View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by ScholarWorks@UNIST

AIP | The Journal of Chemical Physics

A moment closure method for stochastic reaction networks

Chang Hyeong Lee, Kyeong-Hun Kim, and Pilwon Kim

Citation: The Journal of Chemical Physics **130**, 134107 (2009); doi: 10.1063/1.3103264
View online: http://dx.doi.org/10.1063/1.3103264
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/130/13?ver=pdfcov
Published by the AIP Publishing

Articles you may be interested in

Validity conditions for moment closure approximations in stochastic chemical kinetics
J. Chem. Phys. **141**, 084103 (2014); 10.1063/1.4892838

Adaptively biased sequential importance sampling for rare events in reaction networks with comparison to exact solutions from finite buffer dCME method
J. Chem. Phys. **139**, 025101 (2013); 10.1063/1.4811286

A probability generating function method for stochastic reaction networks
J. Chem. Phys. **136**, 234108 (2012); 10.1063/1.4729374

A study of the accuracy of moment-closure approximations for stochastic chemical kinetics
J. Chem. Phys. **136**, 154105 (2012); 10.1063/1.3702848

Fluctuation statistics in networks: A stochastic path integral approach
J. Math. Phys. **45**, 4386 (2004); 10.1063/1.1803927

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 114.70.7.203 On: Mon, 10 Nov 2014 08:41:41

A moment closure method for stochastic reaction networks

Chang Hyeong Lee,[1,a)] Kyeong-Hun Kim,[2,b)] and Pilwon Kim[3,c)]
[1]*Department of Mathematical Sciences, Worcester Polytechnic Institute, 100 Institute Road, Worcester, Massachusetts 01609, USA*
[2]*Department of Mathematics, Korea University, 1 Anam-dong, Sungbuk-gu, Seoul 136-701, South Korea*
[3]*Department of Mathematics, The Ohio State University, 231 W. 18th Ave., Columbus, Ohio 43210, USA*

(Received 26 November 2008; accepted 27 February 2009; published online 6 April 2009)

In this paper we present a moment closure method for stochastically modeled chemical or biochemical reaction networks. We derive a system of differential equations which describes the dynamics of means and all central moments from a chemical master equation. Truncating the system for the central moments at a certain moment term and using Taylor approximation, we obtain explicit representations of means and covariances and even higher central moments in recursive forms. This enables us to deal with the moments in successive differential equations and use conventional numerical methods for their approximations. Furthermore, we estimate the errors in the means and central moments generated by the approximation method. We also find the moments at equilibrium by solving truncated algebraic equations. We show in examples that numerical solutions based on the moment closure method are accurate and efficient by comparing the results to those of stochastic simulation algorithms. © *2009 American Institute of Physics*.
[DOI: 10.1063/1.3103264]

I. INTRODUCTION

A well-mixed biochemical reaction system at constant temperature and volume has traditionally been modeled by a system of ordinary differential equations (ODEs). The general form of the governing equation is

$$\frac{dc}{dt} = V\mathcal{F}(c(t)), \tag{1}$$

where $c(t)=(c_1(t),\ldots,c_n(t))$, each $c_i(t)$ is the concentration of ith species at time t, V is the stoichiometric matrix, and $\mathcal{F}(c)$ is the vector of reaction rate functions, which are often determined by mass-action kinetics. Equation (1) is deterministic since it has no terms corresponding to random events and gives a complete description of the dynamics of the system for a given initial condition.

However, if the system is characterized by interactions of small number of molecular species, the ODE system (1) may not be accurate since the molecular interactions are inherently stochastic and the state variables for describing the dynamics of molecular species are discrete random variables rather than continuous deterministic variables. To describe the random behavior of such a system and investigate how each molecular species evolves, a stochastic description is required.

In the stochastic description, a state variable is defined as $N(t)=(N_1(t),N_2(t),\ldots,N_s(t))$, where each $N_i(t)$ is a random variable that denotes the molecular number of ith species at time t. Under Markovian assumption on the system, a governing equation, the so-called *chemical master equation*, is given by

$$\frac{dp(n,t)}{dt} = \sum_k (\mathcal{R}_k(n-V_k)p(n-V_k,t) - \mathcal{R}_k(n)p(n,t)), \tag{2}$$

where $p(n,t)$ denotes the probability of the state $n=(n_1,\ldots,n_s)$ at time t, i.e., $p(n,t)=\mathrm{Prob}(N(t)=n)$, $\mathcal{R}_k(n)$ denotes the so-called *propensity* of kth reaction at the state n, and V_k denotes the kth column vector of the stoichiometric matrix V.[1]

If one could solve Eq. (2), the solution would completely describe the stochastic dynamics of the system. However, it is generally very difficult, if not impossible, to find the solution of Eq. (2) for most complex chemical systems. For this reason, researchers often resort to simulating the dynamics of system by stochastic simulation algorithms (SSAs) such as Gillespie algorithm or its modifications,[2–4] or turn to finding the first moment(=mean) and the second central moment(=covariance) instead of the probability distribution. It is understandable in that the two moments are computed more easily than the probability distribution and also they give a decent description for the stochastic dynamics such as averaging behavior and evolution of the noise in the system.

For linear systems, an explicit expression for the first two moments can be found.[5] On the contrary, if nonlinear reactions are involved in a given system, it is not the case because equations of moments are not closed, i.e., any differential equation that describes time evolution of a moment includes one or more higher moment terms. To find each moment for such a system, one has to solve an infinite dimensional ODE system. There have been several attempts to solve this problem in the literature of population biology and chemistry, most of which are based on the assumption that

[a)]Author to whom correspondence should be addressed. Electronic mail: changlee@wpi.edu.
[b)]Electronic mail: kyeonghun@korea.ac.kr.
[c)]Electronic mail: pwkim@math.ohio-state.edu.

 © 2009 American Institute of Physics

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 114.70.7.203 On: Mon, 10 Nov 2014 08:41:41

the distribution of the system follows a certain distribution such as normal[6–8] or log normal.[9] Recently, Hespanha and Singh[10] presented a stochastic hybrid method for approximating an infinite ODE system of moments by a finite nonlinear ODE system of the first two moments and Engblom[11] derived a general form of equations for higher order moments.

In this paper, we derive an ODE system of moment equations for stochastic reaction networks and develop a moment closure method (MCM) for efficient approximation. While a similar approach in obtaining the ODE system of moment equations was studied,[11] the derivation of the ODE system in this paper is more explicitly developed in relation to approximation. We further truncate the equations at a certain order of moments so that conventional numerical schemes can be applied to them. To the authors' knowledge, the recursive computation of moments by the truncation of moment equations at high order moments and the error estimation generated by the truncation have not been reported yet.

The outline of the paper is as follows. In Sec. II, we derive a system of ODEs for means and all central moments. In Sec. III, a MCM and numerical approximation based on the method are developed. In Sec. IV, we illustrate its performances through three examples.

II. MOMENT EQUATIONS

We consider a general stochastic chemical or biochemical network with r reactions and s species. As described in Sec. I, the governing equation of the system is

$$\frac{dp(n,t)}{dt} = \sum_{k=1}^{r} (\mathcal{R}_k(n-V_k)p(n-V_k,t) - \mathcal{R}_k(n)p(n,t)). \quad (3)$$

We can also describe the dynamics of the stochastic system by a Markov chain, by considering all possible states and transitions between the states. In this description, the governing equation can be written as a Kolmogorov system

$$\frac{dP(t)}{dt} = KP(t), \quad (4)$$

where $P(t)$ is the vector whose ith entry is the probability of the ith state n^i at time t and K is the transition matrix whose off-diagonal entries are determined by the propensities $\mathcal{R}_k$ and diagonal entries by $K_{jj}=-\Sigma_{i\neq j}K_{ij}$. The formal solution of Eq. (4) is given by

$$P(t) = e^{Kt}P(0),$$

which presents how the system evolves exactly. Note that K is infinite dimensional if the state space of the system is unbounded, and even when the state space is bounded, it is a big and sparse matrix. Due to this reason, one has difficulties in solving Eq. (4) numerically.

A. Derivation of moment equations

In this section we derive an ODE system of any moments for molecular species in reaction networks. Before doing so, we assume that all reactions in the networks are at most bimolecular, since monomolecular, bimolecular, and thermolecular reactions essentially account for all reaction mechanisms in chemical systems and most thermolecular reactions can be considered as two consecutive bimolecular reactions. Thus, hereafter we assume that the propensities $\mathcal{R}_k(N)$ are at most quadratic functions of N.

To obtain the moment equations for the first moments, we multiply n_i to Eq. (3) and sum over all accessible states $n=(n_1,\ldots,n_s)$ and then obtain

$$\sum_n n_i \frac{dp(n,t)}{dt} = \sum_k \sum_n n_i \mathcal{R}_k(n-V_k)p(n-V_k,t) - n_i\mathcal{R}_k(n)p(n,t). \quad (5)$$

By applying a transformation $n-V_k \to n$ in the first term of the right-hand side of the above equation and then rewriting $\mathcal{R}_k$ in terms of $(N-\mu)$ or using the Taylor formula for $\mathcal{R}_k$ at $N=\mu$, one can obtain equations for the first moments and second central moments,

$$\frac{d\mu_i}{dt} = \sum_k V_{k,i}\left(\mathcal{R}_k(\mu) + \frac{1}{2}\sum_{l,m} \frac{\partial^2 \mathcal{R}_k(\mu)}{\partial N_l\, \partial N_m}\sigma_{lm}\right) \quad (6)$$

$$\begin{aligned}\frac{d\sigma_{ij}}{dt} = \sum_k \Bigg[& V_{k,i}\sum_\ell \frac{\partial \mathcal{R}_k(\mu)}{\partial N_\ell}\sigma_{j\ell} + V_{k,j}\sum_\ell \frac{\partial \mathcal{R}_k(\mu)}{\partial N_\ell}\sigma_{i\ell} \\ & + V_{k,i}V_{k,j}\left(\mathcal{R}_k(\mu) + \frac{1}{2}\sum_{\ell,m}\frac{\partial^2 \mathcal{R}_k(\mu)}{\partial N_\ell\, \partial N_m}\sigma_{\ell m}\right) \\ & + V_{k,i}\frac{1}{2}\sum_{\ell,m}\frac{\partial^2 \mathcal{R}_k(\mu)}{\partial N_\ell\, \partial N_m}\sigma_{j\ell m} \\ & + V_{k,j}\frac{1}{2}\sum_{\ell,m}\frac{\partial^2 \mathcal{R}_k(\mu)}{\partial N_\ell\, \partial N_m}\sigma_{i\ell m}\Bigg],\end{aligned} \quad (7)$$

where $\sigma_{ij}=E[(N_i-\mu_i)(N_j-\mu_j)]$ and $\sigma_{ijk}=E[(N_i-\mu_i)(N_j-\mu_j)\times(N_k-\mu_k)]$. Refer to Appendix for a detailed derivation.

