Moment closure approximations for stochastic kinetic models with rational rate laws

Peter Milner, Colin S. Gillespie, Darren J. Wilkinson

School of Mathematics & Statistics, Newcastle University, Newcastle upon Tyne, NE1 7RU

Abstract

Stochastic models are often used when modelling chemical species that have low numbers of molecules (see [1]). However, as these models become large, it can become computationally expensive to simulate even a single realisation of the system. Since even efficient simulation techniques have a high computational cost. One possible technique to approximate the stochastic system is *moment closure*. The moment closure approximation is used to provide analytic approximations to non-linear stochastic models. Until now, this approximation has only been applied to models with *polynomial* rate laws. In this paper we extend the moment closure method to cover models with *rational* rate laws.

Keywords: moment closure, rational, stochastic, Hill, Michaelis-Menten, Holling

1. Introduction

Stochastic population models are becoming increasingly useful in a variety of fields (see for example [2, 3, 4, 1]). For most stochastic models of interest the transition probabilities are non-linear, which makes the resultant stochastic process intractable.

To gain insight into the stochastic process we simulate from the model. Exact realisations from the stochastic model can be obtained using the discrete event simulation method described by Kendall and later by Gillespie [5, 6]. However, for large models this leads to computational problems even for simulating a single realisation, let alone the many repeated realisations

that are needed in order to understand the stochastic variation inherent in the process.

Other simulation methods have been proposed. The next-reaction method [7] is a modification of a variant of the first reaction method [8]. This algorithm can increase the speed of the simulation. This increase in speed is achieved by minimising the recalculation of hazards and reducing the random numbers needed per step. However this increase depends on the model and the random number generator used. Approximate methods are also possible. The τ-Leap Method was developed by Gillespie [9] as an approximate method to accelerate the simulation process with an acceptable loss of accuracy, by moving ahead by a variable amount of time τ and approximating the number of reactions in the interval τ with a Poisson random variable. For each leap, τ is selected to give a trade-off between speed and accuracy, where an increase in speed is given by a large τ and the accuracy is measured by the difference between the reaction hazards at each end of the interval, as this affects the assumption of constant hazard over the interval τ. The original τ-Leap algorithm has been improved over the years [10, 11].

In contrast, moment closure methods derive moment or cumulant ODEs which can be solved numerically to obtain estimates of the moments at a particular time ([12, 13, 14, 15, 16, 17, 18]). Typically for non linear models, the ODEs for the moments form an infinite set, i.e. the i^{th} moment depends on at least the $(i+1)^{th}$ moment. This results in a set of ODEs that can not be solved analytically or numerically. Moment closure is a technique to 'close' this set, by eliminating the dependence on higher order moments, allowing us to get a solvable set of coupled ODEs. This is usually done by setting moments or cumulants above a certain order to zero, and solving the remaining coupled ODEs. For instance [19] and [20] set all third and higher order cumulants to zero, therefore obtaining a normal approximation.

Moment closure yields a large speed up in computational time for the simulation of systems with large populations. For example, in calcium-induced calcium release in cardiac myocytes [21], an algebraic relationship (based on the beta-binomial distribution) is used to express the third moment in terms of lower moments.

However all previous research applies moment closure to models where the rate laws are of *polynomial* form. This is the first paper that generates and analyses systems using *rational* rate laws. In this paper we consider three specific models, which can be used as building blocks in other systems. This extension to the moment closure method now allows us to apply moment

closure techniques to models with more complex propensity functions.

2. Example 1: a simple rational rate law

To illustrate our method we will first look at the simplest rational rate law. In this model, we have a single species and a single reaction

$$X \to \emptyset \tag{1}$$

with propensity function

$$a(x) = \frac{Vx}{k+x}, \tag{2}$$

where $k > 0$, $V > 0$ and x is the number of individuals of X at time t. Thus the probability of (1) occurring in the interval $(t, t+\delta t)$ is $a(x)\delta t + o(\delta t)$.

Holling [22] proposed (1) with propensity function $a(x)$ as the Holling type II response function. Holling used this function form to describe the uptake of substrate. It should be noted that this reaction system also describes the Michaelis-Menten system (see [23]).

The associated forward Kolmogorov equation for (1) with propensity function $a(x)$ is

$$\frac{dP(x;t)}{dt} = P(x+1;t)\frac{V(x+1)}{k+x+1} - P(x;t)\frac{Vx}{k+x}, \tag{3}$$

where $P(x;t)$ is the probability of observing x individuals at time t (see [1]). To find the associated moment equations of (1) we first multiply (3) by $(k+x+1)(k+x)$ to give

$$\begin{aligned}\frac{dP(x;t)}{dt}(k+x+1)(k+x) = {} & P(x+1;t)V(x+1)(k+x) \\ & - P(x;t)Vs(k+x+1).\end{aligned} \tag{4}$$

