J Syst Sci Complex (2010) 23: 896–905

NONLINEAR LANGEVIN MODEL WITH PRODUCT STOCHASTICITY FOR BIOLOGICAL NETWORKS: THE CASE OF THE SCHNAKENBERG MODEL*

Youfang CAO · Jie LIANG

DOI: 10.1007/s11424-010-0213-0
Received: 16 November 2009 / Revised: 13 March 2010

Abstract Langevin equation is widely used to study the stochastic effects in molecular networks, as it often approximates well the underlying chemical master equation. However, frequently it is not clear when such an approximation is applicable and when it breaks down. This paper studies the simple Schnakenberg model consisting of three reversible reactions and two molecular species whose concentrations vary. To reduce the residual errors from the conventional formulation of the Langevin equation, the authors propose to explicitly model the effective coupling between macroscopic concentrations of different molecular species. The results show that this formulation is effective in correcting residual errors from the original uncoupled Langevin equation and can approximate the underlying chemical master equation very accurately.

Key words Langevin equation, master equation, noise, Schnakenberg model.

1 Introduction

Chemical master equation (CME) provides a fundamental theoretical framework for modeling biochemical reaction systems[1]. However, master equation is difficult to solve analytically, except in a few simple special cases. Currently, there are two approaches for obtaining solutions to the master equation. The first is the widely used stochastic simulation algorithm[2−4], which is a Monte Carlo sampling based method with serious limitations: It is difficult to know a priori whether the amount of sampling is sufficient, especially when facing the difficult task of estimating small probabilities of biological events that occur rarely but are critically important. Stochastic simulations can be very inefficient in capturing these rare events of interests.

The second is to solve the chemical master equation computationally. Methods in reducing the state space with controlled error bounds have been developed for achieving this goal[5−7]. Recently, progress has also been made in the optimal enumeration of the state space of the underlying Markov process of the CME, and in obtaining the exact stationary probability

Youfang CAO · Jie LIANG
Ministry of Education Key Laboratory of Systems Biomedicine, Shanghai Center for Systems Biomedicine, Shanghai Jiao Tong University, Shanghai 200240*, China; Department of Bioengineering, University of Illinois at Chicago, Chicago, USA*. Email: youfang@uic.edu; jliang@uic.edu.
*The research is supported by phase II 985 Project under Grant No. T226208001, NIH under Grant No. GM079804-01A1 and GM081682, NSF under Grant No. DBI-0646035 and DMS-0800257, and ONR under Grant No. N00014-09-1-0028.

distribution over these states[8]. It was shown that these techniques can be applied to realistic biological problems, including the well-known system of the epigenetic switch of phage lambda[9].

However, to go beyond biochemical networks of small or moderate size, whose chemical master equations cannot be solved directly, one needs to approximate the stochastic biochemical system. One widely used approach is that of the Langevin equation[10−12]. van Kampen summarized the "Langevin approach" as one consisting of two terms: The drift term represents the macroscopic deterministic part of the system, and the diffusion term represents the intrinsic stochasticity or noise in the system. The basic form of Langevin equation is

$$\frac{d\boldsymbol{X}}{dt} = \mu(\boldsymbol{X}) + \sigma(\boldsymbol{X})\mathcal{N}\Big(0, \frac{1}{dt}\Big), \tag{1}$$

where $\boldsymbol{X}$ is the vector of concentrations of molecular species in the reaction system, $\mu(\boldsymbol{X})$ the deterministic component of the equation, or the drift term. The second term is the diffusion term. Here $\mathcal{N}(0, 1/dt)$ is a vector of 1D Gaussian, with zero mean and $1/dt$ variance. The coefficient $\sigma(\boldsymbol{X})$ controls the amplitude of the Gaussian noise. It can be either a function of $\boldsymbol{X}$ or a constant. The key issue in developing Langevin models for biochemical networks is to determine $\mu(\boldsymbol{X})$ and $\sigma(\boldsymbol{X})$. When $\sigma(\boldsymbol{X})$ is a vector of constants, one adjusts its values so the variance of the Gaussian noise produce the correct fluctuations in the system[13].

