

# Moment Closure for the Stochastic Logistic Model <sup>1</sup>

Abhyudai Singh and João Pedro Hespanha <sup>2</sup>

*Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93101.*

---

## Abstract

Continuous-time birth-death Markov processes serve as useful models in population biology. When the birth-death rates are nonlinear, the time evolution of the first  $n$  order moments of the population is not closed, in the sense that it depends on moments of order higher than  $n$ . For analysis purpose, the time evolution of the first  $n$  order moments is often made to be closed by approximating these higher order moments as a nonlinear function of moments up to order  $n$ , which we refer to as the *moment closure function*.

In this paper, a systematic procedure for constructing moment closure functions of arbitrary order is presented for the stochastic logistic model. We obtain the moment closure function by first assuming a certain *separable* form for it, and then matching time derivatives of the exact (not closed) moment equations with that of the approximate (closed) equations for some initial time and set of initial conditions. The separable structure ensures that the steady-state solutions for the approximate equations are unique, positive and real, while the derivative matching guarantees a good approximation, at-least locally in time. Moreover, the accuracy of the approximation can be improved by increasing the order of the approximate model. To the best of our knowledge, this paper is the first to propose a systematic procedure to construct moment closure functions of arbitrary order that guarantee biologically meaningful equilibria.

A host of other moment closure functions previously proposed in the literature are also investigated. Among these we show that only the ones that achieve derivative matching provide a close approximation to the exact solution. Moreover, we improve the accuracy of several previously proposed moment closure functions by forcing derivative matching. However, for certain ranges of parameter models, moment closure functions that lack the separability property lead to biologically meaningless scenarios of imaginary and even stable negative steady-states of the closed moment equations.

*Key words:* Moment Closure, Stochastic Logistic Model, Stochastic Hybrid Systems

---

<sup>1</sup> This material is based upon work supported by the Institute for Collaborative Biotechnologies through grant DAAD19-03-D-0004 from the U.S. Army Research Office and by the National Science Foundation under Grant No. CCR-0311084.

<sup>2</sup> abhi@engineering.ucsb.edu, hespanha@ece.ucsb.edu

# 1 INTRODUCTION

Continuous-time birth-death Markov processes have been extensively used for modeling stochasticity in population biology (Matis and Kiffe, 2002, 1996; Matis et al., 1998). The time evolution of these processes is typically described by a single equation for a grand probability function, where time and species populations appear as independent variables, called the Master or Kolmogorov equation (Bailey, 1964). However, this equation can only be solved for relatively few, highly idealized cases and a more reasonable goal is to determine the time evolution of a few low-order statistical moments.

In this paper, a method for estimating low-order statistical moments is introduced for the *stochastic logistic model*. This model was first introduced by Barlett et al. (1960) and is a continuous-time birth-death Markov process involving a single specie, with birth and death rates being polynomials of order 2. Although, one can directly use the Kolmogorov equation to derive differential equations for the time evolution of moments of the process, in this paper we use an alternative method. We model the stochastic logistic model as a Stochastic Hybrid System (SHS) whose state  $\mathbf{x}$  is the population of the specie. Then, the time evolution of the moments is obtained using results from the SHS literature (e.g., Hespanha, 2004). Details of the stochastic logistic model and its modeling as a SHS are presented in Section 2.

Table 1

Separable derivative-matching moment closure function  $\varphi_{n+1}^s(\mu)$  for  $n \in \{2, 3, 4\}$ .

	$n = 2$	$n = 3$	$n = 4$
$\varphi_{n+1}^s(\mu)$	$\frac{\mu_2^3}{\mu_1^3}$	$\frac{\mu_1^4 \mu_3^4}{\mu_2^6}$	$\frac{\mu_2^{10} \mu_4^5}{\mu_1^5 \mu_3^{10}}$

Let  $\mu_m$  be the  $m^{th}$  order uncentered moment of  $\mathbf{x}$ , i.e.  $\mu_m(t) = \mathbf{E}[\mathbf{x}(t)^m]$  where  $\mathbf{x}(t)$  denotes the population of the specie at time  $t$ . We will show in Section 3 that for the stochastic logistic model the time derivative of  $\mu_m$  is a linear combination of moments up to order  $m + 1$ . Hence, if one stacks all moments in an infinite vector  $\mu_\infty = [\mu_1, \mu_2, \dots]^T$ , its dynamics can be written as

$$\dot{\mu}_\infty = A_\infty \mu_\infty, \tag{1}$$

for some infinite matrix  $A_\infty$ . As the above infinite dimensional system cannot be solved analytically, we truncate (1) by creating a vector  $\mu = [\mu_1, \dots, \mu_n]^T$ , where  $n$  is the order of the truncation. Its dynamics, given by

$$\dot{\mu} = A\mu + B\mu_{n+1} \tag{2}$$

for some matrices  $A$  and  $B$  is not closed, in the sense that the time evolution of the vector  $\mu$  depends on the  $n + 1^{th}$  order moment  $\mu_{n+1}$ . For analysis purposes, we close the above system by approximating  $\mu_{n+1}$  as a nonlinear function  $\varphi_{n+1}(\mu)$  of

moments up to order  $n$ . This procedure is commonly referred to as *moment closure*. We call  $\varphi_{n+1}(\mu)$  as the *moment closure function* for  $\mu_{n+1}$ . Let  $\mathbf{v} = [v_1, \dots, v_n]^T$  denote the state of the new closed system. Then, its dynamics is given by

$$\dot{\mathbf{v}} = A\mathbf{v} + B\varphi_{n+1}(\mathbf{v}) \quad (3)$$

and is referred to as the *truncated moment dynamics*. We denote the states of (2) and (3) using different symbols because  $\mu$  refers to the actual moment dynamics whereas  $\mathbf{v}$  to an approximated moment dynamics.

In Section 4, we consider moment closure functions which have the following separable form:

$$\varphi_{n+1}^s(\mu) = \mu_1^{\gamma_1} \mu_2^{\gamma_2} \dots \mu_n^{\gamma_n}$$

for appropriately chosen constants  $\gamma_m \in \mathbb{R}$ . These constants are obtained by matching time derivatives of  $\mu_{n+1}$  and  $\varphi_{n+1}^s(\mu)$  in (2) and (3) respectively, at some initial time  $t_0$ , for a deterministic initial condition  $\mathbf{x}(t_0) = x_0$  with probability one. The reason for this lies in the fact that the class of deterministic distributions forms a natural basis for the infinite dimensional space containing the vector  $\mu_\infty$ . We refer to this moment closure as the *separable derivative-matching moment closure*. We show that for all  $n \geq 2$ , this determines the function  $\varphi_{n+1}^s$  uniquely, which is independent of the birth and death rates. Table 1 shows the functions  $\varphi_{n+1}^s$  that we obtained for truncations of order  $n = 2, 3$  and 4. The striking feature of the separable derivative-matching moment closure is that the accuracy of the approximate moment dynamics improves by increasing order of truncation and the dependence of higher order moments on lower order ones is consistent with  $\mathbf{x}(t)$  being *lognormally distributed*, in spite of the fact that the derivative matching procedure used to construct  $\varphi_{n+1}^s$  did not make any assumption on the distribution of the population.

Alternative moment closure methods which have appeared in literature typically construct the moment closure functions  $\varphi$  by directly assuming the probability distribution to be normal (Whittle, 1957), lognormal (Keeling, 2000), poisson or binomial (Nasell, 2003a). We refer to them as *normal, lognormal, poisson and binomial moment closures* respectively and review them in Section 5. In Section 6, they are compared with the separable derivative-matching moment closure. In Section 6.1 the comparisons are done based on how well the moment closure function  $\varphi_{n+1}(\mu)$  approximates  $\mu_{n+1}$ . Towards that end, we introduce the error

$$e_{n+1}(t) := \mu_{n+1}(t) - \varphi_{n+1}(\mu(t)) = \sum_{i=0}^{\infty} \frac{(t-t_0)^i}{i!} \varepsilon_{n+1}^i(x_0)$$

where we expanded the error as a Taylor series with  $\varepsilon_{n+1}^i(x_0)$  defined to be

$$\varepsilon_{n+1}^i(x_0) := \left. \frac{d^i \mu_{n+1}(t)}{dt^i} \right|_{t=t_0} - \left. \frac{d^i \varphi_{n+1}(\mu(t))}{dt^i} \right|_{t=t_0}.$$

We call  $\varepsilon_{n+1}^i(x_0)$  the  $i^{\text{th}}$  order *derivative matching error*. Ideally one would like this error to be zero but this is generally not possible. When  $\mathbf{x}(t_0) = x_0$  with probability one, the derivative matching error is typically a polynomial in  $x_0$ . The lesser the order of this polynomial, the lesser is the error  $e_{n+1}(t)$ , and hence the better is  $\varphi_{n+1}(\mu)$  in approximating  $\mu_{n+1}$ .

We show that for  $n = 2$ , all the above moment closure functions perform derivative matching except the poisson moment closure function by Nasell (2003a). This is because, it has a  $0^{\text{th}}$  order derivative matching error  $\varepsilon_3^0(x_0)$  which grows linearly with  $x_0$  while for separable derivative-matching, lognormal, binomial and normal moment closure functions the  $0^{\text{th}}$  order error is always zero. Hence, Nasell's poisson moment closure function exhibits a larger initial error than the others. We propose an alternative poisson moment closure function, for which  $\varepsilon_3^0(x_0) = 0$ , and show that it performs better than the one proposed by Nasell (2003a).

Although the above moment closures provide good estimates for a second order of truncation ( $n = 2$ ), it is typically beneficial to consider  $n \geq 3$  because they lead to better moment approximations and reduce the errors by a few orders of magnitude. However, for distributions like the normal and the lognormal, which are characterized by less than 2 parameters,  $n \geq 3$  typically leads to multiple normal and lognormal moment closure functions. In Section 6.1.2 we illustrate how derivative matching can be used as a tool for gauging performances of these multiple moment closure functions and choosing the ones among them that yield the least derivative matching errors.

