

Exact Moment Dynamics for Feedforward Nonlinear Chemical Reaction Networks

EDUARDO D. SONTAG^{1,2} (Fellow, IEEE), AND ABHYUDAI SINGH³ (Member, IEEE)

¹Mathematics Department, Rutgers University, New Brunswick, NJ 08901 USA

²Center for Quantitative Biology, Rutgers University, New Brunswick, NJ 08901 USA

³Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19716 USA

CORRESPONDING AUTHOR: E. D. SONTAG (eduardo.sontag@gmail.com)

The work of E. D. Sontag was supported in part by the Air Force Office of Scientific Research under Grant FA9550-14-1-0060 and in part by the Office of Naval Research under Grant 5710003367.

ABSTRACT We introduce a class of chemical reaction networks for which all moments can be computed by finite-dimensional linear differential equations. This class allows second and higher order reactions, subject to certain assumptions on structure and/or conservation laws.

INDEX TERMS Systems biology, stochastic systems.

I. INTRODUCTION

CHEMICAL systems are inherently stochastic, as reactions depend on random (thermal) motion. This motivates the study of stochastic models, and specifically the Chemical Master Equation (CME), a discrete-space continuous-time Markov process that describes stochastic chemical kinetics. Exact studies using the CME are difficult, and several moment closure tools related to “mass fluctuation kinetics” and “fluctuation-dissipation” formulas can be used to obtain approximations of moments [8], [1], [3], [6]. We introduce a class of nonlinear networks for which exact computation is possible. For steady-state distributions, see also [7].

II. PRELIMINARIES

We start by reviewing standard concepts regarding master equations for biochemical networks, see for instance [6]. We assume that temperature and volume Ω are constant, and the system is well-mixed.

We consider a chemical reaction network consisting of m reactions which involve the n species S_i , $i \in \{1, 2, \dots, n\}$. The reactions R_j , $j \in \{1, 2, \dots, m\}$ are specified by combinations of reactants and products: $R_j: \sum_{i=1}^n a_{ij} S_i \rightarrow \sum_{i=1}^n b_{ij} S_i$ where the a_{ij} and b_{ij} are non-negative integers, the *stoichiometry coefficients*, and the sums are understood informally, indicating combinations of elements. The $n \times m$ *stoichiometry matrix* $\Gamma = \{\gamma_{ij}\}$ has entries: $\gamma_{ij} = b_{ij} - a_{ij}$, $i = 1, \dots, n$, $j = 1, \dots, m$. Thus, γ_{ij} counts the net change in the number of units of species S_i each time that reaction R_j takes place. We will denote by γ_j the j th column of Γ : $\gamma_j = b_j - a_j$ where $a_j = (a_{1j}, \dots, a_{nj})'$ and $b_j = (b_{1j}, \dots, b_{nj})'$ (prime indicates transpose) and assume that no $\gamma_j = 0$ (that is, every reaction changes at least one species). In general, for every $v \in \mathbb{Z}_{\geq 0}^n$, we denote $\oplus v = \oplus (v_1, \dots, v_n) := v_1 + \dots + v_n$. In particular, for each $j \in \{1, \dots, m\}$, we define the *order* of reaction R_j as $A_j = \oplus a_j = \sum_{i=1}^n a_{ij}$ (the total number of units of all species participating in the reaction).

Stochastic models of chemical reaction networks are described by a column-vector Markov stochastic process $X = (X_1, \dots, X_n)'$ which is indexed by time $t \geq 0$ and takes values in $\mathbb{Z}_{\geq 0}^n$. Thus, $X(t)$ is a $\mathbb{Z}_{\geq 0}^n$ -valued random variable, for each $t \geq 0$. Abusing notation, we also write $X(t)$ to represent an outcome of this random variable on a realization of the process. The state of the system at time t is: $X_i(t) = k_i =$ number of units of species i at time t . Let $p_k(t) = \mathbb{P}[X(t) = k]$ for each $k \in \mathbb{Z}_{\geq 0}^n$. Then $p(t) = (p_k)_{k \in \mathbb{Z}_{\geq 0}^n}$ is the discrete probability density (also called the “probability mass function”) of $X(t)$. This note is concerned with the computation of moments of this density.

