Available online at www.sciencedirect.com

ScienceDirect

Journal of Computational and Applied Mathematics 205 (2007) 696–707

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

www.elsevier.com/locate/cam

Stochastic delay differential equations for genetic regulatory networks

Tianhai Tian[a,*], Kevin Burrage[a], Pamela M. Burrage[a], Margherita Carletti[b]

[a] *Advanced Computational Modelling Centre, The University of Queensland, Australia*
[b] *Institute of Biomathematics, University of Urbino, Italy*

Received 3 November 2005

Abstract

Time delay is an important aspect in the modelling of genetic regulation due to slow biochemical reactions such as gene transcription and translation, and protein diffusion between the cytosol and nucleus. In this paper we introduce a general mathematical formalism via stochastic delay differential equations for describing time delays in genetic regulatory networks. Based on recent developments with the delay stochastic simulation algorithm, the delay chemical master equation and the delay reaction rate equation are developed for describing biological reactions with time delay, which leads to stochastic delay differential equations derived from the Langevin approach. Two simple genetic regulatory networks are used to study the impact of intrinsic noise on the system dynamics where there are delays.

MSC: 65C05; 92B99; 92C45

Keywords: Stochastic delay differential equations; Genetic regulatory networks; Chemical Langevin equation; Stochastic simulation algorithm

1. Introduction

A genetic regulatory network (GRN) consists of a number of genes that interact and regulate the expression of other genes by the gene derivatives, i.e. proteins. The change in expression of a gene is controlled by the stimulation and inhibition of proteins in transcriptional, translational and post-translational processes. Due to small numbers of transcriptional factors and other key signalling proteins, there is considerable experimental evidence that noise plays a very important role in gene regulation. In recent years there is an accelerating interest in the development of stochastic models and simulation methods for describing the functions of intrinsic noise, due to the uncertainty of biochemical reactions, and extrinsic noise, due to fluctuations in the environment. Although the stochastic simulation method (SSA) is a statistically exact method for simulating the dynamics of GRNs based on detailed biochemical reactions [10], the large computational time of the SSA is a barrier in the applications of this method to large biological systems. In order to improve the efficiency of the SSA, the τ-leap methods have been proposed by using Poisson [10] or binomial [20] random variables. In addition, stochastic differential equations (SDEs) have been used in stochastic simulations based

* Corresponding author. Tel.: +61 7 33656137; fax: +61 7 33651477.
E-mail addresses: tian@maths.uq.edu.au (T. Tian), kb@maths.uq.edu.au (K. Burrage), pmb@maths.uq.edu.au (P.M. Burrage), m.carletti@mat.uniurb.it (M. Carletti).

0377-0427/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2006.02.063

on the chemical master equation [12]. Recently, Burrage et al. [8] have proposed multi-scale methods that include the SSA, τ-leap methods and SDEs in a sophisticated way for simulating chemical systems when there are wide ranges of molecular numbers.

In current approaches for modelling GRNs, it is assumed that each reaction fires in the system immediately with the SSA being used to determine the small waiting time for the next reaction. This assumption is valid for fast reactions such as dimerization, binding reactions and phosphorylations. On the other hand, slow reactions such as transcription and translation involve a number of multi-stage reactions. In an important recent experiment on mouse, Hirata et al. [13] have shown that there is a time lag of about 15 min in the peaks between the mRNA molecules and proteins of the gene *hes*1. For the gene *her*1 in zebrafish, it will take approximately 21 min from the initiation of transcription to the arrival of the mature mRNA molecule in the cytoplasm, and about 2.8 min from the initiation of translation to the emergence of a complete functional protein molecule [15]. However, such time delays were ignored in the current approaches for stochastic modelling of GRNs. In order to deal with these time delays in discrete stochastic systems, Barrio et al. [3] have proposed the delay SSA (DSSA), the counterpart of the SSA for describing the dynamics of GRNs more accurately.

It will not be surprising that we will have the same problems of large computational time in the application of the DSSA. Aimed at improving the efficiency of the DSSA, here we develop the delay chemical master equation and delay reaction rate equation that will be the basis for the development of multi-scale methods for GRNs with time delay. We will introduce stochastic delay differential equations (SDDEs) for GRNs based on the Langevin approach when the propensity functions associated with the reactions are relatively large. We emphasize that these SDDEs represent intrinsic noise in a biological system due to the uncertainty of knowing when and what reaction occurs. This work will enrich the application fields of SDDEs that have already been used in modelling complex systems in population dynamics, epidemiology, immunology, physiology, cell kinetics and mathematical finance [5,14]. Based on the progress of efficient methods for simulating SDDEs [2,6,7,14,16], we will use two simple systems to study the impact of time delay on the dynamics of GRNs and discuss the implementations of variable time delays. The rest of this paper is organized as follows: Section 2 will discuss discrete simulation methods SSA and DSSA. Section 3 will introduce SDDEs for GRNs. Simulation results will be presented in Section 4.

2. Discrete stochastic simulation methods

When applying stochastic simulation algorithms for simulating biochemical reaction systems, it is assumed that a well-stirred chemical reaction system contains N molecular species $\{S_1, \ldots, S_N\}$ with non-negative number $X_i(t)$ of the species S_i at time t. These species of molecules chemically interact inside some volume Ω at a constant temperature through reactions $\{R_1, \ldots, R_M\}$.

For each reaction R_j (j=1, …, M), we define a propensity function $a_j(\mathbf{x})$ in a given state $\mathbf{X}(t)=(X_1(t), \ldots, X_N(t))^\top = \mathbf{x}$ and use $a_j(\mathbf{x})\,\mathrm{d}t$ to represent the probability that one reaction R_j will fire somewhere inside Ω in the infinitesimal time interval $[t, t+\mathrm{d}t)$. In addition, a state change vector v_j is defined to characterize reaction R_j. The element v_{ij} of v_j represents the change in the number of species S_i due to reaction R_j. The $N \times M$ matrix v with elements v_{ij} is called the stoichiometric matrix.