In particular, we show that among the moment closure functions consistent with the lognormal distribution, the separable derivative-matching moment closure function yields the least order polynomial for the derivative matching error, and hence, exhibits the best performance. Based on derivative matching we also propose families of normal moment closure functions which provide good approximations for  $\mu_{n+1}$ . Towards the end, for  $n = 3$ , we propose a new moment closure function, for which, unlike the lognormal or normal moment closure functions both the  $0^{\text{th}}$  and  $1^{\text{st}}$  order derivative matching error are zero, and hence, provides the best estimates for  $\mu_4$  as compared to them, at least locally in time.

In Section 6.2 comparisons are done based on the steady-state solutions of the truncated moment dynamics (3). We show that the separable derivative matching moment closure, always yields a unique non-trivial positive real steady-state  $\forall n \in \mathbb{N}_{\geq 2}$ . Thus, it is preferable to the other moment closures techniques which lack the separable structure and exhibit spurious, imaginary and even stable negative steady-states, which would be biologically meaningless.

## 2 Stochastic Logistic Model

### 2.1 Model formulation

The stochastic logistic model is the stochastic birth-death analogous model of the well-known deterministic Verhulst-Pearl equations (Pielou, 1977) and has been extensively used for modeling stochasticity in population biology (Matis and Kiffe, 2002, 1996; Matis et al., 1998). For this continuous-time birth-death Markov process, the conditional probabilities of a unit increase and decrease, respectively, in an “infinitesimal” time interval  $(t, t + dt]$  is given by

$$\begin{aligned} P\{\mathbf{x}(t + dt) = x + 1 | \mathbf{x}(t) = x\} &= \begin{cases} \eta(x)dt, & \forall x \leq U \\ 0, & \text{otherwise} \end{cases} \\ P\{\mathbf{x}(t + dt) = x - 1 | \mathbf{x}(t) = x\} &= \chi(x)dt, \end{aligned}$$

where  $\mathbf{x}(t) \in \mathbb{N}$  represents the population size at time  $t$ ,

$$\eta(x) := a_1x - b_1x^2 > 0, \quad \chi(x) := a_2x + b_2x^2 > 0, \quad \forall x \in (0, U) \quad (4)$$

and

$$U := a_1/b_1 \in \mathbb{N}, \quad a_1 > 0, \quad a_2 > 0, \quad b_1 > 0, \quad b_2 \geq 0.$$

We assume that the initial condition satisfies  $\mathbf{x}(t_0) \in \{1, 2, \dots, U\}$ , and hence,  $\mathbf{x}(t) \in \{0, 1, \dots, U\}$ ,  $\forall t \in [0, \infty)$  with probability one. We call  $U$  the *population limit*.

### 2.2 Stationary and Quasi-Stationary Distributions

Since the birth and death rates are zero for  $x = 0$  ( $\eta(0) = \chi(0) = 0$ ) we have that  $\mathbf{x} = 0$  is an absorbing state and eventual convergence to the origin is certain. Thus, the stationary distribution is degenerate with probability one at the origin. However, though out this paper we assume the mean time to extinction to be very large, in which case there exists a “quasi-stationary” distribution (Nasell, 2001). Barlett et al. (1960) shows that a good approximate for  $\bar{P}_x$  (the probability that  $\mathbf{x} = x$  at “quasi-equilibrium”) can be numerically obtained using the following recurrence relationship

$$\chi(x)\bar{P}_x = \eta(x-1)\bar{P}_{x-1}, \quad \forall x \in \{2, 3, \dots, U\}. \quad (5)$$

### 2.3 Transient Distributions

Ideally one would like to determine the exact probability distribution of  $\mathbf{x}(t)$  at any time  $t$ . It can be shown that the probability  $P_x(t)$  that  $\mathbf{x}(t) = x$  satisfies the following differential equations

$$\dot{P}_0(t) = \chi(1)P_1(t) \quad (6a)$$

$$\dot{P}_1(t) = \chi(2)P_2(t) - [\eta(1) + \chi(1)]P_1(t) \quad (6b)$$

$$\dot{P}_x(t) = \chi(x+1)P_{x+1}(t) - [\eta(x) + \chi(x)]P_x(t) + \eta(x-1)P_{x-1}(t), \quad (6c)$$

$\vdots$

$$\dot{P}_U(t) = -\chi(U)P_U(t) + \eta(U-1)P_{U-1}(t) \quad (6d)$$

(Bailey, 1964). These differential equations are commonly referred to as the *Master or Kolmogorov equations*. If the population limit  $U$  is small, the above system of equations can be solved numerically. However, for large  $U$ , a more reasonable goal (and one that is of primary interest in applications) is to determine the evolution of the some lower-order moments of  $\mathbf{x}(t)$ .

## 3 Time Evolution of Moments

### 3.1 Modeling the Stochastic Logistic Model

To model the time evolution of  $\mathbf{x}(t)$ , we consider a special class of systems known as Stochastic Hybrid Systems (SHS). These systems were introduced by Hespanha and Singh (2005) to model the stochastic time evolution of the populations of different species involved in a chemical reaction. More specifically, to fit the framework of our problem, these system are characterized by two *reset maps*:

$$\mathbf{x} \mapsto \phi_1(\mathbf{x}) := \mathbf{x} + 1, \quad \mathbf{x} \mapsto \phi_2(\mathbf{x}) := \mathbf{x} - 1 \quad (7)$$

one corresponding to a birth and the other to a death, with associated *transition intensities* given by

$$\lambda_1(\mathbf{x}) := \eta(\mathbf{x}), \quad \lambda_2(\mathbf{x}) := \chi(\mathbf{x}). \quad (8)$$

Between births and deaths the population remains constant and thus  $\dot{\mathbf{x}} = 0$ . In essence, whenever a “birth event” or a “death event” takes place, the corresponding reset  $\phi_i(\mathbf{x})$  is “activated” and  $\mathbf{x}$  is reset accordingly, furthermore, the probability of the activation taking place in an “infinitesimal” time interval  $(t, t + dt]$  is determined by the associated transition intensities  $\lambda_i(\mathbf{x})dt$ .

### 3.2 Moment Dynamics

Given  $m \in \{1, 2, \dots\}$ , we define the  $m^{\text{th}}$  order (*uncentered*) moment of  $\mathbf{x}$  to be

$$\mu_m(t) = \sum_{x=1}^{\infty} x^m P_x(t) := \mathbf{E}[\mathbf{x}(t)^m], \quad \forall t \geq 0. \quad (9)$$

The time evolution of moments is given by the following result, which is a straightforward application of Theorem 1 in Hespanha (2004) to the above SHS.

**Theorem 1:** The time evolution of  $\mu_m$  is given by

$$\frac{d\mu_m}{dt} = \mathbf{E} \left[ \sum_{i=1}^2 [(\phi_i(\mathbf{x}))^m - \mathbf{x}^m] \lambda_i(\mathbf{x}) \right]. \quad (10)$$

■

Using the above Theorem, we show in Appendix A that one can conclude

$$\dot{\mu}_m = \sum_{p=1}^2 \sum_{r=1}^{m+1} C_{m+p-r}^m f(m+p-r, p) \mu_r, \quad (11)$$

where we define  $C_j^m$  and  $f(j, p)$  as follows<sup>3</sup>  $\forall j, m, p \in \mathbb{N}$ .

$$C_j^m := \begin{cases} \frac{m!}{(m-j)!j!} & m \geq j \geq 0 \\ 0 & m < j \end{cases} \quad (12)$$

$$f(j, p) := \begin{cases} 0 & j = 0 \\ a_1 + (-1)^j a_2 & j > 0, p = 1 \\ -b_1 + (-1)^j b_2 & j > 0, p = 2. \end{cases} \quad (13)$$

One can see from the right-hand-side of (11), that the time derivative of  $\mu_m$  is a *linear combination* of the moments  $\mu_r$ , up to order  $r = m + 1$ . Hence, if one stacks all moments in an infinite vector  $\mu_\infty = [\mu_1, \mu_2, \dots]^T$ , its dynamics can be written as

$$\dot{\mu}_\infty = A_\infty \mu_\infty, \quad (14)$$

for some infinite matrix  $A_\infty$ . Let  $\mu = [\mu_1, \mu_2, \dots, \mu_n]^T \in \mathbb{R}^n$  contains the top  $n$  elements of  $\mu_\infty$ . Then, using (11) the evolution of  $\mu$  is given by

$$\dot{\mu} = A\mu + B\mu_{n+1}, \quad (15)$$

<sup>3</sup>  $n!$  denotes the factorial of  $n$ .

for some  $n \times n$  and  $n \times 1$  matrices  $A$  and  $B$  which have the following structure

$$A = \begin{bmatrix} * & * & 0 & \dots & 0 \\ * & * & * & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ * \end{bmatrix}, \quad (16)$$

where the  $*$  denotes non-zero entries. Our goal is to approximate (15) by a finite-dimensional nonlinear ODE of the form

$$\dot{\mathbf{v}} = A\mathbf{v} + B\varphi_{n+1}(\mathbf{v}), \quad \mathbf{v} = [v_1, v_2, \dots, v_n]^T \quad (17)$$

where the map  $\varphi_{n+1} : \mathbb{R}^n \rightarrow \mathbb{R}$  should be chosen so as to keep  $\mathbf{v}(t)$  close to  $\mu(t)$ . This procedure is commonly referred to as *moment closure*. We call (17) the *truncated moment dynamics* and  $\varphi_{n+1}(\mu)$  the *moment closure function* for  $\mu_{n+1}$ .