The univariate moment generating function is defined as

$$M(\theta, t) = \sum_{x=0}^{\infty} P(x;t)e^{x\theta} = \sum_{i=0}^{\infty} \frac{\mu_i(t)\theta^i}{i!}$$

where

$$\mu_i(t) = E[X(t)^i] = \sum_{i=0}^{\infty} P(x;t)x(t)^i \ .$$

Multiplying equation (4) by $e^{x\theta}$ and summing over all possible values of x gives

$$\sum_i h_i \frac{\partial^{i+1} M(\theta,t)}{\partial t \partial \theta^i} = \sum_i f_i \frac{\partial^i M(\theta,t)}{\partial \theta^i} e^{-\theta} - \sum_i g_i \frac{\partial^i M(\theta,t)}{\partial \theta^i}, \tag{5}$$

where

$$h_i = \begin{cases} k^2 + k & \text{for } i = 0; \\ 2k + 1 & \text{for } i = 1; \\ 1 & \text{for } i = 2; \\ 0 & \text{otherwise,} \end{cases} \qquad f_i = \begin{cases} V(k-1) & \text{for } i = 1; \\ V & \text{for } i = 2; \\ 0 & \text{otherwise,} \end{cases}$$

and

$$g_i = \begin{cases} V(k+1) & \text{for } i = 1; \\ V & \text{for } i = 2; \\ 0 & \text{otherwise.} \end{cases}$$

By equating coefficients of θ in equation (5) we get expressions for each of the moments. For example, extracting coefficients of θ^0 gives an expression for the first moment

$$(2k+1)\frac{d\mu_1}{dt} = -2V\mu_1 - \frac{d\mu_2}{dt}. \tag{6}$$

By rewriting equation (6) in terms of cumulants we obtain

$$\frac{d\kappa_1}{dt} = -\frac{V\kappa_1}{k + 1/2 + \kappa_1} - \frac{d\kappa_2}{dt} \tag{7}$$

where κ_1 is the mean and κ_2 is the variance. Setting $\kappa_2 = 0$ in equation (7) gives an approximation to the deterministic equation (2). When we have polynomial rate laws, the deterministic equation is identical to the moment closure equation with κ_2 set equal to zero. However, for rational rate laws

Distribution	Moments	Cumulants
Normal	$\mu_3 = 3\mu_2\mu_1 - 2\mu_1^3$	$\kappa_3 = 0$
Poisson	$\mu_3 = \mu_1 + 3\mu_1^2 + \mu_1^3$	$\kappa_3 = \kappa_2 = \kappa_1$
Log-Normal	$\mu_3 = (\mu_2/\mu_1)^3$	$\kappa_3 = (\kappa_2/\kappa_1)^3 + 3\kappa_2^2/\kappa_1$

Table 1: *Table showing some options for closing second order moment for different distributional assumptions.*

this equality does not hold. Instead, equation (7) has an additional term "1/2", in the denominator.

If we equate coefficients of θ^1 and θ^2 we retrieve expressions for the second and third moments respectively,

$$\begin{aligned}(2k+1)\frac{d\mu_2}{dt} &= -V(k-1)\mu_1 - 3V\mu_2 - k(k+1)\frac{d\mu_1}{dt} - \frac{d\mu_3}{dt}\\ (2k+1)\frac{d\mu_3}{dt} &= V(k-1)\mu_1 + V(3-2k)\mu_2 - 4V\mu_3\\ &\quad - k(k+1)\frac{d\mu_2}{dt} - \frac{d\mu_4}{dt},\end{aligned}$$

where $\mu_i \equiv \mu_i(t)$.

The ODE for $\mu_1(t)$ (see equation (6)) is not closed as it depends $\mu_2(t)$, similarly the ODE for $\mu_2(t)$ depends on $\mu_3(t)$ etc. Whenever a model includes non-linear rate laws this dependence structure appears. Table 1 shows three distributions that could be used to close the moment equations. Throughout this paper we will use the normal approximation. Figure 1 shows the mean with an approximate 95% confidence interval from the moment closure approximation and from 100,000 Monte Carlo simulations. We can see that the moment closure approximation gives a good estimate of the true mean and variance, although between $t = 35$ and $t = 45$ the approximation deviates slightly from the true mean and variance.

3. General form

In this section we develop general multivariate results, for forward Kolmogorov equations that contain rational rate coefficients, allowing the calcu-