A *Chemical Master Equation (CME)* (also known as a *Kolmogorov forward equation*) is a system of linear differential equations for the p_k 's, of the following form. Suppose given m functions $\rho_j: \mathbb{Z}_{\geq 0}^n \rightarrow \mathbb{R}_{\geq 0}$, $j = 1, \dots, m$, with $\rho_j(0) = 0$. These are the *propensity functions* for the respective reactions R_j . An intuitive interpretation is that $\rho_j(k)dt$ is the probability that reaction R_j takes place, in a short interval of length dt , provided that the state was k at the beginning of the interval. The CME is (see [2] for existence and uniqueness results):

$$\frac{dp_k}{dt} = \sum_{j=1}^m \rho_j(k - \gamma_j) p_{k-\gamma_j} - \sum_{j=1}^m \rho_j(k) p_k, \quad k \in \mathbb{Z}_{\geq 0}^n \quad (1)$$

where, for notational simplicity, we omitted the time argument “ t ” from p , and where we make the convention that $\rho_j(k - \gamma_j) = 0$ unless $k \geq \gamma_j$ (coordinatewise inequality). There is one equation for each $k \in \mathbb{Z}_{\geq 0}^n$, so this is an infinite system of linked equations. When discussing the CME, we will assume that an initial probability vector $p(0)$ has been specified, and that there is a unique solution of (1) defined for all $t \geq 0$. A different CME results for each choice of propensity functions, a choice that is dictated by physical chemistry considerations. Here we will restrict attention to the most standard model, mass-action kinetics propensities. For each $k = (k_1, \dots, k_n)' \in \mathbb{Z}_{\geq 0}^n$, we let (recall that a_j denotes

the vector $(a_{1j}, \dots, a_{mj})'$: $\binom{k}{a_j} = \prod_{i=1}^n \binom{k_i}{a_{ij}}$ where $\binom{k_i}{a_{ij}}$ is the usual combinatorial number $k_i!/(k_i - a_{ij})!a_{ij}!$, which we define to be zero if $k_i < a_{ij}$. The most commonly used propensity functions, and the ones best-justified from elementary physical principles, are *ideal mass action kinetics* propensities, defined as follows:

$$\rho_j(k) = \kappa_{a_j} \binom{k}{a_j}, j = 1, \dots, m. \quad (2)$$

The m non-negative constants κ_{a_j} are arbitrary, and they represent quantities related to the volume, shapes of the reactants, chemical and physical information, and temperature. Notice that $\rho_j(k)$ can be expanded into a polynomial in which each variable k_i has an exponent less or equal to a_{ij} . In other words, $\rho_j(k) = \sum_{c_j \leq a_j} \kappa_{c_j} k^{c_j}$ (" \leq " is understood coordinatewise, and by definition $k^{c_j} = k_1^{c_{1j}} \dots k_n^{c_{nj}}$ and $r^0 = 1$ for all integers), for suitably redefined coefficients κ_{c_j} 's. Often one uses the simplification

$$\rho_j(k) = \kappa_j k^{a_j}, j = 1, \dots, m. \quad (3)$$

(Approximate $x(x-1)\dots(x-r+1) \approx x^r$.) Suppose given a function $M: \mathbb{Z}_{\geq 0}^n \rightarrow \mathbb{R}$ (to be taken as a monomial when computing moments). The expectation of the random variable $M(X)$ is $\mathbb{E}[M(X(t))] = \sum_{k \in \mathbb{Z}_{\geq 0}^n} p_k(t) M(k)$ because $\mathbb{P}[X(t) = k] = p_k(t)$. Let us define, for any $\gamma \in \mathbb{Z}_{\geq 0}^n$, the new function $\Delta_\gamma M$ given by $(\Delta_\gamma M)(k) := M(k + \gamma) - M(k)$. With these notations,

$$\frac{d}{dt} \mathbb{E}[M(X(t))] = \sum_{j=1}^m \mathbb{E}[\rho_j(X(t)) \Delta_{\gamma_j} M(X(t))] \quad (4)$$

(see [6] for more details). We next specialize to a monomial function: $M(k) = k^u = k_1^{u_1} k_2^{u_2} \dots k_n^{u_n}$ where $u \in \mathbb{Z}_{\geq 0}^n$. There results $(\Delta_{\gamma_j} M)(k) = \sum_{\nu \in \mathcal{S}(u, j)} d_\nu k^\nu$ for appropriate coefficients d_ν , where

$$\mathcal{S}(u, j) := \left\{ \nu \in \mathbb{Z}_{\geq 0}^n \mid \begin{array}{l} \nu = u - \mu, u \geq \mu \neq 0 \\ \mu_i = 0 \text{ for each } i \text{ such that } \gamma_{ij} = 0 \end{array} \right\}$$