The SSA is a statistically exact procedure for generating the time and index of the next occurring reaction in accordance with the current values of the propensity functions. In the SSA, we draw two independent random numbers r_1 and r_2 from the uniform distribution in the unit interval $U(0, 1)$, and then take the time of the next reaction to be the current time plus μ, where

$$\mu = \frac{1}{a_0(\mathbf{x})} \ln\left(\frac{1}{r_1}\right), \tag{1}$$

and $a_0(\mathbf{x})$ is the sum of the values of all propensity functions

$$a_0(\mathbf{x}) = \sum_{k=1}^{M} a_k(\mathbf{x}).$$

The index of the next reaction is the value of j that satisfies

$$\sum_{k=1}^{j-1} a_k(\mathbf{x}) < r_2 a_0(\mathbf{x}) \leqslant \sum_{k=1}^{j} a_k(\mathbf{x}). \tag{2}$$

Then the system is updated by

$$\mathbf{x}(t+\mu) = \mathbf{x}(t) + v_j. \tag{3}$$

The SSA assumes that the next reaction will fire in the next reaction time interval $[t, t+\tau)$ with small values of τ. For systems including both fast and slow reactions, however, this assumption may be not valid if the slow reactions take a much longer time than the fast reactions. The large reaction time of slow reactions should be realized by time delay if we hope to put both fast and slow reactions in a system consistently and to study the impact of slow reactions on the system dynamics. Most recently Barrio et al. have proposed the DSSA by including the time delay in the simulations of GRNs [3]. This method uses a similar approach as that in the SSA for deciding the next reaction and a temporary next reaction time. The system will be updated at this temporary next reaction time if the next reaction has no time delay, otherwise the update will take place at some time later according to the time delay of the selected reaction and the temporary next reaction time. This method is given in detail as follows.

Step 1: Generate one random number r_1 from the uniform distribution $U(0, 1)$ and calculate a temporary reaction time required for the next reaction

$$\mu = \frac{1}{a_0(\mathbf{x})} \ln\left(\frac{1}{r_1}\right), \tag{4}$$

where $a_0(\mathbf{x}) = \sum_{j=1}^{M} a_j(\mathbf{x})$.

Step 2: Select a reaction based on the current system state at t whose index satisfies

$$\sum_{k=1}^{j-1} a_k(\mathbf{x}) < r_2 a_0(\mathbf{x}) \leqslant \sum_{k=1}^{j} a_k(\mathbf{x}). \tag{5}$$

Step 3: If the selected reaction has no time delay, update the system as

$$\mathbf{x}(t+\mu) = \mathbf{x}(t) + v_j. \tag{6}$$

Otherwise set a record to update the system at $t_d = t + T_j + \mu$ if the selected next reaction has time delay T_j.

Step 4: For all reactions having time delay, check if there is any reaction R_j that is scheduled to fire in the time interval $[t, t+\tau)$, and update the system based on

$$\mathbf{x}(t+\mu) = \mathbf{x}(t) + v_j.$$

3. SDDEs via the master equation

The SSA describes the evolution of a discrete nonlinear Markov process. This stochastic process has a probability density function that is the solution of a differential equation (the chemical master equation). This master equation can be used to write down an ODE that describes the deterministic behaviour of the mean associated with the SSA or a SDE that represents the intrinsic noise in continuous form. In a similar manner the DSSA described in Section 2 will have corresponding representations as delay differential equations or SDDEs.

The discussion in this section is aimed at providing the theoretical background for constructing SDDEs that faithfully represent intrinsic noise (in a continuous form) arising from delayed reactions. Here it is assumed that a well-stirred chemical reaction system contains N molecular species $\{S_1, \ldots, S_N\}$ with number $X_i(t)$ of the species S_i at time t. Among the M reaction channels, the first M_1 reactions $\{R_1, \ldots, R_{M_1}\}$ are assumed to have time delay $\{T_1, \ldots, T_{M_1}\}$, respectively, and the last $M - M_1$ reactions have no time delay. From the description of the DSSA in the previous section, the state vector $X(t)$ is a non-negative N-dimensional jump stochastic process but is not a Markov process

any more due to the time delay. Here we are interested in the conditional probability function based on the initial state $\mathbf{X}(t_0)$ and the states involved with the time delay $\mathbf{X}(t) = \Phi(t)$ $(t \leqslant t_0)$, given by

$$P(\mathbf{x}, t) \equiv \text{Prob}\{\mathbf{X}(t) = \mathbf{x} | \mathbf{X}(t_0) = \mathbf{x}_0, \text{ and } \mathbf{X}(t) = \Phi(t), \ t \leqslant t_0\}. \tag{7}$$

For GRNs with time delay, the master equation should be based on the current system state at time t for reactions without time delay and the system state at $t - T_j$ for the reaction channel R_j with time delay T_j. In order to derive the time evolution equation of the probability function (7), we take a time increment $\mathrm{d}t$ that is so small that the probability for two or more reactions to occur in $\mathrm{d}t$ is negligible compared to the probability for at most one reaction. It is assumed that the reaction time of a reaction without time delay is $\mathrm{d}t$, while for a reaction with time delay T_j the reaction time is $\mathrm{d}t + T_j$. Then the probability of the system being in state $\mathbf{x}$ at $t + \mathrm{d}t$ is given by

$$\begin{aligned} P(\mathbf{x}, t + \mathrm{d}t) = P(\mathbf{x}, t) &- \sum_{j=1}^{M_1} \sum_{\mathbf{x}_i \in I(\mathbf{x})} a_j(\mathbf{x}_i) P(\mathbf{x}, t; \mathbf{x}_i, t - T_j)\, \mathrm{d}t \\ &+ \sum_{j=1}^{M_1} \sum_{\mathbf{x}_i \in I(\mathbf{x})} a_j(\mathbf{x}_i) P(\mathbf{x} - v_j, t; \mathbf{x}_i, t - T_j)\, \mathrm{d}t \\ &- \sum_{j=M_1+1}^{M} a_j(\mathbf{x}) P(\mathbf{x}, t)\, \mathrm{d}t + \sum_{j=M_1+1}^{M} a_j(\mathbf{x} - v_j) P(\mathbf{x} - v_j, t)\, \mathrm{d}t, \end{aligned} \tag{8}$$