When a sufficiently large but *finite* number of derivatives of  $\mu(t)$  and  $\mathbf{v}(t)$  match point-wise, then, the difference between solutions to (15) and (17) remains close on a given compact time interval. This follows from a Taylor series approximation argument. To be more precise, for each  $\delta > 0$  and integer  $N$ , there exists  $T \in \mathbb{R}$ , for which the following result holds: Assume that for every  $t_0 \geq 0$ ,

$$\mu(t_0) = \mathbf{v}(t_0) \quad \text{and} \quad \left. \frac{d^i \mu(t)}{dt^i} \right|_{t=t_0} = \left. \frac{d^i \mathbf{v}(t)}{dt^i} \right|_{t=t_0}, \quad \forall i \in \{1, \dots, N\} \quad (18)$$

where  $\frac{d^i \mu(t)}{dt^i}$  and  $\frac{d^i \mathbf{v}(t)}{dt^i}$  represent the  $i^{\text{th}}$  time derivative of  $\mu(t)$  and  $\mathbf{v}(t)$  along the trajectories of system (14) and (17) respectively. Then,

$$\|\mu(t) - \mathbf{v}(t)\| \leq \delta, \quad \forall t \in [t_0, T], \quad (19)$$

along solutions of (14) and (17), where  $\mu$  denote the first  $n$  elements of  $\mu_\infty$ . In the next section we use (18) to construct *moment closure functions*  $\varphi_{n+1}(\mu)$ .

#### 4 Separable Derivative-Matching Moment Closures

In this section we construct truncated moment dynamics (17) for the stochastic logistic model using (18). After replacing (15) and (17) in (18), equality (18) becomes a PDE on  $\varphi_{n+1}$ . We will seek for solutions  $\varphi_{n+1}$  to this PDE that have the following separable form

$$\varphi_{n+1}^s(\mu) = \prod_{m=1}^n \mu_m^{\gamma_m} \quad (20)$$

for appropriately chosen constants  $\gamma_m \in \mathbb{R}$ . As we will see in Section 6.2, this separable structure ensures that the steady-state solutions for the truncated moment dynamics (17) are unique, positive real, and hence, biologically meaningful. In the sequel we refer to such  $\varphi_{n+1}^s(\boldsymbol{\mu})$  as a *separable derivative-matching moment closure function* for  $\boldsymbol{\mu}_{n+1}$ .

One can see that the infinite vector  $\boldsymbol{\mu}_\infty \in \Omega_\infty$  can be expressed as

$$\boldsymbol{\mu}_\infty = \begin{bmatrix} \mathbf{E}[\mathbf{x}(t)] \\ \mathbf{E}[\mathbf{x}(t)^2] \\ \mathbf{E}[\mathbf{x}(t)^3] \\ \vdots \end{bmatrix} = \sum_{x=1}^{\infty} \begin{bmatrix} x \\ x^2 \\ x^3 \\ \vdots \end{bmatrix} P_x(t).$$

Hence, the infinite vectors  $\boldsymbol{\mu}_\infty = [x, x^2, x^3, \dots]^T$ , which corresponds to a deterministic distribution, i.e.  $\mathbf{x}(t) = x$  with probability one, form a natural basis for  $\Omega_\infty$ . In particular, we will find constants  $\gamma_m$  which satisfy (18) for each vector  $\boldsymbol{\mu}_\infty(t_0)$  belonging to this basis, i.e. for the class of deterministic initial conditions. However, often it is not possible to find  $\gamma_m$  for which (18) holds exactly. We will therefore relax this condition and simply demand the following

$$\boldsymbol{\mu}(t_0) = \mathbf{v}(t_0) \quad \text{and} \quad \left. \frac{d^i \boldsymbol{\mu}(t)}{dt^i} \right|_{t=t_0} = \left. \frac{d^i \mathbf{v}(t)}{dt^i} \right|_{t=t_0} + \mathbf{E}[\boldsymbol{\varepsilon}_i(\mathbf{x}(t_0))], \quad (21)$$

$\forall i \in \{1, 2\}$ , where each element of the vector  $\boldsymbol{\varepsilon}_i(\mathbf{x}(t_0))$  is a polynomial in  $\mathbf{x}(t_0)$ . One can think of (21) as an approximation to (18) that is valid as long as  $\left. \frac{d^i \boldsymbol{\mu}(t)}{dt^i} \right|_{t=t_0}$  dominates  $\mathbf{E}[\boldsymbol{\varepsilon}_i(\mathbf{x}(t_0))]$ .

The following theorem summarizes the main result, the proof of which is given in Appendix B.

**Theorem 2:** Let  $\gamma_m, m \in \{1, \dots, n\}$  be chosen as

$$\gamma_m = (-1)^{n-m} \mathbf{C}_m^{n+1}. \quad (22)$$

Then, for every deterministic initial condition  $\mathbf{v}(t_0) = \boldsymbol{\mu}(t_0) = [x_0, \dots, x_0^n]^T$  which corresponds to  $\mathbf{x}(t_0) = x_0$  with probability one, we have

$$\left. \frac{d\boldsymbol{\mu}(t)}{dt} \right|_{t=t_0} = \left. \frac{d\mathbf{v}(t)}{dt} \right|_{t=t_0} \quad (23a)$$

$$\left. \frac{d^2 \boldsymbol{\mu}(t)}{dt^2} \right|_{t=t_0} = \left. \frac{d^2 \mathbf{v}(t)}{dt^2} \right|_{t=t_0} + \boldsymbol{\varepsilon}_2(x_0), \quad (23b)$$

where  $\left. \frac{d^i \boldsymbol{\mu}(t)}{dt^i} \right|_{t=t_0}$  and  $\left. \frac{d^i \mathbf{v}(t)}{dt^i} \right|_{t=t_0}$  represent the  $i^{\text{th}}$  time derivative of  $\boldsymbol{\mu}(t)$  and  $\mathbf{v}(t)$  along the trajectories of the systems (14) and (17), respectively, and the  $n^{\text{th}}$  element of the vector  $\boldsymbol{\varepsilon}_2(x_0)$  is a polynomial in  $x_0$  of order 2 with all other elements being zero. ■

*Remark 1.* Using (11) it can be shown that  $\frac{d^2\mu_n(t)}{dt^2}$  is a linear combination of moments of  $\mathbf{x}$  up to order  $n+2$ . Thus,  $\frac{d^2\mu_n(t)}{dt^2}|_{t=t_0}$  is a polynomial in  $x_0$  of order  $n+2$ , and hence, for<sup>4</sup>  $x_0 \gg 1$ ,  $\varepsilon_2^n(x_0)/\frac{d^2\mu_n(t)}{dt^2}|_{t=t_0} = O(x_0^{-n})$ , where  $\varepsilon_2^n(x_0)$  is the  $n^{\text{th}}$  element of  $\varepsilon_2(x_0)$ . Hence, the term  $\frac{d^2\mu_n(t)}{dt^2}|_{t=t_0}$  dominates  $\varepsilon_2^n(x_0)$  by  $x_0^{-n}$ . ■

*Remark 2.* It can be verified that the separable derivative-matching moment closures also matches derivatives of order higher than 2 in (23) with small errors. For example for  $n \in \{2, 3, 4\}$  and  $i \in \{2, \dots, 9\}$ , we have

$$\frac{d\mu(t)}{dt}\Big|_{t=t_0} = \frac{dv(t)}{dt}\Big|_{t=t_0} \quad (24a)$$

$$\frac{d^i\mu(t)}{dt^i}\Big|_{t=t_0} = \frac{d^i v(t)}{dt^i}\Big|_{t=t_0} + \varepsilon_i(x_0), \quad (24b)$$

where the  $m^{\text{th}}$  element of  $\varepsilon_i(x_0)$  is a polynomial in  $x_0$  of order  $m-n+i$ , for  $m-n+i \geq 2$  and equal to zero otherwise. We conjecture that the above equality holds  $\forall n \in \mathbb{N}$  and  $\forall i \in \mathbb{N}$  but we only verified it for  $n$  up to 4 and  $i$  up to 9. As above, the elements of vector  $\frac{d^i\mu(t)}{dt^i}|_{t=t_0}$  dominate the corresponding elements of  $\varepsilon_i(x_0)$  by  $x_0^{-n}$ , and hence, with increasing  $n$ , the truncated moment dynamics  $v(t)$  should provide a more accurate approximations to the lower order moments  $\mu(t)$ . ■

## 5 Distribution based Moment Closures

Most moment closure techniques that appeared in the literature start by assuming a specific class of distributions  $\mathbb{D}$  for the population, and use this assumption to express higher order moments as a function of the lower order ones. We refer to such a moment closure function as being consistent with the distribution  $\mathbb{D}$  and define them as follows.

**Definition:** Let  $\mathbb{D}$  be a class of distributions parameterized by  $m$  parameters  $(q_1, \dots, q_m) \in \mathcal{Q}$ , with the  $k^{\text{th}}$  order moment  $\mu_k$  given in terms of the  $q_1, \dots, q_m$  as follows

$$\mu_k = f_k(q_1, \dots, q_m), \quad \forall k \in \{1, 2, \dots\}. \quad (25)$$

The moment closure function  $\varphi_{n+1}^{\mathbb{D}}(\mu)$  for  $\mu_{n+1}$  is said to be *consistent with the distribution*  $\mathbb{D}$  if, for every  $(q_1, \dots, q_m) \in \mathcal{Q}$ , one has that

$$\mu_{n+1} = f_{n+1}(q_1, \dots, q_m) = \varphi_{n+1}^{\mathbb{D}}(\mu) \quad (26)$$

<sup>4</sup>  $O(\cdot)$  denotes order of.

Table 2

Unique Moment Closure Functions for the  $m^{\text{th}}$  order truncation ( $n = m$ ) and different distributions  $\mathbb{D}$ .