Furthermore, after some computations, we can obtain the equation for general mth central moments, denoted by

$$M_{i_1,\ldots,i_s} = E[(N_1-\mu_1)^{i_1}\cdots(N_s-\mu_s)^{i_s}],$$

for $m=i_1+\cdots+i_s\geq 2$ as follows:

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 114.70.7.203 On: Mon, 10 Nov 2014 08:41:41

$$\frac{dM_{i_1,i_2,\ldots,i_s}}{dt} = \frac{d}{dt}E[(N_1-\mu_1)^{i_1}\cdots(N_s-\mu_s)^{i_s}] = \sum_k \mathcal{R}_k(\mu) \sum_{\substack{\ell_1,\ell_2,\ldots,\ell_s \\ \ell_1+\cdots+\ell_s\neq m}} \binom{i_1}{\ell_1}\cdots\binom{i_s}{\ell_s}(V_{k,1})^{i_1-\ell_1}\cdots(V_{k,s})^{i_s-\ell_s}M_{\ell_1,\ldots,\ell_s}$$
$$+\sum_k\sum_q \frac{\partial\mathcal{R}_k(\mu)}{\partial N_q} \sum_{\substack{\ell_1,\ell_2,\ldots,\ell_s \\ \ell_1+\cdots+\ell_s\neq m}} \binom{i_1}{\ell_1}\cdots\binom{i_s}{\ell_s}(V_{k,1})^{i_1-\ell_1}\cdots(V_{k,s})^{i_s-\ell_s}M_{\ell_1,\ldots,\ell_q+1,\ldots,\ell_s}$$
$$+\frac{1}{2}\sum_k\sum_{q,r} \frac{\partial^2\mathcal{R}_k(\mu)}{\partial N_q\,\partial N_r} \sum_{\substack{\ell_1,\ell_2,\ldots,\ell_s \\ \ell_1+\cdots+\ell_s\neq m}} \binom{i_1}{\ell_1}\cdots\binom{i_s}{\ell_s}(V_{k,1})^{i_1-\ell_1}\cdots(V_{k,s})^{i_s-\ell_s}M_{\ell_1,\ldots,\ell_s+e_q+e_r} - \sum_{j=1}^{s} i_j\mu_j' M_{i_1,\ldots,i_j-1,\ldots,i_s}, \quad (8)$$

where $\binom{i}{\ell}=i!/\ell!(i-\ell)!$ and the subscript $\ell_1,\ldots,\ell_s+e_q+e_r$ denotes adding 1 to qth and rth entries of $\ell_1,\ldots,\ell_s$, respectively. Refer to Appendix for a detailed derivation.

If an initial condition is given deterministically, i.e., $p(n,t=0)=1$ for a certain state $n=(n_1,\ldots,n_s)$, then $\mu_i(0)=n_i$ for any $i=1,\ldots,s$, and

$$M_{i_1,\ldots,i_s}(0) = \begin{cases} 1, & \text{if } i_1=\cdots=i_s=0 \\ 0, & \text{otherwise,} \end{cases}$$

for any $i_1,\ldots,i_s$ such that $i_1+\cdots+i_s=m\geq 2$. Thus, one can obtain from Eq. (8),

$$M'_{i_1,i_2,\ldots,i_s}(0) = \sum_k V_{k,1}\cdots V_{k,s}\mathcal{R}_k(\mu(0)), \quad (9)$$

which will be used when $M'_{i_1,i_2,\ldots,i_s}(0)$ is computed in Sec. III.

Furthermore, if one can find the central moments from Eq. (8), the time-dependent probability of the system can be written formally by inversion theorem for characteristic function as follows. If we denote the molecular number of a certain species and its mean by N and μ, the characteristic function $\phi(\lambda,t)$ of $(N-\mu)$ at time t is given by

$$\phi(\lambda,t) = E[e^{i\lambda(N-\mu(t))}] = E\left[\sum_{m=0}^{\infty}\frac{(i(N(t)-\mu(t)))^m}{m!}\lambda^m\right]$$
$$= \sum_{m=0}^{\infty}\frac{i^m\lambda^m}{m!}E[(N(t)-\mu(t))^m] = \sum_{m=0}^{\infty}\frac{i^m\lambda^m}{m!}M_m(t),$$

where $M_m(t)=E[(N(t)-\mu(t))^m]$. By the inversion formula,[12] one can obtain a formal expression of probability distribution

$$p(n=\mu+x,t) = p(n-\mu=x,t) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-ix\lambda}\phi(\lambda,t)d\lambda$$
$$= \frac{1}{2\pi}\sum_{m=0}^{\infty}\frac{i^m}{m!}M_m(t)\int_{-\pi}^{\pi} e^{-ix\lambda}\lambda^m d\lambda \quad (10)$$

for each $x=-\mu,-\mu+1,\cdots$ at time t.

B. Moment equations for special cases

When all reactions are zero or first order, $\mathcal{R}_k$ are constant or linear functions. In this case, one can obtain from Eqs. (6) and (7) closed equations for the means and covariances

$$\frac{d\mu_i}{dt} = \sum_k V_{k,i}\mathcal{R}_k(\mu), \quad (11)$$

$$\frac{d\sigma_{ij}}{dt} = \sum_k\left[V_{k,i}\sum_\ell\frac{\partial\mathcal{R}_k(\mu)}{\partial N_\ell}\sigma_{j\ell} + V_{k,j}\sum_\ell\frac{\partial\mathcal{R}_k(\mu)}{\partial N_\ell}\sigma_{i\ell} + V_{k,i}V_{k,j}\mathcal{R}_k(\mu)\right]. \quad (12)$$

Thus, one can explicitly find the expression for the means and covariances by solving Eqs. (11) and (12). Note that in the deterministic description of zero or first order reaction networks, the governing Eq. (1) is of a form

$$\frac{dc_i}{dt} = \sum_k V_{k,i}\mathcal{F}_k(c), \quad (13)$$

where $\mathcal{F}_k$ are constant or linear functions of the concentration c. Thus, one can see that the equation obtained by dividing Eq. (11) by the volume of the system is the same equation as Eq. (13).

If the system is assumed to have (approximately) a symmetric distribution such as multivariate normal distribution,[13] one can obtain closed mean and covariance equations, since third central moments are zero. In this case, one can rewrite Eq. (7) as

$$\frac{d\sigma_{ij}}{dt} = \sum_k\left[V_{k,i}\sum_\ell\frac{\partial\mathcal{R}_k(\mu)}{\partial N_\ell}\sigma_{j\ell} + V_{k,j}\sum_\ell\frac{\partial\mathcal{R}_k(\mu)}{\partial N_\ell}\sigma_{i\ell} + V_{k,i}V_{k,j}\left(\mathcal{R}_k(\mu)+\frac{1}{2}\sum_{\ell,m}\frac{\partial^2\mathcal{R}_k(\mu)}{\partial N_\ell\,\partial N_m}\sigma_{\ell m}\right)\right]. \quad (14)$$

Thus, one can obtain the means and covariances by solving Eqs. (6) and (14).

Note that if the system has a second-order reaction, Eqs. (6) and (7) are not closed, since Eq. (6) contains second central moment terms and Eq. (7) contains third central moment terms. In general, one can see that $(m+1)$st central

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 114.70.7.203 On: Mon, 10 Nov 2014 08:41:41

moment appears in mth central moment equation due to the existence of a second-order reaction. Thus, in this case, it is difficult, if not impossible, to find the explicit solution for the mean, covariance, and higher central moments.

III. MOMENT CLOSURE METHOD

In this section we develop a MCM to obtain the approximate solutions of means and covariances for molecular numbers of species in general reaction networks. Its numerical performance will be tested in three examples in Sec. IV.

A. Approximation of moments

Let us discuss the approximation based on the Taylor theorem first. Due to the recursive structure of Eqs. (6)–(8), one can derive a numerical solution as accurate as one wants, using Taylor theorem, as follows. First, in the case that a system has bounded state space, the matrix K in Eq. (4) is finite dimensional. Here note that every closed reaction system has bounded state space, but a system with a bounded state space is not necessarily closed, e.g., $S_1+S_2 \rightleftarrows S_3 \to \phi$. In such a system with bounded state space, the solution $P(t)$ of Eq. (4) is given by linear combinations of exponential terms $t^i e^{\lambda_j t}$, where λ_j are eigenvalues of K and i is a nonnegative integer. More precisely, one can write the solution $P(t)$ of Eq. (4) as

$$P(t) = e^{Kt}P(0) = \sum_j \left[e^{\lambda_j t}\Phi_j + \sum_{i=1}^{N-1} \frac{t^i}{(i-1)!} e^{\lambda_j t}(D_j)^i \right] P(0), \tag{15}$$

where Φ_j and D_j are the eigenprojection and the eigen-nilpotent of the eigenvalue λ_j of K, respectively, and N is the order of eigen-nilpotents of K.[15] Thus, each component $P_i(t)$ of the vector $P(t)$ is an analytic function in the set $\mathbb{R}$ of real numbers by which we mean that each $P_i(t)$ is infinitely differentiable and power series expandable. Therefore, for any point $t_0 \in \mathbb{R}$, the Taylor series

$$\sum_{k=0}^{\infty} \frac{P_i^{(k)}(t_0)}{k!}(t-t_0)^k$$

converges to $P_i(t)$ in a neighborhood of t_0. The mean $E[N_i(t)]=\mu_i(t)$ is also an analytic function in $\mathbb{R}$, since $\mu_i(t)$ is the finite sum $\Sigma_{n_i} n_i P_{(i)}(t)$, where $P_{(i)}(t)$ denotes the probability that $N_i(t)=n_i$, which can be found from the probability solution $P(t)$. Similarly, any mth central moments for $m \geq 2$ are analytic functions in $\mathbb{R}$. Thus, according to the Taylor theorem, the exact means and variances are written as

$$\mu_i(t) = \sum_{k=0}^{\infty} \frac{\mu_i^{(k)}(0)}{k!} t^k, \tag{16}$$

$$\sigma_{ii}(t) = \sum_{k=0}^{\infty} \frac{\sigma_{ii}^{(k)}(0)}{k!} t^k, \tag{17}$$

and also one can approximate the means and variances by Taylor polynomials of degree N,

$$\mu_i(t) \approx \sum_{k=0}^{N} \frac{\mu_i^{(k)}(0)}{k!} t^k, \tag{18}$$

$$\sigma_{ii}(t) \approx \sum_{k=0}^{N} \frac{\sigma_{ii}^{(k)}(0)}{k!} t^k. \tag{19}$$

Here N is a positive integer that denotes a truncation degree. The errors in Eqs. (18) and (19) are

$$\frac{\mu_i^{(N+1)}(t_1)}{(N+1)!} t^{N+1}$$

and

$$\frac{\sigma_{ii}^{(N+1)}(t_2)}{(N+1)!} t^{N+1},$$

respectively, for certain $t_1, t_2 \in [0,t]$. Note that the approximations (18) and (19) are valid under the condition that means and moments are infinitely differentiable, which is a weaker condition than analyticity.

In case that a system has unbounded state space, e.g., $\phi \to S_1 \rightleftarrows S_2 \to \phi$, the matrix K is infinite dimensional and therefore there is no guarantee that the solution of Eq. (4) is analytic. Thus, in this case, the Taylor representations (16) and (17) for mean and variance may not be valid. However, the Taylor approximations (18) and (19) are still valid, since it is obvious from the infinite dimensional ODE systems (6)–(8) that the means and variances are infinitely differentiable.

Once the initial distribution of the system is given, one can find $\mu_i^{(k)}(0)$ and $\sigma_{ii}^{(k)}(0)$, $k=0,1,2,\cdots$, and therefore approximate $\mu_i(t)$ and $\sigma_{ii}(t)$ as in Eqs. (18) and (19). Let us describe in detail a recursive procedure to compute $\mu_i^{(k)}$ and $\sigma_{ii}^{(k)}$ from Eqs. (6)–(8). Hereafter, for the sake of simplicity, we denote by bold types $\boldsymbol{\mu}$ and $\mathbf{M}_k, k \geq 2$ the vector of means and the vector of all kth central moments $M_{i_1,\cdots,i_s}$ with $i_1 + \cdots + i_s = k$, respectively, and we denote by $M_k, k \geq 2$ any one of kth central moments.