where $P(\mathbf{x}, t; \mathbf{x}_i, t - T_j)$ is the probability that the system is both in the state $\mathbf{x}$ at t and in the state $\mathbf{x}_i$ at $t - T_j$, and $I(\mathbf{x})$ is the set of all possible system states. The second term on the right-hand side is the probability that no reaction will fire in $[t - T_j, t + \mathrm{d}t)$ for the reaction channel R_j with time delay T_j and the system is in the state $\mathbf{x}$ at $t + \mathrm{d}t$. Here we should consider all the possible states $\mathbf{x}_i$ at $t - T_j$ because the system evolves in the time period $[t - T_j, t)$ based on the reactions of other reaction channels. For the case of one reaction, the third term also considers all the possible system states at $t - T_j$ for the reaction with time delay T_j but the system should be in the state $\mathbf{x} - v_j$ at t. The last two terms are the probabilities for reactions without time delay, in which only the system states at t should be included in the master equation.

When $\mathrm{d}t \to 0$, this leads to the delay chemical master equation

$$\begin{aligned} \frac{\partial}{\partial t} P(\mathbf{x}, t) = &- \sum_{j=1}^{M_1} \sum_{\mathbf{x}_i \in I(\mathbf{x})} a_j(\mathbf{x}_i) P(\mathbf{x}, t; \mathbf{x}_i, t - T_j) + \sum_{j=1}^{M_1} \sum_{\mathbf{x}_i \in I(\mathbf{x})} a_j(\mathbf{x}_i) P(\mathbf{x} - v_j, t; \mathbf{x}_i, t - T_j) \\ &- \sum_{j=M_1+1}^{M} a_j(\mathbf{x}) P(\mathbf{x}, t) + \sum_{j=M_1+1}^{M} a_j(\mathbf{x} - v_j) P(\mathbf{x} - v_j, t). \end{aligned} \tag{9}$$

The value of $a_j(\mathbf{x}_i) P(\mathbf{x} - v_j, t; \mathbf{x}_i, t - T_j)$ is the probability that one reaction of R_j (with time delay T_j) fires at $t - T_j$ and at the same time the system states are $\mathbf{x} - v_j$ at time t and $\mathbf{x}_i$ at time $t - T_j$, respectively. Here the probability $P(\mathbf{x} - v_j, t; \mathbf{x}_i, t - T_j)$ measures the strength of coupling of the system states at t and $t - T_j$. If the reaction number in the time period $[t - T_j, t]$ is relatively small, this probability is critical to the system behaviour and the DSSA must be employed. However, if the time delays are large and there are a relatively large number of reactions in the time interval $[t - T_j, t)$, it is reasonable to assume that the coupling of the system states at t and $t - T_j$ is weak and the probability can be approximated by

$$P(\mathbf{x} - v_j, t; \mathbf{x}_i, t - T_j) \approx P(\mathbf{x} - v_j, t) \times P(\mathbf{x}_i, t - T_j). \tag{10}$$

Here we should re-emphasize that the above assumption is based on the fact that there are a relatively large number of reactions firing in the time period $[t - T_j, t]$. This is an appropriate application area of SDEs. In addition, this assumption is only for the coupling of systems states but does not connect to the importance of time delay in the system dynamics. In fact time delay still plays a very important role in the evolution of system dynamics through reactions that are represented by the first two items in (9). Similar consideration can be applied to the first item in (9).

Based on assumption (10), we can obtain the mean of the propensity functions for reactions with time delay, given by

$$\overline{a_j(\mathbf{x}(t-T_j))} = \sum_{\mathbf{x}_i \in I(\mathbf{x})} a_j(\mathbf{x}_i) P(\mathbf{x}_i, t-T_j). \tag{11}$$

Then the delay chemical master equation can be simplified as

$$\begin{aligned}\frac{\partial}{\partial t} P(\mathbf{x},t) = &-\sum_{j=1}^{M_1} \overline{a(\mathbf{x}(t-T_j))} P(\mathbf{x},t) + \sum_{j=1}^{M_1} \overline{a(\mathbf{x}(t-T_j))} P(\mathbf{x}-v_j,t) \\ &- \sum_{j=M_1+1}^{M} a_j(\mathbf{x}) P(\mathbf{x},t) + \sum_{j=M_1+1}^{M} a_j(\mathbf{x}-v_j) P(\mathbf{x}-v_j,t).\end{aligned} \tag{12}$$

In the case that there is no delay equation (12) reduces to the well-known chemical master equation associated with the SSA, namely

$$\frac{\partial}{\partial t} P(\mathbf{x},t) \sum_{j=1}^{M} a_j(\mathbf{x}-v_j) P(\mathbf{x}-v_j,t) - \sum_{j=1}^{M} a_j(\mathbf{x}) P(\mathbf{x},t). \tag{13}$$

If we multiply the delay chemical master equation (12) by all of the states at t, sum over all these system states, and then re-index the summation on the right-hand side, we can obtain the equations for the mean of $\mathbf{X}(t)$

$$\frac{\mathrm{d}\overline{\mathbf{X}(t)}}{\mathrm{d}t} = \sum_{j=1}^{M_1} v_j \overline{a_j(\mathbf{X}(t-T_j))} + \sum_{j=M_1+1}^{M} v_j \overline{a_j(\mathbf{X}(t))}. \tag{14}$$

When all molecular numbers are very large and fluctuations are not important, we can get the delay reaction rate equation, given by

$$\frac{\mathrm{d}\mathbf{X}(t)}{\mathrm{d}t} = \sum_{j=1}^{M_1} v_j a_j(\mathbf{X}(t-T_j)) + \sum_{j=M_1+1}^{M} v_j a_j(\mathbf{X}(t)). \tag{15}$$