(inequalities " \geq " in $\mathbb{Z}_{\geq 0}^n$ are understood coordinatewise). Thus, for (3) and (2) respectively:

$$\frac{d}{dt} \mathbb{E}[X(t)^u] = \sum_{j=1}^m \sum_{\nu \in \mathcal{S}(u, j)} d_\nu \kappa_j \mathbb{E}[X(t)^{\nu+a_j}]. \quad (5)$$

$$\frac{d}{dt} \mathbb{E}[X(t)^u] = \sum_{j=1}^m \sum_{c_j \leq a_j} \sum_{\nu \in \mathcal{S}(u, j)} d_\nu \kappa_{c_j} \mathbb{E}[X(t)^{\nu+c_j}]. \quad (6)$$

III. NEW RESULTS

For each multi-index $u \in \mathbb{Z}_{\geq 0}^n$, we define $\mathcal{R}^0(u) = \{u\}$,

$$\mathcal{R}^1(u) := \{\nu + a_j, 1 \leq j \leq m, \nu \in \mathcal{S}(u, j)\}$$

if using (5), or

$$\mathcal{R}^1(u) := \{\nu + c_j, 1 \leq j \leq m, c_j \leq a_j, \nu \in \mathcal{S}(u, j)\}$$

if using (6), and, more generally, for any $\ell \geq 1$, $\mathcal{R}^{\ell+1}(u) := \mathcal{R}^1(\mathcal{R}^\ell(u))$ where, for any set U , $\mathcal{R}^\ell(U) := \bigcup_{u \in U} \mathcal{R}^\ell(u)$. Finally, we set $\mathcal{R}(u) := \bigcup_{i=0}^{\infty} \mathcal{R}^i(u)$. Each set $\mathcal{R}^\ell(u)$ is finite, but the cardinality $\#\mathcal{R}(u)$ may be infinite. It is finite

if and only if there is some $L \geq 0$ such that $\mathcal{R}(u) = \bigcup_{i=0}^L \mathcal{R}^i(u)$, or equivalently $\mathcal{R}^{L+1}(u) \subseteq \bigcup_{i=0}^L \mathcal{R}^i(u)$.

Equation (5) (or (6)) says that the derivative of the u -th moment can be expressed as a linear combination of the moments in the set $\mathcal{R}^1(u)$. The derivatives of these moments, in turn, can be expressed in terms of the moments in the set $\mathcal{R}^1(u')$, for each $u' \in \mathcal{R}^1(u)$, i.e., in terms of moments in the set $\mathcal{R}^2(u)$. Iterating, we have the following:

Main Lemma. Suppose that $N := \#\mathcal{R}(u) < \infty$, and $\mathcal{R}(u) = \{u = u_1, \dots, u_N\}$. Then, writing

$$x(t) := (\mathbb{E}[X^{u_1}(t)], \dots, \mathbb{E}[X^{u_N}(t)])',$$

there is an $A \in \mathbb{R}^{N \times N}$ such that $\dot{x}(t) = Ax(t)$ for all $t \geq 0$.

This motivates the following problem: *characterize those chemical reaction networks for which $\#\mathcal{R}(u) < \infty$ for all $u \in \mathbb{Z}_{\geq 0}^n$* . One simple sufficient condition is that all reactions be of order 0 or 1, i.e. $\oplus a_j \in \{0, 1\}$. In that case, since $\mu \neq 0$ in the definition of $\mathcal{S}(u, j)$, it follows that $\oplus a_j \leq \oplus \mu$ for every index j . Therefore, $\oplus(\nu + a_j) = \oplus \nu + \oplus a_j - \oplus \mu \leq \oplus \nu$ for all ν , and the same holds for $\nu + c_j$ if $c_j \leq a_j$. So all elements in $\mathcal{R}(u)$ have degree $\leq \oplus u$, hence $\#\mathcal{R}(u) < \infty$. We will generalize to "weighted L^1 norms" $\beta_1 v_1 + \dots + \beta_n v_n$.

Considering the difference inclusion $z(t+1) \in \mathcal{R}^1(z(t))$, $z(0) = u$, the following definition is natural.

Definition. A function $V: \mathbb{Z}_{\geq 0}^n \rightarrow \mathbb{R}_{\geq 0}$ will be called a Lyapunov-like function with respect to a given chemical network if the following two properties hold:

- 1) for each $u, v \in \mathbb{Z}_{\geq 0}^n$: $v \in \mathcal{R}^1(u) \Rightarrow V(v) \leq V(u)$ [nondecreasing property],
- 2) for each $\alpha \geq 0$: $V_\alpha := \{v \mid V(v) \leq \alpha\}$ is finite [properness].

