{cat} E[N_j(t)] + k_i^s E[N_j]$$

and K^s and M_d are diagonal matrices,

$$\begin{aligned}
\gamma(t) &= col(\Gamma(t) + \Gamma(t)^T) \\
&= col(K^{cat} M_d(t) + (K^{cat} M_d(t))^T) + col(K^s L(t) + (K^s L(t))^T) \\
&= (K^{cat} \otimes I_n + I_n \otimes K^{cat})col M_d + (L^T \otimes I_n + I_n \otimes L^T)col K^s.
\end{aligned}$$

Thus

$$\begin{aligned}
P_{ij}\gamma(t) &= P_{ij}col(K^{cat} M_d + (K^{cat} M_d)^T + K^s L + (K^s L)^T) \\
&= (P_i \otimes P_j)(K^{cat} \otimes I_n + I_n \otimes K^{cat})col M_d \\
&\quad + P_{ij}(L^T \otimes I_n + I_n \otimes L^T)col K^s \\
&= (P_i K^{cat} \otimes P_j + P_i \otimes P_j K^{cat})col M_d \\
&\quad + (P_i L^T \otimes P_j + P_i \otimes P_j L^T)col K^s.
\end{aligned}$$

We have that

$$\begin{aligned}
P_i M(t) &= P_i \left(\left(\sum_i e^{\lambda_i t} P_i \right) M(0) - \sum_i \frac{(1 - e^{\lambda_i t})}{\lambda_i} P_i k^s \right) \\
&= e^{\lambda_i t} P_i M(0) - \frac{(1 - e^{\lambda_i t})}{\lambda_i} P_i k^s
\end{aligned}$$

and

$$\begin{aligned}
P_i L^T &= P_i [M(t) | \cdots | M(t)] \\
&= [P_i M(t) | \cdots | P_i M(t)] \\
&= e^{\lambda_i t} P_i M_0 - \frac{(1 - e^{\lambda_i t})}{\lambda_i} P_i S,
\end{aligned}$$

where $M_0 = [M(0) | M(0) | \cdots | M(0)]$, and $S = [k^s | k^s | \cdots | k^s]$. Thus

$$\begin{aligned}
P_{ij}\gamma(t) &= (P_i K^{cat} \otimes P_j + P_i \otimes P_j K^{cat})col M_d \\
&\quad + \left[e^{\lambda_i t}(P_i M_0 \otimes P_j) - \frac{1 - e^{\lambda_i t}}{\lambda_i}(P_i S \otimes P_j) + e^{\lambda_j t}(P_i \otimes P_j M_0) \right. \\
&\quad \left. - \frac{1 - e^{\lambda_j t}}{\lambda_j}(P_i \otimes P_j S) \right] col K^s
\end{aligned}$$

and therefore

$$v = e^{\mathcal{V}t} v(0) + \int_0^t e^{\mathcal{V}(t-\tau)} \gamma(\tau) d\tau$$

$$= \sum_{i,j} e^{(\lambda_i+\lambda_j)t} P_{ij} v(0) + \int_0^t \sum_{i,j} e^{(\lambda_i+\lambda_j)(t-\tau)} P_{ij} \gamma(\tau) d\tau$$

$$= \sum_{i,j} e^{(\lambda_i+\lambda_j)t} P_{ij} v(0) + \sum_{i,j} \left(\frac{e^{(\lambda_i+\lambda_j)t} - e^{\lambda_i t}}{\lambda_j} (P_i M_0 \otimes P_j) \right.$$

$$+ \left(\frac{1}{\lambda_i \lambda_j} (e^{(\lambda_i+\lambda_j)t} - e^{\lambda_i t}) - \frac{1}{\lambda_i(\lambda_i+\lambda_j)} (e^{(\lambda_i+\lambda_j)t} - 1) \right) (P_i K^s \otimes P_j)$$

$$+ \frac{e^{(\lambda_i+\lambda_j)t} - e^{\lambda_j t}}{\lambda_i} (P_i \otimes P_j M_0) + \left(\frac{1}{\lambda_i \lambda_j} (e^{(\lambda_i+\lambda_j)t} - e^{\lambda_j t}) \right.$$

$$\left. - \frac{1}{\lambda_j(\lambda_i+\lambda_j)} (e^{(\lambda_i+\lambda_j)t} - 1) \right) (P_i \otimes P_j K^s) \bigg) col K^s$$

$$+ \sum_{i,j} \int_0^t e^{(\lambda_i+\lambda_j)(t-\tau)} (P_i K^{cat} \otimes P_j + P_i \otimes P_j K^{cat}) col M_d d\tau.$$

