Math. Model. Nat. Phenom.
Vol. 9, No. 1, 2014, pp. 79–91
DOI: 10.1051/mmnp/20149106

Simulating Stochasticities in Chemical Reactions with Deterministic Delay Differential Equations

H. Jin, J. Lei *

Zhou Pei-Yuan Center for Applied Mathematics, MOE Key Laboratory of Bioinformatics
Tsinghua University, Beijing 100084, China

Abstract. The stochastic dynamics of chemical reactions can be accurately described by chemical master equations. An approximated time-evolution equation of the Langevin type has been proposed by Gillespie based on two explicit dynamical conditions. However, when numerically solve these chemical Langevin equations, we often have a small stopping time–a time point of having an unphysical solution–in the case of low molecular numbers. This paper proposes an approach to simulate stochasticities in chemical reactions with deterministic delay differential equations. We introduce a deterministic Brownian motion described by delay differential equations, and replace the Gaussian noise in the chemical Langevin equations by the solutions of these deterministic equations. This modification can largely increase the stopping time in simulations and regain the accuracy as in the chemical Langevin equations. The novel aspect of the present study is to apply the deterministic Brownian motion to chemical reactions. It suggests a possible direction of developing a hybrid method of simulating dynamic behaviours of complex gene regulation networks.

Keywords and phrases: Master equation, chemical Langevin equation, deterministic Brownian motion, delay differential equation

Mathematics Subject Classification: 60-08, 92-08

1. Introduction

Stochasticities in chemical reactions have received much attentions in recent years due to their applications in the study of molecular dynamics in cell biology [7, 12, 13, 17, 21, 22, 24]. For a spatially homogeneous chemical system, the chemical master equation provides a microscopic description for the time evolution of molecule numbers [14]. The chemical master equation is established following a firm physical basis, but unfortunately is often mathematically intractable because of the huge number of dimension (same as the total number of possible states of the system). The chemical Langevin equation, proposed by Gillespie in 2000 [9], is an alternative approach to model the stochastic chemical reactions system. The chemical Langevin equation is obtained based on the combination of the microphysical premise for the chemical master equation and two explicit dynamical conditions (named as *Gillespie conditions*, to be detailed below), and therefore share the firm physical basis of the stochasticity related to the chemical

*Corresponding author. E-mail: jzlei@tsinghua.edu.cn

© EDP Sciences, 2014

Article published by EDP Sciences and available at http://www.mmnp-journal.org or http://dx.doi.org/10.1051/mmnp/20149106
https://doi.org/10.1051/mmnp/20149106 Published online by Cambridge University Press

master equation. The chemical Langevin equation is an Itô stochastic differential equation of dimension the same as the reaction species, the number of Gaussian noise equals the number of reaction channels. These equations can be solved numerically with schemes of stochastic simulation (to be detailed below). However, the two dynamical conditions for deriving the chemical Langevin equation somehow contradict to each other, and not always satisfied simultaneously. This often happens in intercellular biological processes. Furthermore, in simulating the chemical Langevin equation, we often face the problem of negative population due to the unbounded Gaussian noise and must stop the simulation. Thus, two questions are raised: how can we model the stochasticity of a chemical reaction system that fails to satisfy the Gillespie conditions; and how can we avoid the issue of negative population in simulations. These two issues relate to each other because the problem of negative population often associates with failures in the Gillespie conditions. In this study, we modify the chemical Langevin equation to release conditions on time-constrain of maintaining positive populations, and examine how the modified equations can be applied to describe the stochasticity of a chemical reaction system.

Consider a well-stirred mixture of $N \geq 1$ molecular species $\{S_1, \cdots, S_N\}$ that are chemically interacting, inside some fixed volume Ω and at a constant temperature, through $M \geq 1$ reaction channels $\{R_1, \cdots, R_M\}$. We specify the dynamical state of this system by $\boldsymbol{X}(t) = (X_1(t), \cdots, X_N(t))$, where $X_i(t)$ is the number of S_i in the system at time t. Each reaction channel R_j associates a *propensity function* a_j and a *state-change vector* $\boldsymbol{v}_j = (v_{j1}, \cdots, v_{jN})$, which are defined such that (refer [9] to details) $a_j(\boldsymbol{x})dt$ gives the probability, given $\boldsymbol{X}(t) = \boldsymbol{x}$, that one reaction R_j will occur somewhere inside Ω in the next infinitesimal time interval $[t, t + dt)$, and the state is changed from $\boldsymbol{X}$ to $\boldsymbol{X} + \boldsymbol{v}_j$ if the reaction occurs.

Suppose that the state of a system at current time t is $\boldsymbol{X}(t) = \boldsymbol{x}$. Let $K_j(\boldsymbol{x}, \tau)(\tau > 0)$ be the number of reaction R_j that occur in the subsequent time interval $[t, t+\tau]$. Then the system state at time $t+\tau$ is

$$\boldsymbol{X}(t+\tau) = \boldsymbol{x} + \sum_{j=1}^{M} K_j(\boldsymbol{x}, \tau)\boldsymbol{v}_j. \tag{1.1}$$

Here $K_j(\boldsymbol{x}, \tau)$ is a random variables. Gillespie proposed following two conditions that ate opposing to each other [9]:

(G1) τ is *small* enough such that the change in the state during $[t, t+\tau]$ is slight that none of the propensity functions changes its value "appreciably".

(G2) τ is *large* enough such that the expected number of occurrences of each reaction channel R_j in $[t, t+\tau]$ is much larger than 1.