We first find $\boldsymbol{\mu}'(0)$ from Eq. (6). The first derivative of the second central moment $\mathbf{M}_2'(0)$ is obtained from Eq. (7) and then again, by taking differentiation on Eq. (6), $\boldsymbol{\mu}''(0)$ is derived. Similarly, we compute $\mathbf{M}_3'(0)$, $\mathbf{M}_2''(0)$, and $\boldsymbol{\mu}^{(3)}(0)$ in order, involving successive differentiations on Eqs. (6)–(8). Note that if the initial condition is deterministic, one can use the simplified Eq. (9) for finding $\mathbf{M}_n'(0)$ for any $n \geq 2$. This procedure can be visualized in a diagram,

$$\boldsymbol{\mu}'(0),$$

$$\mathbf{M}_2'(0) \to \boldsymbol{\mu}''(0),$$

$$\mathbf{M}_3'(0) \to \mathbf{M}_2''(0) \to \boldsymbol{\mu}^{(3)}(0),$$

and in general,

$$\mathbf{M}_n'(0) \to \mathbf{M}_{n-1}''(0) \to \cdots \to \mathbf{M}_2^{(n-1)}(0) \to \boldsymbol{\mu}^{(n)}(0).$$

This procedure generates a Taylor polynomial of degree N for the exact mean and variance.

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 114.70.7.203 On: Mon, 10 Nov 2014 08:41:41

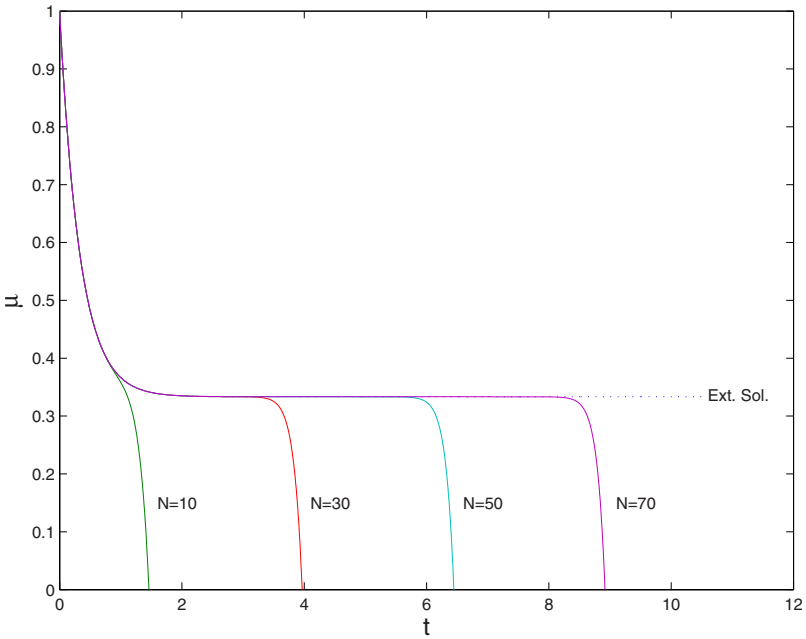

FIG. 1. (Color online) Comparison of the MCM with various degrees of truncations. The example of the reversible bimolecular reaction in Sec. IV A is used with the initial condition $n_1(0)=50$ and k_1, $k_{-1}=1$.

Figure 1 depicts several approximations for the mean with various truncation degree N. One can see that the approximation range gradually widens as N gets increased. A weak point of this approximation is that although the above recursive algorithm based on the Taylor expansion gives as accurate numerical solutions as we want by raising the truncation degree, its computational cost is generally expensive for a practical use. As N increases, the amount of computation involved in symbolic differentiation grows at a rate of N^2, while the approximation range extends linearly with N as in Fig. 1. Moreover, it is not easy to find explicitly *a priori* estimation for the range of the approximation within a certain error; suppose we want to make the errors by Taylor polynomial (18) of the mean μ_i less than a constant C as follows:

$$\left| \frac{\mu_i^{(N+1)}(t_1)}{(N+1)!} t^{N+1} \right| \leq C \tag{20}$$

for some $t_1 \in [0,t]$. Thus, we have

$$0 \leq t \leq \left(\frac{(N+1)!\,C}{|\mu_i^{(N+1)}(t_1)|} \right)^{N+1}.$$

Since we only know the existence of the t_1 from Taylor theorem, if we let $D_i(N+1,t) = \max_{s\in[0,t]} |\mu_i^{(N+1)}(s)|$, we obtain

$$0 \leq t \leq \left(\frac{(N+1)!\,C}{D_i(N+1,t)} \right)^{N+1}.$$

Then, one can obtain *a priori* estimation for the interval of the approximation if one knows an upper bound for $(N+1)$st derivative of the mean μ_i. However, it is not easy to find an upper bound for $(N+1)$st derivative of μ_i when μ_i is unknown, which is the case for nonlinear reaction systems. Thus, an explicit *a priori* estimation is hardly obtainable in general nonlinear reaction systems. The same difficulty arises when one finds *a priori* estimation for the interval of the approximation of the variance σ_{ii}.

This observation turns our attention to alternative methods for numerical integration that gives an explicit error estimation.

B. Truncation of the system

While Eqs. (6)–(8) actually define an infinite dimensional systems of ODEs, one can reduce the system into finite one by adopting an assumption based on the recurrence structure of the equations. If the Nth moments $\mathbf{M}_N$ are assumed to be constantly zero, the system is separated into two parts, finite and infinite dimensional ones. The finite dimensional subsystem only contains $\boldsymbol{\mu}$ and low order moments $\mathbf{M}_2, \cdots \mathbf{M}_{N-1}$ and works independently of the other part. This truncation is well justified in terms of numerical approximation by the following Theorem 3. Once we obtain the corresponding finite subsystem, we can apply conventional numerical schemes to obtain numerical solutions.

The idea of truncation can be seen as a generalization of the normal approximation.[6,14] In the normal approximation, it is assumed that nth moment $\mathbf{M}_n=0$ for all odd n, and so one can close the moment Eq. (8) at $n=3$. The moment Eq. (8) enables us to improve this closure method and close Eq. (8) at an arbitrary $n \geq 3$ without placing any assumption on $\mathbf{M}_3$.

Let us estimate the errors generated by the truncation method. We can write Eqs. (6)–(8) for any means and kth central moment $\mathbf{M}_k$ for $k \geq 2$ as

$$\frac{d\boldsymbol{\mu}}{dt} = \mathbf{f}_1(\boldsymbol{\mu}) + \mathbf{A}_2\mathbf{M}_2, \tag{21}$$

$$\frac{d\mathbf{M}_k}{dt} = \mathbf{f}_k(\boldsymbol{\mu}, \boldsymbol{\mu}', \mathbf{M}_2, \ldots, \mathbf{M}_k) + \mathbf{A}_{k+1}\mathbf{M}_{k+1}, \tag{22}$$

where $\mathbf{f}_k, k \geq 1$ and $\mathbf{A}_k, k \geq 2$ are vector functions and constant matrices determined by Eqs. (6) and (8), respectively. By setting $\mathbf{M}_N=0$, we truncate the infinite system (8) into

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 114.70.7.203 On: Mon, 10 Nov 2014 08:41:41

finite set of equations only with means and first $N-1$ central moments as follows:

$$\frac{d\boldsymbol{\mu}}{dt} = \mathbf{f}_1(\boldsymbol{\mu}) + \mathbf{A}_2\mathbf{M}_2, \tag{23}$$

$$\frac{d\mathbf{M}_i}{dt} = \mathbf{f}_i(\boldsymbol{\mu},\boldsymbol{\mu}',\mathbf{M}_2,\ldots,\mathbf{M}_i) + \mathbf{A}_{i+1}\mathbf{M}_{i+1}, \quad \text{for } i=2,\ldots,N-2, \tag{24}$$

$$\frac{d\mathbf{M}_{N-1}}{dt} = \mathbf{f}_{N-1}(\boldsymbol{\mu},\boldsymbol{\mu}',\mathbf{M}_2,\ldots,\mathbf{M}_{N-1}). \tag{25}$$

In the following theorem, the big O notation $O(h^n)$ is adopted, by which we mean there is a constant $K>0$ such that

$$\lim_{h\to 0}\left|\frac{O(h^n)}{h^n}\right| \le K.$$

Theorem 1: *Suppose that we truncate systems* (21) *and* (22) *by letting* Nth central moment $\mathbf{M}_N(t)=0$ at any t in a time interval $[0,h]$, then the error between the exact mean and the approximated mean obtained from the truncated equations is $O(h^{N-1})$ and the error in second central moments $\mathbf{M}_2(t)$ is $O(h^{N-2})$ for any $t\in[0,h]$.

Proof: Note that since $M_N(t)$ is a continuous function and so it is bounded in $[0,h]$, there is a constant $C>0$ such that for any Nth moments M_N,

$$|M_N(t)| \le C$$

for any $t\ge 0$ in a time interval $[0,h]$. Thus, for any t in $[0,h]$, one can obtain the solution of Eq. (22) for $k=N-1$,

$$\mathbf{M}_{N-1}(t) = \int_0^t \mathbf{f}_{N-1}ds + \mathbf{A}_N\int_0^t \mathbf{M}_N ds \equiv \int_0^t \mathbf{f}_{N-1}ds + \mathbf{C}(N-1,t). \tag{26}$$

Here note that the vector function $\mathbf{C}(N-1,t)$ is the error in $\mathbf{M}_{N-1}$ generated by letting $\mathbf{M}_N=0$ and its ith component $\mathbf{C}^i(N-1,t)$ is bounded by

$$\begin{aligned}|\mathbf{C}^i(N-1,t)| &= \left|\mathbf{A}_N^i\cdot\int_0^t \mathbf{M}_N ds\right| \\ &\le |\mathbf{A}_N^i|\left|\int_0^t \mathbf{M}_N ds\right| \\ &\le |\mathbf{A}_N|hC \equiv hD_N,\end{aligned}$$

where $\mathbf{A}_N^i$ denotes ith row vector of $\mathbf{A}_N$, $|\mathbf{v}|$ denotes the length of the vector $\mathbf{v}$, and $|\mathbf{A}_N|=\max_i\{|\mathbf{A}_N^i|\}$. Moreover, for the vector $\mathbf{M}_{N-2}$ of $(N-2)$nd moments,

$$\begin{aligned}\frac{d\mathbf{M}_{N-2}}{dt} &= \mathbf{f}_{N-2}(\boldsymbol{\mu},\boldsymbol{\mu}',\mathbf{M}_2,\ldots,\mathbf{M}_{N-2}) + \mathbf{A}_{N-1}\mathbf{M}_{N-1} \\ &= \mathbf{f}_{N-2}(\boldsymbol{\mu},\boldsymbol{\mu}',\mathbf{M}_2,\ldots,\mathbf{M}_{N-2}) \\ &\quad + \mathbf{A}_{N-1}\left(\int_0^t \mathbf{f}_{N-1}ds + \mathbf{C}(N-1,t)\right),\end{aligned}$$

and so

$$\mathbf{M}_{N-2}(t) \approx \int_0^t \mathbf{f}_{N-2}ds + \mathbf{A}_{N-1}\int_0^t\left(\int_0^s \mathbf{f}_{N-1}dr\right)ds,$$