In this study, we explore a novel approach for developing stochastic models of biochemical networks, extending the Langevin formulation. Using the well-studied Schnakenberg's model as example, we show that by explicitly incorporating the coupling of macroscopic concentrations of different molecular species, a more accurate model whose solution approaches the original chemical master equation can be obtained. This paper is organized as follows. Section 2 describes the development of the coupling Langevin equation. We then present results of numerical solutions and analysis in Section 3. Section 4 provides discussion and conclusion.

2 Models

2.1 Schnakenberg Model

The Schnakenberg model was originally developed for studying the limit cycle behavior in a simple chemical reaction system[14]. This system consists of only two reacting components. However, such a simple system can already produce complex behavior such as oscillation. There have been many theoretical studies on the Schnakenberg model (see [15–16] for recent examples).

The Schnakenberg model used in this study consists of three reversible reactions:

$$X \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} A, \quad B \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} Y, \quad 2X + Y \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} 3X, \tag{2}$$

where X and Y are reacting species of the system, and A and B are external reactants whose copy numbers or concentrations are fixed constants. Each reaction has a corresponding microscopic reaction rate. The fixed copy numbers or concentrations of A and B can be adjusted, which lead to different behavior of the system.

2.2 Chemical Master Equation (CME)

The chemical master equation for the Schnakenberg model can be written as:

$$\frac{dp(X,Y,t)}{dt}$$

$$= k_{-1}Ap(X-1,Y,t) + k_2Bp(X,Y-1,t) + \frac{k_3}{V^2}(X-1)(X-2)(Y+1)p(X-1,Y+1,t)$$

$$+k_1(X+1)p(X+1,Y,t) + k_{-2}(Y+1)p(X,Y+1,t) + \frac{k_{-3}}{V^2}(X+1)X(X-1)p(X+1,Y-1)$$

$$-(k_{-1}a + k_2b + \frac{k_3}{V^2}X(X-1)Y + k_1X + k_{-2}Y + \frac{k_{-3}}{V^2}X(X-1)(X-2))p(X,Y,t), \qquad (3)$$

where X, Y, A and B are copy numbers of molecular species X, Y, A and B, respectively. $p(X,Y,t)$ is the probability that the system has the state of X copies of X and Y copies of Y at time t. V represents the volume of the system.

2.3 Macroscopic Deterministic Model (ODE)

The macroscopic deterministic model or the ODE model provides the first degree approximation to the master equation. It does not model the stochasticity in the system as it contains no noise-related diffusion term. The deterministic equations of the concentration of X and Y can be written as:

$$\frac{dx}{dt} = k_{-1}a - k_1x + k_3x^2y - k_{-3}x^3 \quad \text{and} \quad \frac{dy}{dt} = k_2b - k_{-2}y - k_3x^2y + k_{-3}x^3, \qquad (4)$$

in which x, y, a and b are concentrations of molecular species X, Y, A and B, respectively. Concentrations can be calculated from copy numbers of corresponding molecular species after division by the system volume V.

2.4 Coupling Langevin Equation

2.4.1 The Drift Term

The noise-related diffusion term is added to the deterministic ODE model to account for the stochasticity in the system. We first derive the differential form of the mean of X and Y from the chemical master equation. After multiplying both sides of Equation (3) with X and sum over X and Y from 0 to ∞, we combine similar terms, and re-index the summations. We obtain

$$\frac{d\sum_{X=0}^{\infty}\sum_{Y=0}^{\infty}Xp(X,Y,t)}{dt} = k_{-1}A\sum_{X=0}^{\infty}\sum_{Y=0}^{\infty}p(X,Y,t) - k_1\sum_{X=0}^{\infty}\sum_{Y=0}^{\infty}Xp(X,Y,t)$$

$$+\frac{k_3}{V^2}\sum_{X=0}^{\infty}\sum_{Y=0}^{\infty}X(X-1)Yp(X,Y,t)$$

$$-\frac{k_{-3}}{V^2}\sum_{X=0}^{\infty}\sum_{Y=0}^{\infty}X(X-1)(X-2)p(X,Y,t), \qquad (5)$$

that is,

$$\frac{d\langle X\rangle}{dt} = k_{-1}A - k_1\langle X\rangle + \frac{k_3}{V^2}\langle X(X-1)Y\rangle - \frac{k_{-3}}{V^2}\langle X(X-1)(X-2)\rangle, \qquad (6)$$

where $\langle \cdots \rangle$ stands for the mean of stochastic variable "$\cdots$". Similarly, we have for $\langle Y \rangle$:

$$\frac{d\langle Y \rangle}{dt} = k_2 B - k_{-2}\langle Y \rangle - \frac{k_3}{V^2}\langle X(X-1)Y \rangle + \frac{k_{-3}}{V^2}\langle X(X-1)(X-2) \rangle. \tag{7}$$