Using the fact that $col(ABC) = (C^T \otimes A) col B$, we obtain

$$V(t) = \sum_{i,j} e^{(\lambda_i+\lambda_j)t} [P_j V(0) P_i^T] + \sum_{i,j} \left[\left(\frac{e^{(\lambda_i+\lambda_j)t} - e^{\lambda_i t}}{\lambda_j} \right) [P_j K^s (P_i M_0)^T] \right.$$

$$+ \left(\frac{e^{(\lambda_i+\lambda_j)t} - e^{\lambda_i t}}{\lambda_i \lambda_j} - \frac{e^{(\lambda_i+\lambda_j)t} - 1}{\lambda_i(\lambda_i+\lambda_j)} \right) [P_j K^s (P_i S)^T]$$

$$+ \frac{e^{(\lambda_i+\lambda_j)t} - e^{\lambda_j t}}{\lambda_i} [P_j M_0 K^s P_i^T]$$

$$+ \left(\frac{e^{(\lambda_i+\lambda_j)t} - e^{\lambda_j t}}{\lambda_i \lambda_j} - \frac{e^{(\lambda_i+\lambda_j)t} - 1}{\lambda_i(\lambda_i+\lambda_j)} \right) [P_j S K^s P_i^T] \bigg]$$

$$+ \sum_{i,j} \int_0^t e^{(\lambda_i+\lambda_j)(t-\tau)} (P_j M_d (P_i K^{cat})^T + P_j K^{cat} M_d P_i^T) d\tau.$$

Thus

$$V_{ll}(t) = \sum_{i,j} e^{(\lambda_i+\lambda_j)t} [P_j V(0) P_i^T]_{ll} + \sum_{i,j} \left[\left(\frac{e^{(\lambda_i+\lambda_j)t} - e^{\lambda_i t}}{\lambda_j} \right) [P_j K^s (P_i M_0)^T]_{ll} \right.$$

$$+ \left(\frac{e^{(\lambda_i+\lambda_j)t} - e^{\lambda_i t}}{\lambda_i \lambda_j} - \frac{e^{(\lambda_i+\lambda_j)t} - 1}{\lambda_i(\lambda_i+\lambda_j)} \right) [P_j K^s (P_i S)^T]_{ll}$$

$$+ \frac{e^{(\lambda_i+\lambda_j)t} - e^{\lambda_j t}}{\lambda_i} [P_j M_0 K^s P_i^T]_{ll}$$

$$+ \left(\frac{e^{(\lambda_i + \lambda_j)t} - e^{\lambda_j t}}{\lambda_i \lambda_j} - \frac{e^{(\lambda_i + \lambda_j)t} - 1}{\lambda_i (\lambda_i + \lambda_j)} \right) [P_j S K^s P_i^T]_{ll} \Bigg]$$

$$+ \sum_{i,j} \int_0^t e^{(\lambda_i + \lambda_j)(t - \tau)} \left(P_j M_d (P_i K^{cat})^T + P_j K^{cat} M_d P_i^T \right)_{ll} \mathrm{d}\tau.$$

We have

$$P_j M_d = \begin{bmatrix} P_{j11} M_1 & P_{j12} M_2 & \cdots & P_{j1n} M_n \\ P_{j21} M_1 & P_{j22} M_2 & \cdots & P_{j2n} M_n \\ \vdots & \ddots & \ddots & \vdots \\ P_{jn1} M_1 & P_{jn2} M_2 & \cdots & P_{jnn} M_n \end{bmatrix}$$

$$(P_i K^{cat})^T = \begin{bmatrix} \sum_j P_{i1j} K_{j1}^{cat} & \sum_j P_{i2j} K_{j1}^{cat} & \cdots & \sum_j P_{inj} K_{j1}^{cat} \\ \sum_j P_{i1j} K_{j2}^{cat} & \sum_j P_{i2j} K_{j2}^{cat} & \cdots & \sum_j P_{inj} K_{j2}^{cat} \\ \vdots & \ddots & \ddots & \vdots \\ \sum_j P_{i1j} K_{jn}^{cat} & \sum_j P_{i2j} K_{jn}^{cat} & \cdots & \sum_j P_{inj} K_{jn}^{cat} \end{bmatrix}.$$