These two conditions allow us to approximate $K_j(\boldsymbol{x}, \tau)$ by a normal random variable with both mean and variance equal $a_j(\boldsymbol{x})\tau$, and hence (1.1) becomes

$$\boldsymbol{X}(t+\tau) = \boldsymbol{x} + \sum_{j=1}^{M} \boldsymbol{v}_j \left(a_j(\boldsymbol{x})\tau + \sqrt{a_j(\boldsymbol{x})\tau}\mathcal{N}_j(0,1) \right), \tag{1.2}$$

where $\mathcal{N}_j(0,1)$ denotes standard normal random variables. The equation (1.2) is approximately the Euler scheme of the following chemical Langevin equation

$$dX_i = \sum_{j=1}^{M} v_{ji} a_j(\boldsymbol{X})dt + \sum_{j=1}^{M} v_{ji}\sqrt{a_j(\boldsymbol{X})}dW_j, \quad (i = 1, \cdots, N). \tag{1.3}$$

The Gillespie conditions (G1) and (G2) are satisfied when the propensity functions $a_j(\boldsymbol{X})$ are large enough so that $1 \gg \tau \gg 1/a_j(\boldsymbol{X})$. However, in many biological systems, the molecule numbers of each species are small, and thus the two conditions are not satisfied simultaneously. In this case, the validity of (1.3) is questionable. Furthermore, it is obvious that the equation (1.3) is valid only when $a_j(\boldsymbol{X}) \geq 0$ for

https://doi.org/10.1051/mmnp/20149106 Published online by Cambridge University Press

any j. This condition defines a domain $\Lambda \in \mathbb{R}^N$ of all physically possible states, and a physically possible solution $\boldsymbol{X}(t)$ satisfies $\boldsymbol{X}(t) \in \Lambda$ for all $t \geq 0$. Nevertheless, when we numerically solve the Langevin equation (1.3), it is always possible to reach a state beyond Λ at some time point $t^* > 0$, i.e., $\boldsymbol{X}(t^*) \notin \Lambda$, and hence the simulation fail. This yields a *stopping time* T in stochastic simulations. A crucial issue in long-term simulation is how can we increase the stopping time.

There are many numerical schemes trying to release the Gillespie conditions and to avoid negative values (refer [11] for a review), including the tau-leaping that approximates K_j by Poisson random numbers [3,10], an adaptive method incorporating postleap checks in tau-leaping [1], the binomial leaping method that approximates K_j by binomial random numbers [4,23], combinations of tau-leaping with the stochastic simulation algorithm (SSA) [2,6], *etc.* Despite successful applications of these methods, there is no simple formalism in terms of differential equations as we have seen in the original chemical Langevin equation.

In the above numerical methods, a key is to replace the random number K_j by a bounded random variable at each step of simulations. Recently, one author of this paper found that a type of delay differential equations can generate a deterministic Brownian motion that show statistical properties akin to those of a classical Brownian motion over a wide time scale, but is truncated with an upper bound depending on an adjustable parameter [18]. In this paper, we seek a way to replace the Brownian motion W_j in the chemical Langevin equation (1.3) by the deterministic Brownian motion generated from the delay differential equations. In this way, we are able to simulate the stochasticity in chemical reactions by deterministic dynamics.

In the rest of this paper, we first introduce the deterministic Brownian motion generated by delay differential equations in Section 2. We also discuss the modelling of chemical reaction systems with delay differential equation and the stopping time in simulations. In Section 3, we study two examples with the delay differential equation model, and show that the deterministic models are able to reproduce stochasticities as in the stochastic simulation algorithm (SSA). This paper is concluded with a discussion in Section 4.

2. Modelling with delay differential equations

2.1. Deterministic Brownian motion generated from delay differential equations

From the deduction of the chemical Langevin equations, the noise terms in (1.3) are introduced to approximate the number of reactions in a finite time interval and must be a finite value. We ask whether it is possible to replace the stochastic terms dW_j in (1.3) with some other noise terms that are bounded but have similar statistical properties such as a near Gaussian distribution signals and diffusive behaviours. In [18], a class of truncated Gaussian distribution noise, called *quasi-Gaussian distribution noise*, are introduced by deterministic Brownian motion. These deterministic Brownian motions are generated from solutions of delay differential equations of form

$$\begin{aligned} &\frac{d\eta}{ds} = -\eta + \sin(2\pi\beta\eta(s-1)), \\ &\eta(\theta) = \phi(\theta), -1 < \theta < 0. \end{aligned} \tag{2.1}$$

Here $s = \omega t$ has a time scale differ from that of the chemical reaction system. In numerical schemes discussed below, we want a weak correlation between $\eta(t)$ and $\eta(t+\Delta t)$ so that $\{\eta(n\Delta t)\}_{n=0}^{+\infty}$ approximates a time series of white noise. According to [18], the correlation time of (2.1) is about $s = 1$ (or $t = 1/\omega$ in terms of the original time scale) and independent to β. In all simulations in the current study, we take a time scale so that $s = 100\,t$ (*i.e.*, $\omega = 100$).