It is easy to see that Equation (6) and Equation (7) are equivalent to the macroscopic deterministic equations Equation (4), after both sides are divided by the system volume V and $\langle X \rangle$ and $\langle Y \rangle$ are transformed into concentration (Equation (8)). We will use these as the drift term of the Langevin equation:

$$\begin{aligned} \mu_x(x,y) &= k_{-1}a - k_1 x + k_3 x^2 y - k_{-3}x^3, \\ \mu_y(x,y) &= k_2 b - k_{-2} y - k_3 x^2 y + k_{-3}x^3. \end{aligned} \tag{8}$$

2.4.2 The Diffusion Term

As in Equation 1, we use Gaussian noise to model the diffusion term. The coefficient of the diffusion term can be derived from the variance of the copy numbers of X and Y. By the definition of $\sigma^2(X) = \langle X^2 \rangle - \langle X \rangle^2$, we have

$$\frac{d\sigma^2(X)}{dt} = \frac{d\langle X^2 \rangle}{dt} - 2\langle\ X\rangle \frac{d\langle X \rangle}{dt}. \tag{9}$$

We note that $\frac{d\langle X \rangle}{dt}$ has already been derived (Equation 6). $\frac{d\langle X^2 \rangle}{dt}$ can be obtained using the same approach as that of Equation (6) and Equation (7). After multiplying both sides of Equation (3) with X^2 and sum over X and Y from 0 to ∞, we combine the terms, and re-index them. We obtain

$$\begin{aligned} \frac{d\langle X^2 \rangle}{dt} &= 2k_{-1}A\langle X \rangle + k_{-1}A - 2k_1\langle X^2 \rangle + k_1\langle X \rangle + \frac{2k_3}{V^2}\langle X^2(X-1)Y \rangle \\ &\quad + \frac{k_3}{V^2}\langle X(X-1)Y \rangle - \frac{2k_{-3}}{V^2}\langle X^2(X-1)(X-2) \rangle + \frac{k_{-3}}{V^2}\langle X(X-1)(X-2) \rangle. \end{aligned} \tag{10}$$

Substituting Equations (6) and (10) into Equation (9), we now have obtained $\frac{d\sigma^2(X)}{dt}$ as

$$\begin{aligned} \frac{d\sigma_X^2(X,Y)}{dt} &= \frac{d\langle X^2 \rangle}{dt} - 2\langle X \rangle \frac{d\langle X \rangle}{dt} \\ &= k_{-1}A + k_1\langle X \rangle + \frac{k_3}{V^2}\langle X(X-1)Y \rangle + \frac{k_{-3}}{V^2}\langle X(X-1)(X-2) \rangle. \end{aligned} \tag{11}$$

In order to convert the copy numbers X and Y into concentrations, Equation (11) is divided by V^2. We arrive at the expression for the variance of the concentration of molecular species X as:

$$\sigma^2(\boldsymbol{x}) = \frac{d\sigma_x^2(x,y)}{dt} = \frac{1}{V}(k_{-1}a + k_1 x + k_3 x^2 y + k_{-3}x^3). \tag{12}$$

This is also the variance in Equation (1). Similarly, we have

$$\sigma^2(\boldsymbol{y}) = \frac{1}{V}(k_2 b + k_{-2} y + k_3 x^2 y + k_{-3}x^3). \tag{13}$$

We can now obtain the Langevin equation for the Schnakenberg model from its underlying chemical master equation:

$$\begin{aligned}\frac{dx}{dt} &= (k_{-1}a - k_1x + k_3x^2y - k_{-3}x^3) + \sqrt{\frac{1}{V}(k_{-1}a + k_1x + k_3x^2y + k_{-3}x^3)}\mathcal{N}\Big(0, \frac{1}{dt}\Big),\\ \frac{dy}{dt} &= (k_2b - k_{-2}y - k_3x^2y + k_{-3}x^3) + \sqrt{\frac{1}{V}(k_2b + k_{-2}y + k_3x^2y + k_{-3}x^3)}\mathcal{N}\Big(0, \frac{1}{dt}\Big).\end{aligned} \tag{14}$$

We call this the uncoupled Langevin equation (ULE).