where the approximation error in any ith component of $\mathbf{M}_{N-2}$ is

$$\begin{aligned}\left|\mathbf{A}_{N-1}^i\cdot\int_0^t \mathbf{C}(N-1,s)ds\right| &\le |\mathbf{A}_{N-1}^i|\left|\int_0^t \mathbf{C}(N-1,s)ds\right| \\ &\le |\mathbf{A}_{N-1}|h^2D_N \equiv h^2D_{N-1},\end{aligned}$$

where $|\mathbf{A}_{N-1}|=\max_i\{|\mathbf{A}_{N-1}^i|\}$. Similarly, one obtains

$$\begin{aligned}\mathbf{M}_{N-3}(t) \approx &\int_0^t \mathbf{f}_{N-3}(\boldsymbol{\mu},\boldsymbol{\mu}',\ldots,\mathbf{M}_{N-3})ds + \mathbf{A}_{N-2}\int_0^t\left(\int_0^s \mathbf{f}_{N-2}dr\right. \\ &\left. + \mathbf{A}_{N-1}\int_0^s\left(\int_0^r \mathbf{f}_{N-1}du\right)dr\right)ds,\end{aligned}$$

where the approximation error in any ith component of $\mathbf{M}_{N-3}$ is less than $|\mathbf{A}_{N-2}|h^3D_{N-1}\equiv h^3D_{N-2}$ and $|\mathbf{A}_{N-2}| =\max_i\{|\mathbf{A}_{N-2}^i|\}$. By applying the same argument repeatedly, one can see that the approximation errors in $\mathbf{M}_2$ and $\boldsymbol{\mu}(t)$ are $O(h^{N-2})$ and $O(h^{N-1})$, respectively. This completes the proof of the theorem.

The theorem guarantees consistency of truncation as a numerical approximation. Raising truncation order N generally improves approximation. However, if the chosen order N is even, one can most likely obtain a better result with the order $N-1$. To explain this interesting phenomenon, one needs to understand in the above proof that the order of magnitude of errors depends proportionally on the bound C of the truncated moment M_N. We argue that truncation at an even order, say k, is generally less beneficial for approximation than truncation at the preceding odd order $k-1$, since M_{k-1} is likely smaller than M_k in magnitude. Particularly,

$$|M_{k-1}(t)| \le |M_k(t)|, \quad t\in[0,h]. \tag{27}$$

This can be roughly justified as follows. Observing a kth moment $M_k=M_{i_1,i_2,\ldots,i_s}=\Sigma_{n_i}(n_1-\mu_1)^{i_1}\cdots(n_s-\mu_s)^{i_s}p(n,t)$, one can see that all summands of M_k are positive. On the contrary, each summand of the M_{k-1} is either positive or negative, depending on the sign of each factor $n_i-\mu_i$. This makes M_{k-1} most probably less than M_k. (For a detailed proof for the case of a single variable, refer to Appendix.) Moreover, for a system with a sufficiently large volume, Eq. (27) can be seen as a consequence of Kurtz's limit theorem[14] mentioned in Sec. II B, that is, the distribution of the system is known to approach a multivariate normal distribution as the volume of the system increases, and the fact that the multivariate normal distribution has zero central moments at all odd orders and positive central moments at all even orders.

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 114.70.7.203 On: Mon, 10 Nov 2014 08:41:41

Now suppose we truncate the system at $n=N$ by letting $\mathbf{M}_N=0$. Although the original infinite dimensional system is reduced to finite dimensional one, in most cases it is still too complicated to find an exact solution. Therefore one has to apply a certain numerical scheme to integrate the truncated system numerically. In this respect, the numerical approximation using the MCM is a combination of two procedures: truncation and numerical integration.

The error generated at the first procedure, which is the difference between the exact solution of the original system $\mu(t)$ and the exact solution of the truncated system $\mu^T(t)$, is shown to be $O(h^{N-1})$ in the previous theorem, where h is a small time step. Application of a numerical scheme also generates the second error between $\mu^T(t)$ and the numerical solution $\mu^R(t)$. Thus, the final error is bounded as

$$|\mu(t)-\mu^R(t)| \le |\mu(t)-\mu^T(t)| + |\mu^T(t)-\mu^R(t)|. \tag{28}$$

If the numerical scheme being used has the local error $O(h^M)$, then the local error of the MCM is generally $O(h^L)$, where $L=\min(M,N-1)$.This means that in order to improve numerical approximation, one should raise not only the order of the numerical scheme but also the order of truncation together. A similar argument can be made for variance σ_{ii}.

C. Approximation of equilibriums

It is generally challenging to find equilibriums of complex stochastic reaction systems, either analytically or computationally. The examples in Sec. IV show that the MCM carries out great performances in tracking the solutions and capturing its steady states numerically. However, if finding steady states is the only concern, that is, if we do not need the transition trajectory of the solution but its equilibriums, we can easily find them by solving algebraic equations instead of performing numerical integration.

One can see that for truncated systems (23)–(25) with the truncation order N, the constant solutions for $\boldsymbol{\mu},\mathbf{M}_2,\ldots,\mathbf{M}_{n-1}$ at an equilibrium satisfy

$$0=\mathbf{f}_1(\boldsymbol{\mu})+\mathbf{A}_2\mathbf{M}_2, \tag{29}$$

$$0=\mathbf{f}_n(\boldsymbol{\mu},0,\mathbf{M}_2,\ldots,\mathbf{M}_n)+\mathbf{A}_{n+1}\mathbf{M}_{n+1}, \quad \text{for } n=2,\ldots,N-2, \tag{30}$$

$$0=\mathbf{f}_{N-1}(\boldsymbol{\mu},0,\mathbf{M}_2,\ldots,\mathbf{M}_{N-1}). \tag{31}$$

Note that the derivatives on the left hand side are all zero, so the original differential equations now turn into the algebraic equations. Solving these equations gives an approximation for the steady-state solutions, which would be obtained by the application of the above MCM with a very small time step. This is rather a shortcut to approximation for the steady-state solutions in that solving algebraic equations is far less expensive in computational cost, especially when the system takes long time to settle down to equilibrium.

Solving the algebraic equations naturally leads to multiple sets of solutions, suggesting existence of multiple equilibriums of the original systems. It is interesting that the number of solution sets that one can find in the approximation depends on the truncation order. In many cases, most of those solutions turn out to be negative or complex numbers, and therefore physically invalid. There is commonly one or two valid solution sets derived from the truncated systems. For multiple equilibriums, one can carry out the linear stability analysis on the system and classify the equilibrium of the system according to the initial conditions.

IV. APPLICATIONS

In the following three examples, we present numerical performances of the MCM developed in Sec. III. To show clearly how the MCM works, we first consider two simple but important reaction systems, a reversible bimolecular reaction and an enzyme-substrate model. The third example compares the results of the MCM with those of the τ-leaping method in computational cost.

A. A reversible bimolecular reaction

We consider a simple nonlinear reversible reaction

$$S_1+S_2 \underset{k_2}{\overset{k_1}{\rightleftarrows}} S_3.$$

Let $N_i(t)$ be the number of molecules of species S_i at time t for each $i=1,2,3$. Note that the reaction network is a closed system, which has two conservation relations $n_1+n_3=A$ and $n_2+n_3=B$, where $N_i(t)=n_i$, $i=1,2,3$ at a time t, and A and B are positive integers. Also, the stoichiometric matrix V is given by

$$V=\begin{bmatrix} -1 & 1 \\ -1 & 1 \\ 1 & -1 \end{bmatrix}$$

and propensity functions $\mathcal{R}_1$ and $\mathcal{R}_2$ are given by

$$\mathcal{R}_1=k_1N_1N_2=k_1N_1(B-A+N_1), \quad \mathcal{R}_2=k_2N_3=k_2(A-N_1).$$

Since $\partial\mathcal{R}_1/\partial N_1=k_1(B-A+2N_1)$, $\partial^2\mathcal{R}_1/\partial N_1^2=2k_1$, $\partial\mathcal{R}_2/\partial N_1=-k_2$, and $\partial^2\mathcal{R}_2/\partial N_1^2=0$, by Taylor series expansion of $\mathcal{R}_1$ and $\mathcal{R}_2$ at $N=\mu$, one can obtain

$$\mathcal{R}_1(N_1)=k_1\mu_1(B-A+\mu_1)+k_1(B-A+2\mu_1)(N_1-\mu_1)+k_1(N_1-\mu_1)^2, \tag{32}$$

$$\mathcal{R}_2(N_1)=k_2(A-\mu_1)-k_2(N_1-\mu_1). \tag{33}$$

Hereafter, for convenience of notation, we drop the subscript "1" of N_1 and n_1 and use only N and n instead of N_1 and n_1, respectively. Using Eqs. (6) and (8), one can find the equation for the mean μ of the species S_1,

$$\frac{d\mu}{dt}=(-1)(k_1\mu(B-A+\mu)+k_1M_2)+k_2(A-\mu), \tag{34}$$

and the equation for mth central moment $M_m=E[(N-\mu)^m]$ for $m\ge 2$,

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 114.70.7.203 On: Mon, 10 Nov 2014 08:41:41

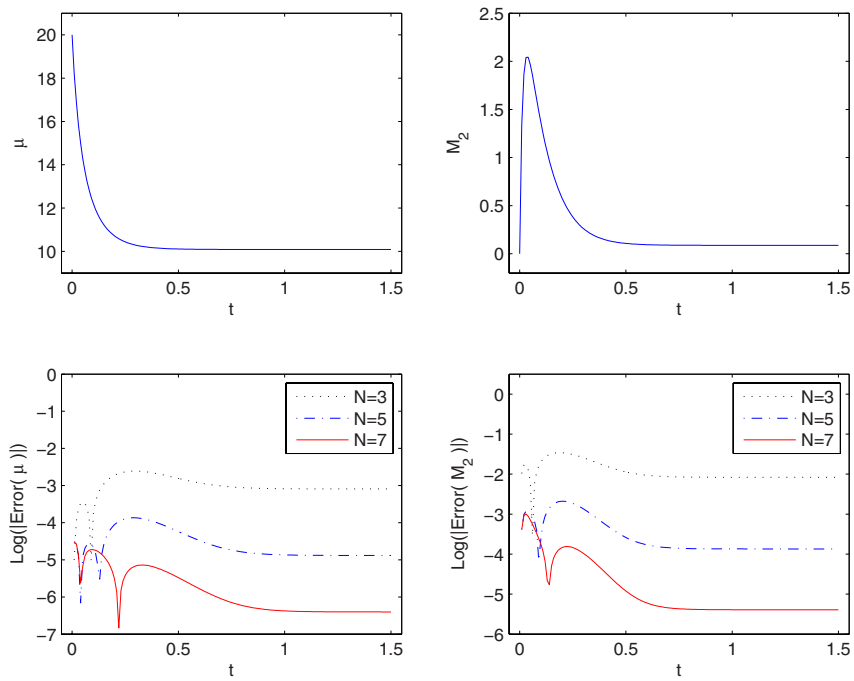

FIG. 2. (Color online) When $N_1(0)=20$, $N_2(0)=10$, $k_1=1$, and $k_2=0.1$ in the dimerization, exact mean and variance of molecular number of species S_1 (upper two graphs) and their errors in $\log_{10}$ generated from RK4 with step size $h=0.01$ and truncation at $N=3,5,7$, that is, $\log_{10}|\text{exact solution}-\text{approximate solution}|$ (lower two graphs). Here exact mean and variance are computed directly by solving the Kolmogorov equation $dp/dt=Kp$.

$$\begin{aligned}\frac{dM_m}{dt} &= \frac{d}{dt}\left(\sum_n (n-\mu)^m p(n,t)\right) = \sum_n (n-\mu)^m \frac{dp(n,t)}{dt} - m\mu' \sum_n (n-\mu)^{m-1} p(n,t) \\ &= k_1\mu(B-A+\mu)\left(\sum_{k=0}^{m-1}\binom{m}{k}(-1)^{m-k}M_k\right) + k_1(B-A+2\mu)\left(\sum_{k=0}^{m-1}\binom{m}{k}(-1)^{m-k}M_{k+1}\right) \\ &\quad + k_1\left(\sum_{k=0}^{m-1}\binom{m}{k}(-1)^{m-k}M_{k+2}\right) + k_2(A-\mu)\left(\sum_{k=0}^{m-1}\binom{m}{k}M_k\right) \\ &\quad - k_2\left(\sum_{k=0}^{m-1}\binom{m}{k}M_{k+1}\right) - m\mu' M_{m-1}, \end{aligned} \tag{35}$$

where $M_0(t)=E[1]=1$ and $M_1(t)=E[N-\mu]=0$ for all $t\geq 0$.