It was shown numerically that for almost all values of β away from 0 (say $\beta > 0.85$) and any initial functions $\phi(\theta) \in [-1, 1]$, the solutions $\eta(s)$ of (2.1) is chaotic, statistically similar to a Gaussian noise with mean 0 and variance

$$\sigma(\beta) = 0.32\beta^{-1/2}, \tag{2.2}$$

https://doi.org/10.1051/mmnp/20149106 Published online by Cambridge University Press

and is bounded with

$$|\eta(s)| \le K(\beta) := \frac{1}{0.68\sqrt{\beta} + 0.60} \tag{2.3}$$

when $s > 0$ (away from the initial value at $s = 0$). Thus, let

$$\zeta(s) = \frac{\eta(s)}{\sigma(\beta)}, \tag{2.4}$$

then $\zeta(s)$ is a time sequence of *standard quasi-Gaussian noise* that has standard Gaussian-like distribution but with a truncate tail so it is supported at a finite interval (refer [18] for details). Explicitly (and approximately), $|\zeta(s)|$ has an upper bound

$$K_0(\beta) = \frac{\sqrt{\beta}}{0.21\sqrt{\beta} + 0.19}, \tag{2.5}$$

and the density function is

$$Pr(\zeta = z|\beta) = \begin{cases} \mathcal{C}_0 e^{-z^2/2}, & |z| \le K_0 \\ 0, & |z| > K_0 \end{cases} \tag{2.6}$$

at the stationary state, where[1]

$$\mathcal{C}_0 = \frac{1}{\int_{-K_0}^{K_0} e^{-z^2/2} dz}. \tag{2.7}$$

Fig. 1 shows a sample trajectory of $\zeta(s)$ and the corresponding density function.

In the rest of this paper, we examine how can we use the above standard quasi-Gaussian noise to simulate stochasticities in chemical reactions.

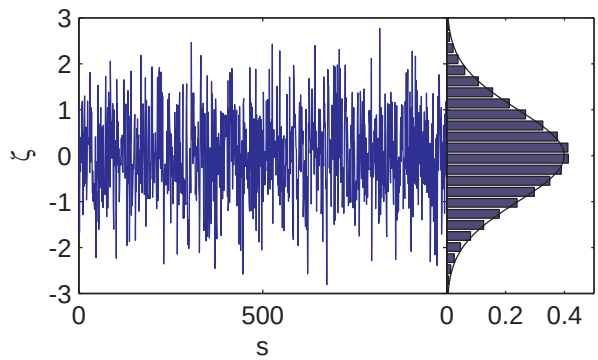


FIGURE 1. Quasi-Gaussian noise. A trajectory of $\zeta(s) = \eta(s)/\sigma(\beta)$ determined by (2.1) (left panel), and the probability density of the trajectory (right panel) (solid line shows the probability density of a standard Gaussian distribution noise). Here $\beta = 5.0$, and the initial function is $\phi(\theta) \equiv -0.134162$.

[1]Here we correct a mistake in [18] in which $\mathcal{C}_0 = \frac{1}{\sqrt{2\pi}\int_{-K_0}^{K_0} e^{-z^2/2} dz}$ by mistake.

https://doi.org/10.1051/mmnp/20149106 Published online by Cambridge University Press

2.2. Deterministic Chemical Langevin Equation

In the chemical Langevin equation (1.3), we divide dt from both sides and replace each stochastic term dW_j/dt with a standard quasi-Gaussian noise (2.4), this yields the following delay differential equations

$$\frac{dx_i}{dt} = \sum_{j=1}^{M} v_{ji} a_j(\boldsymbol{x}) + \sum_{j=1}^{M} v_{ji}\sqrt{a_j(\boldsymbol{x})}\zeta_j(t), (1 \le i \le N) \tag{2.8}$$

where $\zeta_j(t) = \eta_j(\omega t)/\sigma(\beta)$ with η_j determined by delay differential equations of form (2.1)

$$\frac{d\eta_j(\omega t)}{d\,\omega t} = -\eta_j + \sin(2\pi\beta\eta_j(\omega t - 1)), \quad (1 \le j \le M). \tag{2.9}$$

Here a time scale ω is introduced so that $\eta(\omega t)$ satisfies an equation of form (2.1) with $s = \omega t$.

Equations (2.8)-(2.9) are a set of deterministic delay differential equations, and can be solved numerically. For example, by the Euler method

$$x_i(t + \Delta t) = x_i + \Delta t \sum_{j=1}^{M} v_{ji}(a_j(\boldsymbol{x}) + \sqrt{a_j(\boldsymbol{x})}\zeta_j(t)), \tag{2.10}$$

here $\boldsymbol{x}(t) = x$, and $\zeta_j(t) = \eta_j(\omega t)/\sigma(\beta)$ with $\eta_j(\omega t)$ from (2.9). In simulations, we need to chose initial functions for $\eta_j(\theta)$. A simple strategy is to let $\eta_j(\theta)$ equal constants when $-1 \le \theta \le 1$, but differ from each other for each j. Furthermore, there is an adjustable parameter β in (2.9). In a simple example of binary state transition discussed below, simulation results are not sensitive to different values of β.

We note that while solving the equation (2.8) with the scheme (2.10), one need to generate the 'random' terms ζ_j by solving the stiff equation (2.9). To this end, at each iteration step (2.10), we need to solve (2.9) separately with a smaller time step. Thus, it is less efficient to solve the delay differential equations (2.8)-(2.9) than to directly solve the stochastic differential equation (1.3). Nevertheless, as we can see below, solving the equations (2.8)-(2.9) can largely increase the stopping time, and therefore one can use a larger time step in simulations. Thus, our method is still valuable for long time behaviors simulations.

2.3. Stopping time

Now, we discuss the stopping time according to the two different simulation methods: solve the chemical Langevin equation (1.3) with the explicit order 1.0 Runge-Kutta method (also called the Milstein scheme) [15, 20], or solve the deterministic equation (2.8)-(2.9) with the order 1.0 Euler method (2.10). Both methods are of first order approximation.