Theorem. For every chemical network, the following two statements are equivalent:

- There exists a Lyapunov-like function.
- $\#\mathcal{R}(u) < \infty$ for all $u \in \mathbb{Z}_{\geq 0}^n$.

Proof. Sufficiency is clear: pick any u , and let $\alpha := V(u)$; iterating on the nondecreasing property, $V(v) \leq \alpha$ for all $v \in \mathcal{R}(u)$, meaning that $\mathcal{R}(u) \subseteq V_\alpha$, and thus $\#\mathcal{R}(u) < \infty$.

To prove the converse, assume that $\#\mathcal{R}(u) < \infty$ for all u . Define $V(u) := \max_{w \in \mathcal{R}(u)} \oplus w$. Since $\#\mathcal{R}(u) < \infty$, it follows that $V(u) < \infty$. As $u \in \mathcal{R}(u)$, it follows from the definition of V that $\oplus u \leq V(u)$. Now pick any u, v so that $v \in \mathcal{R}^1(u)$. Since $\mathcal{R}(v) \subseteq \mathcal{R}(u)$, it follows that $\{\oplus w, w \in \mathcal{R}(v)\} \subseteq \{\oplus w, w \in \mathcal{R}(u)\}$. Therefore $V(v) \leq V(u)$ (nondecreasing property). Now pick any $\alpha \geq 0$, which we may take without loss of generality to be a nonnegative integer, and any element $v \in V_\alpha$. Since $\oplus v \leq V(v)$, it follows that $\oplus v \leq \alpha$. So V_α is a subset of the set of all nonnegative vectors v such that $\oplus v \leq \alpha$, which has $\binom{\alpha+n}{n}$ elements. ■

The nonincrease requirement means, using the definition of $\mathcal{R}^1(u)$, that

$$V(u - \mu + c_j) \leq V(u) \quad (7)$$

for all $c_j \leq a_j$ ($1 \leq j \leq m$) under definition (2) for propensities, or just for $c_j = a_j$ if propensities have the simplified form (3), and every μ for which $u \geq \mu \neq 0$ and $\mu_i = 0$ for every i such that $\gamma_{ij} = 0$. Pick any reaction index j and for this index pick any species index i such that the

species S_i changes, that is, $\gamma_{ij} \neq 0$. Now pick $u = \mu = e_i$, the canonical unit vector with a “1” in the i th position (this choice of μ is allowed, since it is false that $\gamma_{ij} = 0$) and apply (7). A necessary condition for decrease is:

$$\gamma_{ij} \neq 0 \Rightarrow V(c_j) \leq V(e_i) \quad (8)$$

for all $c_j \leq a_j$, or for $c_j = a_j$ in the case (3).

We now consider the special case of Lyapunov-like functions which can be extended to an additive map $V : \mathbb{Z}^n \rightarrow \mathbb{R}$. In this case, (8) is equivalent to (7). To see that (8) implies (7), pick any u , any reaction index j , and every μ for which $u \geq \mu \neq 0$ and $\mu_i = 0$ for every i such that $\gamma_{ij} = 0$. Since $\mu \neq 0$ and $\mu \leq u$, there is some species index i such that $\gamma_{ij} \neq 0$ and $\mu_i \neq 0$, i.e. $\mu \geq e_i$. Applying (8) with this i :

$$V(u - \mu + c_j) \leq V(u) + V(e_i - \mu) + V(c_j - e_i) \leq V(u),$$

where we used that $V(\mu - e_i) \geq 0$. A map $V : \mathbb{Z}_{\geq 0}^n \rightarrow \mathbb{R}_{\geq 0}$ that extends to an additive function $V : \mathbb{Z}^n \rightarrow \mathbb{R}$ is necessarily of the form $V(u) = \beta u = \beta_1 u_1 + \dots + \beta_n u_n$ for some $\beta = (\beta_1, \dots, \beta_n) \in \mathbb{Z}_{\geq 0}^n$ and it automatically satisfies the properness property provided that all $\beta_i \neq 0$, which we assume from now on. Thus, $\#(\mathcal{R}(u)) < \infty$ will be satisfied for all u if V has this form and satisfies (8). This condition can be made a little more explicit in the linear case. Let $\Delta_j := \{i \mid \gamma_{ij} \neq 0\}$. Then a linear Lyapunov-like function amounts to solving a linear program: find β such that $i \in \Delta_j \Rightarrow \beta a_j \leq \beta_i$.

