


Product-Form Stationary Distributions for Deficiency Zero Networks with Non-mass Action Kinetics

David F. Anderson¹ · Simon L. Cotter² 

Received: 23 May 2016 / Accepted: 4 October 2016 / Published online: 27 October 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract In many applications, for example when computing statistics of fast sub-systems in a multiscale setting, we wish to find the stationary distributions of systems of continuous-time Markov chains. Here we present a class of models that appears naturally in certain averaging approaches whose stationary distributions can be computed explicitly. In particular, we study continuous-time Markov chain models for biochemical interaction systems with non-mass action kinetics whose network satisfies a certain constraint. Analogous with previous related results, the distributions can be written in product form.

Keywords Product-form stationary distributions · Deficiency zero · Constrained averaging · Stochastically modeled reaction network

DFA and SLC would like to thank the Isaac Newton Institute for Mathematical Sciences, Cambridge, for support and hospitality during the program Stochastic Dynamical Systems in Biology: Numerical Methods and Applications, where work on this paper was undertaken. This work was supported by EPSRC Grant No EP/K032208/1.

David F. Anderson: Grant support from NSF-DMS-1318832 and Army Research Office Grant W911NF-14-1-0401.

Simon L. Cotter: Grant support from EPSRC first Grant EP/L023393/1.

✉ Simon L. Cotter
simon.cotter@manchester.ac.uk

David F. Anderson
anderson@math.wisc.edu

¹ Department of Mathematics, University of Wisconsin, Madison, WI, USA

² Department of Mathematics, University of Manchester, Manchester, UK

1 Introduction

Biological interaction systems are typically modeled in one of three ways. If the counts of the constituent species are high, then their concentrations are often modeled via a system of ordinary differential equations with state space $\mathbb{R}_{\geq 0}^d$, where $d > 0$ is the number of species. If the counts are moderate (perhaps order 10^2 or 10^3), then they may be approximated by some form of continuous diffusion process (Gillespie 2000; Van Kampen 2007). If the counts are low, then the system is typically modeled stochastically as a continuous-time Markov chain in $\mathbb{Z}_{\geq 0}^d$ (Anderson and Kurtz 2011, 2015). We often want to understand the stationary behavior of the model under consideration. For deterministic models, understanding the stationary behavior usually entails characterizing the stable fixed points of the system, whereas for stochastic models, we require the calculation of the stationary distribution.

Stationary distributions are also useful in a multiscale setting, where the stationary statistics of a fast subsystem can be utilized in the approximation of the dynamics of the slow variables, which are typically of most interest. When an analytical form for the stationary distribution of the fast subsystem is not known, numerical approximations can be used. However, these computations are often expensive and part of an “inner loop,” typically making this calculation the rate-limiting step of the analysis. In the context of biochemical reaction networks, quasi-equilibrium (QE)-based approximations lead to fast subsystems which preserve mass action kinetics (Goutsias 2005; Janssen 1989; Thomas et al. 2012; Cao et al. 2005; Weinan et al. 2005). However, more recent improvements in stochastic averaging can lead to fast subsystems with non-mass action kinetics, and this observation was the motivation for the present work (Cotter 2016; Cotter and Erban 2016; Cotter et al. 2011).

One class of interaction networks that has been quite successfully analyzed, and that appears ubiquitously as fast subsystems, are those that are weakly reversible and have a deficiency of zero (see Appendix). For this class of models, and under the assumption of mass action kinetics, the fixed points of the deterministic models (Anderson 2011, 2008; Craciun 2015; Feinberg 1979, 1987; Gunawardena 2003) and the stationary distributions for the stochastic models, have been fully characterized (Anderson et al. 2010, 2015; Cappelletti and Wiuf 2016; Van Kampen 1976). In fact, it is the study of this class of networks that is largely responsible for the development of the field of chemical reaction network theory (Feinberg 1972, 1979; Gunawardena 2003; Horn 1972), a branch of applied mathematics in which the dynamical properties of the mathematical model are related to the structural properties of the interaction network.

In this article, we return to stochastically modeled interaction networks that are weakly reversible and have a deficiency of zero, though we consider propensity functions (also called intensity functions or rate functions) that are more general than mass action kinetics. However, we add a certain condition to the rates within the reaction network (see Assumption 1 below). Following Anderson et al. (2010), which was motivated by the work of Kelly (1979) who discovered the product-form stationary distribution of certain stochastically modeled queuing networks, we provide the form of the stationary distribution for this class of models. In particular, and in similarity with the main results of Anderson et al. (2010), in Theorem 2 we show that the distribution

is of product form and that the key parameter of the distribution is a complex-balanced equilibrium value of an associated deterministically modeled system with mass action kinetics. The result of Anderson et al. (2010) can now be viewed as a special case of this new result.

The paper proceeds as follows: In Sect. 2, we introduce the formal mathematical model of interest. In Sect. 3, we provide the main theorem of this article, which characterizes the stationary distribution for the class of models of interest. In Sect. 4, we provide a series of examples which demonstrate the usefulness of the main result. Some brief concluding remarks are given in Sect. 5.

We assume throughout that the reader is familiar with terminology from chemical reaction network theory. However, we provide in Appendix all necessary terminology and results from this field that are used in the present work.

2 Mathematical Model

We consider a system with d chemical species, $\{S_1, \dots, S_d\}$, undergoing reactions which alter the state of the system. For concreteness, we suppose there are $K > 0$ distinct reaction channels. For the k th reaction channel, we denote by $v_k, v'_k \in \mathbb{Z}_{\geq 0}^d$ the vectors representing the number of molecules of each species consumed and created in one instance of that reaction, respectively. Note that $v'_k - v_k \in \mathbb{Z}^d$ is the net change in the system due to one instance of the k th reaction. We associate each such v_k and v'_k with a linear combination of the species in which the coefficient of S_i is v_{ki} , the i th element of v_k . For example, if $v_k = [1, 2]^T$ for a system consisting of two species, then we associate with v_k the linear combination $S_1 + 2S_2$. Under this association, each v_k and v'_k is termed a *complex* of the system. We denote any reaction by the notation $v_k \rightarrow v'_k$, where v_k is the source, or reactant, complex and v'_k is the product complex. We note that each complex may appear as both a source complex and a product complex in the system. The set of all complexes will be denoted by $\{v_k\}$.