First, consider the chemical Langevin equation (1.3). From the Milstein scheme, we have an iteration

$$\begin{aligned} X_i(t + \Delta t) = X_i(t) &+ \sum_{j=1}^{M} v_{ji} a_j(\boldsymbol{X}(t))\Delta t \\ &+ \sum_{j=1}^{M} v_{ji}\sqrt{a_j(\boldsymbol{X}(t))}\Delta W_j \\ &+ \frac{1}{2}\sum_{j=1}^{M}\sum_{l=1}^{N} v_{jl} v_{ji} \frac{\partial a_j(\boldsymbol{X}(t))}{\partial X_l}((\Delta W_j)^2 - \Delta t), \end{aligned} \tag{2.11}$$

where $\Delta W_j \sim \sqrt{\Delta t}\mathcal{N}(0, 1)$. Let $\boldsymbol{x} = \boldsymbol{X}(t)$, and $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_M)$ a M-dimensional standard Gaussian random variable defined as

$$\xi_j = \Delta W_j/\sqrt{\Delta t}, \quad (j = 1, \cdots, M).$$

https://doi.org/10.1051/mmnp/20149106 Published online by Cambridge University Press

Let $\boldsymbol{f}(\boldsymbol{\xi};\boldsymbol{x},\Delta t)=(f_1,\cdots,f_N)$ with

$$f_i = x_i + \Delta t\sum_{j=1}^{M} v_{ji}\left(a_j(\boldsymbol{x}) + \sqrt{a_j(\boldsymbol{x})}\frac{\xi_j}{\sqrt{\Delta t}}\right) + \frac{1}{2}\Delta t\sum_{j=1}^{M}\sum_{l=1}^{N} v_{jl}v_{ji}\frac{\partial a_j(\boldsymbol{x})}{\partial x_l}(\xi_j^2-1),$$

then

$$\boldsymbol{X}(t+\Delta t)=\boldsymbol{f}(\boldsymbol{\xi};\boldsymbol{x},\Delta t). \tag{2.12}$$

Thus, since $\boldsymbol{x}\in\Lambda$, the probability of a valid iteration so that $\boldsymbol{X}(t+\Delta t)\in\Lambda$ is given by a function of $\boldsymbol{x}$ and Δt as

$$p(\boldsymbol{x},\Delta t)=\frac{1}{(\sqrt{2\pi})^M}\int_{\boldsymbol{f}(\boldsymbol{z};\boldsymbol{x},\boldsymbol{\Delta t})\in\Lambda} e^{-\boldsymbol{z}^2/2}d\boldsymbol{z}. \tag{2.13}$$

The probability $p(\boldsymbol{x},\Delta t)$ depends on $\boldsymbol{x}$ and varies from step to step. To obtain an estimation of the maximum probability that only depends on the step size Δt, we define

$$p(\Delta t)=\sup_{\boldsymbol{x}\in\Lambda} p(\boldsymbol{x},\Delta t). \tag{2.14}$$

Then $p(\Delta t)$ gives the maximum probability of one successful iteration, given the step size Δt.

From the probability (2.14), the number S of successful iterations is bounded by a geometric distribution random number so that

$$Pr(S=n|\Delta t)\le p(\Delta t)^n(1-p(\Delta t)). \tag{2.15}$$

Thus, the average number of successful iterations satisfies

$$S_{\rm cle}(\Delta t)\le\frac{p(\Delta t)}{1-p(\Delta t)}, \tag{2.16}$$

and hence the average stopping time satisfies

$$T_{\rm cle}(\Delta t)\le\frac{p(\Delta t)\Delta t}{1-p(\Delta t)}. \tag{2.17}$$

The equation (2.17) gives an upper bound of the average stopping time when we apply the Milstein scheme to the chemical Langevin equation (1.3).

Now, we consider the deterministic equations (2.8) and the corresponding Euler method (2.10). Let $\boldsymbol{\zeta}=(\zeta_1,\cdots,\zeta_M)$, and $\boldsymbol{g}(\boldsymbol{\zeta};\boldsymbol{x},\Delta t)=(g_1,\cdots,g_M)$ with

$$g_i = x_i + \Delta t\sum_{j=1}^{M} v_{ji}(a_j(\boldsymbol{x}) + \sqrt{a_j(\boldsymbol{x})}\zeta_j), \tag{2.18}$$

then

$$\boldsymbol{x}(t+\Delta t)=\boldsymbol{g}(\zeta;\boldsymbol{x},\Delta t).$$

Similar to the previous discussions, the probability of a successful iteration, given $\boldsymbol{x}$ and Δt, is

$$q(\boldsymbol{x},\Delta t)=\mathcal{C}_0^M\int_{g(\boldsymbol{z};\boldsymbol{x},\Delta t)\in\Lambda,|z_i|\le K_0(\beta)} e^{-\boldsymbol{z}^2/2}d\boldsymbol{z}. \tag{2.19}$$

Let

$$q(\Delta t)=\sup_{\boldsymbol{x}\in\Lambda} q(\boldsymbol{x},\Delta t), \tag{2.20}$$

https://doi.org/10.1051/mmnp/20149106 Published online by Cambridge University Press

then the average stopping time satisfies

$$T_{\rm dde} \leq \frac{q(\Delta t)\Delta t}{1-q(\Delta t)}. \tag{2.21}$$

Equations (2.17) and (2.21) give estimations for the upper bound of average stopping times for the two numerical methods. These upper bounds are usually much larger than the real stopping time in simulations. Fig. 2 shows simulating stopping times for an example of binary state transition obtained from these two schemes, respectively. The results indicate that, at least for this simple example, the simulation method based on delay differential equation model can largely increase the stopping time.

Finally, we ask if it is possible to have infinite stopping time, i.e., the numerical scheme is valid forever. Here we discuss a situation that delay differential equation modelling is able to yield an infinite stopping time.