Thus

$$[P_j M_d (P_i K^{cat})^T]_{ll} = P_{jl1} M_1 \left(\sum_j P_{ilj} K_{j1}^{cat} \right)$$

$$+ P_{jl2} M_2 \left(\sum_j P_{ilj} K_{j2}^{cat} \right) + \cdots + P_{jln} M_n \left(\sum_j P_{ilj} K_{jn}^{cat} \right)$$

$$= \sum_k P_{jlk} M_k C(i, k, l)$$

where

$$C(i, k, l) = \sum_j P_{ilj} K_{jk}^{cat}.$$

Similarly

$$[P_j K^{cat} M_d P_i^T]_{ll} = \sum_m P_{ilm} M_m C(j, k, l).$$

Thus

$$\int_0^t e^{(\lambda_i + \lambda_j)(t - \tau)} (P_j M_d (P_i K^{cat})^T + P_j K^{cat} M_d P_i^T)_{ll} \mathrm{d}\tau$$

$$= e^{(\lambda_i + \lambda_j)t} \int_0^t e^{-(\lambda_i + \lambda_j)\tau} \left(\sum_k (P_j)_{lk} M_k C(i, k, l) + \sum_k (P_i)_{lk} M_k C(j, k, l) \right) \mathrm{d}\tau$$

$$= e^{(\lambda_i + \lambda_j)t} \int_0^t e^{-(\lambda_i + \lambda_j)\tau} \left[\sum_k (P_j)_{lk} \left(\sum_r e^{\lambda_r \tau} (P_r M(0))_k \right. \right.$$

$$- \sum_r \frac{1 - e^{\lambda_r \tau}}{\lambda_r} (P_r k^s)_k \Bigg) C(i, k, l)$$

$$+ \sum_k (P_i)_{lk} \left(\sum_r e^{\lambda_r \tau} (P_r M(0))_k - \sum_r \frac{1 - e^{\lambda_r \tau}}{\lambda_r} (P_r k^s)_k \right) C(j, k, l)) \Bigg] d\tau$$

$$= e^{(\lambda_i + \lambda_j)t} \sum_k (P_j)_{lk} \sum_r C(i, k, l) \left[(P_r M(0))_k \int_0^t e^{(\lambda_r - \lambda_i - \lambda_j)\tau} d\tau - \frac{(P_r k^s)_k}{\lambda_r} \right.$$

$$\times \left. \left(\int_0^t e^{-(\lambda_i + \lambda_j)\tau} d\tau - \int_0^t e^{(\lambda_r - \lambda_i - \lambda_j)\tau} d\tau \right) \right]$$

$$+ e^{(\lambda_i + \lambda_j)t} \sum_k (P_i)_{lk} \sum_r C(j, k, l) \left[(P_r M(0))_k \int_0^t e^{(\lambda_r - \lambda_i - \lambda_j)\tau} d\tau \right.$$

$$\left. - \frac{(P_r k^s)_k}{\lambda_r} \left(\int_0^t e^{-(\lambda_i + \lambda_j)\tau} d\tau - \int_0^t e^{(\lambda_r - \lambda_i - \lambda_j)\tau} d\tau \right) \right]$$

$$= \sum_k ((P_j)_{lk} C(i, k, l) + (P_i)_{lk} C(j, k, l))$$

$$\times \left[\sum_{r, \lambda_r \neq \lambda_i + \lambda_j} (P_r M(0))_k \frac{1}{\lambda_r - \lambda_i - \lambda_j} (e^{\lambda_r t} - e^{(\lambda_i + \lambda_j)t}) \right.$$

$$- \frac{(P_r k^s)_k}{\lambda_r} \left(\frac{1}{\lambda_i + \lambda_j} (e^{(\lambda_i + \lambda_j)t} - 1) - \frac{1}{\lambda_r - \lambda_i - \lambda_j} (e^{\lambda_r t} - e^{(\lambda_i + \lambda_j)t}) \right)$$

$$+ \sum_{r, \lambda_r = \lambda_i + \lambda_j} (P_r M(0))_k t e^{(\lambda_i + \lambda_j)t}$$

$$\left. - \frac{(P_r k^s)_k}{\lambda_r} \left(\frac{1}{\lambda_i + \lambda_j} (e^{(\lambda_i + \lambda_j)t} - 1) - t e^{(\lambda_i + \lambda_j)t} \right) \right],$$

where

$$C(i, k, l) = \sum_{m=1}^n P_{ilm} K_{mk}^{cat}.$$

Thus the (l, l) component of V is

$$V_{ll}(t) = \sum_{i,j} e^{(\lambda_i + \lambda_j)t} [P_j V(0) P_i^T]_{ll} + \sum_{i,j} \left[\left(\frac{e^{(\lambda_i + \lambda_j)t} - e^{\lambda_i t}}{\lambda_j} \right) [P_j K^s (P_i M_0)^T]_{ll} \right.$$