2.4.3 The Coupling Terms Between Molecular Species

Although the Langevin equations obtained thus far can approximate the chemical master equation well, effects of higher order coupling intrinsic in complex biochemical systems are not accounted for in this model. For example, the molecular species X and Y in the Schnakenberg model are strongly coupled, but they are treated as independent random variables in the present form of the Langevin equation.

In order to obtain better approximation to the chemical master equation, we now extend the conventional Langevin formulation and add explicitly a coupling term. To account for the coupling effect between X and Y, we derive the $\frac{d\langle XY\rangle}{dt}$ term from the chemical master equation Equation (3) using similar approach as before:

$$\begin{aligned}\frac{d\langle XY\rangle}{dt} &= k_{-1}A\langle Y\rangle + k_2B\langle X\rangle - (k_1 + k_{-2})\langle XY\rangle\\ &\quad -\frac{k_3}{V^2}\langle XY(X-1)(X-Y+1)\rangle + \frac{k_{-3}}{V^2}\langle X(X-1)(X-2)(X-Y-1)\rangle.\end{aligned} \tag{15}$$

In the form of concentrations, we have

$$\frac{dxy}{dt} = k_{-1}ay + k_2bx - (k_1 + k_{-2})xy - k_3x^2y(x-y) + k_{-3}x^3(x-y). \tag{16}$$

To add this coupling term to the Langevin equation with proper dimension, we decompose this term by dividing it with y and x, respectively, to obtain $\frac{1}{y}\frac{dxy}{dt}$ and $\frac{1}{x}\frac{dxy}{dt}$. These are then added to Equation (14) with a scaling coefficient α:

$$\begin{aligned}\frac{dx}{dt} &= (k_{-1}a - k_1x + k_3x^2y - k_{-3}x^3) + \sqrt{\frac{1}{V}(k_{-1}a + k_1x + k_3x^2y + k_{-3}x^3)}\mathcal{N}\Big(0, \frac{1}{dt}\Big) + \alpha\frac{dxy}{ydt},\\ \frac{dy}{dt} &= (k_2b - k_{-2}y - k_3x^2y + k_{-3}x^3) + \sqrt{\frac{1}{V}(k_2b + k_{-2}y + k_3x^2y + k_{-3}x^3)}\mathcal{N}\Big(0, \frac{1}{dt}\Big) + \alpha\frac{dxy}{xdt}.\end{aligned} \tag{17}$$

The coefficient $\alpha = \alpha_{X,Y}$ represents the strength of the coupling between molecular species X and Y in the system. We assumes that the coupling effects are proportional to the number of coupling reactions in the system and the rates of such coupling reactions. It is therefore obtained by following the empirical formula $\alpha_{X,Y} = m\prod_i k_i$, in which k_i is the rate of reactions i that couples species X and Y, and the constant m is the total number of such coupling reactions. For the Schnakenberg model, there are two reactions through which X and Y are coupled, namely, the forward reaction and the backward reaction: $2X + Y \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} 3X$. We have the constant $m = 2$, $k_3 = 1.0$/sec, and $k_{-3} = 0.01$/sec. Therefore, $\alpha_{X,Y} = 2 \times 1.0 \times 0.01 = 0.02$.