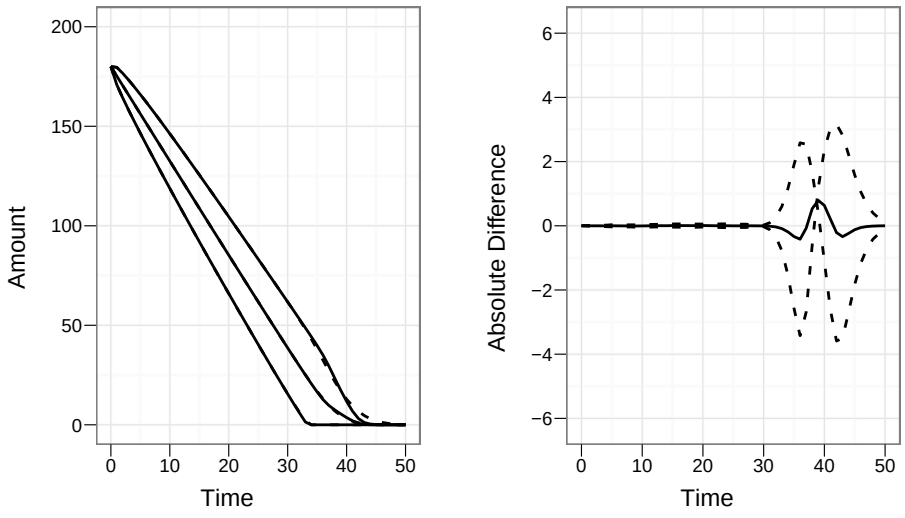

Figure 1: *(a) The mean ± 2 standard deviations obtained from 100,000 simulations (− − −) and the approximate mean ± 2 standard deviations from the moment closure approximation (solid), where* $x(0) = 180$ *and* $\{V, k\} = \{4.8, 5/3\}$. *(b) The error (approximation − exact) for the mean (solid), the upper 95% (− − −) and the lower 95% (− − −).*

lation of moment equations up to any order.

Suppose we have N species $\{y_1, \ldots, y_N\}$ and l reactions $\{R_1, \ldots, R_L\}$, where reaction R_l corresponds to

$$\underline{s}_{l1}y_1 + \ldots + \underline{s}_{lN}y_N \longrightarrow \bar{s}_{l1}y_1 + \ldots + \bar{s}_{lN}y_N,$$

where $\underline{\mathbf{s}}_\mathbf{l}$ and $\bar{\mathbf{s}}_\mathbf{l}$ are the number of reactants and products of each species involved with reaction l.

Let $\mathbf{y}$ be a column vector of species numbers, $\mathbf{s_l} = \bar{\mathbf{s}}_\mathbf{l} - \underline{\mathbf{s}}_\mathbf{l}$ and $s_{li} = \bar{s}_{li} - \underline{s}_{li}$ be the stoichiometric coefficient of species y_i in reaction R_l. Then when reaction R_l occurs $y_i \to y_i + s_{li}$. Thus we can write the propensity function for reaction R_l as, $\alpha_l(\mathbf{y})/\beta_l(\mathbf{y})$, where α and β are polynomials.

Let $p(\mathbf{y})(t) = p(\mathbf{y})$ be the probability of being in state $\mathbf{y}$ at time t, with initial conditions of $\mathbf{y}(\mathbf{0})$. The time evolution of $\mathbf{y}$ can be formulated as the

forward Kolmogorov equation

$$\frac{d}{dt}P(\mathbf{y}) = \sum_{l=1}^{L} P(\mathbf{y}-\mathbf{s_l})\frac{\alpha_l(\mathbf{y}-\mathbf{s_l})}{\beta_l(\mathbf{y}-\mathbf{s_l})} - P(\mathbf{y})\frac{\alpha_l(\mathbf{y})}{\beta_l(\mathbf{y})}, \tag{8}$$

where $\beta_l(\cdot) \neq 0$ and $\beta_l(0,0,\ldots,0) \neq 0$. If we take $\beta_l(\cdot) \equiv 1$ in (8) then we have the standard form for the CME. Multiplying (8) by

$$\prod_{l=1}^{L} \beta_l(\mathbf{y}-\mathbf{s_l})\beta_l(\mathbf{y})$$

yields

$$\frac{d}{dt}P(\mathbf{y})\prod_{l=1}^{L}\beta_l(\mathbf{y}-\mathbf{s_l})\beta_l(\mathbf{y}) = \left[\sum_{l=1}^{L} P(\mathbf{y}-\mathbf{s_l})\alpha_l(\mathbf{y}-\mathbf{s_l})\left(\prod_{r\neq l}^{L}\beta_r(\mathbf{y}-\mathbf{s_r})\right)\left(\prod_{s=1}^{L}\beta_s(\mathbf{y})\right)\right] - \left[\sum_{l=1}^{L} P(\mathbf{y})\alpha_l(\mathbf{y})\left(\prod_{r\neq l}^{L}\beta_r(\mathbf{y})\right)\left(\prod_{s=1}^{L}\beta_s(\mathbf{y}-\mathbf{s_l})\right)\right].$$

On extracting coefficients of $\mathbf{y}^{\mathbf{i}}$ and $(\mathbf{y}-\mathbf{s_l})^{\mathbf{i}}$ for each reaction l, yields

$$\frac{d}{dt}P(\mathbf{y})h(\mathbf{y}) = \sum_{l=1}^{L}\left[P(\mathbf{y}-\mathbf{s_l})f(\mathbf{y}-\mathbf{s_l})\right] - P(\mathbf{y})g(\mathbf{y}). \tag{9}$$

where

$$h(\mathbf{y}) = \prod_{l=1}^{L}\beta_l(\mathbf{y}-\mathbf{s_l})\beta(\mathbf{y}) = \sum_{\mathbf{i}} h_{\mathbf{i}}\mathbf{y}^{\mathbf{i}}, \tag{10}$$