Definition 1 Let $\mathcal{S} = \{S_i\}$, $\mathcal{C} = \{v_k\}$, and $\mathcal{R} = \{v_k \rightarrow v'_k\}$ denote the sets of species, complexes, and reactions, respectively. The triple $\{\mathcal{S}, \mathcal{C}, \mathcal{R}\}$ is called a *chemical reaction network*.

Throughout, we assume that $v_k \neq v'_k$ for each $k \in \{1, \dots, K\}$.

2.1 Deterministic Model

The usual deterministic model for a chemical reaction network $\{\mathcal{S}, \mathcal{C}, \mathcal{R}\}$ assumes that the vector of concentrations for the species satisfies a differential equation of the form

$$\dot{x}(t) = \sum_{k=1}^K r_k(x(t))(v'_k - v_k), \quad (1)$$

where $r_k : \mathbb{R}_{\geq 0}^d \rightarrow \mathbb{R}_{\geq 0}$ is the (state-dependent) rate of the k th reaction channel. If we assume that each function r_k satisfies deterministic mass action kinetics, then

$$r_k(x) = \kappa_k \prod_{i=1}^d x_i^{v_{ki}}.$$

Definition 2 An equilibrium value $c \in \mathbb{R}_{\geq 0}^d$ of (1) is said to be *complex balanced* if for each $\eta \in \mathcal{C}$,

$$\sum_{k:v_k=\eta} r_k(c) = \sum_{k:v'_k=\eta} r_k(c),$$

where the sum on the left, respectively right, is over those reactions with η as source complex, respectively product complex. In the special case of mass action kinetics, c is complex balanced if and only if

$$\sum_{k:v_k=\eta} \kappa_k \prod_{i=1}^d c_i^{v_{ki}} = \sum_{k:v'_k=\eta} \kappa_k \prod_{i=1}^d c_i^{v_{ki}}. \tag{2}$$

2.2 Stochastic Model, Previous Results, and Assumptions

The usual stochastic model for a reaction network $\{\mathcal{S}, \mathcal{C}, \mathcal{R}\}$ treats the system as a continuous-time Markov chain for which the rate of transition from state $x \in \mathbb{Z}_{\geq 0}^d$ to state $x + v'_k - v_k$ is $\lambda_k(x)$, where $\lambda_k : \mathbb{Z}_{\geq 0}^d \rightarrow \mathbb{R}_{\geq 0}$ is a suitably chosen intensity function (the intensity functions are also termed *propensity* functions in the literature). This stochastic process can be characterized in a variety of useful ways (Anderson and Kurtz 2011, 2015). For example, it is the stochastic process with state space $\mathbb{Z}_{\geq 0}^d$ and infinitesimal generator

$$Af(x) = \sum_{k=1}^K \lambda_k(x)(f(x + v'_k - v_k) - f(x)),$$

where $f : \mathbb{Z}_{\geq 0}^d \rightarrow \mathbb{R}$. Kolmogorov’s forward equation for this model, termed the chemical master equation in the biology literature, is

$$\begin{aligned} \frac{d}{dt} p_\mu(t, x) &= \sum_{k=1}^K \lambda_k(x - v'_k + v_k) p_\mu(t, x - v'_k + v_k) \mathbb{1}_{\{x - v'_k + v_k \geq 0\}} \\ &\quad - \sum_{k=1}^K \lambda_k(x) p_\mu(t, x), \end{aligned} \tag{3}$$

$$= A^* p_\mu(t, x) \mathbb{1}_{\{x \geq 0\}}, \tag{4}$$

where $p_\mu(t, x)$ is the probability the process is in state $x \in \mathbb{Z}_{\geq 0}^d$ at time $t \geq 0$, given an initial distribution of μ , and A^* is the adjoint of A . Note that (3) implies that a stationary distribution for the model, π , must satisfy

$$\sum_{k=1}^K \lambda_k(x - v'_k + v_k)\pi(x - v'_k + v_k) = \sum_{k=1}^K \lambda_k(x)\pi(x). \tag{5}$$

Note also that (4) implies that π is in the null space of A^* . Thus, and as is well known, finding π reduces to solving $\pi A = 0$, with $\sum_x \pi(x) = 1$, where A is reinterpreted as a generator matrix.

Of particular interest to us are the rate functions, λ_k . Under the assumption of stochastic mass action kinetics, we have

$$\lambda_k(x) = \kappa_k \prod_{i=1}^d \frac{x_i!}{(x_i - v_{ki})!} = \kappa_k \prod_{i=1}^d \prod_{j=0}^{v_{ki}-1} (x_i - j),$$

where $\kappa_k > 0$ are the *reaction rate constants*. Here and throughout, we interpret any product of the form $\prod_{i=0}^{-1} a_i$ to be equal to one. In Anderson et al. (2010), the authors considered a class of stochastically modeled reaction networks with mass action kinetics (those with a deficiency of zero and are weakly reversible, see Appendix) and characterized their stationary distributions as products of Poisson distributions. However, they did not just consider mass action kinetics in Anderson et al. (2010), but any kinetics satisfying the functional form

$$\lambda_k(x) = \kappa_k \prod_{i=1}^d \prod_{j=0}^{v_{ki}-1} \theta_i(x_i - j), \tag{6}$$

so long as $\theta_i(0) = 0$ for each i . In particular, they proved the following.

Theorem 1 (Anderson et al. 2010) *Let $\{S, C, R\}$ be a reaction network, and let $(\kappa_1, \dots, \kappa_k)$ be a choice of positive rate constants. Suppose that modeled deterministically with mass action kinetics and rate constants κ_k , the system is complex balanced with complex-balanced equilibrium $c \in \mathbb{R}_{>0}^d$. Then, the stochastically modeled system with intensity functions (6) admits the invariant measure*

$$\pi(x) = \prod_{i=1}^d \frac{c_i^{x_i}}{\prod_{j=0}^{x_i-1} \theta_i(x_i - j)}. \tag{7}$$

The measure π can be normalized to a stationary distribution so long as it is summable.

The connection between weakly reversible, deficiency zero networks, and complex-balanced equilibria is given in Appendix.