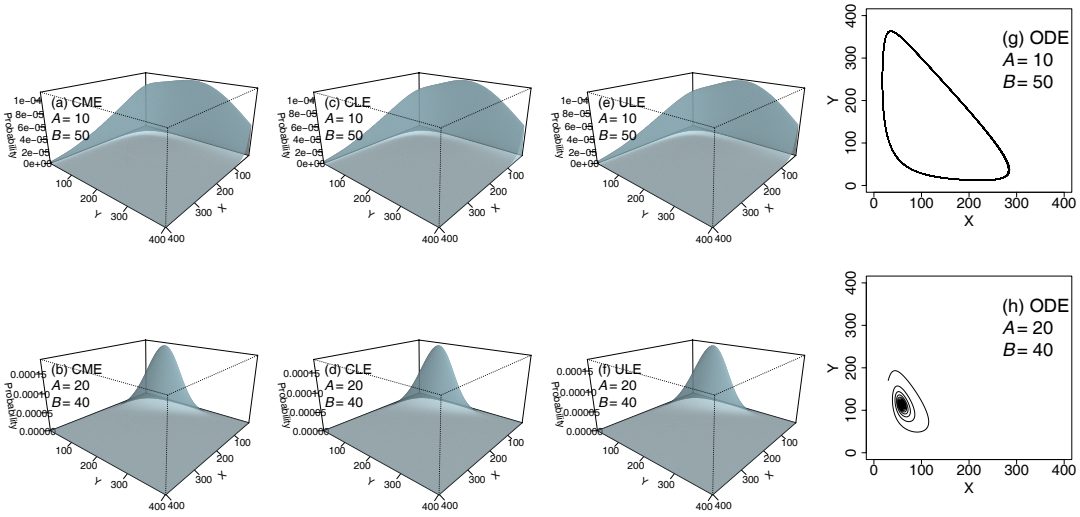

Figure 1 Calculated probability distributions over different copy numbers of X and Y and the trajectories of evolving concentrations of X and Y. (a) and (b): Exact probability distributions over copy numbers X and Y obtained by directly solving the chemical master equation (CME). Two sets of copy numbers of (A, B) at $(10, 50)$ and $(20, 40)$ are used for the fixed parameters A and B; (c) and (d): Reconstructed probability distributions over X and Y obtained from $200,000$ simulations of the coupling Langevin equation (CLE); (e) and (f): Reconstructed probability distributions over X and Y obtained from $200,000$ simulations of the original uncoupled Langevin equation (ULE); and (g) and (h): Trajectories of evolving concentrations of X and Y according to the deterministic ordinary differential equation (ODE). Here (g) shows the well-known oscillating limit cycle behavior of the Schnakenberg model, and (h) shows the convergence towards a fixed point.

It is possible this model can be further improved. The analysis of the computed results, and comparisons to other Langevin equation-based models are discussed below.

For a Gaussian probability distribution, the first and second order cumulants corresponds to its means and variances. Since we assume the noise term to be a Gaussian distribution in our Langevin equation, the drift and diffusion terms therefore will exactly correspond to the first and second cumulants (mean and variance) of the Gaussian distribution.

3 Numerical Simulations and Analysis

We have computed the exact steady state probability distributions of the system at different concentrations of X and Y by solving the chemical master equation using the method described in [8] (Figure 1a and 1b, CME for chemical master equation). The reaction rates are set at fixed values of $k_1 = k_2 = k_3 = k_{-1} = 1.0$ and $k_{-2} = k_{-3} = 0.01$. The system volume is set to a constant of $V = 100$. We use two sets of copy numbers of (A, B) of $(10, 50)$ and $(20, 40)$ for the fixed parameters A and B in the Schnakenberg model in our calculations. We have also reconstructed the probability distributions based on $200,000$ simulations of the improved coupling Langevin equation (Figure 1c and 1d, CLE for coupling Langevin Equation), and of the original uncoupled Langevin equation (Figure 1e and 1f, ULE for uncoupled Langevin