A. SPECIAL CASE: MULTI-LAYER FEEDFORWARD NETWORKS

A general class for which there is a linear Lyapunov-like function, and hence $\#(\mathcal{R}(u)) < \infty$ for all u , is that of multi-layer feedforward networks with linear reactions in the first layer. These are defined as follows. We find it convenient to separate degradations from more general reactions. So we will assume that there are reactions R_j , $j \in \{1, 2, \dots, m\}$, which are partitioned into $p \geq 1$ layers: $\mathbf{R}_1, \dots, \mathbf{R}_p$. Species S_i , $i \in \{1, 2, \dots, n\}$, are also partitioned into p layers $\mathbf{S}_1, \dots, \mathbf{S}_p$. In addition, we allow additional “pure degradation” reactions $D_j : S_{i_j} \rightarrow 0$, $j \in \{1, \dots, d\}$ (so the total number of reactions is $m' = m + d$).

We assume that the reactions R_j that belong to the first layer \mathbf{R}_1 are all of order zero or one, i.e. they have $\oplus a_j \in \{0, 1\}$. (This first layer might model several independent separate chemical subnetworks; we collect them all as one larger network.) More generally, for reactions at any given layer π the only species that appear as reactants in nonlinear reactions are those in layers $< \pi$ and the only ones that can change are those in layer π , that is:

$$\text{if } R_j \in \mathbf{R}_\pi : \begin{cases} a_{ij} \neq 0 \text{ and } \oplus a_j > 1 \Rightarrow S_i \in \bigcup_{1 \leq s < \pi} \mathbf{S}_s \\ \gamma_{ij} \neq 0 \Rightarrow S_i \in \mathbf{S}_\pi \end{cases} \quad (9)$$

This means that except for order zero or one reactions, every reaction R_j at layer $1 < \pi \leq p$ has the form:

$$\begin{aligned} & a_{i_1 j} S_{i_1} + \dots + a_{i_q j} S_{i_q} \\ \rightarrow & a_{i_1 j} S_{i_1} + \dots + a_{i_q j} S_{i_q} + b_{i_{q+1} j} S_{i_{q+1}} + \dots + b_{i_{q+q'} j} S_{i_{q+q'}} \end{aligned}$$

with S_{i_1}, \dots, S_{i_q} in layers $< \pi$ and $S_{i_{q+1}}, \dots, S_{i_{q+q'}}$ in layer π . We claim that there is a linear Lyapunov-like function

for any such network. Note that, for a degradation reaction $D_j : S_{i_j} \rightarrow 0$, the entry γ_{ij} of the stoichiometry vector is nonzero (and equal to -1) only when $i = i_j$, and for this index we have $a_{ij} = 1$. Thus the linear Lyapunov condition simply requires $\beta_i \leq \beta_i$ and hence is automatically satisfied no matter what is the choice of β . Thus we may ignore degradations and assume from now on that only the reactions R_j are present. We prove the claim by induction on the number of layers p . If $p = 1$, all reactions have order 0 or 1, so we can take $\beta_i = 1$ for all i . Arrange the species indices so that S_{r+1}, \dots, S_n are the species in \mathbf{S}_p ; these do not appear any reactions belonging to \mathbf{R}_π for $\pi < p$. So layers \mathbf{R}_π for $\pi < p$ and species in \mathbf{S}_π for $\pi < p$ define a network with $p - 1$ layers, and we may assume by induction that a linear V_0 has been defined for that network. This means that we have a vector of positive numbers $\beta^0 = (\beta_1, \dots, \beta_{n-r})$ such that the linear Lyapunov condition holds for this subnetwork, which means, for any extension to a vector $\beta = (\beta^0, \star)$ with n components (since the coefficients of a_j are zero for indices $r+1, \dots, n$) that $\beta a_j \leq \beta_i$ whenever $i \in \Delta_j$, when j, i index reactions and species in the first $p - 1$ layers.

So all that is needed is to define the additional coefficients β_i , $i \in \{r+1, \dots, n\}$, such that the inequality $\beta a_j \leq \beta_i$ holds for all pairs (i, j) such that (1) $R_j \in \mathbf{R}_p$ or $S_i \in \mathbf{S}_p$ and (2) $\gamma_{ij} \neq 0$. We show that it suffices to pick all these β_i equal to a common value $\bar{\beta} := \text{maximum of } \beta^0 a_j^0 \text{ over all reactions } R_j \in \mathbf{R}_p$, where a_j^0 is the restriction of the vector a_j to its first r components.