In this paper, we generalize Theorem 1 by showing how to find the stationary distributions for models with rates that do not seem to satisfy the form (6). However,

we add an assumption pertaining to the form of the rates within the reaction network, which we describe now. We begin by partitioning the set of species into two sets, $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2$. We will say that $S_i \in \mathcal{S}_2$ if $\alpha_i = \text{gcd}\{v_{1i}, v_{2i}, \dots, v_{Ki}\} > 1$. Otherwise, we say that $S_i \in \mathcal{S}_1$, noting that $\alpha_i = 1$. We now assume that the intensity functions for the reaction network are of the form

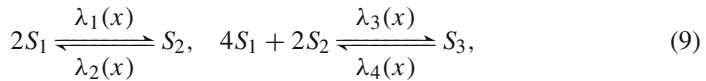
$$\lambda_k(x) = \kappa_k \prod_{i=1}^d \prod_{j=0}^{\frac{v_{ki}}{\alpha_i} - 1} \theta_i(x_i - j\alpha_i), \tag{8}$$

where $\kappa_k > 0$ and $\theta_i(x_i) = 0$ if and only if $x_i \leq \alpha_i - 1$.

Assumption 1 A stochastically modeled reaction network satisfies this assumption if it satisfies the partition described above and has intensity functions of the form (8).

We provide an example to clarify the notation.

Example 1 Consider the stochastically modeled system with reaction network



where the intensity functions are placed next to the reaction arrows. Here $S_1 \in \mathcal{S}_2$, with $\alpha_1 = 2$, and $S_2, S_3 \in \mathcal{S}_1$. The assumption (8) then supposes that for appropriate functions $\theta_1, \theta_2, \theta_3 : \mathbb{Z}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$, we have

$$\begin{aligned} \lambda_1(x) &= \kappa_1 \theta_1(x_1) \\ \lambda_2(x) &= \kappa_2 \theta_2(x_2) \\ \lambda_3(x) &= \kappa_3 \theta_1(x_1) \theta_1(x_1 - 2) \theta_2(x_2) \theta_2(x_2 - 1) \\ \lambda_4(x) &= \kappa_4 \theta_3(x_3). \end{aligned}$$

For example, valid choices include $\theta_1(x_1) = x_1(x_1 - 1) + 1(x_1 \geq 2)$, $\theta_2(x_2) = x_2$, and $\theta_3(x_3) = \frac{x_3}{1+x_3}$, in which case the form for the stationary distribution does not follow immediately from Theorem 1.

3 Main Result

Here we state and prove our main result.

Theorem 2 Let $\{\mathcal{S}_1 \cup \mathcal{S}_2, \mathcal{C}, \mathcal{R}\}$ be a reaction network satisfying Assumption 1, and let $(\kappa_1, \dots, \kappa_K)$ be a choice of positive rate constants. Suppose that modeled deterministically with mass action kinetics and rate constants κ_k the system is complex balanced with complex-balanced equilibrium $c \in \mathbb{R}_{>0}^d$. Then the stochastically modeled system with intensity functions (8) admits the invariant measure

$$\pi(x) = \prod_{i=1}^d \frac{c_i^{x_i}}{\prod_{j=0}^{\lfloor x_i/\alpha_i \rfloor - 1} \theta_i(x_i - j\alpha_i)}. \tag{10}$$

The measure π can be normalized to a stationary distribution so long as it is summable.

Note that the theorem applies to models with reaction networks satisfying Assumption 1 and that are weakly reversible and have a deficiency of zero. See Appendix.

Proof First note that if $\mathcal{S}_2 = \emptyset$, then Theorem 2 is the same as Theorem 1 and there is nothing to show. Thus, we suppose $\mathcal{S}_2 \neq \emptyset$.

The proof proceeds in the following manner. First, for each $S_i \in \mathcal{S}_2$ we will demonstrate the existence of a function $\varphi_i : \mathbb{Z}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ for which

$$\theta_i(x_i) = \prod_{\ell=0}^{\alpha_i-1} \varphi_i(x_i - \ell) = \varphi_i(x_i) \cdots \varphi_i(x_i - \alpha_i + 1). \tag{11}$$

Next, we will apply Theorem 1 and prove that the resulting distribution is indeed given by (10).

Let $S_i \in \mathcal{S}_2$. We begin by setting

$$\varphi_i(0) = 0, \quad \text{and} \quad \varphi_i(z) = 1, \quad \text{for } 1 \leq z \leq \alpha_i - 1.$$

For $z \geq \alpha_i$ an integer, we may define φ_i recursively via the formula

$$\varphi_i(z) = \frac{\theta_i(z)}{\varphi_i(z - 1) \cdots \varphi_i(z - \alpha_i + 1)}. \tag{12}$$

Note that φ_i is a well-defined function since $\theta_i(z) > 0$ for each $z \geq \alpha_i$ by assumption. It is clear that (11) is satisfied with this choice of φ_i .

For $x \in \mathbb{Z}_{\geq 0}^d$, we may now write

$$\begin{aligned} \lambda_k(x) &= \kappa_k \prod_{i=1}^d \prod_{j=0}^{\frac{v_{ki}}{\alpha_i} - 1} \theta_i(x_i - j\alpha_i) = \kappa_k \prod_{i=1}^d \prod_{j=0}^{\frac{v_{ki}}{\alpha_i} - 1} \prod_{\ell=0}^{\alpha_i-1} \varphi_i(x_i - j\alpha_i - \ell) \\ &= \kappa_k \prod_{i=1}^d \prod_{b=0}^{v_{ki}-1} \varphi_i(x_i - b). \end{aligned}$$

Hence, we may apply Theorem 1 and conclude that

$$\pi(x) = \prod_{i=1}^d \frac{c_i^{x_i}}{\prod_{j=0}^{x_i-1} \varphi_i(x_i - j)}$$

is an invariant measure for the system, where c is a complex-balanced fixed point for the deterministic system. It remains to show that

$$\prod_{j=0}^{x_i-1} \varphi_i(x_i - j) = \prod_{j=0}{\lfloor x_i/\alpha_i \rfloor - 1} \theta_i(x_i - j\alpha_i). \tag{13}$$

First note that if $x_i < \alpha_i$, then both sides of (13) are equal to one. For the time being, assume that $\alpha_i \leq x_i < 2\alpha_i$. Under this assumption, the right-hand side of (13) is

$$\theta_i(x_i) = \varphi_i(x_i) \cdots \varphi_i(x_i - \alpha_i + 1) = \varphi_i(x_i) \cdots \varphi_i(\alpha_i),$$

where in the final equality we used that $\varphi_i(\ell) = 1$ for $1 \leq \ell < \alpha_i$ (and that $x_i - \alpha_i < \alpha_i$), and the left-hand side is

$$\varphi_i(x_i) \cdots \varphi_i(1) = \varphi_i(x_i) \cdots \varphi(\alpha_i),$$

where we again used that $\varphi_i(\ell) = 1$ when $1 \leq \ell < \alpha_i$. Hence, (13) is verified when $x_i < 2\alpha_i$.