$\mathbb{D}$	$m$	Unique moment closure function
Normal	2	$\varphi_3^g(\mu) = 3\mu_2\mu_1 - 2\mu_1^3$
Lognormal	2	$\varphi_3^l(\mu) = \frac{\mu_2^3}{\mu_1^3}$
Poisson	1	$\varphi_2^p(\mu) = \mu_1^2 + \mu_1$
Binomial	2	$\varphi_3^b(\mu) = 2\frac{(\mu_2 - \mu_1^2)^2}{\mu_1} - (\mu_2 - \mu_1^2) + 3\mu_1\mu_2 - 2\mu_1^3$

where

$$\mu := \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix} = \begin{bmatrix} f_1(q_1, \dots, q_m) \\ \vdots \\ f_n(q_1, \dots, q_m) \end{bmatrix}. \quad (27)$$

For well known classes of distributions — such as lognormal, normal, poisson, or binomial — we simply say that  $\varphi_{n+1}^{\mathbb{D}}$  is the lognormal, normal, poisson, or binomial moment closure function.

### 5.1 Techniques for obtaining $\varphi_{n+1}^{\mathbb{D}}$

Generically, when the dimension  $m$  of the parameter space  $\mathcal{Q}$  is the same as the dimension  $n$  of the domain of the moment closure function  $\varphi_{n+1}^{\mathbb{D}}(\mu)$ , the functional equation (26)–(27) in the “unknown”  $\varphi_{n+1}^{\mathbb{D}}(\cdot)$  has a unique solution<sup>5</sup>. In fact, to determine  $\varphi_{n+1}^{\mathbb{D}}(\mu)$  one can start by solving (27) for  $q_1, \dots, q_m$  in terms of  $\mu_1, \dots, \mu_m$ , and then substituting these back in (26) to obtain a unique moment closure function  $\varphi_{n+1}^{\mathbb{D}}(\cdot)$ . As we choose  $\mathbb{D}$  to be normal (Whittle, 1957), log-normal (Keeling, 2000), poisson, or binomial (Nasell, 2003a), this procedure results in the different moment closure functions shown in Table 2.

Difficulties arise when the dimension  $m$  of the parameter space  $\mathcal{Q}$  is strictly smaller than the dimension  $n$  of the domain of the moment closure function  $\varphi_{n+1}^{\mathbb{D}}(\mu)$ , because in this case the functional equation (26)–(27) does not have a unique solution<sup>6</sup> and one can find infinitely many moment closure functions consistent with

<sup>5</sup> The reader is invited to convince herself of this by imagining that all functions are locally linear, in which case  $\varphi_{n+1}^{\mathbb{D}}(\mu)$  is represented by a vector with  $n$  unknowns. Since (26) must hold for a (local) basis of  $\mathcal{Q}$ , we have exactly  $m$  equations to determine the  $n = m$  unknowns.

<sup>6</sup> In contrast to the previous case, we now have  $n$  unknowns to represent the local lin-

the same family of distributions. However, there is a strong incentive to consider this case because, as well shall see shortly, large values for  $n$  generally lead to significantly more accurate moment closures. In the sequel, we illustrate some options for moment closures with  $n > m$ ,

*Example 1:* Consider the class of poisson distributions characterized by their expected value  $\theta$  ( $m = 1$ ). Their moments are given by

$$\begin{aligned}\mu_1 &= f_1(\theta) := \theta \\ \mu_2 &= f_2(\theta) := \theta(1 + \theta) \\ \mu_3 &= f_3(\theta) := \theta(1 + 3\theta + \theta^2), \dots\end{aligned}$$

Nasell (2003a) proposed the following poisson moment closure function for  $n = 2$ :

$$\varphi_3^{p1}(\mu) = \mu_1 + 3\mu_1\mu_2 - 2\mu_1^3, \quad (28)$$

for which it is straightforward to verify that (26)–(27) holds because

$$\mu_3 = \theta(1 + 3\theta + \theta^2) = \varphi_3^{p1}(\mu), \quad \mu = [\theta, \theta(1 + \theta)]^T.$$

However, an alternative choice for the poisson moment closure function that also satisfies (26)–(27) is given by

$$\varphi_3^{p2}(\mu) = \mu_2 - \mu_1^2 + 3\mu_1\mu_2 - 2\mu_1^3, \quad (29)$$

which, as we will see in the next section, performs better than (28). The explanation for this lies in the fact that (29) has better derivative matching properties than (28), in the sense of (21). In the sequel we refer to (28) and (29) as the Nasell-poisson and new-poisson moment closure function, respectively.

*Example 2:* Consider now the class of normal distributions parameterized by their mean  $\omega$  and variance  $\sigma^2$  ( $m = 2$ ). Their moments are given by

$$\begin{aligned}\mu_1 &= f_1(\theta, \sigma) := \omega \\ \mu_2 &= f_2(\theta, \sigma) := \omega^2 + \sigma^2 \\ \mu_3 &= f_3(\theta, \sigma) := \omega(\omega^2 + 3\sigma^2) \\ \mu_4 &= f_4(\theta, \sigma) := \omega^4 + 6\omega^2\sigma^2 + 3\sigma^4, \dots\end{aligned}$$

For  $n = 3$ , any function of the following form is a normal moment closure function for  $\mu_4$

$$\varphi_4(\mu) := 4\mu_1\mu_3 + 3\mu_2^2 - 12\mu_2\mu_1^2 + 6\mu_1^4 + h(\mu_1, \mu_2 - \mu_1^2, \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3)$$

parameterization of  $\varphi_{n+1}^{\mathbb{D}}(\mu)$  but a (local) basis of  $\mathcal{Q}$  only provides  $m < n$  equations.

where  $h(x, y, z)$  is any function with the property that  $h(x, y, 0) = 0$ . To verify that this is so, we note that (26)–(27) holds because for

$$\mu = [\omega, \quad \omega^2 + \sigma^2, \quad \omega(\omega^2 + 3\sigma^2)]^T,$$

we obtain  $h(\mu_1, \mu_2 - \mu_1^2, \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3) = h(\omega, \sigma^2, 0) = 0$  and therefore

$$\varphi_4(\mu) = \omega^4 + 6\omega^2\sigma^2 + 3\sigma^4 = \mu_4.$$

*Example 3:* Finally consider the class of lognormal distributions characterized by the parameters  $\alpha > 0$ ,  $\beta > 0$  ( $m = 2$ ), whose moments are given by

$$\mu_k = f_k(\alpha, \beta) := \alpha^k \beta^{k^2}, \quad \forall k \in \{1, 2, \dots\}. \quad (30)$$

For every  $n \geq 2$ , the separable derivative-matching moment closure functions defined by (20), with coefficient given by (22) in Theorem 2 are lognormal moment closure functions for  $\mu_{n+1}$ . We can verify this by noting that (26)–(27) holds because, from (30) and (B.1) used in the proof of Theorem 2 (Appendix B) we conclude that

$$\begin{aligned} \varphi_{n+1}^s(\mu) &= \alpha^{(\sum_{m=1}^n \gamma_m m)} \beta^{(\sum_{m=1}^n \gamma_m m^2)} \\ &= \alpha^{(\sum_{m=1}^n \gamma_m C_1^m)} \beta^{(\sum_{m=1}^n \gamma_m \{2C_2^m + C_1^m\})}, \\ &= \alpha^{n+1} \beta^{2C_2^{n+1} + C_1^{n+1}} = \alpha^{n+1} \beta^{(n+1)^2} = \mu_{n+1}, \end{aligned}$$

where we used the facts that  $k = C_1^k$ ,  $k^2 = 2C_2^k + C_1^k$ ,  $\forall k \in \mathbb{N}$ .

## 5.2 Cumulant closure functions

Most of the literature on moment closure prefers to work with a vector  $\kappa = [\kappa_1, \dots, \kappa_n]$  where  $\kappa_n(t)$  is the  $n^{\text{th}}$  order cumulant<sup>7</sup>, instead of the previously introduced vector  $\mu$  of uncentered moments in (15). Then, instead of doing moment closure one performs cumulant closure by approximating  $\kappa_{n+1}$  by a nonlinear function  $\phi_{n+1}(\kappa)$  of  $\kappa_1, \dots, \kappa_n$ , which we refer to as the *cumulant closure function*. The disadvantage of working with  $\kappa$  instead of  $\mu$  is that the dynamics of  $\kappa$  is always nonlinear. However, for ease of comparison with other papers, we provide in Table 2 the cumulant closure functions corresponding to the different moment closure functions

<sup>7</sup> The  $n^{\text{th}}$  order cumulant,  $\kappa_n$  is given as follows in terms of the uncentered moments

$$\begin{aligned} \kappa_1 &= \mu_1, & \kappa_2 &= \mu_2 - \mu_1^2 \\ \kappa_3 &= \mu_3 - 3\mu_1\mu_2 + 2\mu_1^3, & \kappa_4 &= \mu_4 - 4\mu_1\mu_3 - 3\mu_2^2 + 12\mu_2\mu_1^2 - 6\mu_1^4, \dots \end{aligned}$$

Table 3

Moment Closure Functions (MCF) for second order truncation ( $n = 2$ ) and corresponding Cumulant Closure Functions (CCF) for  $\mu_3$  and  $\kappa_3$ , respectively, corresponding to the different Moment Closure Techniques (MCT) discussed in this paper. SDM refers to separable derivative-matching.