If $R_j \in \mathbf{R}_p$ and $S_i \notin \mathbf{S}_p$, the second condition in (9) (with $\pi = p$) says that (2) is not satisfied. Thus, we only need to consider $S_i \in \mathbf{S}_p$, i.e. $i \in \{r+1, \dots, n\}$. Suppose first that $\oplus a_j > 1$. The first condition in (9) (with $\pi = p$) insures that $a_{ij} = 0$ for all such i . Thus, $a_j = (a_j^0, 0)$ where the vector 0 has length $n - r$. It follows that $\beta a_j = \beta^0 a_j^0 \leq \bar{\beta}$. Next, suppose that $\oplus a_j \leq 1$. If $\oplus a_j = 0$, then $a_j = 0 \leq \bar{\beta}$. So assume $\oplus a_j = 1$ and pick the unique index i' such that $a_{i'j} = 1$. If $S_{i'} \in \mathbf{S}_\pi$, with $\pi < p$, then once again $a_{ij} = 0$ for all $i \in \{r+1, \dots, n\}$ and $\beta a_j = \beta^0 a_j^0 \leq \bar{\beta}$. Finally, assume that $a_j = e_i$ with $i \in \{r+1, \dots, n\}$. Now $\bar{\beta} = \beta a_j \leq \beta_i$ is trivially satisfied. ■

B. EXAMPLES

Let us start with the system shown in [4] to have moment closure: $R_1 : 0 \rightarrow S_1$, $R_2 : S_1 \rightarrow S_1 + S_2$, $R_3 : S_1 + S_2 \rightarrow S_1 + S_2 + S_3$, $D_1 : S_1 \rightarrow 0$, $D_2 : S_2 \rightarrow 0$, $D_3 : S_3 \rightarrow 0$. This is a three-layer system with one reaction in each layer, plus degradations. As we said, we may ignore degradations, so we consider: $a_1 = (0, 0, 0)'$, $a_2 = (1, 0, 0)'$, $a_3 = (1, 1, 0)'$, and we have $\Delta_1 = \{1\}$, $\Delta_2 = \{2\}$, $\Delta_3 = \{3\}$. We must find a positive vector $\beta = (\beta_1, \beta_2, \beta_3)$ such that $\beta a_i \leq \beta_i$, $i \in \{1, 2, 3\}$, i.e., $\beta_1 \leq \beta_2$ and $\beta_1 + \beta_2 \leq \beta_3$. We may pick $\beta = (1, 1, 2)$.

Here is a more complicated example involving several reversible first order reactions as well as some dimeric and trimeric reactions: $R_1 : S_1 \rightarrow S_2$, $R_2 : S_2 \rightarrow S_1$, $R_3 : S_3 \rightarrow S_4$, $R_4 : S_4 \rightarrow S_3$, $R_5 : 2S_1 + S_2 + S_3 \rightarrow 2S_1 + S_2 + S_3 + S_5$, $R_6 : S_1 + 3S_5 \rightarrow S_1 + 3S_5 + S_6$. We have three layers: $\mathbf{R}_1 = \{R_1, R_2, R_3, R_4\}$, $\mathbf{R}_2 = \{R_5\}$, $\mathbf{R}_3 = \{R_6\}$, and $\mathbf{S}_1 = \{S_1, S_2, S_3, S_4\}$, $\mathbf{S}_2 = \{S_5\}$, $\mathbf{S}_3 = \{S_6\}$. Using e_i to denote canonical unit vectors: $a_j = e_j$, $j \in \{1, 2, 3, 4\}$, $a_5 = (2, 1, 1, 0, 0, 0)'$, $a_6 = (1, 0, 0, 0, 3, 0)'$, and

$\Delta_1 = \Delta_2 = \{1, 2\}$, $\Delta_3 = \Delta_4 = \{3, 4\}$, $\Delta_5 = \{5\}$, $\Delta_6 = \{6\}$. We must find a positive vector β such that $\beta a_1 \leq \beta_1$, $\beta a_1 \leq \beta_2$, $\beta a_2 \leq \beta_1$, $\beta a_2 \leq \beta_2$, $\beta a_3 \leq \beta_3$, $\beta a_3 \leq \beta_4$, $\beta a_4 \leq \beta_3$, $\beta a_4 \leq \beta_4$, $\beta a_5 \leq \beta_5$, $\beta a_6 \leq \beta_6$, i.e. so that $\beta_1 = \beta_2$, $\beta_3 = \beta_4$, $2\beta_1 + \beta_2 + \beta_3 \leq \beta_5$, and $\beta_1 + 3\beta_5 \leq \beta_6$. These constraints can be satisfied with $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 1$, $\beta_5 = 4$, $\beta_6 = 13$.