We will now prove that (13) holds in general by induction. We suppose that (13) holds for all $z \leq x_i$, where $x_i \geq 2\alpha_i - 1$, and will show it to hold at $x_i + 1$. Using (12), the left-hand side of (13) evaluated at $x_i + 1$ is

$$\begin{aligned} & \prod_{j=0}^{(x_i+1)-1} \varphi_i(x_i + 1 - j) \\ &= \prod_{j=1}^{x_i+1} \varphi_i(j) = \varphi_i(x_i + 1) \cdots \varphi_i(x_i + 1 - \alpha_i + 1) \prod_{j=1}^{x_i+1-\alpha_i} \varphi_i(j) \\ &= \theta_i(x_i + 1) \prod_{j=1}^{x_i+1-\alpha_i} \varphi_i(j), \end{aligned}$$

where the final equality is an application of (12) with $x_i + 1$ in place of the variable z . Continuing, we have

$$\begin{aligned} & \theta_i(x_i + 1) \prod_{j=1}^{x_i+1-\alpha_i} \varphi_i(j) \\ &= \theta_i(x_i + 1) \prod_{j=0}^{x_i+1-\alpha_i-1} \varphi_i(x_i + 1 - \alpha_i - j) \quad (\text{by a rearrangement}) \\ &= \theta_i(x_i + 1) \prod_{j=0}^{\lfloor (x_i+1-\alpha_i)/\alpha_i \rfloor - 1} \theta_i((x_i + 1 - \alpha_i) - j\alpha_i) \quad (\text{by the inductive hypothesis}) \\ &= \theta_i(x_i + 1) \prod_{j=0}^{\lfloor (x_i+1)/\alpha_i \rfloor - 2} \theta_i(x_i + 1 - (j + 1)\alpha_i) \end{aligned}$$

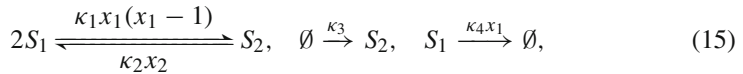
$$\begin{aligned}
 &= \theta_i(x_i + 1) \prod_{j=1}^{\lfloor (x_i+1)/\alpha_i \rfloor - 1} \theta_i(x_i + 1 - j\alpha_i) \\
 &= \prod_{j=0}^{\lfloor (x_i+1)/\alpha_i \rfloor - 1} \theta_i(x_i + 1 - j\alpha_i),
 \end{aligned} \tag{14}$$

where the final equalities are straightforward. Note that (14) is the right-hand side of (13) evaluated at $x_i + 1$, and so the proof is complete. \square

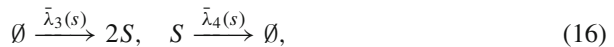
4 Examples

4.1 Example 1: Motivating Example

First, we consider a motivating example arising from model reduction, through constrained averaging (Cotter 2016; Cotter and Erban 2016; Cotter et al. 2011), of the following system:

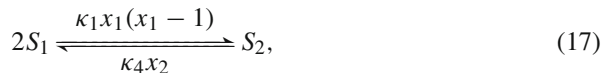


where the intensity functions are placed next to the reaction arrows. Note that the intensities of all the reactions follow mass action kinetics. We consider this system in a parameter regime where the reversible dimerization reactions $2S_1 \rightleftharpoons S_2$ are occurring more frequently than the production of S_2 and the degradation of S_1 . Both S_1 and S_2 are changed by the fast reactions, but the quantity $S = S_1 + 2S_2$ is invariant with respect to the fast reactions, and as such is the slow variable in this system. We wish to reduce the dynamics of this system to a model only concerned with the possible changes in S :



where $\bar{\lambda}_3(s)$ and $\bar{\lambda}_4(s)$ are the effective rates of the system.

Using the QE approximation (QEA), $\bar{\lambda}_3(s) = \kappa_3$ and $\bar{\lambda}_4(s) = \kappa_4 \mathbb{E}_{\pi_{\text{QEA}}(s)}[X_1]$, where $\pi_{\text{QEA}}(s)$ is the stationary distribution for the system



under the assumption that $X_1(0) + 2X_2(0) = s$. Since the system (17) satisfies the necessary conditions of the results of Anderson et al. (2010) (weak reversibility and deficiency of zero), the invariant distribution $\pi_{\text{QEA}}(s)$ is known exactly.

In comparison, the constrained approach requires us to find the invariant distribution π_{Con} of the following system:

$$2S_1 \xrightleftharpoons[(k_2 + k_4)x_2]{k_1x_1(x_1 - 1) + k_3\mathbb{1}_{\{x_1>1\}}} S_2, \tag{18}$$

subject to $X_1(0) + 2X_2(0) = s$. Readers interested in seeing how this is derived should refer to [Cotter \(2016\)](#). This network is weakly reversible and has a deficiency of zero. However, the form of the rates in this system does not satisfy the conditions specified in [Anderson et al. \(2010\)](#). In the context of constrained averaging, this lack of a closed form for the stationary distribution would result in the need for some form of approximation of the stationary distribution. There are two common methods utilized for performing this approximation. One possibility would be to perform exhaustive stochastic simulation of the system (18). Another option involves finding the distribution by finding the null space of the adjoint of the generator (see the discussion in and around (5)). However, as the state space of (18) will typically be huge, the latter method often involves truncating the state space and approximating the actual distribution with that of the stationary distribution of the truncated system ([Cotter 2016](#)). Both approaches will lead to approximation errors and varying amounts of computational cost. However, note that the system (18) does satisfy Assumption 1, with $\alpha_1 = 2$ and $\alpha_2 = 1$. We denote the rate of dimerization by λ_D and its reverse by λ_{-D} . Therefore,

$$\begin{aligned} \lambda_D(x) &= k_1x_1(x_1 - 1) + k_3\mathbb{1}_{\{x_1>1\}}, \\ &= k_1 \left(x_1(x_1 - 1) + \frac{k_3}{k_1}\mathbb{1}_{\{x_1>1\}} \right), \\ &= k_1\theta_1(x_1), \end{aligned}$$

with θ_1 defined in the final equality. The form of the rate of the reverse reaction is much simpler and is given by

$$\lambda_{-D}(x) = (k_2 + k_4)x_2 = (k_2 + k_4)\theta_2(x),$$

which defines θ_2 .