$$f(\mathbf{y}-\mathbf{s_l}) = \sum_{l=1}^{L}\alpha_l(\mathbf{y}-\mathbf{s_l})\prod_{r\neq l}^{L}\beta_r(\mathbf{y}-\mathbf{s_r})\prod_{s=1}^{L}\beta_s(\mathbf{y}) = \sum_{l=1}^{L}\sum_{\mathbf{i}} f_{l,\mathbf{i}}(\mathbf{y}-\mathbf{s_l})^{\mathbf{i}}, \tag{11}$$

$$g(\mathbf{y}) = \sum_{l=1}^{L}\alpha_l(\mathbf{y})\prod_{r\neq l}^{L}\beta_r(\mathbf{y})\prod_{s=1}^{L}\beta_s(\mathbf{y}-\mathbf{s_l}) = \sum_{\mathbf{i}} g_{\mathbf{i}}\mathbf{y}^{\mathbf{i}}, \tag{12}$$

with $h_{\mathbf{i}}$ and $g_{\mathbf{i}}$ being the respective coefficients of $\mathbf{y}^{\mathbf{i}} = y_1^{i_1}\times\ldots\times y_N^{i_N}$, and $f_{l\mathbf{i}}$

being the coefficients of $(\mathbf{y}-\mathbf{s_l})^{\mathbf{i}} = (y_1 - s_{l1})^{i_1} \times \ldots \times (y_N - s_{lN})^{i_N}$ for each reaction l. On multiplying (9) by $e^{\mathbf{y}\theta}$ and summing over all values of $\mathbf{y}$, we get a PDE for the m.g.f, viz

$$\sum_{\mathbf{i}} h_{\mathbf{i}} \frac{\partial^{\mathbf{i+1}} M(\theta,t)}{\partial t \partial \theta^{\mathbf{i}}} = \sum_{l=1}^{L} \left[\sum_{\mathbf{i}} f_{l,\mathbf{i}} \frac{\partial^{\mathbf{i}} M(\theta,t)}{\partial \theta^{\mathbf{i}}} \exp(\mathbf{s_l}\theta) \right] - \sum_{\mathbf{i}} g_{\mathbf{i}} \frac{\partial^{\mathbf{i}} M(\theta,t)}{\partial \theta^{\mathbf{i}}}. \tag{13}$$

Since h and g are coefficients of $P(\mathbf{y})$, they are independent of l. On expanding the partial derivatives in (13), we get

$$\begin{aligned} \frac{\partial}{\partial t} \sum_{\mathbf{i}} h_{\mathbf{i}} \sum_{\mathbf{j=i}}^{\infty} \frac{\theta^{\mathbf{j-i}}}{(\mathbf{j}-\mathbf{i})!} \mu_{\mathbf{j}} = \sum_{l=1}^{L} \left[\sum_{\mathbf{i}} f_{l,\mathbf{i}} \sum_{\mathbf{j=i}}^{\infty} \frac{\theta^{\mathbf{j-i}}}{(\mathbf{j}-\mathbf{i})!} \mu_{\mathbf{j}} \sum_{\mathbf{k=0}}^{\infty} \frac{(\mathbf{s_l}\theta)^{\mathbf{k}}}{\mathbf{k}!} \right] \\ - \sum_{\mathbf{i}} g_{\mathbf{i}} \sum_{\mathbf{j=i}}^{\infty} \frac{\theta^{\mathbf{j-i}}}{(\mathbf{j}-\mathbf{i})!} \mu_{\mathbf{j}}, \end{aligned} \tag{14}$$

where

$$\frac{(\mathbf{s_l}\theta)^{\mathbf{k}}}{\mathbf{k}!} = \frac{(s_{l1}\theta_1)^{k_1}}{k_1!} \times \ldots \times \frac{(s_{lN}\theta_N)^{k_N}}{k_N!}$$

and

$$\mu_{\mathbf{j}} = \mu_{j_1, j_2, \ldots, j_N}(t).$$

Extracting the coefficients of θ from (14) yields

$$\sum_{\mathbf{n=0}}^{\infty} \theta^{\mathbf{n}} \sum_{\mathbf{i}} h_{\mathbf{i}} \frac{\partial}{\partial t} \mu_{\mathbf{i+n}} = \sum_{\mathbf{n=0}}^{\infty} \theta^{\mathbf{n}} \left\{ \sum_{l=1}^{L} \left[\sum_{\mathbf{i}} f_{l,\mathbf{i}} \sum_{\mathbf{k=0}}^{\mathbf{n}} \mathbf{s_l}^{\mathbf{k}} \binom{\mathbf{n}}{\mathbf{k}} \mu_{\mathbf{n-k+i}} \right] - \sum_{\mathbf{i}} g_{\mathbf{i}} \mu_{\mathbf{n+i}} \right\}, \tag{15}$$

where

$$\mathbf{s_l}^{\mathbf{k}} \binom{\mathbf{n}}{\mathbf{k}} = s_{l1}^{k_1} \binom{n_1}{k_1} \times \ldots \times s_{lN}^{k_N} \binom{n_N}{k_N}$$