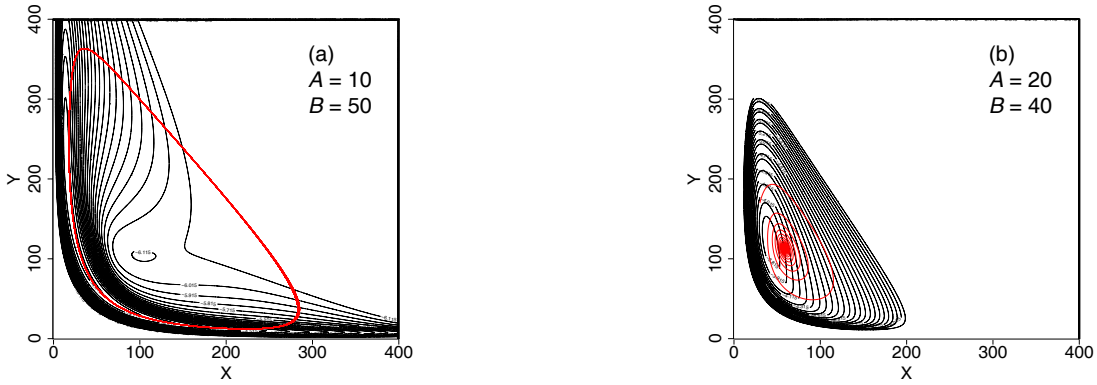

Figure 2 The contour maps of probability landscapes of the chemical master equation (CME). (a) The probability landscape contour map of the chemical master equation (CME) with parameters $A = 10$, and $B = 50$. The red line shows the limit cycle of the Schnakenberg ODE model at the same parameters. (b) The chemical master equation probability landscape contour map at parameters $A = 20$, and $B = 40$. The red line shows the fixed point of the Schnakenberg ODE model at the same parameters. The fixed point of ODE model is $X = 59$, and $Y = 113$. The black solid spot shows the location of maximum probability according to CME ($X = 41$, $Y = 118$)

Equation). For numerical simulation of the Langevin Equations, we follow Ito's formula[17] and use the Euler-Maruyama method[18]. In addition, we have computed the trajectories of the concentrations of X and Y evolving at different time according to the deterministic model (Figure 1g and 1h, ODE for Ordinary Differential Equation).

3.1 Comparison of Results From Models of Chemical Master Equations, Langevin Equations, and ODE Models

At the parameter values of $A = 10$ and $B = 50$, the well-known oscillating limit cycle behavior of the Schnakenberg model can be seen in Figure 1g. At $A = 20$ and $B = 40$, the behavior of the system converges towards a fixed point (Figure 1h). The landscape of the steady state probability distribution obtained from solving the chemical master equation shows a shallow crater surrounded by a mountainous ridge for the parameter set of $A = 10$ and $B = 50$ (Figure 2a). This corresponds well with the limit cycle behavior observed in the ODE model (Figure 1g and the circle line in Figure 2a). But the crater disappears in the landscapes of two Langevin equations (ULE and CLE) (detailed data not shown), due to residual errors between CME and Langevin equations. At $A = 20$ and $B = 40$, the landscapes of CME and Langevin equations show a single peak, which again corresponds well with the fixed-point behavior observed in the ODE model (Figure 1b,d,f and Figure 2b, in which the converging line shows the fixed point of the ODE model). It is obvious that there is a difference between the limit cycle in the ODE model and the CME ridge in Figure 2a, and the difference becomes more apparent for the case of single peak (fixed point) behavior, in which the ODE fixed point ($X = 58$, $Y = 113$) deviates significantly from the CME peak point ($X = 41$, $Y = 118$).

From these comparisons, we find that the model of Langevin equation generally approximates well the probability landscape obtained directly from chemical mater equation. From the comparisons of results from 25 different parameter sets (detailed data not shown), we find