Assume that the region Λ is given by

$$\Lambda = \{\boldsymbol{x}|\ \alpha_i \leq x_i \leq \beta_i\}, \tag{2.22}$$

where α_i, β_i are constants. This assumption indicates that when $x_i = \alpha_i$ (and $\boldsymbol{x} \in \Lambda$), the reactions might produce the molecule S_i, but cannot contain S_i as a reactant, and hence

$$v_{ji}a_j(\boldsymbol{x}) \geq 0. \tag{2.23}$$

When $x_i = \beta_i$ (and $\boldsymbol{x} \in \Lambda$), the reactions might contain molecule S_i as a reactant, but can not produce S_i, and hence

$$v_{ji}a_j(\boldsymbol{x}) \leq 0. \tag{2.24}$$

In this case, an iteration step is successful if the time step Δt is taken so that

$$\alpha_i \leq x_i + \Delta t \sum_{j=1}^{M} v_{ji}(a_j(\boldsymbol{x}) + \sqrt{a_j(\boldsymbol{x})}\zeta_j) \leq \beta_i, \quad (i = 1, \cdots, N). \tag{2.25}$$

Since $|\zeta_j| \leq K_0(\beta)$ and (2.23)-(2.24), (2.25) is satisfied when $\Delta t < h(\boldsymbol{x})$ at each step with

$$h(\boldsymbol{x}) = \min_{1\leq i\leq N} \left\{ \frac{x_i - \alpha_i}{\max\{0, \sum_{j=1}^{M}(K_0(\beta)|v_{ji}\sqrt{a_j(\boldsymbol{x})}| - v_{ji}a_j(\boldsymbol{x}))\}}, \right. \\ \left. \frac{\beta_i - x_i}{\max\{0, \sum_{j=1}^{M}(v_{ji}a_j(\boldsymbol{x}) + K_0(\beta)|v_{ji}\sqrt{a_j(\boldsymbol{x})}|)\}} \right\}. \tag{2.26}$$

Thus, if the time step Δt is taken so that

$$\Delta t < \min_{\boldsymbol{x}} h(\boldsymbol{x}), \tag{2.27}$$

the Euler scheme (2.10) is always valid.

3. Examples

Here, we consider two examples to illustrate the validity of the application of deterministic Brownian motion in modelling stochastic chemical reactions. The two examples include a simple model of binary state transition and a simple circuit of gene regulation with positive feedback.

https://doi.org/10.1051/mmnp/20149106 Published online by Cambridge University Press

3.1. A simple model of binary state transition

Consider a simple example of binary state transition of a molecule

$$\mathrm{A} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \mathrm{I}. \tag{3.1}$$

This simple reaction is often involved in many biological processes, such as state transitions of promoters, ion channel gates, protein activities, nucleosome states, *etc.* When there are low number of molecules, it is easy to simulate the dynamics of this reaction with SSA. Here we show that the delay differential equation model can reproduce long time dynamics obtained from SSA, and numerical schemes according to the chemical Langevin equation fail to procedure long time simulation due to a small stopping time.

Let n the total number of molecules, and x the number of active ones (A), then a physically possible state must have $0 \le x \le n$. The chemical Langevin equation model (CLE) is given by

$$dx = (-k_1 x + k_2(n-x))dt - \sqrt{k_1 x}dW_1 + \sqrt{k_2(n-x)}dW_2, \tag{3.2}$$

and the delay differential equations model (DDE) is

$$\frac{dx}{dt} = -k_1 x + k_2(n-x) - \sqrt{k_1 x}\zeta_1(t) + \sqrt{k_2(n-x)}\zeta_2(t) \tag{3.3}$$

where $\zeta_j(t) = \eta_j(\omega t)/\sigma(\beta)$ with $\eta_j(\omega t)$ satisfies (2.9). When n is large, simulations based on both models are consistent with each other, and in agreement with simulations based on SSA in terms of their statistical properties (data not shown). However, there are significant differences when n is small.

When $n = 1$, Fig. 2 shows stopping times as functions of the step size Δt obtained from different numerical schemes according to the two equations (3.2) and (3.3) (we always set $\omega = 100$ in simulations). It is obvious that the DDE models yield much larger stopping time than those obtained from the CLE model. Furthermore, varying the parameter β would not affect the stopping time. Finally, if DDE and the Brownian motion have same statistical properties, both DDE and CLE method would give the same stopping time value for large values of Δt. This is confirmed from Fig. 2.

Fig. 3 shows sample trajectories obtained from DDE and CLE models. The trajectories obtained from the DDE model (each solution start from different initial condition) show fluctuations (in a way similar to stochastic fluctuation) around the mean obtained from the exact SSA (Fig. 3a). While the solutions obtained from the CLE model have much larger fluctuations, and are very easy to go beyond the physically possible state so that the simulations are stop (Fig. 3b).

A comparison of mean and variance obtained from the two methods is given at Table 1. Here we modified the Milstein scheme (CLE models) in two different ways so that we always have $0 \le x \le n$. The first method, absorbed BM in Table 1, introduces an absorbing boundary condition that sets $x = 0$ whenever $x < 0$ and $x = n$ when $x > n$ in simulations. The second method, truncated BM in Table 1, introduces truncated Brownian motions so that the numerical scheme becomes

$$x(t+\Delta t) = x(t) + (-k_1 x(t) + k_2(n - x(t)))\Delta t - \sqrt{k_1 x}\Delta\tilde{W}_1 + \sqrt{k_2(n-x(t))}\Delta\tilde{W}_2, \tag{3.4}$$

where $\Delta\tilde{W}_i$ are Gaussian variables (approximately $\Delta\tilde{W}_i \sim \Delta t\mathcal{N}(0,1)$) but are restricted with $|\Delta\tilde{W}_i/\Delta t| < K_0'$ (here we take $K_0' = 5$ in simulations). These two modifications are natural to avoid the issue of negative population. From Table 1, the DDE models (with varying β values) agree well with SSA in the mean, but not the variance (3 order smaller than the value obtained from SSA), and are not sensitive to β. The absorbed BM method gives a larger mean value (about double of the SSA result), and also a smaller variance. The truncated BM method gives a mean value consistent with SSA, and an even smaller variance (4 order smaller than the SSA value).