and

$$\mu_{\mathbf{n-k+i}} = \mu_{n_1-k_1+i_1,\ldots,n_N-k_N+i_N}(t).$$

To obtain an equation for a particular moment, we extract the coefficients of $\theta_1, \ldots, \theta_N$ from each side of equation (15). Hence equation (15) is a general result that can be applied to a wide class of stochastic models. For example, to obtain the moment equation for the marginal mean, we extract the coefficients of θ_j to get an equation for the first moment of species j. If we required an equation for $\mu_{1,0,\ldots,0} = \mu_{\mathbf{j}}$, we extract coefficients of θ_1 only, yielding

$$\sum_{\mathbf{i}} h_{\mathbf{i}} \frac{\partial}{\partial t} \mu_{\mathbf{i+n}} = \sum_{l=1}^{L} \left(\sum_{\mathbf{i}} f_{l,\mathbf{i}} \mathbf{s_l}^{\mathbf{n}} \mu_{\mathbf{n-k+i}} \right) - \sum_{\mathbf{i}} g_{\mathbf{i}} \mu_{\mathbf{n+i}}, \tag{16}$$

where $\mathbf{n} = 1, 0, \ldots, 0$. On extracting $\partial \mu_{\mathbf{j}} / \partial t$ we gain the following expression

$$h_{\mathbf{j}} \frac{\partial \mu_{\mathbf{j}}}{\partial t} = \sum_{l=1}^{L} \sum_{\mathbf{i}} f_{l,\mathbf{i}} \mu_{\mathbf{i}} - \sum_{\mathbf{i}} g_{\mathbf{i}} \mu_{\mathbf{i}} - \sum_{\mathbf{i} \neq \mathbf{j}} h_{\mathbf{i}} \frac{\partial \mu_{\mathbf{i}}}{\partial t}, \tag{17}$$

where $\mu_{\mathbf{j}} = \mu_{1,0,\ldots,0}$ and $h_{\mathbf{j}} = h_{1,0,\ldots,0} \neq 0$.

Single species case

If there is only a single species, equation (15) simplifies to

$$\sum_{n=0}^{\infty} \theta_1^n \sum_{i=0}^{\infty} h_i \frac{\partial}{\partial t} \mu_{i+n} = \sum_{n=0}^{\infty} \theta_1^n \left\{ \sum_{l=1}^{L} \left[\sum_{i=0}^{\infty} f_{l,i} \sum_{k=0}^{n} s_l^k \binom{n}{k} \mu_{n-k+i} \right] - \sum_{i=0}^{\infty} g_i \mu_{n+i} \right\}. \tag{18}$$

This allows us to gain a simple approximation for the first moment, which cannot be done for multiple species. In the single species case to find the equation for the first moment we extract the coefficients of θ_1^0, hence the equation for the first moment is

$$h_1 \frac{\partial \mu_1}{\partial t} = \left[\sum_{l=1}^{L} \sum_{i=0}^{\infty} f_{l,i} \mu_i \right] - \sum_{i=0}^{\infty} g_i \mu_i - \sum_{i \neq 1} h_i \frac{\partial \mu_i}{\partial t}, \tag{19}$$

and the equation for the n^{th} moment ($n \geq 2$) is

$$h_1 \frac{\partial \mu_n}{\partial t} = \left[\sum_{l=1}^{L} \sum_{i=0}^{\infty} f_{l,i} \sum_{k=0}^{n} s_l^k \binom{n}{k} \mu_{n-k+i} \right] - \sum_{i=0}^{\infty} g_i \mu_{n+i} - \sum_{i \neq 1}^{\infty} h_i \frac{\partial \mu_{n-1+i}}{\partial t}. \tag{20}$$

This simplification is not possible for models with more than one species as this approximation to the first moments from the coefficients of θ_1^0, gives one equation with N (number of species) unknowns.

4. Further Examples

In this section we will apply our method to two further examples which have rational rate laws. To test the speed and accuracy of our moment closure approximations, we have simulated 'exact' means and variances for each of the examples. The means and variances are calculated from 100,000 simulations of a Gillespie algorithm coded in C. Each of the moment closure approximations was also coded in C. The moment equations were solved using the Runge Kutta Cash Karp (rkck) method from the GSL ODE libraries. All examples are run on a Intel Xeon E5345 2.33GHz processor.

4.1. Example 2: rate laws involving powers

In this example we consider rational rate laws involving powers. This model contains a single species Y and a single reaction:

$$Y \to \emptyset$$

with rate law

$$\frac{Vy^2}{k + y^2}.$$

This is a fairly standard reaction rate and is known as Holling Type III response function for modelling vertebral predators. The rate has in a variety of scenarios, such as modelling the outbreak of budworms in forests and predator-prey systems (see [22, 24, 25]).