$$+ \left(\frac{e^{(\lambda_i + \lambda_j)t} - e^{\lambda_i t}}{\lambda_i \lambda_j} - \frac{e^{(\lambda_i + \lambda_j)t} - 1}{\lambda_i (\lambda_i + \lambda_j)} \right) [P_j K^s (P_i S)^T]_{ll}$$

$$+ \frac{e^{(\lambda_i + \lambda_j)t} - e^{\lambda_j t}}{\lambda_i} [P_j M_0 K^s P_i^T]_{ll}$$

$$+ \left(\frac{e^{(\lambda_i + \lambda_j)t} - e^{\lambda_j t}}{\lambda_i \lambda_j} - \frac{e^{(\lambda_i + \lambda_j)t} - 1}{\lambda_i (\lambda_i + \lambda_j)} \right) [P_j S K^s P_i^T]_{ll} \bigg]$$

$$+ \sum_{i,j} \sum_k ((P_j)_{lk} C(i,k,l) + (P_i)_{lk} C(j,k,l))$$

$$\times \Bigg[\sum_{r, \lambda_r \neq \lambda_i + \lambda_j} (P_r M(0))_k \frac{1}{\lambda_r - \lambda_i - \lambda_j} (e^{\lambda_r t} - e^{(\lambda_i + \lambda_j)t})$$

$$- \frac{(P_r k^s)_k}{\lambda_r} \left(\frac{1}{\lambda_i + \lambda_j} (e^{(\lambda_i + \lambda_j)t} - 1) - \frac{1}{\lambda_r - \lambda_i - \lambda_j} (e^{\lambda_r t} - e^{(\lambda_i + \lambda_j)t}) \right)$$

$$+ \sum_{r, \lambda_r = \lambda_i + \lambda_j} (P_r M(0))_k t e^{(\lambda_i + \lambda_j)t}$$

$$- \frac{(P_r k^s)_k}{\lambda_r} \left(\frac{1}{\lambda_i + \lambda_j} (e^{(\lambda_i + \lambda_j)t} - 1) - t e^{(\lambda_i + \lambda_j)t} \right) \Bigg].$$

Note that

$$E[N_l(t)] = [M(t)]_l = \sum_i e^{\lambda_i t} (P_i M(0))_l - \sum_i \frac{1 - e^{\lambda_i t}}{\lambda_i} (P_i k^s)_l$$

and that

$$E[N_l(t)]^2 = ([M(t)]_l)^2 = \sum_{i,j} e^{(\lambda_i + \lambda_j)t} (P_i M(0))_l (P_j M(0))_l$$

$$+ \sum_{i,j} \frac{(1 - e^{\lambda_i t})(1 - e^{\lambda_j t})}{\lambda_i \lambda_j} (P_i k^s)_l (P_j k^s)_l$$

$$- 2 \sum_{i,j} \frac{e^{\lambda_i t} (1 - e^{\lambda_j t})}{\lambda_j} (P_i M(0))_l (P_j k^s)_l.$$

Therefore we have the variance of lth species

$$\sigma_l^2(t) = V_{ll}(t) + [M(t)]_l - ([M(t)]_l)^2$$

$$= \sum_{i,j} e^{(\lambda_i + \lambda_j)t} (P_j V(0) P_i^T)_{ll} + \sum_{i,j} \Bigg[\left(\frac{e^{(\lambda_i + \lambda_j)t} - e^{\lambda_i t}}{\lambda_j} \right) [P_j K^s (P_i M_0)^T]_{ll}$$

$$+ \left(\frac{e^{(\lambda_i + \lambda_j)t} - e^{\lambda_i t}}{\lambda_i \lambda_j} - \frac{e^{(\lambda_i + \lambda_j)t} - 1}{\lambda_i (\lambda_i + \lambda_j)} \right) [P_j K^s (P_i S)^T]_{ll}$$

$$+ \frac{e^{(\lambda_i + \lambda_j)t} - e^{\lambda_j t}}{\lambda_i} [P_j M_0 K^s P_i^T]_{ll} + \frac{e^{(\lambda_i + \lambda_j)t} - e^{\lambda_j t}}{\lambda_i \lambda_j}$$

$$- \frac{e^{(\lambda_i + \lambda_j)t} - 1}{\lambda_j (\lambda_i + \lambda_j)} [P_j S K^s P_i^T]_{ll} \Bigg]$$