It is not surprise to have smaller variance in DDE and CLE models. This is because at each step simulation, the state x is digitalized and equals 0 or 1 at SSA, but is continuous and $0 \le x \le 1$ in CLE

https://doi.org/10.1051/mmnp/20149106 Published online by Cambridge University Press

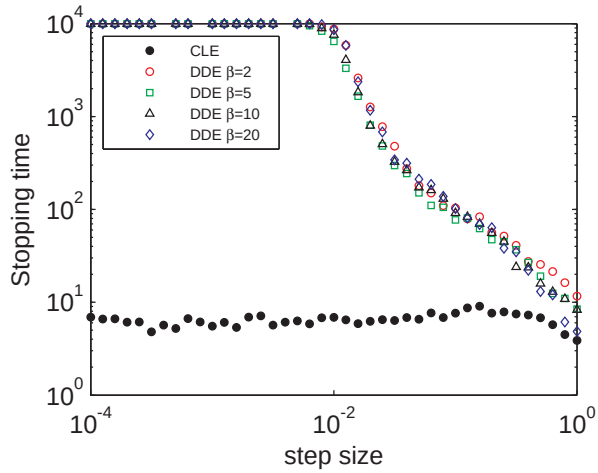


FIGURE 2. Stopping times obtained from different numerical methods. Markers show the stopping time in simulating the example of binary state transition with different methods: Solving the CLE model (3.2) with Milstein method (solid circles), and the DDE model (3.3) with varying values of β (hollow markers). In simulations, we vary Δt from 10^{-4} to 1, and the maximum simulation time is 10^4 in each simulation. To obtain the stopping time, for each method (for given parameter and step size), we solve the solution for 100 independent trajectories (with different initial functions while solving the DDE model), and each solution gives a stopping time. The average stopping time, shown by the figure, is the average of these 100 stopping times. We take $k_1 = 0.03, k_2 = 0.005$ and $n = 1$ in all simulations.

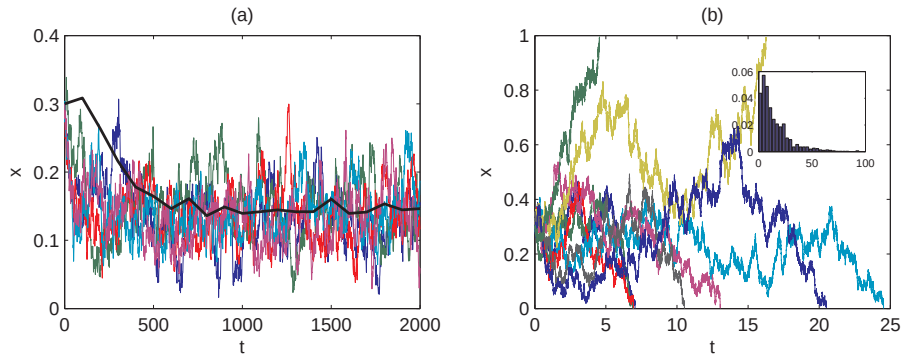


FIGURE 3. Trajectories obtained from different numerical schemes with $n = 1$. (a) Sample trajectories obtained by solving the delay differential equation model (3.3) with Euler method with $\Delta t = 0.01$ (shown by different colour curves). Here $\beta = 5.0$, and the initial functions are taken randomly for each trajectory. The black curve shows the average time course from SSA of 100 sample paths. (b) Sample trajectories obtained by solving the corresponding Chemical Langevin Equation with the Milstein scheme with $\Delta t = 0.001$. Inset shows the distribution of the stopping time of 1000 sample paths. Parameters used are the same as in Fig. 2.

https://doi.org/10.1051/mmnp/20149106 Published online by Cambridge University Press

or DDE models. This inconsistent is originated from the continuation of the state valuable in deriving the chemical Langevin equations, and become worse when we replace the Gaussian noise with quasi-Gaussian noise (which is bounded). It is more reasonable to interpret the solution $x(t)$ of DDE models as the average of finite trajectories obtained by SSA with fluctuations originated from uncertainties due to the variable continuation (refer Fig. 3 and Fig. 5 below). If we want to refer the solution $x(t)$ as a single pathway, we often need to introduce additional fluctuations as compensation. For example, for a particular solution $x(t)$ of the DDE model (3.3), if we define the molecule state as $X(t) = 1$ with a probability $x(t)$, then we can reconstruct a sample path with state X in $\{0, 1\}$. The random pathways reconstructed in this way are consistent with SSA in their mean and variance at the stationary state (data not shown). Nevertheless, this method of reconstruction is only valid for this particular example, and it requires further studies for a common way to recompense the fluctuations for general reaction systems.