Note, if there are no delays, (15) is an ODE describing the standard chemical kinetics rate equations. However, if intrinsic noise is still important but there are large enough numbers of molecules we would expect a representation described by an SDDE. This would be then a completely natural generalization of the SSA case where a SDE describes the evolution of the molecular concentrations. In order to see how this comes about we apply the explicit Euler method with constant stepsize h to the delay reaction rate equation (15), giving

$$\mathbf{X}(t_n+h) = \mathbf{X}(t_n) + h \sum_{j=1}^{M_1} v_j a_j(\mathbf{X}(t_n-T_j)) + h \sum_{j=M_1+1}^{M} v_j a_j(\mathbf{X}(t_n)). \tag{16}$$

In the SSA setting Gillespie [10] introduced the concept of Poisson τ-leap methods in order to speed up the computational performance of the SSA. But it can also be used as a convergent technique to describe the evolution of molecular concentrations when noise is still important but the representation is continuous rather than discrete. In this method a larger time step than that taken in the SSA is used and a number of reactions are allowed to fire within this step with a frequency drawn from a Poisson distribution. We can adopt the same approach for the DSSA and this gives rise to the delay Poisson τ-leap method, given by

$$\mathbf{X}(t_n+h) = \mathbf{X}(t_n) + \sum_{j=1}^{M_1} v_j P(a_j(\mathbf{X}(t_n-T_j))h) + \sum_{j=M_1+1}^{M} v_j P(a_j(\mathbf{X}(t_n))h), \tag{17}$$

where $P(\lambda)$ is a Poisson random variable with mean λ.

In order to understand the behaviour of this scheme as $a_j(\mathbf{X}(t))h \gg 0$, we note that if the mean of the Poisson random variables in (17) is very large, a Poisson random variable $P(a_j(\mathbf{X}(t))h)$ can be approximated by the Gaussian random variable with mean and variance $a_j(\mathbf{X}(t))h$ [10], namely

$$\begin{aligned} P(a_j(\mathbf{X}(t))h) &\approx N(a_j(\mathbf{X}(t))h, a_j(\mathbf{X}(t))h) \\ &= a_j(\mathbf{X}(t))h + \sqrt{a_j(\mathbf{X}(t))h}N(0,1). \end{aligned}$$

Then the discrete stochastic processes $\mathbf{X}(t)$ in (17) can be approximated by the following continuous processes with Gaussian random variables

$$\begin{aligned} \mathbf{X}(t_n + h) = \mathbf{X}(t_n) &+ h\sum_{j=1}^{M_1} v_j a_j(\mathbf{X}(t_n - T_j)) + h\sum_{j=M_1+1}^{M} v_j a_j(\mathbf{X}(t_n)) \\ &+ \sum_{j=1}^{M_1} v_j\sqrt{a_j(\mathbf{X}(t_n - T_j))h}I_j + \sum_{j=M_1+1}^{M} v_j\sqrt{a_j(\mathbf{X}(t_n))h}I_j, \end{aligned} \tag{18}$$

where $I_j \sim N(0,1)$. When $h \to 0$, this gives rise to systems of SDDEs that describe the evolution of chemical concentrations for GRNs with time delay

$$\begin{aligned} \mathrm{d}\mathbf{X} = &\sum_{j=1}^{M_1} v_j a_j(\mathbf{X}(t - T_j))\,\mathrm{d}t + \sum_{j=M_1+1}^{M} v_j a_j(\mathbf{X}(t))\,\mathrm{d}t \\ &+ \sum_{j=1}^{M_1} v_j\sqrt{a_j(\mathbf{X}(t - T_j))}\,\mathrm{d}W_j(t) + \sum_{j=M_1+1}^{M} v_j\sqrt{a_j(\mathbf{X}(t))}\,\mathrm{d}W_j(t), \end{aligned} \tag{19}$$

where $W_j(t)$ is the Wiener process whose increment is a Gaussian process $\Delta W_j(t) = W_j(t+\Delta t) - W_j(t) \sim \sqrt{\Delta t}N(0,1)$.

In the case that there are no delays this corresponds to the so-called chemical Langevin equation (see [9])

$$\mathrm{d}\mathbf{X} = \sum_{j=1}^{M} v_j a_j(\mathbf{X}(t))\,\mathrm{d}t + \sum_{j=1}^{M} v_j\sqrt{a_j(\mathbf{X}(t))}\,\mathrm{d}W_j(t). \tag{20}$$

Rather than perturbing the $a_j(\mathbf{X}(t))h$ in the above ansatz we perturb the $\mathbf{X}(t)$ itself. In the case of no delays this leads to a modified form of the chemical Langevin equation

$$\mathrm{d}\mathbf{X} = \sum_{j=1}^{M} v_j a_j(\mathbf{X}(t))\,\mathrm{d}t + \sum_{j=1}^{N} b_j(\mathbf{X}(t))\,\mathrm{d}W_j(t), \tag{21}$$

where the b_j are the columns of $B = \sqrt{vDv^\top}$ and D is a diagonal matrix

$$D = \mathrm{diag}\{a_1(\mathbf{X}(t)), \ldots, a_M(\mathbf{X}(t))\}.$$

A similar approach in a delay setting leads to the SDDE

$$\mathrm{d}\mathbf{X} = \sum_{j=1}^{M_1} v_j a_j(\mathbf{X}(t - T_j))\,\mathrm{d}t + \sum_{j=M_1+1}^{M} v_j a_j(\mathbf{X}(t))\,\mathrm{d}t + \sum_{j=1}^{N} b_j(\mathbf{X}(t))\,\mathrm{d}W_j(t), \tag{22}$$

where the matrix $B = (b_{ik})_{N\times N}$ is the principal square root of the matrix $vD_Av^\top$ satisfying $B(t) = \sqrt{vD_Av^\top}$, and matrix D_A is a diagonal matrix defined by

$$D_A = \mathrm{diag}\{a_1(\mathbf{X}(t - T_1)), \ldots, a_{M_1}(\mathbf{X}(t - T_{M_1})), a_{M_1+1}(\mathbf{X}(t)), \ldots, a_M(\mathbf{X}(t))\}.$$

4. Modelling and numerical implementations

Before discussing implementational aspects an issue that we should address is that of the time delays in GRNs being variable. For example, the time delay in the transcription of the gene *her*1 in zebrafish is known to vary over a relatively wide range of values [15]. Variable time delays have already been considered in both a deterministic setting [18] and a stochastic setting [19]. In the deterministic setting, it has been assumed that a slow reaction with time delay can fire at any time in a specific interval and never fire outside it. The time delay is uniformly distributed and realized deterministically by taking the average of function values in a specific time interval [18]. The following four types of implementations for time delays will be tested in this paper.