By Theorem 2, we can write down the stationary distribution of this system. The complex-balanced equilibrium of the associated deterministically modeled system with $c_1 + 2c_2 = 1$ is given by $(c_1, c_2) = \left(\frac{\sqrt{(k_2+k_4)(k_2+8k_1+k_4)}-k_2-k_4}{4k_1}, \frac{1-c_1}{2} \right)$. Then by Theorem 2, and by recalling that all states (x_1, x_2) in the domain satisfy $s = x_1 + 2x_2$, the stationary distribution for S_2 is given by

$$\pi_{\text{Con}}(x_2) = \frac{1}{\Gamma_{\text{Con}}} \frac{c_1^{(s-2x_2)}}{\prod_{j=0}^{\lfloor (s-2x_2)/2 \rfloor - 1} \left((s - 2x_2 - 2j)(s - 2x_2 - 2j - 1) + \frac{k_3}{k_1} \right)} \frac{c_2^{x_2}}{x_2!}, \tag{19}$$

where Γ_{Con} is a normalization constant and s is the conserved quantity. Note that the indicator function in θ_1 (in the denominator) has disappeared since it is always equal to one over the domain of the product.

We can compare (19) with the distribution of (17), which arises from the QEA, and also with the distribution of the full system (15) conditioned on $S_1 + 2S_2 = s$ (which can be approximated by finding the null space of the adjoint of the generator of the full system on the truncated domain). First, we consider the QEA approximation. The invariant distribution of the fast subsystem (17) can be found using Theorem 1 and is given by

$$\pi_{\text{QEA}}(x) = \frac{1}{\Gamma_{\text{QEA}}} \frac{d_1^{s-2x_2}}{(s - 2x_2)!} \frac{d_2^{x_2}}{x_2!}, \tag{20}$$

where $(d_1, d_2) = \left(\frac{\sqrt{k_4(8k_1+k_4)}-k_4}{4k_1}, \frac{1-d_1}{2} \right)$ is the complex-balanced equilibrium for this system satisfying $d_1 + 2d_2 = 1$ and Γ_{QEA} is a normalizing constant.

Since the full system (15) does not have a deficiency of zero, we are not able to find its invariant distribution directly. However, by truncating the state space appropriately, we are able to approximate the full distribution by constructing the generator on this truncated state space and finding the null space of the adjoint.

Once we have approximated the null space of the truncated generator, we can find the approximation of $\mathbb{P}(X_2 = x_2 | X_1 + 2X_2 = s)$ by taking the probabilities of all states with $x_1 + 2x_2 = s$ and renormalizing. In what follows, we truncated the domain of the generator to $x \in \{0, 1, \dots, 1000\} \times \{0, 1, \dots, 500\}$.

We consider the system (15) with parameters given by:

$$k_1 = 1, \quad k_2 = 100, \quad k_3 = 1500, \quad k_4 = 30. \tag{21}$$

Note that it is not obvious from these rates that the reactions with rates k_3 and k_4 are in fact the slow reactions in this system. The invariant density is largely concentrated in a small region centered close to the point $x = (99, 114)$. By using the approximation of the invariant density that we have computed on the truncated domain, we can compute the expected ratio between occurrences of the fast reactions with rates k_1 and k_2 with the slow reactions with rates k_3 and k_4 . For this choice of parameters, the expected proportion of the total reactions which are fast reactions (dimerization/disassociation) is 82.68 %. This indicates a difference in timescales between these reactions, but the difference is not particularly stark, and as such we would expect there to be significant error in any approximation relying on the QEA.

Figure 1 shows the three approximations of the distribution $\mathbb{P}(X_2 = x_2 | X_1 + 2X_2 = 300)$ for the system (15) with parameters given by (21). The constrained and QEA approximations are computed using (19) and (20), respectively, with the normalizing constants computed numerically. As would be expected in this parameter regime, the constrained approximation is far more accurate than the QEA.

We can quantify the accuracy of each of the approximations by computing the relative l^2 differences with the distribution computed using the full generator on the truncated domain. This relative difference was 4.464×10^{-1} for the QEA, in com-

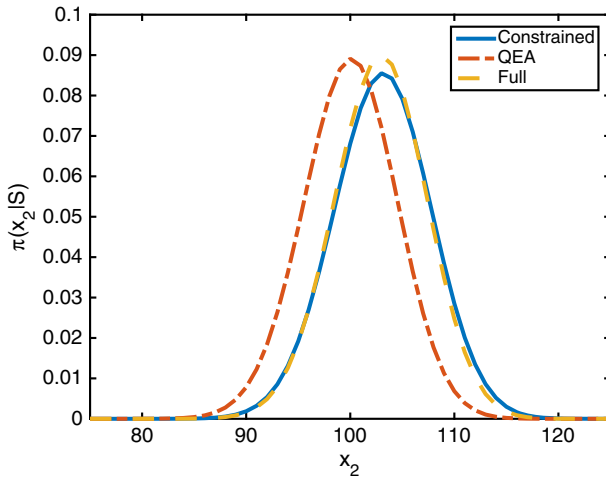
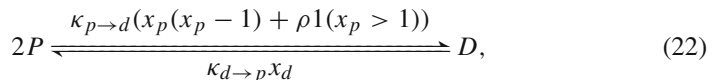


Fig. 1 (Color figure online) Approximations of the distribution $\mathbb{P}(X_2 = x_2|X_1 + 2X_2 = 300)$ for the system (15) with parameters given by (21) using constrained averaging, QEA averaging, and through approximation of the invariant distribution of the full system on $x \in \{0, 1, \dots, 1000\} \times \{0, 1, \dots, 500\}$.

parison with 5.2337×10^{-2} for the constrained approximation. This demonstrates the improvement in approximation that can be achieved by using constrained averaging, and which motivates the need for results like Theorem 2 which take non-mass action kinetics into account.