For approximation, we truncate the system of Eqs. (34) and (35) at various order, $N=3,5,7$. On each truncated system, we apply the Runge–Kutta method of order 4 (RK4), which is one of the most popular numerical schemes for ODEs.

Figure 2 depicts the approximate solutions for the mean μ and the variance M_2. One can see that the numerical result improves gradually as the truncation order rises. However, it does not generate much improvement at $N=9$, which is understandable considering that the order of approximation using the MCM is constrained by both the order of truncation and that of the numerical scheme. Refer to Eq. (28).

Although the above numerical solutions successfully capture the equilibrium solutions for μ and M_2, one can derive even better approximation for equilibriums using the method mentioned in Sec. III C. We turn Eqs. (34) and (35) into algebraic equations by setting $d\mu/dt=dM_m/dt=0$. Solving the corresponding algebraic equations about $\mu, M_2, \ldots, M_{N-1}$ gives the steady-state solutions. Note that this result is the same as what we would obtain by the application of the MCM with infinitely small step size h. Table I shows the errors at equilibrium obtained from the method in Sec. III C.

The errors seem to agree well with the final errors appearing in Fig. 2. However, they are much better in that we obtained the same results even without numerical integration. If the system would take longer time to settle down to equilibrium, this difference in computational cost would be huge. Therefore, especially when our concern is to approximate only equilibriums and not the whole trajectories of the system evolution, the suggested method provides a definite shortcut.

TABLE I. Errors at equilibrium of the reversible bimolecular reaction.

	μ	M_2
$N=3$	8.1053×10^{-4}	8.3307×10^{-3}
$N=5$	1.3115×10^{-5}	1.3481×10^{-4}
$N=7$	3.9524×10^{-7}	4.0627×10^{-6}

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 114.70.7.203 On: Mon, 10 Nov 2014 08:41:41

B. Enzyme-substrate kinetics

We consider the enzyme-substrate model

$$E + S \underset{k_2}{\overset{k_1}{\rightleftarrows}} ES \overset{k_3}{\rightarrow} E + P, \tag{36}$$

where E is an enzyme, S is a substrate, ES is an enzyme-substrate complex, and P is a product. Let $N(t) = (N_1(t), N_2(t), N_3(t), N_4(t))$ be the vector of numbers of molecules of E, S, ES, and P at time t, respectively. Note that this reaction network has two conservation relations, $n_1+n_3 \equiv A$ and $n_2+n_3+n_4=B$, where A and B are positive integers and $N_i(t)=n_i$, $i=1,2,3,4$ at time t. Since $n_3=A-n_1$ and $n_4 = B-n_2-n_3=B-A+n_1-n_2$, one can see that the system is characterized by two independent variables n_1 and n_2. The stoichiometric matrix V is given by

$$V = \begin{bmatrix} -1 & 1 & 1 \\ -1 & 1 & 0 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

and propensity functions $\mathcal{R}_1(N)$, $\mathcal{R}_2(N)$, and $\mathcal{R}_3(N)$ are given by

$$\mathcal{R}_1(N) = k_1 N_1 N_2, \quad \mathcal{R}_2(N) = k_2 N_3 = k_2(A - N_1),$$

$$\mathcal{R}_3(N) = k_3 N_3 = k_3(A - N_1).$$

Thus,

$$\frac{\partial \mathcal{R}_1}{\partial N_1} = k_1 N_2, \quad \frac{\partial \mathcal{R}_1}{\partial N_2} = k_1 N_1,$$

$$\frac{\partial^2 \mathcal{R}_1}{\partial N_1^2} = \frac{\partial^2 \mathcal{R}_1}{\partial N_2^2} = 0, \quad \frac{\partial^2 \mathcal{R}_1}{\partial N_1 N_2} = k_1,$$

$$\frac{\partial \mathcal{R}_2}{\partial N_1} = -k_2, \quad \frac{\partial \mathcal{R}_2}{\partial N_2} = 0, \quad \frac{\partial^2 \mathcal{R}_2}{\partial N_i N_j} = 0 \quad \text{for } i,j = 1,2,$$

$$\frac{\partial \mathcal{R}_3}{\partial N_1} = -k_3, \quad \frac{\partial \mathcal{R}_3}{\partial N_2} = 0, \quad \frac{\partial^2 \mathcal{R}_3}{\partial N_i N_j} = 0 \quad \text{for } i,j = 1,2.$$

By Taylor series expansion of $\mathcal{R}_1$, $\mathcal{R}_2$, and $\mathcal{R}_3$ at $N=\mu$, one can obtain

$$\mathcal{R}_1(N) = k_1\mu_1\mu_2 + k_1\mu_2(N_1-\mu_1) + k_1\mu_1(N_2-\mu_2) + k_1(N_1-\mu_1)(N_2-\mu_2), \tag{37}$$

$$\mathcal{R}_2(N) = k_2(A-\mu_1) - k_2(N_1-\mu_1), \tag{38}$$

$$\mathcal{R}_3(N) = k_3(A-\mu_1) - k_3(N_1-\mu_1), \tag{39}$$

where $\mu_1=E[N_1]$ and $\mu_2=E[N_2]$. For second or higher order central moments, we let $M_{i,j}(t)=E[(N_1-\mu_1)^i(N_2-\mu_2)^j]$ for $i+j\geq 2$. Note that $M_{2,0}(t)=E[(N_1-\mu_1)^2]$ is the variance of N_1 and $M_{0,2}(t)=E[(N_2-\mu_2)^2]$ is the variance of N_2.

From Eqs. (6) and (8), one can find the equations for the means μ_1 and μ_2,

$$\frac{d\mu_1}{dt} = -k_1\mu_1\mu_2 + (k_2+k_3)(A-\mu_1) - k_1 M_{1,1}, \tag{40}$$

$$\frac{d\mu_2}{dt} = -k_1\mu_1\mu_2 + k_2(A-\mu_1) - k_1 M_{1,1}, \tag{41}$$

and the equation for $M_{i,j}$,

$$\begin{aligned}
\frac{dM_{i,j}}{dt} &= \frac{d}{dt}E[(N_1-\mu_1)^i(N_2-\mu_2)^j] \\
&= k_1\mu_1\mu_2\left[\sum_{k=0}^{i-1}\sum_{\ell=0}^{j}\binom{i}{k}\binom{j}{\ell}(-1)^{i-k}(-1)^{j-\ell}M_{k,\ell} + \sum_{\ell=0}^{j-1}\binom{j}{\ell}(-1)^{j-\ell}M_{i,\ell}\right] \\
&\quad + k_1\mu_2\left[\sum_{k=0}^{i-1}\sum_{\ell=0}^{j}\binom{i}{k}\binom{j}{\ell}(-1)^{i-k}(-1)^{j-\ell}M_{k+1,\ell} + \sum_{\ell=0}^{j-1}\binom{j}{\ell}(-1)^{j-\ell}M_{i+1,\ell}\right] \\
&\quad + k_1\mu_1\left[\sum_{k=0}^{i-1}\sum_{\ell=0}^{j}\binom{i}{k}\binom{j}{\ell}(-1)^{i-k}(-1)^{j-\ell}M_{k,\ell+1} + \sum_{\ell=0}^{j-1}\binom{j}{\ell}(-1)^{j-\ell}M_{i,\ell+1}\right] \\
&\quad + k_1\left[\sum_{k=0}^{i-1}\sum_{\ell=0}^{j}\binom{i}{k}\binom{j}{\ell}(-1)^{i-k}(-1)^{j-\ell}M_{k+1,\ell+1} + \sum_{\ell=0}^{j-1}\binom{j}{\ell}(-1)^{j-\ell}M_{i+1,\ell+1}\right] \\
&\quad + k_2(A-\mu_1)\left[\sum_{k=0}^{i-1}\sum_{\ell=0}^{j}\binom{i}{k}\binom{j}{\ell}M_{k,\ell} + \sum_{\ell=0}^{j-1}\binom{j}{\ell}M_{i,\ell}\right] - k_2\left[\sum_{k=0}^{i-1}\sum_{\ell=0}^{j}\binom{i}{k}\binom{j}{\ell}M_{k+1,\ell} + \sum_{\ell=0}^{j-1}\binom{j}{\ell}M_{i+1,\ell}\right] \\
&\quad + k_3(A-\mu_1)\left[\sum_{k=0}^{i-1}\binom{i}{k}M_{k,j}\right] - k_3\left[\sum_{k=0}^{i-1}\binom{i}{k}M_{k+1,j}\right] - i\mu_1' M_{i-1,j} - j\mu_2' M_{i,j-1}.
\end{aligned} \tag{42}$$

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 114.70.7.203 On: Mon, 10 Nov 2014 08:41:41

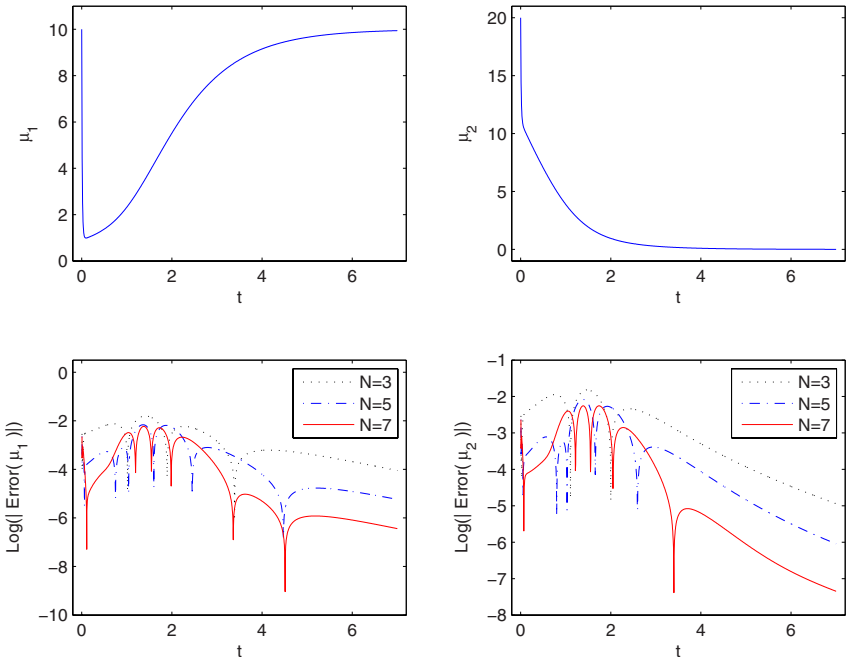

FIG. 3. (Color online) When $n_1=10$, $n_2=20$, $n_3=0$, and $n_4=0$ at time $t=0$ and $k_1=5$, $k_2=5$, $k_3=1$ exact means of molecular numbers of species E and S (upper two figures) and errors in $\log_{10}$ generated by RK4 with $h=0.005$ and each truncation at $N=3,5,7$, that is, $\log_{10}|\text{exact mean}-\text{approximate mean}|$ (lower two graphs).