MCT	MCF	CCF
SDM	$\varphi_3^s(\mu) = \frac{\mu_2^3}{\mu_1^3}$	$\phi_3^s(\kappa) = 3\frac{\kappa_2^2}{\kappa_1} + \frac{\kappa_2^3}{\kappa_1^3}$
Normal	$\varphi_3^g(\mu) = 3\mu_2\mu_1 - 2\mu_1^3$	$\phi_3^g(\kappa) = 0$
Lognormal	$\varphi_3^l(\mu) = \frac{\mu_2^3}{\mu_1^3}$	$\phi_3^l(\mu) = 3\frac{\kappa_2^2}{\kappa_1} + \frac{\kappa_2^3}{\kappa_1^3}$
Nasell-Poisson	$\varphi_3^{p1}(\mu) = \mu_1 + 3\mu_1\mu_2 - 2\mu_1^3$	$\phi_3^{p1}(\kappa) = \kappa_1$
New-Poisson	$\varphi_3^{p2}(\mu) = \mu_2 - \mu_1^2 + 3\mu_1\mu_2 - 2\mu_1^3$	$\phi_3^{p2}(\kappa) = \kappa_2$
Binomial	$\varphi_3^b(\mu) = 2\frac{(\mu_2 - \mu_1^2)^2}{\mu_1} - (\mu_2 - \mu_1^2) + 3\mu_1\mu_2 - 2\mu_1^3$	$\phi_3^b(\mu) = 2\frac{\kappa_2^2}{\kappa_1} - \kappa_2$

discussed so far for  $n = 2$ . We use superscripts  $s$ ,  $l$ ,  $g$ ,  $p1$ ,  $p2$  and  $b$  to denote separable derivative-matching, lognormal, normal, Nasell-poisson, new-poisson and binomial moment closure functions, respectively.

## 6 Comparison of Moment closures

In this section, we introduce two criteria to compare the different moment closure techniques. The first criterion is the error

$$e_{n+1}(t) := \mu_{n+1}(t) - \varphi_{n+1}(\mu(t)) = \sum_{i=0}^{\infty} \frac{(t-t_0)^i}{i!} \varepsilon_{n+1}^i(x_0), \quad (31)$$

where

$$\varepsilon_{n+1}^i(x_0) := \left. \frac{d^i \mu_{n+1}(t)}{dt^i} \right|_{t=t_0} - \left. \frac{d^i \varphi_{n+1}(\mu(t))}{dt^i} \right|_{t=t_0}. \quad (32)$$

We call  $\varepsilon_{n+1}^i(x_0)$  the *derivative matching error*. Ideally, one would like to have  $\varepsilon_{n+1}^i(x_0) = 0$ , but as already pointed out in Section 4, this is generally not possible. With deterministic initial conditions  $\mu_{\infty}(t_0) = [x_0, x_0^2, \dots]^T$  as in Theorem 2, the derivative matching error is typically a polynomial in  $x_0$ . The lesser the order of this polynomial, the better is  $\varphi_{n+1}(\mu)$  in approximating  $\mu_{n+1}$ .

The second criterion is the steady-state solution of the truncated moment dynamics (17). In particular, we are interested in determining if there exists a unique non-trivial positive real steady-state which is physically meaningful. This is important,

because it is well-known that normal moment closures can have spurious, imaginary and sometimes even stable negative steady-states, which lead to biologically meaningless solutions (Keeling, 2000).

## 6.1 Derivative Matching Error

### 6.1.1 Moment Closures for $n = 2$

We recall from Table 3 that  $\varphi_3^l(\mu) = \varphi_3^s(\mu)$ , and therefore we do not need to discuss lognormal moment closure separately. By substituting  $\varphi_3^s(\mu)$ ,  $\varphi_3^g(\mu)$ ,  $\varphi_3^{p1}(\mu)$ ,  $\varphi_3^{p2}(\mu)$  and  $\varphi_3^b(\mu)$  from Table 3 in (31)–(32), one obtains the corresponding derivative matching errors, which will be denoted using the appropriate superscripts.

Using Table 3 and symbolic manipulation in *Mathematica*, we can show that

$${}^s\mathcal{E}_3^0(x_0) = {}^g\mathcal{E}_3^0(x_0) = {}^{p2}\mathcal{E}_3^0(x_0) = {}^b\mathcal{E}_3^0(x_0) = 0 \quad (33a)$$

$${}^{p1}\mathcal{E}_3^0(x_0) = -x_0. \quad (33b)$$

$${}^*\mathcal{E}_3^i(x_0) \in P_{x_0}(i+1), \quad * = \{s, g, p1, p2, b\}, \quad \forall i \in \{1, 2, \dots\} \quad (33c)$$

where  $P_{x_0}(j)$  denotes the set of polynomials in  $x_0$  of order  $j$ . Since  ${}^{p1}\mathcal{E}_3^0(x_0) = -x_0$ , the Nasell-poisson moment closure function will have a large initial error, especially for large initial conditions, when compared to all other moment closure functions. For all,  $i \in \{1, 2, \dots\}$ , all of these moment closure functions match derivatives, with the derivative matching error being of the same order in  $x_0$ . The simulation results discussed below show that with the exception of Nasell-poisson moment closure function, which consistently provides the worst estimates, all other moment closure functions perform fairly well.

*Example:* We consider the stochastic logistic model with

$$a_1 = .30, \quad a_2 = .02, \quad b_1 = .015, \quad b_2 = .001, \quad (34)$$

which is used by Matis et al. (1998) to model the population dynamics of the African Honey Bee. Using (17) with the matrices  $A$  and  $B$  computed in (11), we have the following truncated moment dynamics

$$\begin{bmatrix} \dot{v}_1 \\ \dot{v}_2 \end{bmatrix} = \begin{bmatrix} 0.28 & -0.016 \\ .32 & .546 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} + \begin{bmatrix} 0 \\ -0.032 \end{bmatrix} \varphi_3(v).$$

The time evolution of the moments corresponding to different moment closure techniques is obtained by substituting the appropriate moment closure function from Table 3 in place of  $\varphi_3(v)$ . Figure 1 plots the different errors (31) during the time interval  $[0, 2.5]$  for  $x_0 = 5$ . The error is approximated by the first nine terms of

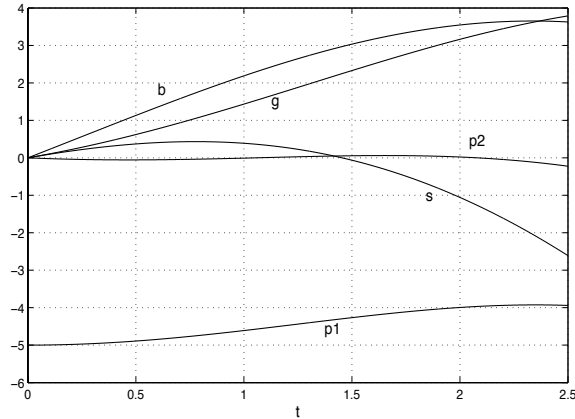


Fig. 1. Evolutions of the error  $e_{n+1}$  in (31) for the different moment closure techniques in Table 3, with parameters as in (34) and  $x_0 = 5$ .

the series in (31). The letters  $s$ ,  $g$ ,  $p1$ ,  $p2$  and  $b$  are used to denote the errors corresponding to separable derivative-matching, normal, Nasell-poisson, new-poisson and binomial moment closure functions, respectively. In order to evaluate the performance of these moment closures techniques past the initial period, we also compute the exact evolution of the moments. This is only possible because the population limit  $U = 25$  is small and one can obtain the exact solution by numerically solving the Kolmogorov equation (6). Figure 2 contains plots of the mean and variance errors, respectively, for the different moment closure functions with  $x_0 = 5$  and  $x_0 = 20$ . The letters  $s$ ,  $g$ ,  $p1$ ,  $p2$  and  $b$  are used to denote the errors corresponding to separable derivative-matching, normal, Nasell-poisson, new-poisson and binomial moment closure functions, respectively. For  $x_0 = 20$  the binomial moment closure function provides the best estimate both initially and at steady-state, whereas for  $x_0 = 5$  the new-poisson moment closure function does best initially, but the binomial moment closure function continues to provide the most accurate steady-state estimate. As one would expect from (33), the Nasell-poisson moment closure function performs the worst.

### 6.1.2 Moment Closures for $n \geq 3$

All distributions  $\mathbb{D}$  discussed in Section 5 are characterized by at most 2 parameters ( $m \leq 2$ ), hence the corresponding distribution-based moment closure functions  $\varphi_{n+1}^{\mathbb{D}}$  are not unique for  $n \geq 3$ . In this section we illustrate for lognormal and normal moment closures, how derivative matching can be used as a tool for gauging their performances and choosing the ones that yield the least derivative matching errors. We also propose a new moment closure function for  $n = 3$ , which as we will see guarantees better approximation at least locally in time, as compared to other ones.

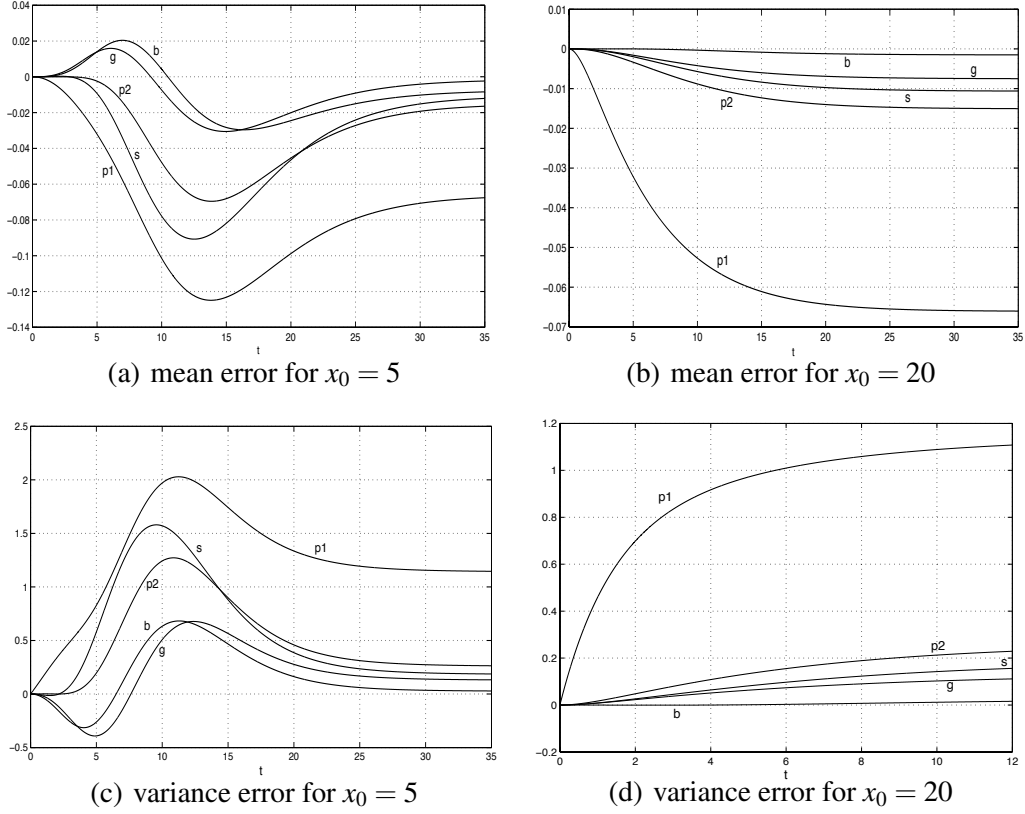


Fig. 2. Evolutions of the mean error  $\mu_1 - v_1$  and of the variance error  $(\mu_2 - \mu_1^2) - (v_2 - v_1^2)$  for the different moment closures in Table 3 for  $n = 2$ , with parameters as in (34) and  $x_0 = 5$  (left)  $x_0 = 20$  (right).