$$+ \sum_{i,j} \sum_k ((P_j)_{lk} C(i,k,l) + (P_i)_{lk} C(j,k,l))$$

$$\times \left[\sum_{r, \lambda_r \neq \lambda_i + \lambda_j} (P_r M(0))_k \frac{1}{\lambda_r - \lambda_i - \lambda_j} (e^{\lambda_r t} - e^{(\lambda_i + \lambda_j)t}) \right.$$

$$- \frac{(P_r k^s)_k}{\lambda_r} \left(\frac{1}{\lambda_i + \lambda_j} (e^{(\lambda_i + \lambda_j)t} - 1) - \frac{1}{\lambda_r - \lambda_i - \lambda_j} (e^{\lambda_r t} - e^{(\lambda_i + \lambda_j)t}) \right)$$

$$+ \sum_{r, \lambda_r = \lambda_i + \lambda_j} (P_r M(0))_k t e^{(\lambda_i + \lambda_j)t} - \frac{(P_r k^s)_k}{\lambda_r}$$

$$\times \left. \left(\frac{1}{\lambda_i + \lambda_j} (e^{(\lambda_i + \lambda_j)t} - 1) - t e^{(\lambda_i + \lambda_j)t} \right) \right]$$

$$- \sum_{i,j} e^{(\lambda_i + \lambda_j)t} (P_i M(0))_l (P_j M(0))_l$$

$$- \sum_{i,j} \frac{(1 - e^{\lambda_i t})(1 - e^{\lambda_j t})}{\lambda_i \lambda_j} (P_i k^s)_l (P_j k^s)_l$$

$$+ 2 \sum_{i,j} \frac{e^{\lambda_i t} (1 - e^{\lambda_j t})}{\lambda_j} (P_i M(0))_l (P_j k^s)_l.$$

After some computation we get

$$\sigma_l^2(t) = - \sum_{i,j} e^{(\lambda_i + \lambda_j)t} (P_i \odot P_j M(0))_l + \sum_i e^{\lambda_i t} \left[(P_i M(0))_l + \frac{(P_i k^s)_l}{\lambda_i} \right]$$

$$- \sum_i \frac{1}{\lambda_i} (P_i k^s)_l$$

$$+ \sum_{i,j} \sum_k D(i,j,k,l) \left[\sum_r \left((P_r M(0))_k + \frac{(P_r k^s)_k}{\lambda_r} \right) f_{r,i,j}(t) \right.$$

$$\left. - \frac{(P_r k^s)_k}{\lambda_r (\lambda_i + \lambda_j)} (e^{(\lambda_i + \lambda_j)t} - 1) \right],$$

where

$$C(i,k,l) = \sum_{m=1}^n (P_i)_{lm} K_{mk}^{cat},$$

$$D(i,j,k,l) = ((P_j)_{lk} C(i,k,l) + (P_i)_{lk} C(j,k,l)),$$

$$f_{r,i,j}(t) = \begin{cases} \dfrac{1}{\lambda_r - \lambda_i - \lambda_j} (e^{\lambda_r t} - e^{(\lambda_i + \lambda_j)t}) & \text{if } \lambda_r \neq \lambda_i + \lambda_j \\ t e^{(\lambda_i + \lambda_j)t} & \text{if } \lambda_r = \lambda_i + \lambda_j, \end{cases}$$

and '$\odot$' is componentwise matrix multiplication.

If there is no catalysis in the system, then

$$D(i, j, k, l) = 0$$

and we have

$$\sigma_l^2(t) = -\sum_{i,j} e^{(\lambda_i + \lambda_j)t} (P_i \odot P_j M(0))_l$$

$$+ \sum_i e^{\lambda_i t} \left[(P_i M(0))_l + \frac{(P_i k^s)_l}{\lambda_i} \right] - \sum_i \frac{1}{\lambda_i} (P_i k^s)_l.$$