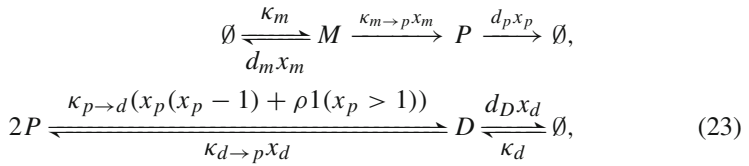
4.2 Example 2: Dimerization

We begin by considering a model consisting of only proteins, denoted P , and dimers, denoted D . We suppose there are two mechanisms by which the proteins are dimerized: random interactions between the protein molecules, and via a catalyst. The rate at which random interactions lead to the formation of dimers can be taken to be of mass action form. Assuming the concentrations are such that the catalyst is acting at capacity, the rate of formation of the dimers due to the catalyst can be faithfully modeled as a constant (so long as there are at least two proteins to make the reaction happen). Thus, letting x_p and x_d denote the numbers of proteins and dimers in the model, respectively, the reaction system can be represented as



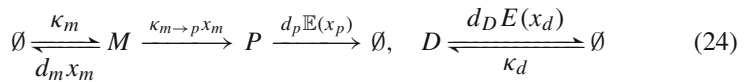
where $\kappa_{p \rightarrow d}$, $\kappa_{d \rightarrow p}$ and ρ are given parameters. Note that after an obvious change of variables the stationary distribution of this particular model is provided in (19).

The reactions in (22) are typically a subset of the reactions in a larger system. For example, the actual model of interest may be



where M represents an mRNA molecule. This is a standard model for dimer production. Depending upon the relevant timescales in the system, we may want to take $d_D = \kappa_d = 0$. If the reactions $2P \rightleftharpoons D$ are appreciably faster than those of (23), then an obvious path simulation strategy presents itself in the mold of the slow-scale SSA (Cao et al. 2005) (i.e., stochastic averaging):

1. for the current values of P and D , find the stationary distribution of the fast subsystem analytically,
2. determine the effective rates of the reduced model



where the expectations are with respect to the distribution found in step 1.

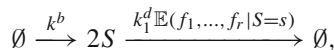
3. simulate forward in time using the stochastic simulation algorithm (Gillespie 1976) or the next reaction method (Anderson 2007; Gibson and Bruck 2000), and return to step 1.

Being able to analytically calculate the stationary distribution in step 1 allows us to bypass the need to numerically approximate the stationary distribution, as is commonly done (Weinan et al. 2005; Weinan and Vanden-Eijnden 2007).

4.3 Example 3

In Sects. 4.1 and 4.2, we applied Theorem 2 on the common motif $2S_1 \rightleftharpoons S_2$. In this example, we present another common motif, $2S \rightleftharpoons \emptyset$, for which Theorem 2 is also useful. As opposed to the specific models we considered in the previous examples, here we present a more general framework in which the specific form of the propensity function for the reaction $2S \rightarrow \emptyset$ is arbitrary. This situation is common when undertaking certain types of averaging arguments (Cotter 2016).

Let us suppose that the effective dynamics of a slow variable in a larger system can be modeled as



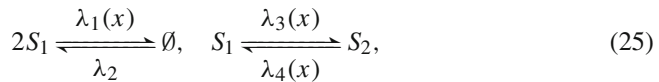
where k^b and k_1^d are positive constants and where $\mathbb{E}(f_1, \dots, f_r | S = s)$ is a conditional expectation of the fast variables f_1, \dots, f_r , conditioned on $S = s$. The conditional expectation could in general be highly nonlinear, and not of the form (6) required by Theorem 1. Supposing that $\mathbb{E}(f_1, \dots, f_r | S = s) = 0$ if $s \in \{0, 1\}$, Theorem 2 says that, up to a normalization constant, the invariant distribution of S is given by

$$\pi(s) \propto \frac{\left(\frac{k^b}{k^d}\right)^s}{\prod_{j=0}^{\lfloor s/2 \rfloor - 1} \mathbb{E}(f_1, \dots, f_r | S = s - 2j)}.$$

In practice, the normalization constant could be approximated by summing over an appropriate domain.

4.4 Example 4

This example will demonstrate the difficulties that can arise when a reaction is added that involves a species in the set \mathcal{S}_2 with multiplicity not equal to α_i . In particular, consider the system



where

$$\lambda_1(x) = \kappa_1 \theta_1(x_1), \quad \lambda_2(x) = \kappa_2, \quad \lambda_3(x) = \kappa_3 \varphi(x_1), \quad \lambda_4(x) = \kappa_4 x_2,$$

where $\theta_1(x_1) = (\mathbb{1}_{\{x_1 > 1\}}(10 + x_1 + 6 \sin(\pi x_1/5)))$. Note that we have a non-mass action kinetics rate λ_1 for the reaction $2S_1 \rightarrow \emptyset$. There is another reaction involving S_1 , but the amount of S_1 molecules involved in this reaction is not a multiple of 2. This means that we cannot apply the result of Theorem 2 to this system unless φ satisfies the recurrence relation detailed in the proof of Theorem 2. In particular, it must satisfy $\varphi(x_1)\varphi(x_1 - 1) = \theta_1(x_1)$, or

$$\begin{aligned} \varphi(0) &= 0, \\ \varphi(1) &= C \in \mathbb{R}_{>0}, \\ \varphi(n) &= \frac{\theta_1(n)}{\varphi(n - 1)}. \end{aligned}$$

This recurrence relation defines a unique function $\varphi : \mathbb{Z}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ for each $C \in \mathbb{R}_{\geq 0}$. In general, the function φ can oscillate wildly. Let us consider, for example, φ when θ_1 is given as above. In this case,

$$\varphi(x_1) = \frac{C^{2\text{mod}(x_1, 2) - 1} \prod_{i=0}^{\lfloor x_1/2 \rfloor - 1} \theta_1(x_1 - 2i)}{\prod_{i=0}^{\lfloor (x_1 - 1)/2 \rfloor - 1} \theta_1(x_1 - 2i - 1)}. \tag{26}$$

Figure 2 demonstrates how $\varphi(x_1)$ and the amplitude of the oscillations grow with x_i , for the case $C = 1$. It is clear that this function does not represent any physical reaction rate arising from chemistry, but we can still write down the invariant distribution of the system (25).