TABLE 1. Comparison of mean and variance at stationary states obtained from different method. Here the absorbed BM means that we numerically solve the CLE (3.2), but with a truncated in the state valuable so that $x = 0$ whenever $x < 0$, and $x = 1$ whenever $x > 1$ in the simulations. The truncated BM means that we apply the numerical scheme (3.4) with truncated Brownian motions (see the text for detail). Parameters are same as in Fig. 2.

method	mean	variance
SSA	0.1419	0.1217
DDE ($\beta = 2$)	0.139	0.0014
DDE ($\beta = 5$)	0.1298	0.0015
DDE ($\beta = 10$)	0.1458	0.0015
DDE ($\beta = 20$)	0.1531	0.0015
Absorbed BM	0.2814	0.0572
Truncated BM	0.1425	0.0001

3.2. Gene regulation circuit with positive feedback

In the previous example, we have seen that the DDE approximation is able to reproduce the mean of dynamics of a simple model of binary state transition, but with much smaller variance. The motif of binary state transition is often included in many complex gene regulation networks. We ask if it is possible to apply the DDE approximation to those state transition motifs, and leave other components modelled by traditional ordinary differential equation (which means that we omit the stochastic effects in these reactions). Here, we consider a simple circuit with feedback regulation (Fig. 4) and examine how it works. In this circuit (and in many complex gene regulation networks), the transition of promoter state (either ON or OFF) is random and with low number. Thus, we often need to model the promoter state transition with SSA. On the other hand, the protein number is often large so that its dynamics can be modelled by deterministic chemical rate reactions. Here, we simulate the circuit dynamics with two methods, stochastic simulation (SSA) or deterministic modelling (DDE modelling), and compare our results.

In the deterministic modelling, we apply the DDE approximation to the promoter state transition as in the previous example. Moreover, the stochasticity in gene transcription is also modelled with DDE approximation because of a low promoter number, and other reactions are described by ordinary differential equations. Let x_1, x_2, x_3, x_4 the numbers of ON state promoter, mRNA and protein molecules,

https://doi.org/10.1051/mmnp/20149106 Published online by Cambridge University Press

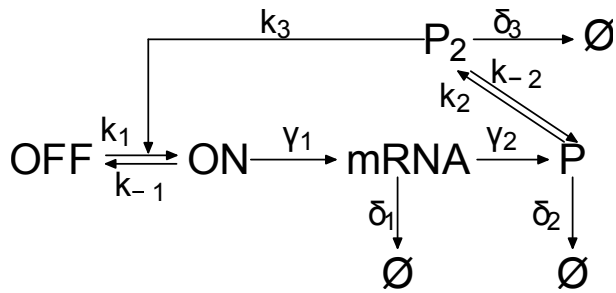


FIGURE 4. A simple model of gene regulation circuit with positive feedback. In this circuit, proteins (P) form dimers (P_2) that can bind to the promoter of their own gene to increase the gene expression level.

and the dimers. The model equation is given by

$$\begin{cases} \dfrac{dx_1}{dt} = (k_1 + k_3 x_4)(1 - x_1) - k_{-1} x_1 \\ \qquad\qquad + \sqrt{(k_1 + k_3 x_4)(1 - x_1)}\zeta_1 \\ \qquad\qquad - \sqrt{k_{-1} x_1}\zeta_2 \\ \dfrac{dx_2}{dt} = \gamma_1 x_1 - \delta_1 x_2 + \sqrt{\gamma_1 x_1}\zeta_3 \\ \dfrac{dx_3}{dt} = \gamma_2 x_2 - \delta_2 x_3 - 2k_2 x_3^2 + 2k_{-2} x_4 \\ \dfrac{dx_4}{dt} = k_2 x_3^2 - k_{-2} x_4 - \delta_3 x_4. \end{cases} \tag{3.5}$$

Again, $\zeta_i(t) = \eta_j(\omega t)/\sigma(\beta)$ with $\eta_j(\omega t)$ given by (2.9).

In simulations, we first use SSA to obtain 100 sample trajectories independently, and then examine the time course of the averages of mRNA and protein numbers of these trajectories. This process is to mimic the dynamics of gene transcriptions of a pool of 100 cells. Results are shown by black curves in Fig. 5. Next, we solve the DDE model (3.5) (with initial conditions consist with SSA simulations) for a single trajectories (shown by red curves in Fig. 5). The results show that deterministic model can well reproduce the dynamics of mRNA and protein numbers of averages of a pool of cells.

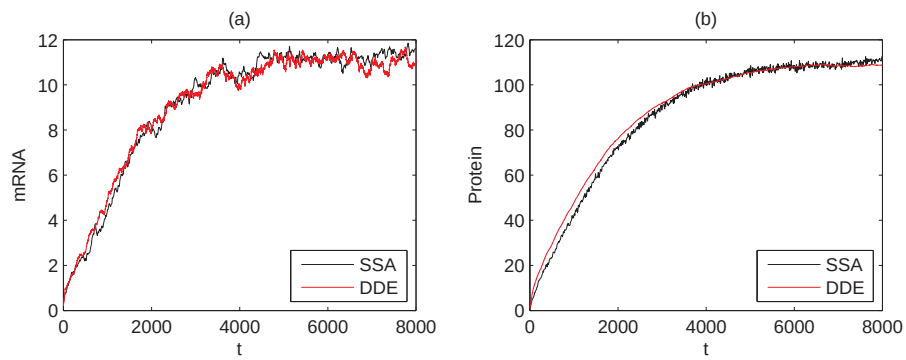


FIGURE 5. Simulation results for the model at Fig. 4. (a) Time course of the protein numbers. (b) Distribution of protein numbers at stationary state. Parameters used are: $k_1 = 0.005, k_{-1} = 0.03, \gamma_1 = 0.07, \delta_1 = 0.005, \gamma_2 = 0.3, \delta_2 = 0.0004, k_2 = 0.1, k_{-2} = 0.001, \delta_6 = 0.001, \beta = 5$, and the time step $\Delta t = 0.01$.

https://doi.org/10.1051/mmnp/20149106 Published online by Cambridge University Press

4. Discussion

In this paper, we have proposed an attempt to simulate stochasticities in chemical reactions with deterministic delay differential equations. Our approach is based on a modification in the chemical Langevin equation introduced by Gillespie [9]. In the chemical Langevin equation, the Brownian motions are unbounded, and therefore often yield small stopping time in simulation. In our approaches, we introduce a deterministic Brownian motion described by delay differential equations [18], and replace the Brownian motion in the chemical Langevin equations by the solutions of these delay differential equations. With this modification, we are able to largely increase the stopping time in simulations. We apply our method to two examples, showing that the deterministic models are able to reproduce the average behaviour in comparing with stochastic simulations using SSA.