Details for the proof of the binomial distribution

$$P(N_1(t) = m, N_2(t) = N - m)$$

$$= \sum_{i=0}^{N} P(N_1(t) = m, N_2(t) = N - m \mid N_1(0) = i,$$

$$N_2(0) = N - i) \cdot P(N_1(0) = i, N_2(0) = N - i)$$

$$= \sum_{k=0}^{m} \sum_{i=k}^{N-m+k} \binom{i}{k} (p_t^k(1, 1) p_t^{i-k}(1, 2)) \binom{N-i}{m-k}$$

$$(p_t^{m-k}(2, 1) p_t^{N-i-m+k}(2, 2)) \binom{N}{i} p_1(0)^i p_2(0)^{N-i}$$

$$= \sum_{k=0}^{m} \sum_{i=k}^{N-m+k} \binom{i}{k} \binom{N-i}{m-k} \binom{N}{i}$$

$$\cdot (p_1(0) p_t(1, 1))^k (p_1(0) p_t(1, 2))^{i-k} (p_2(0) p_t(2, 1))^{m-k}$$

$$\times (p_2(0) p_t(2, 2))^{N-i-m+k}$$

$$= \sum_{k=0}^{m} \sum_{l=0}^{N-m} \binom{k+l}{k} \binom{N-(k+l)}{m-k} \binom{N}{k+l}$$

$$\cdot (p_1(0) p_t(1, 1))^k (p_1(0) p_t(1, 2))^l (p_2(0) p_t(2, 1))^{m-k}$$

$$\times (p_2(0) p_t(2, 2))^{N-m-l}, \text{ by letting } l = i - k$$

$$= \sum_{k=0}^{m} \sum_{l=0}^{N-m} \binom{N}{m} \binom{m}{k} \binom{N-m}{l} (p_1(0) p_t(1, 1))^k (p_1(0) p_t(1, 2))^l$$

$$\times (p_2(0) p_t(2, 1))^{m-k} (p_2(0) p_t(2, 2))^{N-m-l}$$

$$= \binom{N}{m} \left(\sum_{k=0}^{m} \binom{m}{k} (p_1(0) p_t(1, 1))^k (p_2(0) p_t(2, 1))^{m-k} \right)$$

$$\cdot \left(\sum_{l=0}^{N-m} \binom{N-m}{l} (p_1(0) p_t(1, 2))^l (p_2(0) p_t(2, 2))^{N-m-l} \right)$$

$$= \binom{N}{m} (p_1(0)p_t(1, 1) + p_2(0)p_t(2, 1))^m$$

$$\times (p_1(0)p_t(1, 2) + p_2(0)p_t(2, 2))^{N-m}$$

$$= \binom{N}{m} (p_1(t))^m (p_2(t))^{N-m}, \text{ by } (52)$$

where we used the fact that

$$\binom{k+l}{k}\binom{N-(k+l)}{m-k}\binom{N}{k+l}$$

$$= \frac{(k+l)!}{k!l!} \frac{(N-(k+l))!}{(m-k)!(N-l-m)!} \frac{N!}{(k+l)!(N-(k+l))!}$$

$$= \frac{N!}{m!(N-m)!} \frac{m!}{(m-k)!k!} \frac{(N-m)!}{l!(N-m-l)!}$$

$$= \binom{N}{m}\binom{m}{k}\binom{N-m}{l}.$$

Queuing theory and chemical reaction networks

There are formal similarities between chemical reaction networks and queuing networks (Arazi et al., 2004) that can be used to translate results between the two contexts. For example, the following shows that an open conversion network is equivalent to an $M/M/\infty$ queuing network.

Consider an irreducible or strongly connected network consisting of $M/M/\infty$ queues with s stations. The notation $M/M/\infty$ means Poisson arrivals, exponential service time and an infinite number of servers. The scheme of this queuing network is as follows:
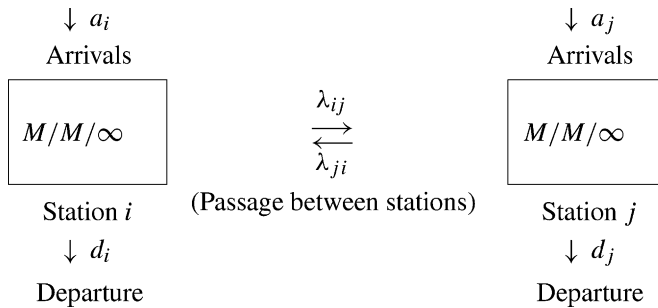
- Each station has $M/M/\infty$ queuing scheme.
- a_i = rate of arrivals from outside the network into station i.
- When station i has n occupants, then individuals may depart the system at rate $d_i n$.
- A customer leaving station i goes to station j with probability λ_{ij}, for $i \neq j$.

The connection between the network of $M/M/\infty$ queue and the open conversion network is as follows.