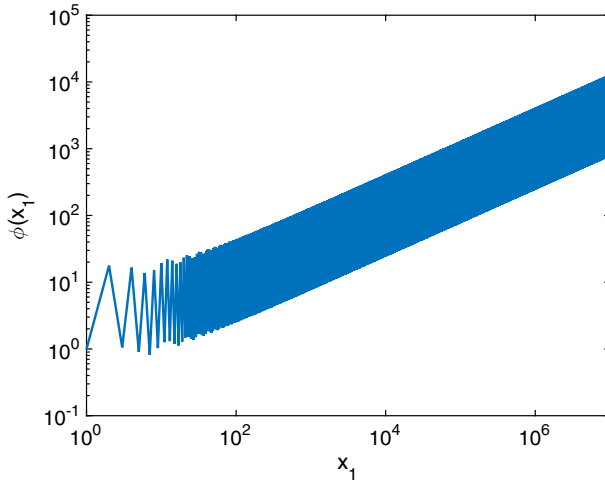


Fig. 2 (Color figure online) φ as given in (25) with $C = 1$.

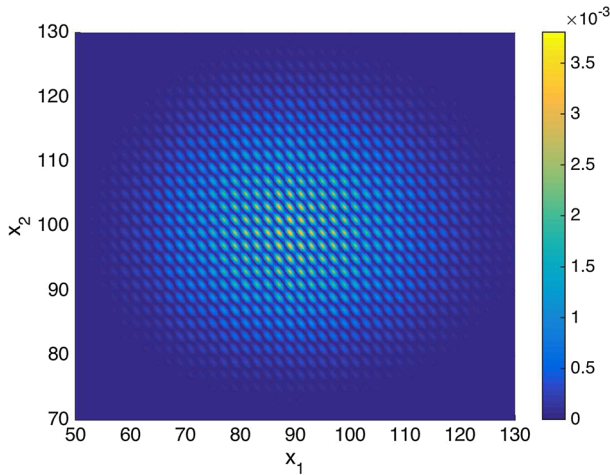


Fig. 3 (Color figure online) Stationary distribution (27) of the system (25) with parameters given by (28), with the assumption that the initial value of $x_1 + x_2$ is even. The normalization constant Γ was approximated by summing the value of all states $x \in \{0, 1, \dots, 1000\} \times \{0, 1, \dots, 1000\}$.

The complex balance equilibrium of the associated mass action kinetics system to (25) is given by $(c_1, c_2) = \left(\sqrt{\frac{k_2}{k_1}}, \frac{k_3}{k_4} \sqrt{\frac{k_2}{k_1}}\right)$. Therefore, the stationary distribution is given by:

$$\pi(x) = \frac{1}{\Gamma} \frac{c_1^{x_1}}{\prod_{i=1}^{x_1} \varphi(x_1)} \frac{c_2^{x_2}}{x_2!}, \quad x \in \{(x_1, x_2) \mid \text{mod}(x_1 + x_2, 2) = a\} \text{ for } a \in \{0, 1\}, \tag{27}$$

where Γ is a normalizing constant and φ is given by (26) with $C = 1$. Note that the value of a here dictates the oddness or evenness of the quantity $x_1 + x_2$, which is preserved by each of the reactions.

Figure 3 shows the invariant distribution (27) of the system (25) with an even initial value of $x_1 + x_2$, and for the following parameters:

$$k_1 = 1, \quad k_2 = 10^2, \quad k_3 = 10, \quad k_4 = 1. \quad (28)$$

The normalization constant was approximated by summing the values of all states $x \in \{0, 1, \dots, 1000\} \times \{0, 1, \dots, 1000\}$.

5 Conclusions

In this paper, we provided the stationary distributions for the class of stochastically modeled reaction networks with non-mass action kinetics that satisfy our Assumption 1 and that admit a complex-balanced equilibrium when modeled deterministically with mass action kinetics. Similarly to the results of Anderson et al. (2010), we showed that the stationary distributions are of product form.

We motivated the need for such results through consideration of modern averaging techniques. In particular, Theorem 2 significantly reduces the computational cost of finding the invariant distribution of most fast subsystems in a multiscale setting, therefore making accurate approximations very cheap to compute. This in turn opens up more possibilities, for example approximation of likelihoods via multiscale reductions in the context of parameter inference for biochemical networks.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix: Terminology and Results from Chemical Reaction Network Theory

For each reaction network, $\{\mathcal{S}, \mathcal{C}, \mathcal{R}\}$, there is a unique directed graph constructed in the following manner. The nodes of the graph are the complexes, \mathcal{C} . A directed edge is then placed from complex v_k to complex v'_k if $v_k \rightarrow v'_k \in \mathcal{R}$. Each connected component of the resulting graph is termed a *linkage class*. We denote the number of linkage classes by ℓ . For example, in the reaction network (9) there are two linkage classes.

Definition 3 A chemical reaction network, $\{\mathcal{S}, \mathcal{C}, \mathcal{R}\}$, is called *weakly reversible* if each linkage class is strongly connected. A network is called *reversible* if $v'_k \rightarrow v_k \in \mathcal{R}$ whenever $v_k \rightarrow v'_k \in \mathcal{R}$.

Note that a network is weakly reversible if and only if for any reaction $v_k \rightarrow v'_k$, there is a sequence of directed reactions beginning with v'_k as a source complex and

ending with v_k as a product complex. That is, there exist complexes v_1, \dots, v_r such that $v'_k \rightarrow v_1, v_1 \rightarrow v_2, \dots, v_r \rightarrow v_k \in \mathcal{R}$.