The forward Kolmogorov equation for this reaction is

$$\frac{dP(y)}{dt} = P(y - s_l) \frac{\alpha(y - s_l)}{\beta(y - s_l)} - P(y) \frac{\alpha(y)}{\beta(y)}, \tag{21}$$

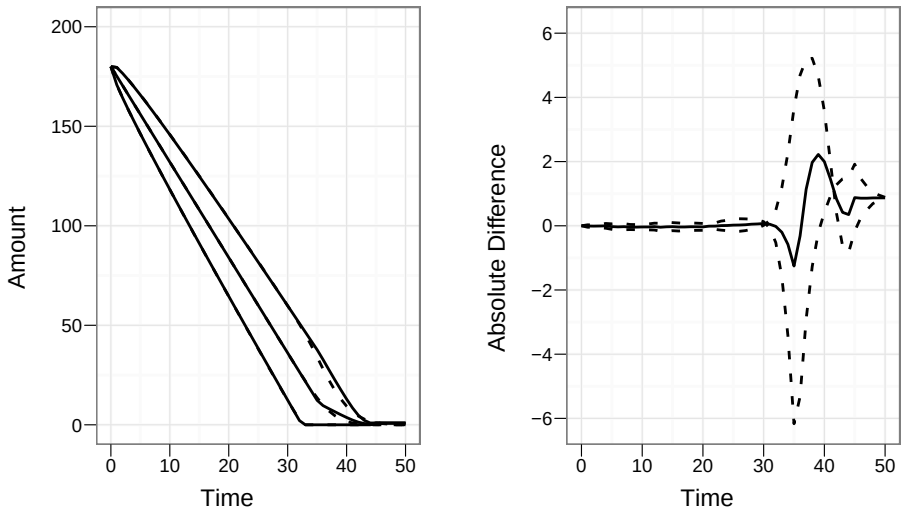

Figure 2: *(a) The mean ± 2 standard deviations for Example 2 obtained from 100,000 simulations (− − −) and the approximate mean ± 2 standard deviations from the moment closure approximation (solid), where* $y(0) = 180$ *and* $\{V, k\} = \{4.8, 5/3\}$*. (b) The error (approximation − exact) for the mean (solid), the upper 95% (− − −) and the lower 95% (− − −).*

where $\alpha(y) = Vy^2$, $\beta(y) = k + y^2$ and $s_l = -1$. From relationship (19), we obtain an equation for the first moment is

$$\begin{aligned}\frac{d\mu_1(t)}{dt} &= \frac{1}{h_1}\left(\sum_{i=2}^{4} f_i\mu_i(t) - \sum_{i=2}^{4} g_i\mu_i(t) - \sum_{i=2}^{4} h_i\frac{d\mu_i(t)}{dt}\right) \\ &= \frac{1}{2k}\left[-4V\mu_3(t) - (2k+1)\frac{d\mu_2(t)}{dt} - 2\frac{d\mu_3(t)}{dt} - \frac{d\mu_4(t)}{dt}\right]. \qquad (22)\end{aligned}$$

Similarly, from relationship (20) we obtain an equation for the second moment

$$\begin{aligned}2k\frac{d\mu_2(t)}{dt} =& -V((k+1)\mu_2(t) - 2\mu_3(t) + 5\mu_4(t)) \\ & - (k^2+k)\frac{d\mu_1(t)}{dt} - (2k+1)\frac{d\mu_3(t)}{dt} - 2\frac{d\mu_4(t)}{dt} - \frac{d\mu_5(t)}{dt}.\end{aligned}$$

This slightly more complicated rate law leads to the ODE for the mean

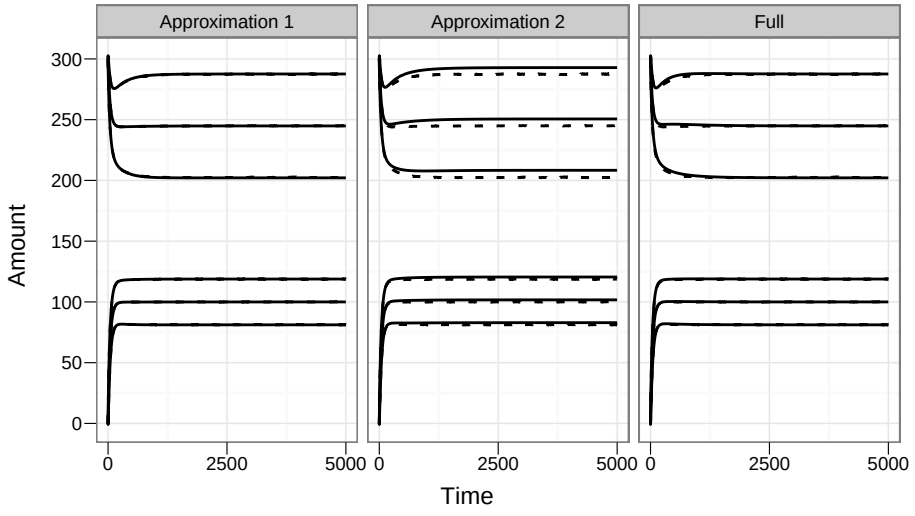

Figure 3: *The exact mean ± from 100,000 simulations 2 standard deviations (− − −) and the approximate mean ± 2 standard deviations from the moment closure approximations (—) for Example 3. The initial conditions are* $\mathbf{y}(0,0) = \{300, 0\}$ *and parameters* $\{\lambda, k, V_1, V_2, C_s, C_r\} = \{1, 0.01, 5, 5, 2, 2\}$.