Many studies of deterministic Brownian motion has been published in the past several decades, and there are numerous investigations have documented the existence of Brownian-like motion from deterministic dynamics [5, 8, 16, 18, 19, 25]. However, to our best knowledge, the current study is the first try to apply the deterministic Brownian motion to chemical reactions which are intrinsically random. This idea suggests a possible direction of developing a hybrid method of simulating dynamic behaviours of complex gene regulation networks.

Acknowledgements. This work is supported by the National Natural Science Foundation of China (NSFC 11272169).

References

[1] D.F. Anderson. *Incorporating postleap checks in tau-leaping*. J. Chem. Phys., 128 (2008), 054103.
[2] Y. Cao, D.T. Gillespie, L.R. Petzold. *Avoiding negative populations in explicit poisson tau-leaping*. J. Chem. Phys., 123 (2005), 054104.
[3] Y. Cao, D.T. Gillespie, L.R. Petzold. *Efficient stepsize selection for the tau-leaping simulation method*. J. Chem. Phys., 124 (2006), 044109.
[4] A. Chatterjee, D. Vlachos, M. Katsoulakis. *Binomial distribution based τ-leap accelerated stochastic simulation*. J. Chem. Phys., 122 (2005), 024112.
[5] L.Y. Chew, C. Ting. *Microscopic chaos and gaussian diffusion processes*. Physica A., 307 (2002), 275–296.
[6] S.L. Cotter, K.C. Zygalakis, I.G. Kevrekidis, R. Erban. *A constrained approach to multi scale stochastic simulation of chemically reacting systems*. J. Chem. Phys., 135 (2011), 094102.
[7] M.B. Elowitz, A.J. Levine, E.D. Siggia, and P.S. Swain, *Stochastic gene expression in a single cell*. Science, 297 (2002), 1183–1186.
[8] P. Gaspard, M.E. Briggs, M.K. Francis, J.V. Sengers, R.W. Gammon, J.R. Dorfman, R.V. Calabrese. *Experimental evidence for microscopic chaos*. Nature, 394 (1998), 865–868.
[9] D.T. Gillespie. *The chemical langevin equation*. J. Chem. Phys., 113 (2000), 297–306.
[10] D.T. Gillespie. *Approximate accelerated stochastic simulation of chemically reacting systems*. J. Chem. Phys., 115 (2001), 1716–1733.
[11] D.T. Gillespie. *Stochastic simulation of chemical kinetics*. Annu. Rev. Phys. Chem., 58 (2007), 35–55.
[12] I. Golding, J. Paulsson, S.M. Zawilski, E.C. Cox, *Real-time kinetics of gene activity in individual bacteria*. Cell, 123 (2005), 1025–1036.
[13] M. Kærn, T.C. Elston, W.J. Blake, J.J. Collins, *Stochasticity in gene expression: from theories to phenotypes*. Nat. Rev. Genet., 6 (2005), 451–464.
[14] N.G. Van Kampen. *Stochastic processes in physics and chemistry*. 3rd ed., Elsevier, 2007.
[15] P.E. Kloeden, E. Platen. *Numerical solution of stochastic differential equations*. Springer, New York, 1992.
[16] A. Lasota, M.C. Mackey. *Probabilistic properties of deterministic systems*. Cambridge University Press, Cambridge, 2008.
[17] J. Lei. *Stochasticity in gene expression with both intrinsic noise and fluctuation in kinetic parameters*. J. Theor. Biol., 256 (2009), 485–492.
[18] J. Lei, M.C. Mackey. *Deterministic brownian motion generated from differential delay equation*. Phy. Rev. E. 84 (2011), 041105.
[19] M.C. Mackey, M. Tyran-Kamińska. *Deterministic brownian motion: The effects of perturbing a dynamical system by a chaotic semi-dynamical system*. Physics Reports, 422 (2006), 167–222.
[20] G.N. Milstein. *Numerical integration of stochastic differential equations*. Kluwer, Dordrecht, 1995.
[21] B. Munsky, G. Neuert, A. van Oudenaarden. *Using gene expression noise to undertand gene regulation*. Science, 336 (2012), 183–187.

https://doi.org/10.1051/mmnp/20149106 Published online by Cambridge University Press

[22] J. Paulsson. *Summing up the noise in gene networks*. Nature, 427 (2004), 415–418.
[23] T. Tian, K. Burrage. *Binomial leap methods for simulating stochastic chemical kinetics*. J. Chem. Phys., 121 (2004), 10356–10364.
[24] T-L.To, N. Maheshri. *Noise can induce bimodality in positive transcriptional feedback loops without bistability*. Science, 327 (2010), 1142–1145.
[25] G. Trefán, P. Grigolini, B.J. West. *Deterministic brownian motion*. Phys. Rev. A., 45 (1992), no. 2, 1249–1252.

https://doi.org/10.1051/mmnp/20149106 Published online by Cambridge University Press