(A) The time delay has a fixed value.
(B) The time delay is a uniformly distributed random variable in a specific time window $[T_j - d_j, T_j + d_j]$ with mean T_j. The value of $\mathbf{X}(t - T_j)$ in (19) or (22) is replaced by the average of function values over this specific window.
(C) We use the same uniform distribution as above, but will generate a sample r at each step from the uniform distribution $U(0, 2d_j)$ and the function value of $\mathbf{X}(t - T_j)$ in (19) or (22) is replaced by $\mathbf{X}(t - T_j - d_j + r)$.
(D) The time delay is a Gaussian random variable $N(T_j, \sigma^2)$, where the variance σ^2 is chosen so that all the generated samples should lie in the specific time window $[T_j - d_j, T_j + d_j]$. In this approach, the values of the generated time delay is focused more about T_j, the mean of the time delay.

In deterministic implementations such as (A) and (B), we choose a stepsize h in order that both T_j/h and $(T_j + d_j)/h$ are integers. Then $t_n - T_j$ is a grid point. However, $t_n - T_j$ may be not a grid point if the time delay T_j is a sample generated from random variables in variable implementations such as in (C) or (D). In this case we use interpolation to approximate function values at $t_n - T_j$. Suppose r is a generated time delay and satisfies $t_k < t_n - r < t_{k+1}$, then the function value $y(t_n - r)$ is approximated by

$$y(t_n - r) \approx y(t_k) + \frac{y(t_{k+1}) - y(t_k)}{t_{k+1} - t_k}(t_n - r - t_k).$$

Note that since our methods are low order and the trajectories are noisy there is no need to use higher order interpolational schemes.

Next, we consider numerical methods for solving d-dimensional SDDEs with multiple time delays in the form

$$\begin{aligned} \mathrm{d}\mathbf{X} &= g_0(t, X(t), X(t - T_1), \ldots, X(t - T_m))\,\mathrm{d}t \\ &\quad + \sum_{j=1}^{d} g_j(t, X(t), X(t - T_1), \ldots, X(t - T_m))\,\mathrm{d}W_j(t), \end{aligned} \tag{23}$$

where the $W_j(t)$ are the Wiener process. We choose stepsize h in order that the $n_k = T_k/h$ $(k = 1, \ldots, m)$ are integers—note this puts a restriction on the nature of the T_k. Let

$$N_0 = \max_{1 \leqslant k \leqslant m} \{n_i\}$$

and consider the equidistant time discretization

$$t_h = \{t_n : n = -N_0, -N_0 + 1, \ldots, 0, 1, \ldots, N\}.$$

The class of Euler methods for solving (23) is given by

$$\begin{aligned} X_{n+1} &= X_n + \theta g_0(t_{n+1}, X_{n+1}, X_{n+1-n_1}, \ldots, X_{n+1-n_m})h + (1 - \theta)g_0(t_n, X_n, X_{n-n_1}, \ldots, X_{n-n_m})h \\ &\quad + \sum_{j=1}^{d} g_j(t_n, X_n, X_{n-n_1}, \ldots, X_{n-n_m})\Delta W_{nj}. \end{aligned} \tag{24}$$

This class contains the explicit Euler method if $\theta = 1$ [1,6] and the semi-implicit method if $\theta = 0$ [14]. The convergence order of this class of methods is of strong order 0.5 [1,6,14] and this class can be applied to solve non-stiff SDDEs or

SDDEs with a stiff drift term. However, it is difficult to design and/or to implement high-order methods or fully implicit methods. For example, we must calculate multiple stochastic Itô integrals when using the explicit method with strong order 1.0, even for solving SDDEs with a one-dimensional Wiener process [14]. Thus, in the next section the explicit Euler method is used for solving SDDEs arising from GRNs. Based on a quick preliminary simulation of the stiffness of our problems, we use a stepsize $h = 0.01$ or 0.005 in order to ensure good accuracy and stability properties.

5. Simulation results

This section presents numerical simulations of the SDDE models for two GRNs. The first model describes the dynamics of the transcriptional factor TF-A in mammalian cells [18]. A DDE model was developed for the time delay due to the protein transport from the cytosol to nucleus, given by

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = g(x(t-T)) - k_d x(t) + R_{\mathrm{bas}}, \tag{25}$$

with

$$g(x) = \frac{k_f x^2}{x^2 + K_d},$$

where x is the concentration of the transcriptional factor, $k_d = 0.1/\min$ is the degradation rate, $R_{\mathrm{bas}} = 0.01\,\mu\mathrm{M}/\min$ is the basal transcriptional rate, $K_d = 10\,\mu\mathrm{M}^2$, and a delay $T = 120\,\min$ is the averaged time for the transport of the transcriptional factor. Unlike the bistable systems generated from double-negative feedbacks [21], this is a bistable system generated by a positive feedback. A small synthesis rate $k_f = 0.1\,\mu\mathrm{M}/\min$ is used when $t \in [0, 200]$ and this system only has basal activities. After $t = 200$, this synthesis rate is increased to $k_f = 20\,\mu\mathrm{M}/\min$ and the system jumps to the other steady state but exhibits a "staircase" transition between the two steady states due to the time delay. Similar to the discussion in [11], we use the factors $\Omega = 10$, 50, 100 and 500 for transferring concentration to molecular numbers. Then the reaction rates are $R_{\mathrm{bas}} = 0.01\Omega$ molecule/min, $k_f = 0.1\Omega$ (20Ω) molecule/min and $K_d = 10\Omega^2$. The initial condition is $x(t) = 0$ $(t \in [-120, 0])$. Fig. 1 gives four simulations of the deterministic model based on different values of Ω.