C. SPECIAL CASE: CONSERVED VARIABLES

In some applications, one is interested in computing the moments $\mathbb{E}[X(t)^u]$ only for trajectories $X(t)$ which remain in some specified subset $\mathcal{C} \subseteq \mathbb{Z}_{\geq 0}^n$. When this subset has the form

$$\mathcal{C} = \mathcal{C}_1 \times \mathbb{Z}_{\geq 0}^{n_2} \subset \mathbb{Z}_{\geq 0}^{n_1} \times \mathbb{Z}_{\geq 0}^{n_2} \quad (10)$$

and the subset \mathcal{C}_1 is finite, the right-hand side of equation (6) (or (5)) can be simplified.

For example, suppose that the first two species S_1 and S_2 indicate the activity of a specified gene (inactive and active, respectively), with S_1 and S_2 reacting according to $S_1 \rightarrow S_2$, $S_2 \rightarrow S_1$, and no other reactions involving a change in S_1 and S_2 . (This does not rule out reactions such as $S_1 \rightarrow S_1 + S_3$ which would model transcription from the active conformation, since such a reaction does not change S_1 nor S_2 .) It is the case that $X_1(t) + X_2(t)$ remains constant in time, so $X_1(t) + X_2(t) = X_1(0) + X_2(0)$ for all t . Moreover, given the biological motivation for these equations, we are only interested in the cases where $(X_1(0), X_2(0)) = (1, 0)$ or $(0, 1)$. Thus, we have that $X_1(t) + X_2(t) = 1$ for all t . This restricts the components $(X_1(t), X_2(t))$ of $X(t)$ to take values in the finite set $\mathcal{C}_1 = \{(1, 0), (0, 1)\}$, and hence all moments can be assumed to have the first two exponents equal to one:

$$\mathbb{E}[X_1^{u_1} X_2^{u_2} X_3^{u_3} \dots X_n^{u_n}] = \mathbb{E}[X_1 X_2 X_3^{u_3} \dots X_n^{u_n}].$$

More generally, for any positive integers r and L , let $\mathcal{L}_{L,r} := \{0, \dots, L\}^r$. Then, for any finite subset $\mathcal{C} \subset \mathbb{Z}_{\geq 0}^r$, there is some integer L with the property that for each $v \in \mathbb{Z}_{\geq 0}^r$ there exists $\{p_d \in \mathbb{R}, d \in \mathcal{L}_{L,r}\}$ so that $k^v = \sum_{d \in \mathcal{L}_{L,r}} p_d k^d$ for all $k \in \mathcal{C}$. In other words, every monomial can be expressed as a linear combination of monomials with exponents $\leq L$. To prove this, observe that the set \mathcal{F} of functions $\mathcal{C} \rightarrow \mathbb{R}$ is a finite-dimensional vector space (canonically identified with $\mathbb{R}^{\#\mathcal{C}}$), where $\#\mathcal{C}$ is the cardinality of \mathcal{C} . Introduce for each i the subspace $\mathcal{F}_{i,r}$ of \mathcal{F} spanned by the monomial functions $k \mapsto k^d$, $d \in \mathcal{L}_{i,r}$. Since $\mathcal{F}_{0,r} \subseteq \mathcal{F}_{1,r} \subseteq \mathcal{F}_{2,r} \subseteq \dots$ is a nondecreasing sequence of subspaces, there is some L such that $\mathcal{F}_{L',r} = \mathcal{F}_{L,r}$ for all $L' > L$ (in fact, one may take $L = \#\mathcal{C} - 1$). Now given any set as in (10) with $\#\mathcal{C}_1 < \infty$, we apply above observation to \mathcal{C}_1 , and so all moments $X_1^{u_1} X_2^{u_2} X_3^{u_3} \dots X_n^{u_n}$ can be written as a linear combination of moments for which the first n_1 exponents are $\leq L$. The remaining reactions could be a feedforward network, and now moments are all determined by a finite set of linear differential equations, so long as we only care about initial conditions in a finite invariant set.