Definition 4 $S = \text{span}_{\{v_k \rightarrow v'_k \in \mathcal{R}\}} \{v'_k - v_k\}$ is the *stoichiometric subspace* of the network. For $z \in \mathbb{R}^d$, we say $z + S$ and $(z + S) \cap \mathbb{R}_{>0}^d$ are the *stoichiometric compatibility classes* and *positive stoichiometric compatibility classes* of the network, respectively. Denote $\dim(S) = s$.

The final definition is that of the *deficiency* of a network (Feinberg 1987).

Definition 5 The *deficiency* of a chemical reaction network, $\{\mathcal{S}, \mathcal{C}, \mathcal{R}\}$, is $\delta = |\mathcal{C}| - \ell - s$, where $|\mathcal{C}|$ is the number of complexes, ℓ is the number of linkage classes of the network graph, and s is the dimension of the stoichiometric subspace of the network.

We state a classical result that can be found in Feinberg (1972, 1979, 1987), Gunawardena (2003), Horn (1972), which relates networks that are weakly reversible and have a deficiency of zero to those that admit complex-balanced equilibria.

Theorem 3 *If the reaction network $\{\mathcal{S}, \mathcal{C}, \mathcal{R}\}$ is weakly reversible and has a deficiency of zero, then for any choice of rate constants κ_k the deterministic mass action system admits a complex-balanced equilibrium, c , satisfying (2). Moreover, within each positive stoichiometric compatibility class, there is precisely one equilibrium, and it is complex balanced.*

References

- Anderson DF (2007) A modified next reaction method for simulating chemical systems with time dependent propensities and delays. *J Chem Phys* 127(21):214107
- Anderson DF (2008) Global asymptotic stability for a class of nonlinear chemical equations. *SIAM J Appl Math* 68(5):1464–1476
- Anderson DF, Craciun G, Kurtz TG (2010) Product-form stationary distributions for deficiency zero chemical reaction networks. *Bull Math Biol* 72(8):1947–1970
- Anderson DF (2011) A proof of the global attractor conjecture in the single linkage class case. *Siam J Appl Math* 71(4):1487–1508
- Anderson DF, Craciun G, Gopalkrishnan M, Wiuf C (2015) Lyapunov functions, stationary distributions, and non-equilibrium potential for reaction networks. *Bull Math Bio* 77(9):1744–1767
- Anderson DF, Kurtz TG (2011) Continuous time Markov chain models for chemical reaction networks. In: Koeppl H, Densmore D, Setti G, di Bernardo M (eds) *Design and analysis of biomolecular circuits: engineering approaches to systems and synthetic biology*. Springer, Berlin, pp 3–42
- Anderson DF, Kurtz TG (2015) *Stochastic analysis of biochemical systems*, vol 1.2, 1st edn. Springer, Berlin
- Cao Y, Gillespie DT, Petzold LR (2005) The slow-scale stochastic simulation algorithm. *J Chem Phys* 122(1):014116
- Cappelletti D, Wiuf C (2016) Product-form poisson-like distributions and complex balanced reaction systems. *Siam J Appl Math* 76(1):411–432
- Cotter SL, Zygalakis KC, Kevrekidis IG, Erban R (2011) A constrained approach to multiscale stochastic simulation of chemically reacting systems. *J Chem Phys* 135(9):094102
- Cotter SL (2016) Constrained approximation of effective generators for multiscale stochastic reaction networks and application to conditioned path sampling. *J Comput Phys* 323:265–282
- Cotter SL, Erban R (2016) Error analysis of diffusion approximation methods for multiscale systems in reaction kinetics. *SIAM J Sci Comput* 38(1):144–163
- Craciun G (2015) Toric differential inclusions and a proof of the global attractor conjecture. [arXiv:1501.02860](https://arxiv.org/abs/1501.02860)

- Feinberg M (1972) Complex balancing in general kinetic systems. *Arch Ration Mech Anal* 49:187–194
- Feinberg M (1979) Lectures on chemical reaction networks. Delivered at the Mathematics Research Center, University of Wisconsin, Madison. <http://www.che.eng.ohio-state.edu/~feinberg/LecturesOnReactionNetworks>
- Feinberg M (1987) Chemical reaction network structure and the stability of complex isothermal reactors—I. The deficiency zero and deficiency one theorems, review article 25. *Chem Eng Sci* 42:2229–2268
- Gibson MA, Bruck J (2000) Efficient exact stochastic simulation of chemical systems with many species and many channels. *J Phys Chem A* 105:1876–1889
- Gillespie DT (1976) A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J Comput Phys* 22:403–434
- Gillespie DT (2000) The chemical Langevin equation. *J Chem Phys* 113(1):297–306
- Goutsias J (2005) Quasiequilibrium approximation of fast reaction kinetics in stochastic biochemical systems. *J Chem Phys* 122(18):184102
- Gunawardena J (2003) Chemical reaction network theory for in-silico biologists. Notes available for download at <http://vcp.med.harvard.edu/papers/crnt.pdf>
- Horn F (1972) Necessary and sufficient conditions for complex balancing in chemical kinetics. *Arch Ration Mech Anal* 49:172–186
- Janssen JAM (1989) The elimination of fast variables in complex chemical reactions. iii. mesoscopic level (irreducible case). *J Stat Phys* 57(1–2):187–198
- Kelly FP (1979) Reversibility and stochastic networks, Wiley series in probability and mathematical statistics. Wiley, Chichester
- Thomas P, Straube AV, Grima R (2012) The slow-scale linear noise approximation: an accurate, reduced stochastic description of biochemical networks under timescale separation conditions. *BMC Syst Biol* 6(1):1
- Van Kampen NG (1976) The equilibrium distribution of a chemical mixture. *Phys Lett A* 59(5):333–334
- Van Kampen NG (2007) Stochastic processes in physics and chemistry, 3rd edn. Elsevier, Amsterdam
- Weinan E, Liu D, Vanden-Eijnden E (2005) Nested stochastic simulation algorithm for chemical kinetic systems with disparate rates. *J Chem Phys* 123(19):194107
- Weinan E, Vanden-Eijnden E (2007) Nested stochastic simulation algorithms for chemical kinetic systems with multiple time scales. *J Comput Phys* 221(1):158–180