To approximate the system, we truncate Eqs. (40)–(42) by setting $M_{i,j}(t)=0$ for $i+j=3,5,7$ and then apply RK4 on the truncated system. In Figs. 3 and 4, the numerical results for the means and the variances are compared with the exact solutions, which are obtained by solving the governing equation $dp/dt=Kp$. Truncation at higher order $i+j$ generates smaller errors as expected. Figure 5 shows that the error decreases by reducing the step size h. Although the error at the steady state decreases for smaller h, the convergence rate gradually slows down as $h\to 0$. However, one can obtain a better result for the steady state by using the method in Sec. III C, which happens to yield the exact equilibrium in this case as $\mu_1=10$, $\mu_2=0$.

C. A decaying-dimerizing reaction network

In this section, we consider a decaying-dimerizing reaction network and present the computational efficiency of our method on it in comparison to SSAs. The decaying-dimerizing reaction network is known as a stiff system and studied in Refs. 3, 10, and 16. This network contains three distinct species interacting through four reaction channels as follows:

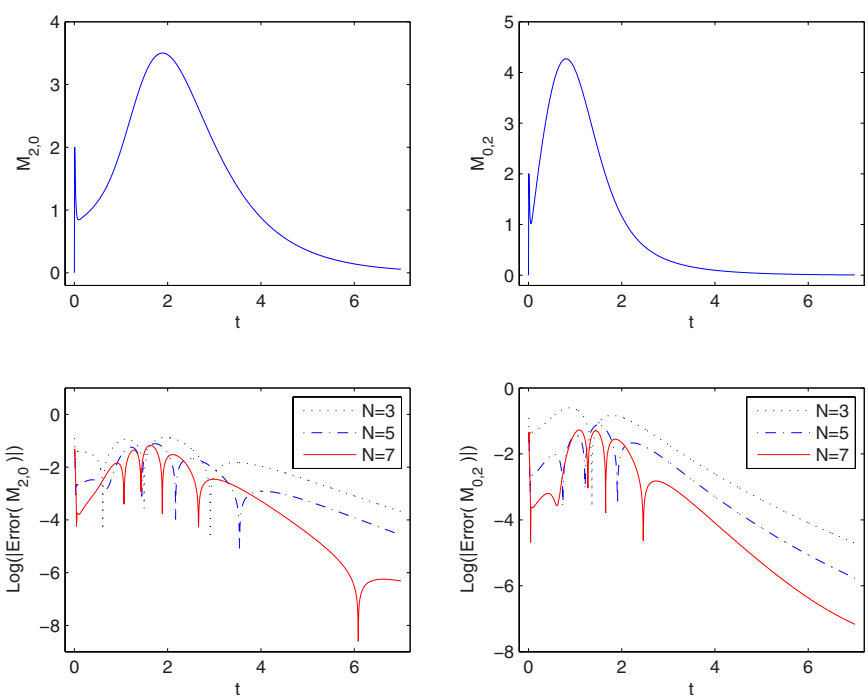

FIG. 4. (Color online) When $n_1=10$, $n_2=20$, $n_3=0$, and $n_4=0$ at time $t=0$ and $k_1=5$, $k_2=5$, $k_3=1$ variances of molecular numbers of species E and S (upper two figures) and the errors $\log_{10}|\text{exact mean}-\text{approximate mean}|$ generated by RK4 with $h=0.005$ and each truncation at $N=3,5,7$ (lower two figures).

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 114.70.7.203 On: Mon, 10 Nov 2014 08:41:41

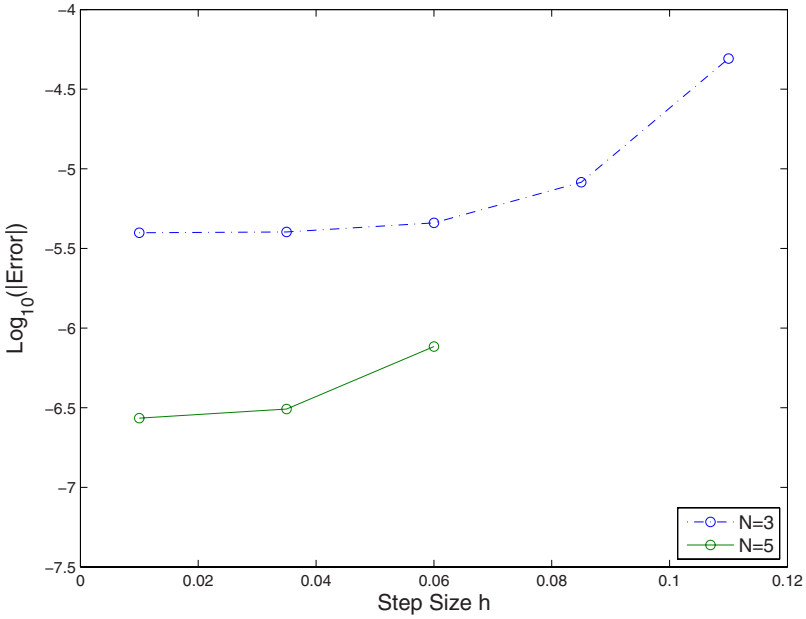

FIG. 5. (Color online) $\log_{10}|\text{error}|$ for μ_2 vs a step size h where k_1, $k_2=1$, and $k_3=0.1$ and the initial condition $n_1=10$, $n_2=20$, and $n_4=0$. The errors are measured at $t=7$. The numerical solution diverges for $h>0.06$ when $N=5$.

$$2A_1 \underset{k_2}{\overset{k_1}{\rightleftarrows}} A_2 \overset{k_3}{\longrightarrow} A_3, \quad A_1 \overset{k_4}{\longrightarrow} \phi.$$

We let N_1, N_2, and N_3 as random variables that denote the number of molecules of A_1, A_2, and A_3, respectively. As in Refs. 16 and 10, we assume reaction rate constants $k_1=10$, $k_2=1000$, $k_3=0.1$, and $k_4=1$ and an initial condition $(N_1,N_2,N_3)=(400,798,0)$, so that the system is stiff. The stoichiometric matrix V and propensity functions $\mathcal{R}_1(N)$, $\mathcal{R}_2(N)$, $\mathcal{R}_3(N)$, and $\mathcal{R}_4(N)$ are given by

$$V=\begin{bmatrix} -2 & 2 & 0 & -1 \\ 1 & -1 & -1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \mathcal{R}(N)=\begin{bmatrix} \mathcal{R}_1(N) \\ \mathcal{R}_2(N) \\ \mathcal{R}_3(N) \\ \mathcal{R}_4(N) \end{bmatrix}=\begin{bmatrix} \frac{k_1}{2}N_1(N_1-1) \\ k_2N_2 \\ k_3N_2 \\ k_4N_1 \end{bmatrix}.$$

We let $\mu_i(t)$ as the mean of N_i at time t and let $M_{i_1,i_2,i_3}(t)$ as the central moment $E[(N_1-\mu_1)^{i_1}(N_2-\mu_2)^{i_2}(N_3-\mu_3)^{i_3}]$ at time t. Note that $M_{0,0,0}(t)=1$, and when $i_1+i_2+i_3=1$, $M_{i_1,i_2,i_3}(t)=0$ for any time t.

Using Eqs. (6) and (8), one can find the equations for means

$$\frac{d\mu_1}{dt}=-k_1\mu_1(\mu_1-1)+2k_2\mu_2-k_4\mu_1-2k_1M_{2,0,0},$$

$$\frac{d\mu_2}{dt}=\frac{k_1}{2}\mu_1(\mu_1-1)-k_2\mu_2-k_3\mu_2+k_1M_{2,0,0},$$

$$\frac{d\mu_3}{dt}=k_3\mu_2,$$

and the equation of the mth moment for $m=i_1+i_2+i_3\geq 2$,

$$\begin{aligned}
\frac{dM_{i_1,i_2,i_3}}{dt} &= \frac{k_1}{2}\mu_1(\mu_1-1)\sum_{\ell_1+\ell_2\neq i_1+i_2}\binom{i_1}{\ell_1}\binom{i_2}{\ell_2}(-2)^{i_1-\ell_1}M_{\ell_1,\ell_2,i_3}+k_2\mu_2\sum_{\ell_1+\ell_2\neq i_1+i_2}\binom{i_1}{\ell_1}\binom{i_2}{\ell_2}2^{i_1-\ell_1}(-1)^{i_2-\ell_2}M_{\ell_1,\ell_2,i_3} \\
&+k_3\mu_2\sum_{\ell_2+\ell_3\neq i_2+i_3}\binom{i_2}{\ell_2}\binom{i_3}{\ell_3}(-1)^{i_2-\ell_2}M_{i_1,\ell_2,\ell_3}+k_4\mu_1\sum_{\ell_1=0}^{i_1-1}\binom{i_1}{\ell_1}(-1)^{i_1-\ell_1}M_{\ell_1,i_2,i_3}+\frac{k_1}{2}(2\mu_1-1)\sum_{\ell_1+\ell_2\neq i_1+i_2}\binom{i_1}{\ell_1} \\
&\times\binom{i_2}{\ell_2}(-2)^{i_1-\ell_1}M_{\ell_1+1,\ell_2,i_3}+k_2\sum_{\ell_1+\ell_2\neq i_1+i_2}\binom{i_1}{\ell_1}\binom{i_2}{\ell_2}2^{i_1-\ell_1}(-1)^{i_2-\ell_2}M_{\ell_1,\ell_2+1,i_3}+k_3\sum_{\ell_2+\ell_3\neq i_2+i_3}\binom{i_2}{\ell_2}\binom{i_3}{\ell_3} \\
&\times(-1)^{i_2-\ell_2}M_{i_1,\ell_2+1,\ell_3}+k_4\sum_{\ell_1=0}^{i_1-1}\binom{i_1}{\ell_1}(-1)^{i_1-\ell_1}M_{\ell_1+1,i_2,i_3}+k_1\sum_{\ell_1+\ell_2\neq i_1+i_2}\binom{i_1}{\ell_1}\binom{i_2}{\ell_2}(-2)^{i_1-\ell_1}M_{\ell_1+2,\ell_2,i_3} \\
&-i_1\mu_1'M_{i_1-1,i_2,i_3}-i_2\mu_2'M_{i_1,i_2-1,i_3}-i_3\mu_3'M_{i_1,i_2,i_3-1}.
\end{aligned}$$

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 114.70.7.203 On: Mon, 10 Nov 2014 08:41:41