A simple example is as follows. We consider the following set of chemical reactions: $R_1 : S_1 \xrightarrow{u} S_2$, $R_2 : S_2 \xrightarrow{k_2} S_1$, $R_3 : 0 \xrightarrow{k_3 u} S_3$, $R_4 : S_2 + S_3 \xrightarrow{k_4} S_2$ where we think of “ u ” as an external input. This is basically the incoherent feedforward loop considered in [5] to study adaptation and

the fold-change detection property in stochastic systems. The only difference is that there we used separate creation and degradation reactions $0 \rightarrow S_2 \rightarrow 0$ (the first with rate u), but here, in order to impose a conservation law, we think of S_2 as being an active form of a kinase (the input controlling the change to active form), which can be constitutively de-activated by a reverse reaction. The effect of u on S_3 is incoherent, in the sense that u promotes formation of S_3 , as well as degradation, because the larger u , the larger the active concentration of S_2 , which degrades S_3 . We have $\rho_1 = u S_1$, $\rho_2 = k_2 S_2$, $\rho_3 = k_3 u$, $\rho_4 = k_4 S_2 S_3$, and $a_1 = (1, 0, 0)$, $a_2 = (0, 1, 0)$, $a_3 = (0, 0, 0)$, $a_4 = (0, 1, 1)$, $\gamma_1 = (-1, 1, 0)$, $\gamma_2 = (1, -1, 0)$, $\gamma_3 = (0, 0, 1)$, $\gamma_4 = (0, 0, -1)$, and $X_1(t) + X_2(t)$ is constant along all solutions. Suppose e.g. $X_1(0) + X_2(0) = 2$. Let us obtain a linear differential equation for the mean of $X_3(t)$. Using these notations: $x_i = \mathbb{E}[X_2^i X_3]$, $i = 0, 1, 2$, $y_i = \mathbb{E}[X_2^i]$, $i = 1, 2$. We are interested in $x_0(t)$. We conclude that:

$$\begin{aligned} \dot{y}_1 &= 2u - k_2 y_1 - y_1 u \\ \dot{y}_2 &= k_2 y_1 + 2u + 3y_1 u - 2y_2 u - 2k_2 y_2 \\ \dot{x}_0 &= k_3 u - k_4 x_1 \\ \dot{x}_1 &= k_3 y_1 u - k_4 x_2 - k_2 x_1 + 2x_0 u - x_1 u \\ \dot{x}_2 &= \alpha x_1 + 2x_0 u - \beta x_2 + k_3 y_2 u + 3x_1 u - 2x_2 u. \end{aligned}$$

where $\alpha = k_2 + 2k_4$ and $\beta = 3k_4 + 2k_2$.

More abstractly, given any finite continuous-time Markov chain with n_1 states q_i and transition rates λ_{ij} , we may introduce n_1 species S_i and reactions $S_i \rightarrow S_j$ with rate λ_{ij} . The stoichiometric matrix consists of columns with exactly one entry equal to 1 and one entry equal to -1 , so the sum $X_1(t) + \dots + X_{n_1}(t)$ is conserved (see e.g. Section 4.8 in [6]). Thus, starting from an initial condition with $X_1(0) + \dots + X_{n_1}(0) = 1$ we have that at all times we have precisely one $X_i(t) = 1$. This provides an embedding of the Markov Chain: state is q_i at time t if $X_i(t) = 1$. This construction is of interest when reaction parameters κ_i in a network are described by functions of finite Markov chains (Hidden Markov Models) and the network is of a feedforward type, to conclude that finite-dimensional ODE's exist for moments.

REFERENCES

- [1] C. A. Gómez-Urbe and G. C. Verghese, “Mass fluctuation kinetics: Capturing stochastic effects in systems of chemical reactions through coupled mean-variance computations,” *J. Chem. Phys.*, 126:024109, 2007.
- [2] S. P. Meyn and R. L. Tweedie, “Stability of Markovian processes III: Foster-Lyapunov criteria for continuous-time processes,” *Adv. Appl. Probab.*, 25:518548, 1993.
- [3] A. Singh and J. P. Hespanha, “A derivative matching approach to moment closure for the stochastic logistic model,” *Bull. Math. Biol.*, 69:1909–1925, 2007.
- [4] A. Singh and M. Soltani, “Quantifying intrinsic and extrinsic variability in stochastic gene expression models,” *PLoS ONE*, 8(12):e84301, 2013.
- [5] M. Skataric, E. V. Nikolaev, and E. D. Sontag, “Scale-invariance in singularly perturbed systems,” in *Proc. IEEE 53rd Annu. Conf. Decision Control*, Los Angeles, CA, USA, Dec. 2014, pp. 3035–3040.
- [6] E. D. Sontag, *Lecture Notes on Mathematical Systems Biology, 2002–2015*. [Online]. Available: http://www.math.rutgers.edu/~sontag/FTP_DIR/systems_biology_notes.pdf, accessed Sep. 14, 2015.
- [7] E. D. Sontag and D. Zeilberger, “A symbolic computation approach to a problem involving multivariate Poisson distributions,” *Adv. Appl. Math.*, 44:359–377, 2010.
- [8] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry*. Amsterdam, The Netherlands: Elsevier, 2001.