**Lognormal moment closure:** One can see from Example 3 of Section 5, that a family of lognormal moment closure functions for  $\mu_{n+1}$  is given by

$$\varphi_{n+1}^l(\mu) = \mu_1^{t_1} \mu_2^{t_2} \dots \mu_n^{t_n}, \quad \sum_{m=1}^n C_1^m t_m = C_1^{n+1}, \quad \sum_{m=1}^n C_2^m t_m = C_2^{n+1}. \quad (35)$$

The derivative matching errors for these moment closure functions are given by

$${}^l \varepsilon_{n+1}^0(x_0) = 0, \quad {}^l \varepsilon_{n+1}^i(x_0) \in P_{x+0}(n+i+1-w), \quad \forall i \in \{1, 2, \dots\} \quad (36)$$

where the constant  $w \in \{1, 2, \dots, n\}$  is such that

$$\begin{aligned} \sum_{m=1}^n C_k^m t_m &= C_k^{n+1}, & \forall k \in \{1, \dots, w\}, \\ \sum_{m=1}^n C_k^m t_m &\neq C_k^{n+1}, & \forall k \in \{w+1, \dots, n\} \end{aligned}$$

with higher values of  $w$  correspond to better approximations. From (35) and (36), we conclude that the lognormal moment closure guarantees that  $w \geq 2$ . But from (B.1) we see that this can be improved and we actually set  $w = n$  for the separable

derivatives-matching moment closure, for which

$${}^s\epsilon_{n+1}^0(x_0) = 0, \quad {}^s\epsilon_{n+1}^i(x_0) \in P_{x+0}(i+1), \quad \forall i \in \{1, 2, \dots\}. \quad (37)$$

Thus, the separable derivative-matching moment closure function leads to the best estimate for  $\mu_{n+1}$ , in the sense that it has the least derivative matching error among all the moment closure functions consistent with the lognormal distribution.

**Normal moment closure:** We first restrict our attention to the case  $n = 3$ . We recall from Example 2 that a family of normal moment closure functions for  $\mu_4$  were given by

$$4\mu_1\mu_3 + 3\mu_2^2 - 12\mu_2\mu_1^2 + 6\mu_1^4 + h(\mu_1, \mu_2 - \mu_1^2, \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3).$$

Its straight forward to see that the function  $h(\kappa_1, \kappa_2, \kappa_3)$  corresponds to the cumulant closer function for  $\kappa_4$ . However, one needs to be careful in picking the functions  $h$ . To demonstrate this we consider the following normal moment closure functions

$$\varphi_4^g(\mu) = 4\mu_1\mu_3 + 3\mu_2^2 - 12\mu_2\mu_1^2 + 6\mu_1^4 \quad (38)$$

$$\varphi_4^{g1}(\mu) = 4\mu_1\mu_3 + 3\mu_2^2 - 12\mu_2\mu_1^2 + 6\mu_1^4 + \mu_3 - 3\mu_1\mu_2 + 2\mu_1^3 \quad (39)$$

$$\varphi_4^{g2}(\mu) = 4\mu_1\mu_3 + 3\mu_2^2 - 12\mu_2\mu_1^2 + 6\mu_1^4 + (\mu_3 - 3\mu_1\mu_2 + 2\mu_1^3)\mu_1 \quad (40)$$

which correspond to cumulant closure functions  $h(\kappa_1, \kappa_2, \kappa_3) = 0$ ,  $\kappa_3$  and  $\kappa_3\kappa_1$  respectively. Using symbolic manipulation in Mathematica we have the following derivative matching errors for the normal moment closure functions (38)–(40):

$${}^*\epsilon_4^i(x_0) \in P_{x_0}(i+1), \quad * = \{g, g1\}$$

$${}^{g2}\epsilon_4^i(x_0) \in P_{x_0}(i+2)$$

for all  $i \in \{1, 2, \dots\}$  and zero for  $i = 0$ . Thus, the normal moment closure functions (38)–(39) have derivative matching errors of the same form as (37), and provide better approximates for  $\mu_4$  as compared to (40), for which the derivative matching errors are an order higher.

In general for  $n \geq 3$ , a family of normal moment closure functions for  $\mu_{n+1}$  is obtained using the cumulant closure function

$$\kappa_{n+1} = h(\kappa_1, \kappa_2, \dots, \kappa_n)$$

where the function  $h$  is zero if any of  $\kappa_3, \dots, \kappa_n$  is zero. Using symbolic manipulation in Mathematica we found that choosing cumulant closure functions as

$$h(\kappa_1, \kappa_2, \dots, \kappa_n) = \sum_{d=3}^n f_d \kappa_d$$

for some constants  $f_d$ , typically lead to derivative matching error of the form (37) and provides good estimates for  $\mu_{n+1}$ .

**Zero first-order error moment closure:** We now propose a slight modification of the normal moment closure function (38) which is given by

$$\varphi_4^z(\mu) = 4\mu_1\mu_3 + 3\mu_2^2 - 12\mu_2\mu_1^2 + 6\mu_1^4 + \mu_2 - \mu_1^2 \quad (41)$$

and corresponds to the cumulant closure function  $\kappa_2$  for  $\kappa_4$ . It yields the following derivative matching errors

$${}^z\mathcal{E}_4^i(x_0) = 0, \quad i \in \{0, 1\} \quad {}^z\mathcal{E}_4^i(x_0) \in P_{x_0}(i+1), \quad \forall i \geq 2.$$

Hence, unlike the separable derivative-matching and normal moment closure functions,  $\varphi_4^z(\mu)$  yields zero 0<sup>th</sup> and 1<sup>st</sup> order derivative matching errors, and hence, provides the best estimates for  $\mu_4$  as compared to them, at least near  $t = 0$ .

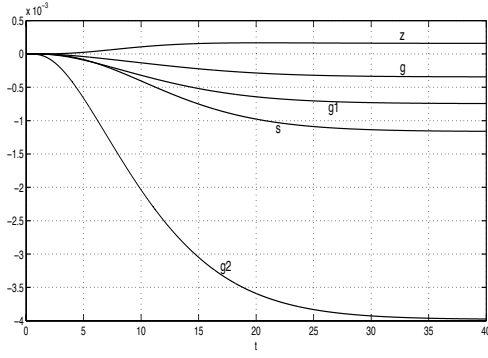
In order to gauge the performances of the above moment closure functions we consider the stochastic logistic model with parameters as in (34),  $n = 3$  and  $x_0 = 20$ . We recall from Table 1 that for  $n = 3$ , we have the following separable derivative-matching moment closure function

$$\varphi_4^s(\mu) = \frac{\mu_1^4\mu_3^4}{\mu_2^6}. \quad (42)$$

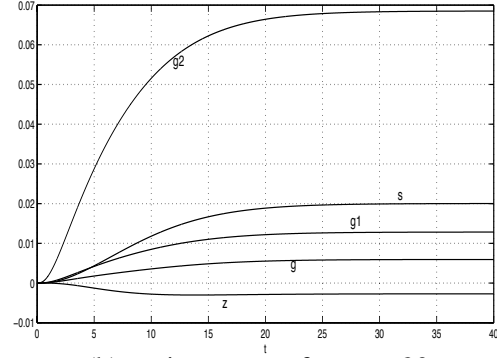
Using (11), we have the following truncated moment dynamics

$$\begin{bmatrix} \dot{v}_1 \\ \dot{v}_2 \\ \dot{v}_3 \end{bmatrix} = \begin{bmatrix} 0.28 & -0.016 & 0 \\ .32 & .546 & -0.032 \\ .28 & .944 & .798 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ -0.048 \end{bmatrix} \varphi_3(v).$$

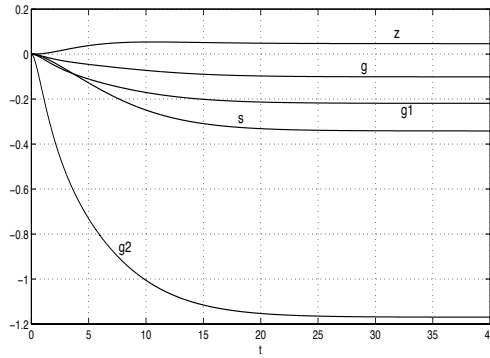
Substituting the moment closure functions (38)–(42) in place of  $\varphi_3(v)$ , we obtain the corresponding approximate time evolution of moments. Figure 3 contains plots of the mean, variance, and third cumulant errors. The letters  $g$ ,  $g1$ ,  $g2$ ,  $z$  and  $s$  represent these errors for the moment closure functions (38)–(42), respectively. As expected the normal moment closure functions (38)–(39) perform much better than the normal moment closure function (40). The separable derivative-matching moment closure function (42) also provide good estimates. One can also see that the zero first-order error moment closure function (41), which guarantees the best approximation near  $t = 0$  actually provides in this case the most accurate estimate for  $\mu_4$  for all time. As can be seen from Figure 2 and 3, the mean and variance errors for  $x_0 = 20$  with  $n = 3$  are an order of magnitude smaller as compared to the ones with  $n = 2$ .