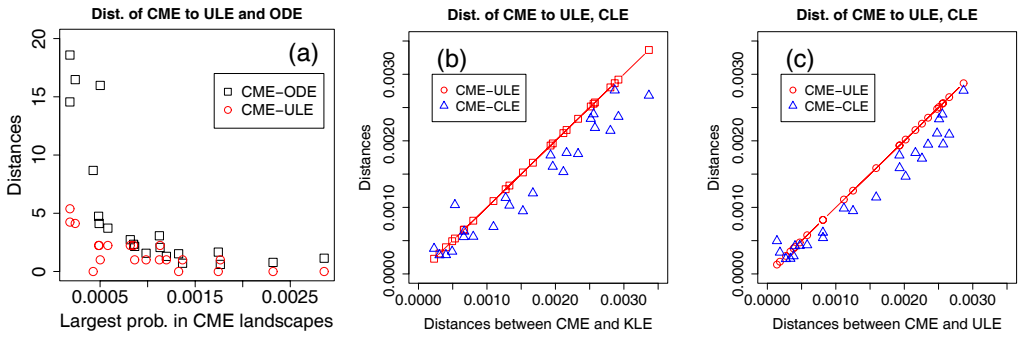

Figure 3 Differences among the methods and the improvements due to incorporation of the coupling effects. (a): Distances between the locations of the maxima of the probability distribution landscapes obtained from the chemical master equation (CME), the uncoupled Langevin equation (ULE), and the locations of the fixed points derived from the ODE model. Here the X-axis is the maximum probability value of the landscape derived from the CME, and Y-axis is the Euclidean distance between the locations of the maximum points by CME and ULE, and the locations of the maximum point by CME and the fixed point derived from the ODE model. Locations have coordinates in copy numbers of X and Y. Squares are distances between the CME and the ODE models, and the circles are distances between the CME and the ULE models. (b): Comparisons of distances from the CME landscape to the landscape computed from the uncoupled Langevin equation (ULE), and to that of the coupling Langevin equation (CLE). Here the distance between CME and ULE are plotted on the x-axis, and the distances beween CME and ULE (circles) as well as CME and CLE (triangle) are both plotted on the y-axis. (c): Same comparisons as in (b), but only for the low copy number region, namely, with $(X, Y) \in ([0, 100] \times [0, 100])$. The scaling coefficient for the coupling terms is calcualted as $\alpha = 0.02$

that overall the location of the maximum probabilities in copy numbers of X and Y from the Langevin equation is very close to that obtained from the chemical master equation (Figure 3a, circles). This indicates that the noise term derived from the chemical master equation works very well for the Schnakenberg model.

By comparisons of locations of CME peak points and ODE fixed points in different values of parameters, we find the locations of ODE fixed points can deviate significantly from the locations of maximum probability derived from the CME model, especially when the landscapes are relatively flat and the maximum probability is less pronounced (Figure 3a, squares).

Overall, our results show that the noise term plays essential role in determining the behavior of the Schnakenberg system, as the distance of maxima between solutions from the Langevin equation and the chemical master equation is usually less than the distance of maxima and fixed point between the solutions from the chemical master equation and the ODE. In Figure 3a, most of the circles are below the squares.

3.2 Effects of the Coupling Between Molecular Species

We now examine the landscapes reconstructed from the coupling Langevin equation. We compare the difference between two landscapes a and b by integrating the overall differences in

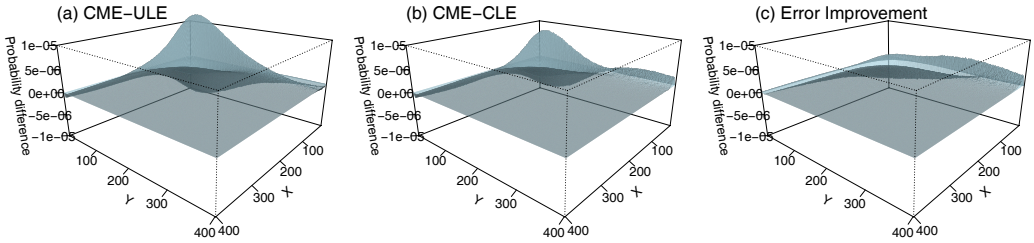

Figure 4 Comparison of errors between different models. (a): Difference in probability landscape of the uncoupled Langevin equation (ULE) and that of the chemical master equation (CME). This represents errors in ULE. (b): Difference in probability landscape of the coupling Langevin equation (CLE) and that of the chemical master equation (CME). This represents errors in CLE. One can apparently see that the errors in CLE has been reduced in comparison with ULE. (c): The amount of the errors in (a) that are corrected by the Coupling Langevin Equation (CLE) model. This shows the improvement due to the coupling term

probabilities: $d_{a,b} = \Big[\sum_{(X,Y)} [p_a(X,Y) - p_b(X,Y)]^2 \Big]^{1/2}$. For each of the 25 parameter sets, we calculate the difference $d_{\mathrm{CME,ULE}}$ between CME and the uncoupled Langevin equation (ULE), as well as the difference $d_{\mathrm{CME,CLE}}$ between CME and the coupling Langevin equation (CLE). The results are plotted in Figure 3b. Figure 3c depicts the errors between solutions obtaind from the chemical master equation and the coupling Langevin equation in the low copy number region of $[0, 100] \times [0, 100]$.