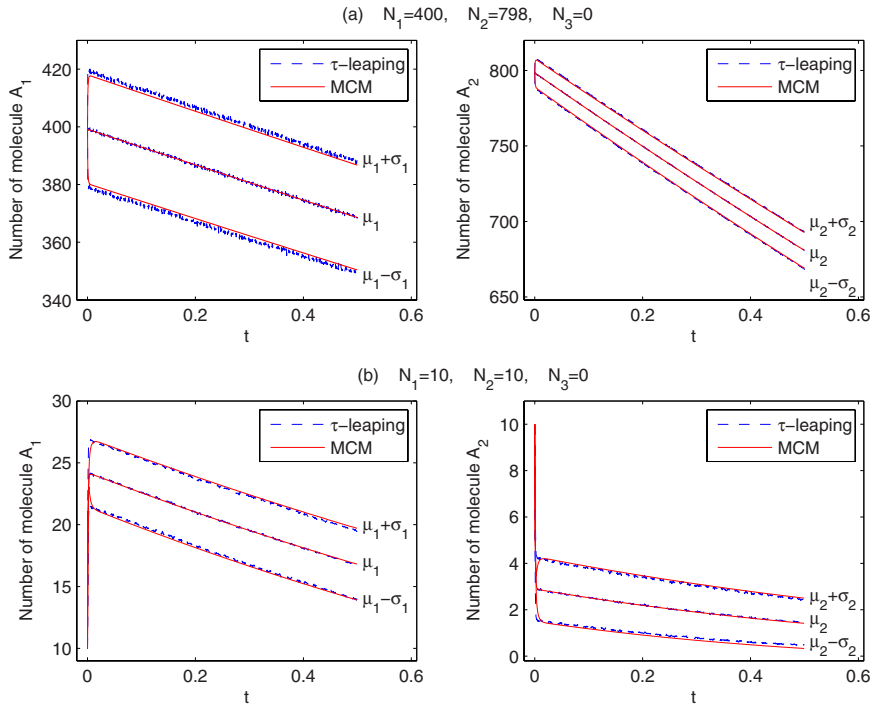

FIG. 6. (Color online) Comparison of MCM and the τ-leaping method in two cases with different initial conditions. For the τ-leaping method, the illustrated results were obtained by averaging 3000 realizations. The parameter ε is set at 0.1 in both (a) and (b). The step sizes for MCM are $h=0.000\,15$ and $h=0.0009$ in (a) and (b), respectively, and truncation order is $N=3$ for both (a) and (b).

The numerical results are illustrated in Fig. 6. We compare the results to those of the SSA. Since simulation of the stiff system by the exact SSA requires a large amount of time, we utilize a more efficient SSA, the so-called τ-leaping method.[3] The tests were carried out with two different initial conditions to study their ability to cope with stiffness of the system. In Fig. 6(a), we set the initial condition as $N_1=400$, $N_2=780$, and $N_3=0$, which renders more stiffness and long time for the system to converge to equilibrium.

One can see that the results of two methods agree well in both Figs. 6(a) and 6(b). However, the comparison in CPU time shows that the MCM is computationally more efficient than the τ-leaping method in Table II.

In Fig. 6(a), the MCM is more than 15 times faster than the τ-leaping method of 1000 realizations and 50 times faster than that of 3000 realizations. However, as the results from 1000 realizations suffer from severe fluctuation, we adopt only the results from 3000 realizations to compare with MCM in Fig. 6. The ratio between two methods decreases from 50:1 to 35:1 in Fig. 6(b) as the stiffness of the system becomes relatively low. This suggests not only that MCM is computationally more efficient than the τ-leaping method in computation of mean and standard deviation, but also that the former is less sensitive against stiffness.

TABLE II. Comparison of moment closure and τ-leaping methods in CPU time.

	MCM	τ-leaping (1000)	τ-leaping (3000)
$(N_1,N_2,N_3)=(400,798,0)$	3.1420×10^2	5.0892×10^3	1.5975×10^4
$(N_1,N_2,N_3)=(10,10,0)$	5.1921×10^1	5.5916×10^2	1.8250×10^3

V. DISCUSSION

In this paper, we presented a MCM for stochastic reaction networks. We first derived an infinite dimensional system of ODEs for mean and all central moments from the chemical master equation. This enabled us to construct the mean and variance up to the order that we want by a Taylor series expansion. Each derivative term of the Taylor series was obtained recursively from the moment equations derived. Furthermore, we extracted a finite ODE system of the moments by truncations at second or higher order central moments. Applying conventional numerical schemes to the truncated system, we obtained numerical approximations for the mean and second or higher central moments. The errors in this approximation were estimated in a rigorous way and shown to be controllable by adjusting the truncation order. Moreover, by simple algebraic manipulation on the truncated system, we found the means and moments at equilibriums even without numerical integration.

We have shown in examples that the approximation based on the MCM provides an efficient and accurate alternative way to simulate stochastic reaction networks, especially stiff systems for which simulation by SSAs requires intensive and expensive computations.

ACKNOWLEDGMENTS

C.H.L. would like to thank Professor Hans G. Othmer and Professor Roger Lui for early discussions on this subject. K.-H.K. would like to acknowledge that this work was supported by the Korea Science and Engineering Foundation (KOSEF) grant funded by the Korea government (MEST) (Grant No. R01-2008-000-20010-0).

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 114.70.7.203 On: Mon, 10 Nov 2014 08:41:41

APPENDIX: DERIVATION OF EQUATIONS (6)–(8) AND PROOF OF EQUATION (27)

1. Derivation of Equations (6) and (7)

By applying a transformation $n-V_k \to n$ in the first term of the right side of Eq. (5),

$$\sum_n n_i \frac{dp(n,t)}{dt} = \sum_k \sum_n n_i \mathcal{R}_k(n-V_k)p(n-V_k,t) - n_i \mathcal{R}_k(n)p(n,t),$$

we obtain

$$\frac{dE[N_i]}{dt} = \sum_k \sum_n ((n_i + V_{k,i})\mathcal{R}_k(n)p(n,t) - n_i \mathcal{R}_k(n)p(n,t)) = \sum_k \sum_n V_{k,i}\mathcal{R}_k(n)p(n,t) = \sum_k V_{k,i}E[\mathcal{R}_k(N)],$$

where we utilize the fact that $E[\mathcal{R}_k(N)]=\sum_n \mathcal{R}_k(n)p(n,t)$ and $V_{k,i}$ denotes ith component of the vector V_k. Let $\mu_i(t)$ denote the mean $E[N_i(t)]$ of the ith species at time t and let $\mu(t) =(\mu_1(t),\ldots,\mu_s(t))$.

To obtain second central moments, we multiply $(n_i-\mu_i)(n_j-\mu_j)$ to Eq. (3) and sum over all accessible states n and then obtain

$$\sum_n (n_i-\mu_i)(n_j-\mu_j)\frac{dp(n,t)}{dt} = \sum_k \sum_n ((n_i-\mu_i)(n_j-\mu_j)\mathcal{R}_k(n-V_k)p(n-V_k,t) - (n_i-\mu_i)(n_j-\mu_j)\mathcal{R}_k(n)p(n,t)).$$

Applying the transformation $n-V_k \to n$ again in the first term of the right-hand side gives

$$\begin{aligned}\frac{dE[(N_i-\mu_i)(N_j-\mu_j)]}{dt} &= \sum_k \sum_n ((n_i+V_{k,i}-\mu_i)(n_j+V_{k,j}-\mu_j)\mathcal{R}_k(n)p(n,t) - (n_i-\mu_i)(n_j-\mu_j)\mathcal{R}_k(n)p(n,t)) \\ &= \sum_k \sum_n ((V_{k,i}(n_j-\mu_j) + V_{k,j}(n_i-\mu_i) + V_{k,i}V_{k,j})\mathcal{R}_k(n)p(n,t)) \\ &= \sum_k (V_{k,i}E[(N_j-\mu_j)\mathcal{R}_k(N)] + V_{k,j}E[(N_i-\mu_i)\mathcal{R}_k(N)] + V_{k,i}V_{k,j}E[\mathcal{R}_k(N)]).\end{aligned}$$

Up to now, we can obtain equations for mean and covariance as

$$\frac{dE[N_i]}{dt} = \sum_k V_{k,i}E[\mathcal{R}_k(N)], \tag{A1}$$

$$\frac{dE[(N_i-\mu_i)(N_j-\mu_j)]}{dt} = \sum_k (V_{k,i}E[(N_j-\mu_j)\mathcal{R}_k(N)] + V_{k,j}E[(N_i-\mu_i)\mathcal{R}_k(N)] + V_{k,i}V_{k,j}E[\mathcal{R}_k(N)]), \tag{A2}$$

for each $i, j=1,\ldots,s$.

Since each $\mathcal{R}_k$ is at most quadratic, we derive by writing $\mathcal{R}_k$ in terms of $(N-\mu)$ or using the Taylor formula at $N=\mu$,

$$\mathcal{R}_k(N) = \mathcal{R}_k(\mu) + \sum_\ell \frac{\partial \mathcal{R}_k(\mu)}{\partial N_\ell}(N_\ell-\mu_\ell) + \frac{1}{2}\sum_{\ell,m}\frac{\partial^2 \mathcal{R}_k(\mu)}{\partial N_\ell\,\partial N_m}(N_\ell-\mu_\ell)(N_m-\mu_m). \tag{A3}$$

Since $E[N_\ell-\mu_\ell]=0$ and μ is nonrandom, from Eq. (A3) we obtain

$$E[\mathcal{R}_k(N)] = \mathcal{R}_k(\mu) + \frac{1}{2}\sum_{\ell,m}\frac{\partial^2 \mathcal{R}_k(\mu)}{\partial N_\ell\,\partial N_m}E[(N_\ell-\mu_\ell)(N_m-\mu_m)], \tag{A4}$$

and for each i,

$$E[\mathcal{R}_k(N)(N_i-\mu_i)] = \sum_\ell \frac{\partial \mathcal{R}_k(\mu)}{\partial N_\ell}E[(N_i-\mu_i)(N_\ell-\mu_\ell)] + \frac{1}{2}\sum_{\ell,m}\frac{\partial^2 \mathcal{R}_k(\mu)}{\partial N_\ell\,\partial N_m}E[(N_\ell-\mu_\ell)(N_m-\mu_m)(N_i-\mu_i)]. \tag{A5}$$

Substitution of Eqs. (A4) and (A5) into Eqs. (A1) and (A2) gives

$$\frac{dE[N_i]}{dt} = \sum_k V_{k,i}\left(\mathcal{R}_k(\mu) + \frac{1}{2}\sum_{\ell,m}\frac{\partial^2 \mathcal{R}_k(\mu)}{\partial N_\ell\,\partial N_m}E[(N_\ell-\mu_\ell)\times(N_m-\mu_m)]\right) \tag{A6}$$

and

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 114.70.7.203 On: Mon, 10 Nov 2014 08:41:41

$$\frac{dE[(N_i-\mu_i)(N_j-\mu_j)]}{dt}=\sum_k\left(V_{k,i}\sum_\ell\frac{\partial\mathcal{R}_k(\mu)}{\partial N_\ell}E[(N_j-\mu_j)(N_\ell-\mu_\ell)]+V_{k,j}\sum_\ell\frac{\partial\mathcal{R}_k(\mu)}{\partial N_\ell}E[(N_i-\mu_i)(N_\ell-\mu_\ell)]+V_{k,i}V_{k,j}\left(\mathcal{R}_k(\mu)\right.\right.$$
$$\left.+\frac{1}{2}\sum_{\ell,m}\frac{\partial^2\mathcal{R}_k(\mu)}{\partial N_\ell\,\partial N_m}E[(N_\ell-\mu_\ell)(N_m-\mu_m)]\right)+V_{k,i}\frac{1}{2}\sum_{\ell,m}\frac{\partial^2\mathcal{R}_k(\mu)}{\partial N_\ell\,\partial N_m}E[(N_j-\mu_j)(N_\ell-\mu_\ell)(N_m-\mu_m)]$$
$$\left.+V_{k,j}\frac{1}{2}\sum_{\ell,m}\frac{\partial^2\mathcal{R}_k(\mu)}{\partial N_\ell\,\partial N_m}E[(N_i-\mu_i)(N_\ell-\mu_\ell)(N_m-\mu_m)]\right).$$