(a) mean error for  $x_0 = 20$



(b) variance error for  $x_0 = 20$



(c) third cumulant error for  $x_0 = 20$

Fig. 3. Evolutions of the mean error  $\mu_1 - v_1$ , variance error  $(\mu_2 - \mu_1^2) - (v_2 - v_1^2)$ , and third cumulant error  $(\mu_3 - 3\mu_2\mu_1 + 2\mu_1^3) - (v_3 - 3v_2v_1 + 2v_1^3)$  for the different moment closure functions (38)–(42) for  $n = 3$ , with parameters as in (34) and  $x_0 = 20$ .

## 6.2 Steady-State Solutions of the Truncated Moment Dynamics

We now look at the steady-state solutions of the truncated moment dynamics (17) for the different moment closure techniques.

### 6.2.1 Separable Derivative-Matching Moment closure

Consider the truncated moment dynamics of order  $n \in \mathbb{N}_{\geq 2}$  with moment closure functions as given in Table 1. From (17), at steady-state we have

$$0 = A\mathbf{v}^s(\infty) + B\boldsymbol{\varphi}_{n+1}^s(\mathbf{v}^s(\infty)) \quad (43)$$

where  $\mathbf{v}^s(\infty)$  denotes the steady-state solution. Using (16) at steady-state we have

$$\mathbf{v}_2^s(\infty) = c_1 \mathbf{v}_1^s(\infty) \quad (44a)$$

$$\vdots$$

$$\mathbf{v}_n^s(\infty) = c_{n-1} \mathbf{v}_1^s(\infty) \quad (44b)$$

$$\varphi_{n+1}(\mathbf{v}^s(\infty)) = c_n v_1^s(\infty) \quad (44c)$$

for some positive real numbers  $c_1, \dots, c_n$ . Using the above equalities and Table 1 we conclude that

$$\varphi_3^s(\mathbf{v}^s(\infty)) = c_1^3, \quad \varphi_4^s(\mathbf{v}^s(\infty)) = \frac{c_2^4}{c_1^6} v_1^s(\infty)^2, \quad \varphi_5^s(\mathbf{v}^s(\infty)) = \frac{c_1^{10} c_3^5}{c_2^{10}}, \dots \quad (45)$$

for  $n \in \{2, 3, 4, \dots\}$  respectively. Substituting (45) in (44) yields the following unique non-trivial solution for  $v_1^s(\infty)$

$$v_1^s(\infty) = \begin{cases} \frac{c_1^3}{c_2} & n = 2 \\ \frac{c_1^6 c_3}{c_2^4} & n = 3 \\ \frac{c_1^{10} c_3^5}{c_2^{10} c_4} & n = 4 \end{cases}$$

and the corresponding  $v_2^s(\infty), \dots, v_n^s(\infty)$  can be calculated from (44). Hence, the separable derivative-matching moment closure always yields a unique non-trivial positive real steady-state  $\forall n \in \mathbb{N}_{\geq 2}$ . In terms of the parameters  $a_1, b_1, a_2$  and  $b_2$ , the constants  $c_1, c_2$  and  $c_3$  are given as follows

$$\begin{aligned} c_1 &= K, & c_2 &= K^2 + \sigma^2, & c_3 &= K^3 + 3K\sigma^2 + \bar{\sigma}\sigma^2 \\ K &= \frac{a_1 - a_2}{b_1 + b_2}, & \sigma^2 &= \frac{a_1 b_2 + b_1 a_2}{(b_1 + b_2)^2}, & \bar{\sigma} &= \frac{b_2 - b_1}{b_2 + b_1}, \end{aligned}$$

and hence,

$$v_1^s(\infty) = \begin{cases} \frac{K}{1 + \frac{\sigma^2}{K^2}} & n = 2 \\ \frac{K \left(1 + \frac{3\sigma^2}{K^2} + \frac{\bar{\sigma}\sigma^2}{K^3}\right)}{\left(1 + \frac{\sigma^2}{K^2}\right)^4} & n = 3. \end{cases}$$

### 6.2.2 Other Moment closure techniques

In this section we will see that the moment closure functions which lack the “separable structure” of (20), can lead to scenarios of biologically meaningless steady-state solutions for the truncated moment dynamics (17). Substituting the moment closure functions from Table 2 in (43), we get the following non-trivial steady-state solutions for  $n = 2$ :

$$\begin{aligned} v_1^g(\infty) &= K \left[ \frac{3}{4} \pm \frac{1}{4} \left(1 - \frac{8\sigma^2}{K^2}\right)^{\frac{1}{2}} \right] \\ v_1^{p1}(\infty) &= K \left[ \frac{3}{4} \pm \frac{1}{4} \left(1 - \frac{8(\sigma^2 - 1)}{K^2}\right)^{\frac{1}{2}} \right] \end{aligned}$$

$$v_1^{p2}(\infty) = \left[ \frac{3K-1}{4} \pm \frac{1}{4} \left( (K+1)^2 - \frac{8\sigma^2}{K^2} \right)^{\frac{1}{2}} \right]$$

$$v_1^b(\infty) = K - \frac{\sigma^2}{K-1}.$$

From the above steady-states we conclude the following for  $n = 2$ :

- (1) The binomial moment closure function leads to a unique non-trivial attracting real steady-state, which can be negative for a range of parameters.
- (2) Normal, Nasell-poisson and new-poisson moment closure functions, yield two non-trivial steady-states, with the one with the  $-$  sign being a “spurious steady-state”. Following the definition by Nasell (2003b), a steady state is “spurious”, if  $\lim_{M \rightarrow \infty} v_1(\infty) \neq K$ , where  $\sigma^2$  and  $K$  are both  $O(M)$ . For  $n = 2$ , all these “spurious steady-states” happen to be unstable, and hence, the truncated model will not converge to them.

When the parameters are chosen, such that the term under the square root sign is negative, then both the non-trivial steady-states would be imaginary, and hence, biologically meaningless.

In general for  $n \geq 3$ , the normal moment closure functions introduced in Section 6.1.2 yield  $K \geq 3$  non-trivial steady-states with  $K - 1$  of them being “spurious”. For example, for  $n = 3$ , the normal and the zero first-order error moment closure functions (38) and (41) yield 3 non-trivial steady-state solutions given as the roots of the third order polynomials

$$(3c_1^2 + 4c_2)v_1(\infty) - 12c_1v_1(\infty)^2 + 6v_1(\infty)^3 = c_3,$$

and

$$(3c_1^2 + 4c_2 - 1)v_1(\infty) - 12c_1v_1(\infty)^2 + 6v_1(\infty)^3 = c_3 - c_1,$$

respectively, with 2 of them being “spurious”. Also for a range of parameter values, biologically meaningless scenarios of a combination of imaginary and negative steady states can happen. On the other hand, the separable derivative-matching moment closure function with a unique non-trivial positive real steady-state  $\forall n \in \mathbb{N}_{\geq 2}$  has a clear advantage.

## 7 Conclusion and Future Work

A procedure for constructing moment closures for the stochastic logistic model was presented. This was done by first assuming a separable form for the moment closure function  $\varphi_{n+1}(v)$ , and then, matching its time derivatives with  $\mu_{n+1}$ , at some initial time  $t_0$ . We showed that for initial conditions  $\mathbf{x}(t_0) = x_0$  with probability one, there exists a unique separable derivative-matching moment closure

function for which the  $i^{\text{th}}$  derivative matching error is a polynomial in  $x_0$  of order  $i + 1$  for all  $i \in \{1, 2, \dots\}$  and zero for  $i = 0$ . Explicit formulas to construct these moment closure functions were provided. Comparisons with alternative moment closure techniques available in literature were carried out based on transient performance and steady-state solutions of the truncated moment dynamics, which led to the following conclusions:

- (1) Derivative matching can be used as an effective tool for gauging the performance of moment closure functions. We showed that for  $n = 2$ , with the exception of the Nasell-poisson, all other moment closure functions in Table 2 perform derivative matching and give fairly good estimates of  $\mu_3$ . For  $n \geq 3$ , we showed that among the moment closure functions consistent with the log-normal distribution, the separable derivatives-matching moment closure function provides the best estimate for  $\mu_{n+1}$ . We also proposed families of normal moment closure functions that perform derivative matching and result in good estimates. For  $n = 3$ , a new zero first-order error moment closure function was also proposed which guaranteed better approximations, at least locally in time, as compared to the other moment closure functions discussed in this paper.
- (2) The separable derivative-matching moment closure, always yields a unique non-trivial positive real steady-state  $\forall n \in \mathbb{N}_{\geq 2}$ , and hence, in some sense superior to the other moment closures, which can have spurious, imaginary and even stable negative steady-states.

Possible directions for future research include the development of a systematic procedure to construct moment closure functions of the form (41) which can yield zero derivative matching errors up to any order and the extension of the results in this paper to multi-specie birth-death Markov processes. Primary results regarding the latter can be found in Hespanha and Singh (2005).