In order to realize the functions of noise in the system behaviours, we use the framework (19) from Section 3 that gives rise to the following SDDE:

$$\mathrm{d}x = (g(x(t-\tau)) - k_d x + R_{\mathrm{bas}})\,\mathrm{d}t + \sqrt{g(x(t-\tau))}\,\mathrm{d}W_1(t) - \sqrt{|k_d x|}\,\mathrm{d}W_2(t). \tag{26}$$

We note that for this system with one single equation, the SDDE based on the approach (22) is given by

$$\mathrm{d}x = (g(x(t-\tau)) - k_d x + R_{\mathrm{bas}})\,\mathrm{d}t + \sqrt{g(x(t-\tau)) + k_d x}\,\mathrm{d}W(t). \tag{27}$$

Fig. 1 also gives stochastic simulations with different values of Ω based on the SDDE (26). These stochastic simulations are based on the same Wiener path in each case. This stochastic system is stabilized at the two steady states with relatively smaller fluctuations than those at the intermediate states where the system is unstable. As with the stochastic simulations in [11], noise has larger impact on the system dynamics if the molecular numbers in the system are smaller. When $\Omega = 10$, fluctuations are large even when the system is in the steady state with high gene expression levels.

We also simulated the two stochastic models (26) and (27) and calculated the mean and variance of the simulations at $t = 600$ where the system is in an unstable intermediate state and at $t = 1000$ where the system is in a steady state based on 10 000 simulations. We found a little difference between the means and variances from these two stochastic models. Hereafter we just use the Langevin approach (19) for testing different implementations of time delays and for simulating the second GRN.

Fig. 2 gives simulations of the stochastic model (26) with different implementations of variable time delays. These simulations are obtained from the same Wiener path and $\Omega = 50$. Fig. 2A gives a stochastic simulation based on the fixed time delay, which is similar to that in Fig. 2B where the time delay at each step is a sample from the Gaussian random variable $N(120, (10/6)^2)$. Based on the uniform distribution, Fig. 2C gives a simulation in which the function value $g(t - 120)$ was replaced by the average of the function values in the time interval $[t - 130, t - 110]$. This simulation is slightly different from that in Fig. 2D in which the time delay at each step is a sample from the uniformly distributed

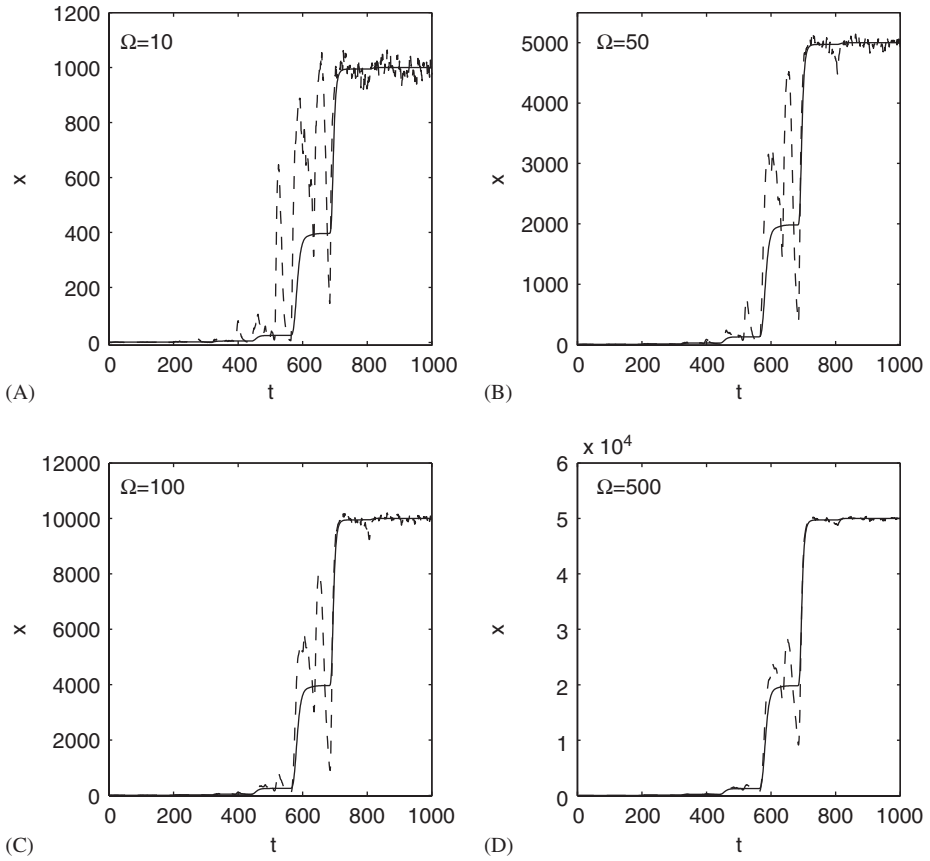

Fig. 1. Simulation of (25) and (26) with time delay based on different values of Ω (solid line: deterministic simulations; dash-line: stochastic simulations based on SDDE (26)).

random variable [110, 130]. Fluctuations in simulations based on the uniformly distributed random variable (Figs. 2C and D) are smaller than those obtained by using either the fixed time delay or the Gaussian random variable time delay (Figs. 2A and B).