(equation (22)) depending on $\mu_2(t)$, $\mu_3(t)$ and $\mu_4(t)$, whereas in Example 1 (§2), the ODE for $\mu_1(t)$ (equation (6)) only depends on $\mu_2(t)$. We can again apply a normal approximation to close the set of ODEs, allowing us to solve numerically for given initial conditions and rate constants.

Figure 2 shows a plot of the exact mean and variance against the moment closure approximations of the mean and variance. We can see that the moment closure approximation gives a good estimate of the true values. As we saw in Example 1 the moment closure approximation is furthest from the true values as the population of species Y approaches zero.

4.2. Example 3: rate laws with multiple species

We now consider a Michaelis-Menten type system where the product (P) can reform the substrate enzyme complex (SE). We also consider that the product can be removed from the system and the substrate immigrates into the system. Thus we have the following reactions,

$$R_1 : \emptyset \to S, \quad R_2 : P \to \emptyset, \quad R_3 : S \to P, \quad R_4 : P \to S,$$

with rates λ, kP, $(V_1 S)/(C_s + S + C_r P)$ and $(C_r V_2 P)/(C_s + S + C_r P)$ respectively. The associated forward Kolmogorov equation is

$$\frac{dP(\mathbf{y})}{dt} = \sum_{\mathbf{l}=\mathbf{1}}^{\mathbf{4}} P(\mathbf{y} - \mathbf{s_l}) \frac{\alpha_l(\mathbf{y} - \mathbf{s_l})}{\beta_l(\mathbf{y} - \mathbf{s_l})} - P(\mathbf{y}) \frac{\alpha_l(\mathbf{y})}{\beta_l(\mathbf{y})}, \tag{23}$$

where $\mathbf{y} = (y_1, y_2)$ are the numbers of molecules of substrate and product respectively. From equations (10-12) we can calculate the coefficients $h_{\mathbf{i}}$, $f_{l,\mathbf{i}}$ and $g_{\mathbf{i}}$. Equation (16) thus allows us to form expressions for moments of our choice. We have once again used a normal approximation to close the moment equations, yielding five coupled ODEs (two means, two variances and covariance). These can be solved numerically to give estimates of the means and variances. We will consider three approximations to the means, variances and covariance,

Full Model (Full)

All terms for the means, variances and covariance in the moment equations are included.

Deterministic means (Approximation 1)

By assuming the variance and covariance are zero in the moment equations for the means, we can gain an approximation to the deterministic solution. We can combine this approximation with the equations for the variances and covariance to get a first approximation.

Assume all differentials on the LHS equal zero (Approximation 2)

We assume that the rate of change for each moment is small compared to the moment itself, equivalently all $h_i = 0 \; \forall i \neq 1$. That is we only consider the differential on the LHS that we are interested in.

For each of these schemes we must solve five moment ODEs. Figure 3 shows the mean and variance estimates for each scheme, along with a Monte Carlo estimate of the true values. We observe that the full model and both approximations estimate match the Monte Carlo estimates fairly closely. As we would expect, Approximation 2 performs the poorest of the three schemes.

Table 2 shows the CPU time taken (in seconds) for each of the three schemes for different simulation times. Overall, using approximations $\{1, 2\}$ compared to the full moment closure ODES gives a speed increase of about

Simplification	Simulation time	CPU time(s)
Full	50	0.57
Approximation 1	50	0.088
Approximation 2	50	0.004
Full	5000	48.75
Approximation 1	5000	7.52
Approximation 2	5000	0.14

Table 2: *Comparison of simulation times for different moment closure schemes, where* $\mathbf{y}(0,0) = \{300, 0\}$ *and* $\{\lambda, k, V_1, V_2, C_s, C_r\} = \{1, 0.01, 5, 5, 2, 2\}$.

a factor of $\{10, 100\}$ respectively. The CPU time scales linearly with the increase in simulation time.

5. Conclusions

It is becoming increasingly recognised that to accurately capture the fundamental characteristics of many systems, the associated mathematical model should contain stochastic features. For example, Proctor et al. [26] modelled DNA damage-response mechanisms in *Saccharomyces Cerevisiae*. This model highlighted the interaction between DNA repair and checkpoint pathways. Cellular damage, by its very nature, is a stochastic event. Proctor *et al.* highlights that while an ODE representation of the model captures the average behaviour, it fails to highlight the variability of the number of cell divisions in the cdc13-1 mutant strains. Even though this model was based on mass-action kinetics, it was necessary to make some limiting assumptions leading to rational laws.

Waldherr et al. [27] considered stochastic and deterministic models which linked cellular decisions taking place on a time scale of years to decades to cellular decisions that were measured in minutes to hours. By using stochastic bistable switch models (involving rational rate laws) they demonstrated that population traits could be predicted from noisy stochastic systems. In their paper they considered only basic models, so future work involving stochastic models will, by necessity, involve approximations (possibly using moment closure) of the system.