If we denote $\sigma_{ij}=E[(N_i-\mu_i)(N_j-\mu_j)]$ and $\sigma_{ijk}=E[(N_i-\mu_i)(N_j-\mu_j)(N_k-\mu_k)]$, the equations for the first moments and second central moments can be found as

$$\frac{d\mu_i}{dt}=\sum_k V_{k,i}\left(\mathcal{R}_k(\mu)+\frac{1}{2}\sum_{\ell,m}\frac{\partial^2\mathcal{R}_k(\mu)}{\partial N_\ell\,\partial N_m}\sigma_{\ell m}\right),$$

$$\frac{d\sigma_{ij}}{dt}=\sum_k\left[V_{k,i}\sum_\ell\frac{\partial\mathcal{R}_k(\mu)}{\partial N_\ell}\sigma_{j\ell}+V_{k,j}\sum_\ell\frac{\partial\mathcal{R}_k(\mu)}{\partial N_\ell}\sigma_{i\ell}+V_{k,i}V_{k,j}\left(\mathcal{R}_k(\mu)+\frac{1}{2}\sum_{\ell,m}\frac{\partial^2\mathcal{R}_k(\mu)}{\partial N_\ell\,\partial N_m}\sigma_{\ell m}\right)\right.$$
$$\left.+V_{k,i}\frac{1}{2}\sum_{\ell,m}\frac{\partial^2\mathcal{R}_k(\mu)}{\partial N_\ell\,\partial N_m}\sigma_{j\ell m}+V_{k,j}\frac{1}{2}\sum_{\ell,m}\frac{\partial^2\mathcal{R}_k(\mu)}{\partial N_\ell\,\partial N_m}\sigma_{i\ell m}\right].$$

2. Derivation of the moment Equation (8)

To obtain Eq. (8), we first denote the central moments by

$$M_{i_1,\ldots,i_s}=E[(N_1-\mu_1)^{i_1}\cdots(N_s-\mu_s)^{i_s}]$$

for $i_1+\cdots+i_s\geq 2$. Then one can compute

$$\frac{dM_{i_1,i_2,\ldots,i_s}}{dt}=\frac{d}{dt}E[(N_1-\mu_1)^{i_1}\cdots(N_s-\mu_s)^{i_s}]$$
$$=\sum_n(n_1-\mu_1)^{i_1}\cdots(n_s-\mu_s)^{i_s}\frac{dp(n,t)}{dt}-\sum_n\sum_{j=1}^s i_j\mu_j'(n_1-\mu_1)^{i_1}\cdots(n_j-\mu_j)^{i_j-1}\cdots(n_s-\mu_s)^{i_s}p(n,t)$$
$$=\sum_k\sum_n[(n_1-\mu_1+V_{k,1})^{i_1}\cdots(n_s-\mu_s+V_{k,s})^{i_s}-(n_1-\mu_1)^{i_1}\cdots(n_s-\mu_s)^{i_s}]\mathcal{R}_k(\mu)p(n,t)-\sum_{j=1}^s i_j\mu_j'M_{i_1,\ldots,i_j-1,\ldots,i_s}$$
$$=\sum_k\sum_n\left[\sum_{\ell_1=0}^{i_1}\binom{i_1}{\ell_1}(n_1-\mu_1)^{\ell_1}(V_{k,1})^{i_1-\ell_1}\cdots\sum_{\ell_s=0}^{i_s}\binom{i_s}{\ell_s}(n_s-\mu_s)^{\ell_s}(V_{k,s})^{i_s-\ell_s}-(n_1-\mu_1)^{i_1}\cdots(n_s-\mu_s)^{i_s}\right]$$
$$\times\left[\mathcal{R}_k(\mu)+\sum_q\frac{\partial\mathcal{R}_k(\mu)}{\partial n_q}(n_q-\mu_q)+\frac{1}{2}\sum_{q,r}\frac{\partial^2\mathcal{R}_k(\mu)}{\partial n_q\,\partial n_r}(n_q-\mu_q)(n_r-\mu_r)\right]p(n,t)-\sum_{j=1}^s i_j\mu_j'M_{i_1,\ldots,i_j-1,\ldots,i_s}$$
$$=\sum_k\sum_n\sum_{\substack{\ell_1,\ell_2,\ldots,\ell_s\\ \ell_1+\cdots+\ell_s\neq m}}\binom{i_1}{\ell_1}\cdots\binom{i_s}{\ell_s}(V_{k,1})^{i_1-\ell_1}\cdots(V_{k,s})^{i_s-\ell_s}(n_1-\mu_1)^{\ell_1}\cdots(n_s-\mu_s)^{\ell_s}$$
$$\times\left[\mathcal{R}_k(\mu)+\sum_q\frac{\partial\mathcal{R}_k(\mu)}{\partial n_q}(n_q-\mu_q)+\frac{1}{2}\sum_{q,r}\frac{\partial^2\mathcal{R}_k(\mu)}{\partial n_q\,\partial n_r}(n_q-\mu_q)(n_r-\mu_r)\right]p(n,t)-\sum_{j=1}^s i_j\mu_j'M_{i_1,\ldots,i_j-1,\ldots,i_s}$$

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 114.70.7.203 On: Mon, 10 Nov 2014 08:41:41

$$= \sum_k \mathcal{R}_k(\mu) \sum_{\substack{\ell_1,\ell_2,\ldots,\ell_s \\ \ell_1+\cdots+\ell_s\neq m}} \binom{i_1}{\ell_1}\cdots\binom{i_s}{\ell_s}(V_{k,1})^{i_1-\ell_1}\cdots(V_{k,s})^{i_s-\ell_s}M_{\ell_1,\ldots,\ell_s} + \sum_k\sum_q \frac{\partial\mathcal{R}_k(\mu)}{\partial n_q}\sum_{\substack{\ell_1,\ell_2,\ldots,\ell_s \\ \ell_1+\cdots+\ell_s\neq m}}\binom{i_1}{\ell_1}\cdots\binom{i_s}{\ell_s}$$

$$\times(\times V_{k,1})^{i_1-\ell_1}\cdots(V_{k,s})^{i_s-\ell_s}M_{\ell_1,\ldots,\ell_q+1,\ldots,\ell_s} + \frac{1}{2}\sum_k\sum_{q,r}\frac{\partial^2\mathcal{R}_k(\mu)}{\partial n_q\,\partial n_r}\sum_{\substack{\ell_1,\ell_2,\ldots,\ell_s \\ \ell_1+\cdots+\ell_s\neq m}}\binom{i_1}{\ell_1}\cdots\binom{i_s}{\ell_s}$$

$$\times(\times V_{k,1})^{i_1-\ell_1}\cdots(V_{k,s})^{i_s-\ell_s}M_{\ell_1,\ldots,\ell_s+e_q+e_r} - \sum_{j=1}^{s} i_j\mu_j' M_{i_1,\ldots,i_j-1,\ldots,i_s},$$

where the subscript $\ell_1,\ldots,\ell_s+e_q+e_r$ denotes adding 1 to qth and rth entries of $\ell_1,\ldots,\ell_s$, respectively.

3. Proof of Equation (27) for a system with a single variable

Suppose that there is only one single independent variable n in the system such as the reversible bimolecular reaction system in Sec. IV A. Further suppose that the largest possible value of n is A for any time $t\in[0,h]$. Let $a=A-\mu$ and $b=[\mu]-\mu$, where $[\mu]$ denotes the largest integer less than or equal to μ. Then clearly, $-1<b\leq 0$ and $0<b+1\leq 1$. If we let $x=n-\mu$, x ranges from $-\mu$ to a, that is, $x=-\mu,\ldots,b,\ b+1,\ldots,a$. We can write $M_{k-1}(t)$ as the sum of positive and negative terms, $M_{k-1}(t)=\Sigma_{x=-\mu}^{b}x^{k-1}p(x,t)+\Sigma_{x=b+1}^{a}x^{k-1}p(x,t)$, where the first summation part is negative and the second one is positive. For the case $M_{k-1}\geq 0$,

$$|M_{k-1}(t)|-|M_k(t)| = \sum_{x=-\mu}^{a} x^{k-1}p(x,t) - \sum_{x=-\mu}^{a} x^k p(x,t), \quad \text{(A7)}$$

$$= \sum_{x=-\mu}^{b} x^{k-1}(1-x)p(x,t) + \sum_{x=b+1}^{a} x^{k-1}(1-x)p(x,t). \quad \text{(A8)}$$

In Eq. (A8), the only positive term is $(b+1)^{k-1}(-b)p(b+1,t)$. Therefore we have

$$|M_{k-1}(t)|-|M_k(t)|\leq 0,$$

unless

$$(b+1)^{k-1}(-b)p(b+1,t) > \left|\sum_{x=b+2}^{a} x^{k-1}(1-x)p(x,t) + \sum_{x=-\mu}^{b} x^{k-1}(1-x)p(x,t)\right|. \quad \text{(A9)}$$

Note that inequality (A9) would be valid only in an exceptional case that the probability of the state $n=[\mu]+1$ is much greater than that of all other states. Thus we can conclude that Eq. (27) holds in most cases. Likewise, the other case $M_{k-1}<0$ leads to the same conclusion.

[1] D. T. Gillespie, Physica A **188**, 404 (1992).
[2] D. T. Gillespie, J. Phys. Chem. **81**, 2340 (1977).
[3] D. T. Gillespie, J. Chem. Phys. **115**, 1716 (2001).
[4] M. A. Gibson and J. Bruck, J. Chem. Phys. **105**, 1876 (2000).
[5] C. Gadgil, C. H. Lee, and H. G. Othmer, Bull. Math. Biol. **67**, 901 (2005).
[6] P. Whittle, J. R. Stat. Soc. Ser. B (Methodol.) **19**, 268 (1957).
[7] I. H. Matis and T. R. Kiffe, Biometrics **52**, 980 (1996).
[8] I. Nasell, Theor Popul. Biol. **63**, 159 (2003).
[9] M. J. Keeling, J. Theor. Biol. **205**, 269 (2000).
[10] J. P. Hespanha and A. Singh, Int. J. Robust Nonlinear Control **15**, 669 (2005).
[11] S. Engblom, Appl. Math. Comput. **180**, 498 (2006).
[12] R. Durrett, *Probability: Theory and Examples* (Duxbury, Belmont, 1996).
[13] Kurtz proved that if a system is sufficiently large, the distribution of the system is approximately multivariate normal (Ref. 14).
[14] T. G. Kurtz, Stochastic Proc. Appl. **6**, 223 (1978).
[15] T. Kato, *Perturbation Theory for Linear Operators* (Springer-Verlag, Berlin, 1976).
[16] M. Rathinam, L. R. Petzold, Y. Cao, and D. T. Gillespie, J. Chem. Phys. **119**, 12784 (2003).

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 114.70.7.203 On: Mon, 10 Nov 2014 08:41:41