There is significant improvement by using the coupling Langevin equation versus the uncoupled Langevin equation, as can be seen in Figures 3b and c. Figure 4a and b show the residual errors in landscapes between CME and ULE (Figure 4a) and between CME and CLE (Figure 4b), which are constructed by subtracting ULE (CLE) probability landscape from the CME one. Figure 4c shows the error that is improved by incorporating the coupling effects (Figure 4c), which is constructed by subtracting the landscape of Figure 4a from Figure 4b. The parameter values are $A = 10$ and $B = 50$ for the case used in this comparison. We can clearly see that the error between CME and CLE is much smaller than that between CME and ULE. These results suggest that the introduction of the coupling effects of the macroscopic concentrations of the molecular species in the formulation of models of stochastic differential equation can significantly improve the accuracy of the model of biochemical networks.

4 Discussions

In this study, we have developed a novel formulation to account for the stochasticity in the biochemical networks. We introduce the noise term to the Langevin equation from the master equation formulation of the Schnakenberg model. An important development is the additional incorporation of the nonlinear coupling effect between molecular species. We find with simulations that for the Schnakenberg model, our coupling Langevin equation model provides a better approximation to the probability landscape of the underlying chemical master equation. The errors between Langevin equations and the chemical master equation (CME) have been significantly reduced in comparisons with the uncoupled Langevin equation by adding the coupling terms. Although our study is based on the Schnakenberg model, we believe the importance of

coupling effects is general. When the copy numbers of molecular species are small, and when they are changed simultaneously by one or more reactions in a system, such coupling effect need to be accounted for explicitly.

Acknowledgements

We thank Profs Ping Ao and Hong Qian for helpful discussions.

References

[1] D. T Gillespie, A rigorous derivation of the chemical master equation, *Physica A*, 1992, **188**(1–3): 404–425.
[2] D. T. Gillespie, Exact stochastic simulation of coupled chemical reactions, *Journal of Physical Chemistry*, 1997, **81**: 2340–2361.
[3] D. T. Gillespie, Approximate accelerated stochastic simulation of chemically reacting systems, *The Journal of Chemical Physics*, 2001, **115**(4): 1716–1733.
[4] D. T. Gillespie, Stochastic simulation of chemical kinetics, *Annual Review of Physical Chemistry*, 2007, **58**: 35–55.
[5] B. Munsky and M. Khammash, The finite state projection algorithm for the solution of the chemical master equation, *The Journal of Chemical Physics*, 2006, **124**(4): 044104.
[6] B. Munsky and M. Khammash, A multiple time interval finite state projection algorithm for the solution to the chemical master equation, *Journal of Computational Physics*, 2007, **226**(1): 818–835.
[7] S. Peles, B. Munsky, and M. Khammash, Reduction and solution of the chemical master equation using time scale separation and finite state projection, *The Journal of Chemical Physics*, 2006, **125**(20): 204104.
[8] Y. Cao and J. Liang, Optimal enumeration of state space of finitely buffered stochastic molecular networks and exact computation of steady state landscape probability, *BMC Systems Biology*, 2008, **2**: 30.
[9] Y. Cao, H. M. Lu, and J. Liang, Stochastic probability landscape model for switching efficiency, robustness, and differential threshold for induction of genetic circuit in phage lambda, *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, 2008, **1**: 611–614.
[10] D. T. Gillespie, The chemical langevin equation, *Journal of Chemical Physics*, 2000, **113**: 297–306.
[11] D. T. Gillespie, The Chemical Langevin and Fokker-Planck Equations for the Reversible Isomerization Reaction, *J. Phys. Chem. A*, 2002, **106**(20): 5063–5071.
[12] Robert Zwanzig, Nonlinear generalized Langevin equations, *Journal of Statistical Physics*, 1973, **9**(3): 215–220.
[13] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry*, 3rd Edition, Elsevier Science and Technology Books, 2007.
[14] J. Schnakenberg, Simple chemical reaction systems with limit cycle behaviour, *Journal of Theoretical Biology*, 1979, **81**: 389–400.
[15] H. Qian, Open-system nonequilibrium steady state: Statistical thermodynamics, fluctuations, and chemical oscillations, *The Journal of Physical Chemistry B*, 2006, **11**: 15063–15074.
[16] H. Qian, S. Saffarian, and E. L. Elson, Concentration fluctuations in a mesoscopic oscillating chemical reaction system, *PNAS*, 2002, **99**: 10376–10381.
[17] B. Oksendal, *Stochastic Differential Equations*, Springer, Berlin, 2003.
[18] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer, Berlin, 1999.