Stochastic hybrid systems (SHS) are increasingly being used by modellers (see [28] for an review). SHS decompose the state space of the model into

a continuous part and a discrete part. Singh and Hespanha highlight that moment closure approximations would be ideal in this setup. The general idea is that fast reactions could be simulated using a fairly accurate moment closure approximation, whilst the slow reaction be simulated separately. The key point is that we are not ignoring the stochasticity of the fast reactions. The method described in this paper provides another tool for stochastic modellers, allowing moment closure techniques to be applied to a wider class of models.

References

[1] D. J. Wilkinson, Stochastic Modelling for Systems Biology, Chapman & Hall/CRC Press, 2006.

[2] V. Isham, Assessing the variability of stochastic epidemics, Mathematical Biosciences 107 (1991) 209–224.

[3] E. Renshaw, Modelling Biological Populations in Space and Time, Cambridge University Press, 1991.

[4] B. Bolker, S. W. Pacala, Using moment equations to understand stochastically driven spatial pattern formation in ecological systems, Theoretical Population Biology 52 (1997) 179 – 197.

[5] D. G. Kendall, An artificial realisation of a simple birth and death process, Journal of the Royal Statisical Society B 12 (1950) 116–119.

[6] D. T. Gillespie, Exact stochastic simulation of coupled chemical reactions, The Journal of Chemical Physics 81 (1977) 2340.

[7] M. A. Gibson, J. Bruck, Efficient exact stochastic simulation of chemical systems with many species and many channels, Journal of Physical Chemistry 104 (2000) 1876–1889.

[8] D. T. Gillespie, A general method for numerically simulating the stochastic time evolution of coupled chemical reactions, Journal of Computational Physics 22 (1976) 403–434.

[9] D. T. Gillespie, Approximate accelerated stochastic simulation of chemically reacting systems, The Journal of Chemical Physics 115 (2001) 1716–1733.

[10] Y. Cao, D. T. Gillespie, L. R. Petzold, Avoiding negative populations in explicit poisson tau-leaping, The Journal of Chemical Physics 123 (2005) 054104.

[11] Y. Cao, D. T. Gillespie, L. R. Petzold, Efficient step size selection for the tau-leaping simulation method, The Journal of Chemical Physics 124 (2006) 044109.

[12] A. Gandhi, S. Levin, S. Orszag, Moment closure in spatial ecological models and moment closure through gaussian approximation, Bulletin of Mathematical Biology 62 (2000) 595–632.

[13] J. H. Matis, T. R. Kiffe, Effects of immigration on some stochastic logistic models: a cumulant truncation analysis, Theoretical Population Biology 56 (1999) 139–161.

[14] M. J. Keeling, Multiplicative moments and measures of persistence in ecology, Journal of Theorical Biology 205 (2000) 269–281.

[15] G. Marion, E. Renshaw, G. Gibson, Stochastic effects in a model of nematode infection in ruminants, IMA Journal of Mathematical Applied in Medicine and Biology 15 (1998) 97–116.

[16] I. Krishnarajah, A. Cook, G. Marion, G. Gibson, Novel moment closure approximations in stochastic epidemics, Bulletin of Mathematical Biology 67 (2005) 855–873.

[17] I. Krishnarajah, G. Marion, G. Gibson, Novel bivariate moment-closure approximations, Mathematical Biosciences 208 (2007) 621–643.

[18] C. S. Gillespie, Moment closure approximations for mass-action models, IET Systems Biology 3 (2009) 52–58.

[19] L. A. Goodman, Population growth of the sexes, Biometrics 9 (1953) 212–225.

[20] P. Whittle, On the use of the normal approximation in the treatment of stochastic processes, Journal of the Royal Statistical Society. Series B (Methodological) 19 (1957) 268–281.

[21] G. S. Williams, M. A. Huertas, E. A. Sobie, M. S. Jafri, G. D. Smith, Moment closure for local control models of calcium-induced calcium release in cardiac myocytes, Biophysical Journal 95 (2008) 1689 – 1703.

[22] C. S. Holling, The functional response of invertebrate predators to prey density, Mem. Entomol. Soc. Can. 48 (1966) 1–86.

[23] L. Michaelis, M. L. Menten, Die kinetik der invertinwirkung, The Journal of Biochemistry 49 (1913) 333.

[24] D. Ludwig, D. D. Jones, C. S. Holling, Qualitative analysis of insect out-break systems: the spruce budworm and forest, Journal of Animal Ecology 47 (1978).

[25] S. B. Hsu, T. W. Huang, Global stability for a class of predator-prey systems, SIAM Journal of Applied Mathematics 55 (1995) 763–783.

[26] C. J. Proctor, D. A. Lydall, R. J. Boys, C. S. Gillespie, D. P. Shanley, D. J. Wilkinson, T. B. L. Kirkwood, Modelling the checkpoint response to telomere uncapping in budding yeast., Journal of the Royal Society, Interface / the Royal Society 4 (2007) 73–90.

[27] S. Waldherr, J. Wu, F. Allgöwer, Bridging time scales in cellular decision making with a stochastic bistable switch., BMC systems biology 4 (2010) 108.

[28] A. Singh, J. P. Hespanha, Stochastic hybrid systems for studying biochemical processes, Philosophical Transactions of the Royal Society A 368 (2010).