The second system that we will consider here describes the oscillation in gene expression and translation. Biological experiments have reported the time delays in the transcription and translation in the *Hes*1 and *Her*1 genes in mouse and zebrafish, respectively, and these delays have been used in deterministic models to realize oscillations in gene expression levels [4,13,15,17]. Here we will consider the deterministic model for the gene *her*1 in zebrafish [15], given by

$$\frac{\mathrm{d}p(t)}{\mathrm{d}t} = am(t - T_p) - bp(t), \tag{28}$$

$$\frac{\mathrm{d}m(t)}{\mathrm{d}t} = g(p(t - T_m)) - cm(t), \tag{29}$$

with

$$g(p) = \frac{k}{1 + p^2/p_0^2},$$

where $p(t)$ and $m(t)$ are the protein number and mRNA molecule number, respectively, $T_m = 12$ and $T_p = 2.8$ are the time delays for transcription and translation, respectively, $a = 4.5$ and $k = 33$ are the maximal synthesis rates, $b = 0.23$ and $c = 0.23$ are degradation rates, and $p_0 = 10$. The autorepression of the gene product to the gene expression

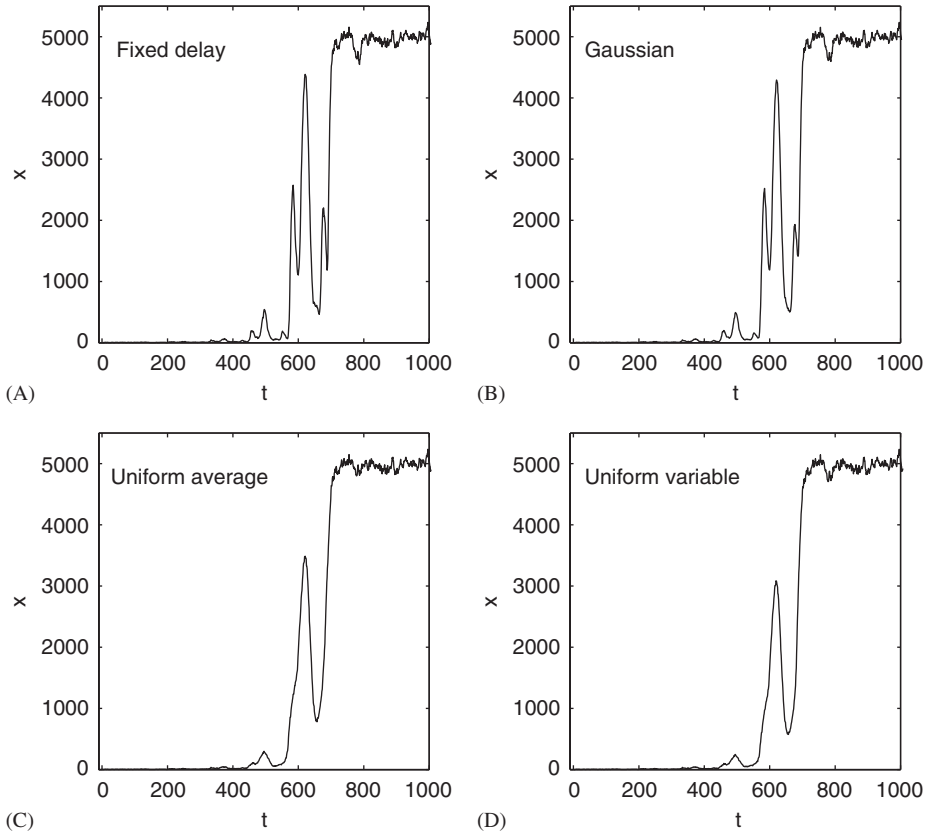

Fig. 2. Simulation of the first stochastic model: (A) fixed time delay with $T = 120$; (B) uniformly distributed time delay that is averaged over a window $t \in [110, 130]$; (C) variable time delay generated by a uniformly distributed random variable $U(110, 130)$; (D) variable time delay generated by the Gaussian random variable $N(120, (10/6)^2)$. Here $x(t) = 0$ ($t \in [-120, 0]$).

together with the time delays will generate oscillations in the numbers of proteins and mRNA molecules. Monk [17] and Bernard et al. [4] showed that sustainable oscillations need the Hill factor to be greater than 4.1

Based on the Langevin approach (19), we can construct the SDDE for the second system, given by

$$\begin{aligned} \mathrm{d}p(t) &= [am(t - T_p) - bp(t)]\,\mathrm{d}t + \sqrt{|am(t - T_p)|}\,\mathrm{d}W_1(t) - \sqrt{|bp(t)|}\,\mathrm{d}W_2(t), \\ \mathrm{d}m(t) &= [g(p(t - T_m)) - cm(t)]\,\mathrm{d}t + \sqrt{g(p(t - T_m))}\,\mathrm{d}W_3(t) - \sqrt{|cm(t)|}\,\mathrm{d}W_4(t). \end{aligned} \tag{30}$$

Both the deterministic and stochastic simulations (with fixed time delay) are presented in Fig. 3 with different synthesis rate a but with all other parameters the same. Deterministic simulations indicate that the oscillations generated from both time delay and autorepression are robust to the variation of the synthesis rate a. The system can still exhibit oscillatory behaviour even when the value of a decreases by 10-fold. However, a 20-fold decrease in this synthesis rate will generate a damped oscillation.

When the stochastic nature in gene expression is taken into account, oscillation still can be observed in the dynamics but with variable amplitudes and periods. Fluctuations will be relatively larger when the value of the synthesis rate a is smaller. Thus, intrinsic noise will have significant impact on the system dynamics when protein numbers in the system are low. A very interesting observation is that the stochastic system can still maintain oscillatory behaviour with a 20-fold decrease in the synthesis rate a, although the deterministic model can only generate a damped oscillation with this very small synthesis rate.

Finally, we note that the simulations are quite quick. One simulation for models 1 and 2 takes $t_1 = 6.9$ and $t_2 = 20.7$ s, respectively, on a Sun workstation with a 500 MHz CPU. Programs are written in MATLAB.