Network of $M/M/\infty$ queues	Open conversion network
Station	Species
Customer	Molecule
Number of customers in station	Number of molecules of species
Arrival of a customer from outside the system	Production of a molecule from source
Departure out of system	Degradation
Transition from ith to jth station	Conversion from ith species to jth species

From this table we can see that $N_i(t)$, the number of customers in the ith station at time t, corresponds to the number of molecules of the ith species at time t, and the rate a_i can be considered as k_i^s, d_i as k_i^d and λ_{ij} as K_{ji}^{con}.

It can be shown that in a network of $M/M/\infty$ queues the stationary distribution is Poisson. Furthermore, the distributions of $N_1(t), \ldots, N_s(t)$ for each time t are independent Poisson if the system is empty initially (Durrett, 1999).

References

Arazi, A., Ben-Jacob, E., Yechiali, U., 2004. Bridging genetic networks and queuing theory. Physica A 332, 585–616.

Athreya, K., Ney, P., 1972. Branching Processes. Springer-Verlag.

Austin, R.H., Beeson, K.W., Eisenstein, L., Frauenfelder, H., Gunsalus, I.C., 1975. Dynamics of ligand binding to myoglobin. Biochemistry 14 (24), 5355–5373.

Bartholomay, A.F., 1958. Stochastic models for chemical reactions: I. theory of the unimolecular reaction process. Math. Biophys. 20, 175–190.

Bartholomay, A.F., 1959. Stochastic models for chemical reactions: II. the unimolecular rate constant. Math. Biophys. 21, 363–373.

Blake, W.J., Kaern, M., Cantor, C.R., Collins, J.J., 2003. Noise in eukaryotic gene expression. Nature 422 (6932), 633–637.

Bodewig, E., 1959. Matrix Calculus. Interscience Publishers, Inc., New York.

Bokinsky, G., Rueda, D., Misra, V.K., Rhodes, M.M., Gordus, A., Babcock, H.P., Walter, N.G., Zhuang, X., 2003. Single-molecule transition-state analysis of RNA folding. Proc. Natl. Acad. Sci. USA 100 (16), 9302–9307.

Brown, F.L.H., 2003. Single-molecule kinetics with time-dependent rates: a generating function approach. Phys. Rev. Lett. 90 (2), 028302.

Darvey, I.G., Ninham, B.W., Staff, P.J., 1966. Stochastic models for second-order chemical reaction kinetics, the equilibrium state. J. Chem. Phys. 45 (6), 2145–2155.

Darvey, I.G., Staff, P.J., 1966. Stochastic approach to first-order chemical reaction kinetics. J. Chem. Phys. 44 (3), 990.

Delbruck, M., 1940. Statistical fluctuations in autocatalytic reactions. J. Chem. Phys. 8, 120–124.

Durrett, R., 1999. Essentials of Stochastic Processes. Springer-Verlag, New York.

Elowitz, M.B., Levine, A.J., Siggia, E.D., Swain, P.S., 2002. Stochastic gene expression in a single cell. Science 297 (5584), 1183–1186.

Fredrickson, A.G., 1966. Stochastic triangular reactions. Chem. Engg. Sci. 21, 687–691.

Gani, J., 1965. Stochastic models for bacteriophage. J. Appl. Prob. 2, 225–268.

Gans, P.J., 1960. Open first-order stochastic processes. J. Chem. Phys. 33 (3), 691.

Gardiner, C.W., 1983. Handbook of Stochastic Methods. Springer, Berlin, Heidelberg.

Gillespie, D.T., 1976. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. J. Comput. Phys. 22, 403–434.

Harris, T., 1963. The Theory of Branching Processes. Springer-Verlag, Berlin.

Horn, F., Jackson, R., 1972. General mass action kinetics. Arch. Ration. Mech. Anal. 48, 81.

Iorio, E.E.D., Hiltpold, U.R., Filipovic, D., Winterhalter, K.H., Gratton, E., Vitrano, E., Cupane, A., Leone, M., Cordone, L., 1991. Protein dynamics. comparative investigation on heme-proteins with different physiological roles. Biophys. J 59 (3), 742–754.

Kelly, F.P., 1979. Reversibility and Stochastic Networks. In: Wiley Series in Probability and Mathematical Statistics. John Wiley and Sons, New York, NY, USA, London, UK, Sydney, Australia.

Kendall, D.G., 1948. On the generalized "birth-and-death" process. Ann. Math. Stat. 19 (1), 1–15.

Kepler, T.B., Elston, T.C., 2001. Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations. Biophys. J. 81 (6), 3116–3136.

Kim, S.K., 1958. Mean first passage time for a random walker and its application to chemical kinetics. J. Chem. Phys. 28 (6), 1057–1067.