## A Appendix

Substituting (4), (7) and (8) in (10) and doing a binomial expansion we have

$$\begin{aligned}
\frac{d\mu_m}{dt} &= \mathbf{E}[(\mathbf{x} + 1)^m - \mathbf{x}^m]\eta(\mathbf{x}) + (\mathbf{x} - 1)^m - \mathbf{x}^m)\chi(\mathbf{x}) \\
&= \mathbf{E}\left[\sum_{j=1}^m C_j^m [(a_1x - b_1x^2) + (-1)^j(a_2x + b_2x^2)]\mathbf{x}^{m-j}\right] \\
&= \mathbf{E}\left[\sum_{p=1}^2 \sum_{j=1}^m C_j^m f(j, p)\mathbf{x}^{m-j+p}\right]
\end{aligned} \tag{A.1}$$

where  $C_j^m$  and  $f(j, p)$  are defined in (12) and (13). Using (A.1), (12) and (13) we conclude that the evolution of  $\mu_m, \forall m \in \mathbb{N}$ , can be written as

$$\begin{aligned}
\dot{\mu}_m &= \sum_{p=1}^2 \sum_{j=1}^m C_j^m f(j, p) \mu_{m-j+p} \\
&= \sum_{p=1}^2 \sum_{r=p}^{m+p-1} C_{m+p-r}^m f(m+p-r, p) \mu_r, \quad r = m-j+p \\
&= \sum_{p=1}^2 \sum_{r=1}^{m+1} C_{m+p-r}^m f(m+p-r, p) \mu_r.
\end{aligned} \tag{A.2}$$

## B Appendix

### Proof of Theorem 2:

*STEP 1:* We first show that with  $\gamma_m$  chosen as the solution to the linear system of equations

$$C_k^{n+1} = \sum_{m=1}^n \gamma_m C_k^m, \quad \forall k = \{1, \dots, n\}. \tag{B.1}$$

equalities (23) hold. Using (15), (16) and (17) one can see that the following equalities imply (23):

$$\mu_{n+1}(t_0) = \varphi_{n+1}^s(\mathbf{v}(t_0)) \tag{B.2}$$

$$\left. \frac{d\mu_{n+1}(t)}{dt} \right|_{t=t_0} = \left. \frac{d\varphi_{n+1}^s(\mathbf{v}(t))}{dt} \right|_{t=t_0} + {}^s\boldsymbol{\varepsilon}_{n+1}^1(x_0), \tag{B.3}$$

where  ${}^s\boldsymbol{\varepsilon}_{n+1}^1(x_0)$  is a polynomial in  $x_0$  of order 2.

Let  $\gamma_m$  be chosen as the solution of (B.1). We show next that equalities (B.2) and (B.3) hold. With initial conditions starting on the set of deterministic distributions, we have  $\mu_m(t_0) = x_0^m$ . From  $\mu(t_0) = \mathbf{v}(t_0)$ , (20) and (B.1) we have

$$\mu_{n+1}(t_0) = x_0^{n+1} = x_0^{\sum_{m=1}^n \gamma_m C_1^m} = \varphi_{n+1}^s(\mu(t_0)) = \varphi_{n+1}^s(\mathbf{v}(t_0))$$

which proves (B.2). From (A.2) we have

$$\begin{aligned}
\dot{\mu}_{n+1}(t) &= \sum_{p=1}^2 \sum_{r=1}^{n+2} C_{n+1+p-r}^{n+1} f(n+1+p-r, p) \mu_r \\
\dot{\mu}_{n+1}(t_0) &= \sum_{p=1}^2 \sum_{r=3}^{n+2} C_{n+1+p-r}^{n+1} f(n+1+p-r, p) x_0^r + {}^s\boldsymbol{\varepsilon}_{n+1}^1(x_0),
\end{aligned} \tag{B.4}$$

where  ${}^s\boldsymbol{\varepsilon}_{n+1}^1(x_0)$  is a polynomial in  $x_0$  of order 2. Using (A.2) and (17) we have

$$\dot{v}_m = \sum_{p=1}^2 \sum_{r=1}^{m+1} C_{m+p-r}^m f(m+p-r, p) v_r, \quad v_{n+1} = \varphi_{n+1}^s(v). \quad (\text{B.5})$$

Thus, from (20), (B.1), (B.5) and  $v(t_0) = [x_0, \dots, x_0^n]^T$

$$\begin{aligned} \frac{d\varphi_{n+1}^s(v(t))}{dt} &= \sum_{m=1}^n \gamma_m v_1^{\gamma_1} \dots v_m^{\gamma_m-1} \dots v_n^{\gamma_n} \dot{v}_m \\ &= \sum_{p=1}^2 \sum_{m=1}^n \sum_{r=1}^{m+1} \gamma_m v_1^{\gamma_1} \dots v_m^{\gamma_m-1} \dots v_n^{\gamma_n} C_{m+p-r}^m f(m+p-r, p) v_r, \\ \frac{d\varphi_{n+1}^s(v(t))}{dt} \Big|_{t=t_0} &= \sum_{p=1}^2 \sum_{m=1}^n \sum_{r=1}^{m+1} \gamma_m C_{m+p-r}^m f(m+p-r, p) x_0^{(\sum_{m=1}^n \gamma_m m) + r - m} \\ &= \sum_{p=1}^2 \sum_{m=1}^n \sum_{r=1}^{m+1} \gamma_m C_{m+p-r}^m f(m+p-r, p) x_0^{n+1+r-m}. \end{aligned}$$

With a change of variable  $q = n + 1 + r - m$ , the last equality becomes

$$\begin{aligned} \frac{d\varphi_{n+1}^s(v(t))}{dt} \Big|_{t=t_0} &= \sum_{p=1}^2 \sum_{m=1}^n \sum_{q=n+2-m}^{n+2} \gamma_m C_{n+1+p-q}^m f(n+1+p-q, p) x_0^q \\ &= \sum_{p=1}^2 \sum_{q=3}^{n+2} \sum_{m=1}^n \gamma_m C_{n+1+p-q}^m f(n+1+p-q, p) x_0^q + {}_2^s \varepsilon_{n+1}^1(x_0), \quad (\text{B.6}) \end{aligned}$$

where  ${}_2^s \varepsilon_{n+1}^1(x_0)$  is a polynomial in  $x_0$  of order 2. Using (B.1), equality (B.6) reduces to

$$\frac{d\varphi_{n+1}^s(v(t))}{dt} \Big|_{t=t_0} = \sum_{p=1}^2 \sum_{q=3}^{n+2} C_{n+1+p-q}^{n+1} f(n+1+p-q, p) x_0^q + {}_2^s \varepsilon_{n+1}^1(x_0). \quad (\text{B.7})$$

Comparing (B.7) with (B.4) one can see that (B.3) holds.

*STEP 2:* We now show that solution to (B.1) is unique and given by (22). Towards that end, we observe that for all  $z \in \mathbb{R}$ , one can write using binomial expansion,

$$\begin{aligned} [1 - (1+z)]^{n+1} &= \sum_{v=0}^{n+1} C_v^{n+1} (-1)^v (1+z)^v \\ &= 1 + \sum_{v=1}^{n+1} C_v^{n+1} (-1)^v \sum_{w=0}^v C_w^v z^w. \quad (\text{B.8}) \end{aligned}$$

Equating coefficients for  $z^s$ ,  $s \in \{1, \dots, n\}$  on both sides of (B.8) we have

$$0 = \sum_{v=1}^{n+1} C_v^{n+1} (-1)^v C_s^v$$

$$\Rightarrow (-1)^n C_s^{n+1} = \sum_{v=1}^n C_v^{n+1} (-1)^v C_s^v. \quad (\text{B.9})$$

Comparing (B.9) with (B.1) one can see that a solution to (B.1) will be (22). Also the system of  $n$  linear equations (B.1) can be put into the form

$$C = \Pi \gamma$$

where  $\gamma = [\gamma_1, \dots, \gamma_n]^T$ ,  $C = [C_1^{n+1}, \dots, C_n^{n+1}]^T$  and

$$\Pi = \begin{bmatrix} C_1^1 & C_1^2 & \dots & C_1^n \\ 0 & C_2^2 & \dots & C_2^n \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & C_m^m \end{bmatrix}.$$

As the upper triangular matrix  $\Pi$  is non-singular, the above solution to  $\gamma_m$  is unique. ■

## References

- Bailey, N. T. J., 1964. *The Elements of Stochastic Processes*. Wiley, New York.
- Barlett, M. S., Gower, J. S., Leslie, P. H., 1960. A comparison of theoretical and empirical results for some stochastic models. *Biometrika* 47, 1–11.
- Hespanha, J. P., Mar. 2004. Stochastic hybrid systems: Applications to communication networks. In: Alur, R., Pappas, G. J. (Eds.), *Hybrid Systems: Computation and Control*. No. 2993 in *Lect. Notes in Comput. Science*. Springer-Verlag, Berlin, pp. 387–401.
- Hespanha, J. P., Singh, A., 2005. Stochastic models for chemically reacting systems using polynomial stochastic hybrid systems. *Int. J. of Robust and Nonlinear Control* 15, 669–689.
- Keeling, M. J., 2000. Multiplicative moments and measures of persistence in ecology. *J. of Theoretical Biology* 205, 269–281.
- Matis, J. H., Kiffe, T. R., 1996. On approximating the moments of the equilibrium distribution of a stochastic logistic model. *Biometrics* 52, 155–166.
- Matis, J. H., Kiffe, T. R., 2002. On interacting bee/mite populations: a stochastic model with analysis using cumulant truncation. *Environmental and Ecological Statistics* 9, 237–258.
- Matis, J. H., Kiffe, T. R., Parthasarathy, P. R., 1998. On the cumulant of population size for the stochastic power law logistic model. *Theoretical Population Biology* 53, 16–29.
- Nasell, I., 2001. Extinction and quasi-stationarity in the verhulst logistic model. *J. of Theoretical Biology* 211, 11–27.
- Nasell, I., 2003a. An extension of the moment closure method. *Theoretical Population Biology* 64, 233–239.

- Nasell, I., 2003b. Moment closure and the stochastic logistic model. *Theoretical Population Biology* 63, 159–168.
- Pielou, E. C., 1977. *Mathematical Ecology*. Wiley, New York.
- Whittle, P., 1957. On the use of the normal approximation in the treatment of stochastic processes. *J. Roy. Statist. Soc., Ser. B* 19, 268–281.