Klein, M.J., 1956. Generalization of the Ehrenfest urn model. Phys. Rev. 103 (1), 17–20.

Krieger, I.M., Gans, P.J., 1960. First-order stochastic processes. J. Chem. Phys. 32 (1), 247.

Kuthan, H., 2001. Self-organisation and orderly processes by individual protein complexes in the bacterial cell. Prog. Biophys. Mol. Biol. 75 (1-2), 1–17.

Laurenzi, I.J., 2000. An analytical solution of the stochastic master equation for reversible biomolecular reaction kinetics. J. Chem. Phys. 113 (8), 3315–3322.

Levsky, J.M., Singer, R.H., 2003. Gene expression and the myth of the average cell. Trends Cell Biol. 13 (1), 4–6.

Mayor, U., Guydosh, N.R., Johnson, C.M., Grossmann, J.G., Sato, S., Jas, G.S., Freund, S.M., Alonso, D.O., Daggett, V., Fersht, A.R., 2003. The complete folding pathway of a protein from nanoseconds to microseconds. Nature 421 (6925), 863–867.

McQuarrie, D.A., 1963. Kinetics of small systems. J. Chem. Phys. 38 (2), 433–436.

McQuarrie, D.A., Jachimowski, C.J., Russell, M.E., 1964. Kinetics of small systems. II. J. Chem. Phys. 40 (10), 2914.

Montroll, E.W., Shuler, K.E., 1958. The application of the theory of stochastic processes to chemical kinetics. Adv. Chem. Phys. 1, 361–399.

Nicolis, G., Prigogine, I., 1977. Self-organization in nonequilibrium systems: from dissipative structures to order through fluctuations. John Wiley and Sons, New York, NY, USA, London, UK, Sydney, Australia, A Wiley–Interscience Publication.

Othmer, H.G., 1969. Interactions of reaction and diffusion in open systems. Ph.D. Thesis, University of Minnesota, Minneapolis.

Othmer, H.G., 1979. A graph-theoretic analysis of chemical reaction networks, Lecture Notes, Rutgers University.

Othmer, H.G., 1981. The interaction of structure and dynamics in chemical reaction networks. In: Ebert, K.H., Deuflhard, P., Jager, W. (Eds.), Modelling of Chemical Reaction Systems. Springer-Verlag, New York, pp. 1–19.

Othmer, H.G., Scriven, L.E., 1971. Instability and dynamic pattern in cellular networks. J. Theor. Biol. 32, 507–537.

Ozbudak, E.M., Thattai, M., Kurtser, I., Grossman, A.D., van Oudenaarden, A., 2002. Regulation of noise in the expression of a single gene. Nat. Genet. 31 (1), 69–73.

Rao, C.V., Arkin, A.P., 2003. Stochastic chemical kinetics and the quasi-steady-state assumption: Application to the Gillespie algorithm. J. Chem. Phys. 118 (11), 4999–5010.

Shuler, K.F., 1960. Relaxation processes in multistate systems. Phys. Fluids 2 (4), 442–448.

Siegert, A.J.F., 1949. On the approach to statistical equilibrium. Phys. Rev. 76 (11), 1708–1714.

Singer, K., 1953. Application of the theory of stochastic processes to the study of irreproducible chemical reactions and nucleation processes. J. Roy. Stat. Soc. Ser. B 15 (1), 92–106.

Spudich, J.L., Koshland, D.E., 1976. Non-genetic individuality: chance in the single cell. Nature 262 (5568), 467–471.

Stundzia, A.B., Lumsden, C.J., 1996. Stochastic simulation of coupled reaction–diffusion processes. J. Comput. Phys. 127 (0168), 196–207.

Swain, P.S., Elowitz, M.B., Siggia, E.D., 2002. Intrinsic and extrinsic contributions to stochasticity in gene expression. Proc. Natl. Acad. Sci. USA 99 (20), 12795–12800.

Thattai, M., van Oudenaarden, A., 2001. Intrinsic noise in gene regulatory networks. Proc. Natl. Acad. Sci. USA 98 (15), 8614–8619.

Turing, A.M., 1952. The chemical basis of morphogenesis. Phil. Trans. Roy. Soc. Lond. B 237, 37–72.
Tyson, J.J., Othmer, H.G., 1978. The dynamics of feedback control circuits in biochemical pathways. Prog. Theor. Biol. 5, 1–62.
Wei, J., Prater, C.D., 1962. The structure and analysis of complex reaction systems. Adv. Catal. 13, 203.