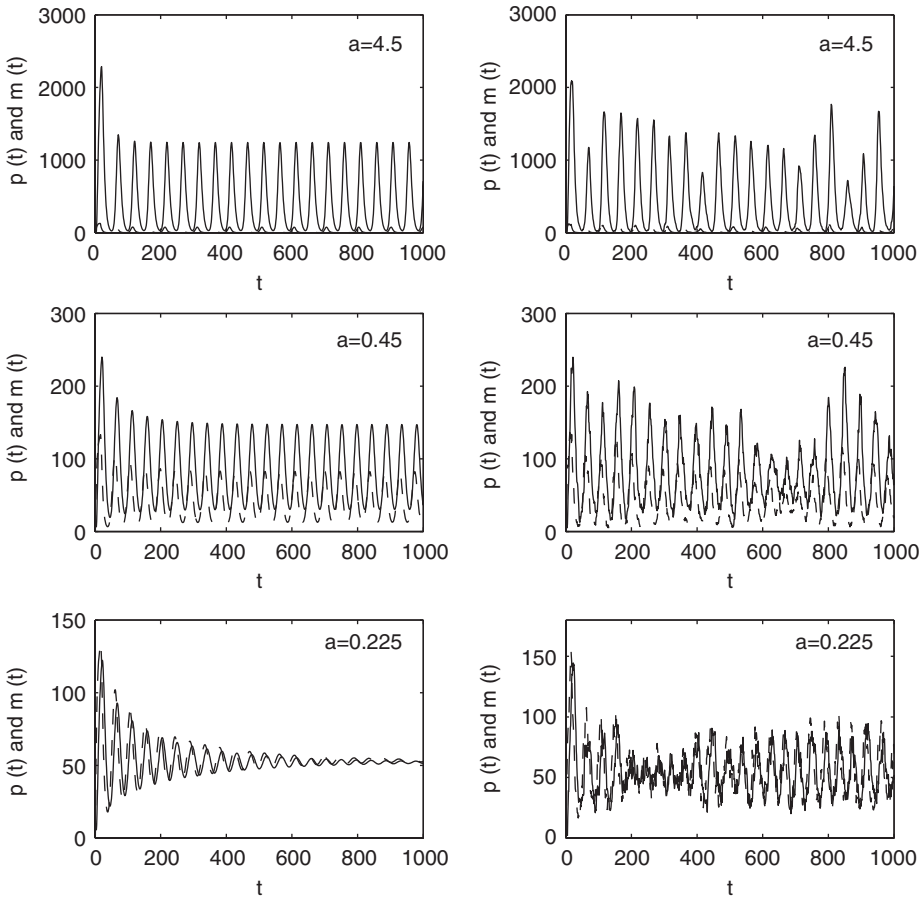

Fig. 3. Simulation of the second system with time delay (solid line: $p(t)$; dash-line: $m(t)$). Here $p(t)=10$ ($t \in [-2.8, 0]$) and $m(t)=1$ ($t \in [-12, 0]$).

6. Conclusion

In this paper we have studied SDDEs for describing the dynamics of GRNs with time delay. Based on the recent progress in the DSSA, we have developed the delay chemical master equation for describing biological reactions. The delay reaction rate equation has been derived based on the assumption of the independence of system states. SDDEs have been introduced from the delay reaction rate equation based on two types of the Langevin approach. In addition, different types of variable time delays have been simulated. Using two GRNs, we have studied the impact of intrinsic noise on the system dynamics. For a bistable network maintained by positive feedback, noise has significant impact on the unstable intermediate state. For an oscillating system generated by time delays, noise can increase the robustness properties of the system to maintain the oscillating expression pattern.

As both fast reactions and slow reactions coexist in biological systems with time delays, fully implicit methods should be developed in order to address the stiffness of such stochastic systems. In addition, the model systems discussed in this paper are very simple and the simulation outputs from two different Langevin approaches are similar. More work is needed to compare the difference between the two types of the Langevin approach. Finally, it would be important to develop stochastic models for GRNs by using SDDEs and to emphasis the importance of time delay in the system evolution. All of these are the topics for future work.

Acknowledgements

One of the authors (K.B.) would like to acknowledge support of the Australian Research Council for the award of a Federation Fellowship.

References

[1] C.T.H. Baker, E. Buckwar, Numerical analysis of explicit one-step methods for stochastic delay differential equations, LMS J. Comput. Math. 3 (2000) 315–335.

[2] C.T.H. Baker, J.M. Ford, N.J. Ford, Bifurcations in approximate solutions of stochastic delay differential equations, Internat. J. Bifurcation Chaos 14 (2004) 2999–3021.

[3] M. Barrio, K. Burrage, A. Leier, T. Tian, Oscillatory regulation of Hes1: discrete stochastic delay modelling and simulation, PLos Comput. Biol., in press.

[4] S. Bernard, B. Čajavec, L. Pujo-Menjouet, M.C. Mackey, H. Herzel, Modelling transcriptional feedback loops: the role of gro/tle1 in hes1 oscillations, Philos. T. Roy. Soc. A 364 (2006) 1155–1170.

[5] G.A. Bocharov, F.A. Rihan, Numerical modelling in bioscience using delay differential equations, J. Comput. Appl. Math. 125 (2000) 183–199.

[6] E. Buckwar, Introduction to the numerical analysis of stochastic delay differential equations, J. Comput. Appl. Math. 125 (2000) 297–307.

[7] E. Buckwar, T. Shardlow, Weak approximation of stochastic differential delay equations, IMA J. Numer. Anal. 25 (2005) 57–86.

[8] K. Burrage, T. Tian, P. Burrage, A multi-scaled approach for simulating chemical reaction systems, Progr. Biophys. Mol. Biol. 85 (2004) 217–234.

[9] D.T. Gillespie, Markov Processes: An Introduction for Physical Scientists, Academic Press, Boston, 1992.

[10] D.T. Gillespie, Approximate accelerated stochastic simulation of chemically reacting systems, J. Chem. Phys. 115 (2001) 1716–1733.

[11] D. Gonze, J. Halloy, A. Goldbeter, Robustness of circadian rhythms with respect to molecular noise, Proc. Nat. Acad. Sci. 99 (2002) 673–678.

[12] E.L. Haseltine, J.B. Rawlings, Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics, J. Chem. Phys. 117 (2002) 6959–6969.

[13] H. Hirata, S. Yoshiura, T. Ohtsuka, Y. Bessho, T. Harada, K. Yoshikawa, R. Kageyama, Oscillatory expression of the bHLH factor Hes1 regulated by a negative feedback loop, Science 298 (2002) 840–843.

[14] U. Kücheler, E. Platen, Strong discrete time approximation of stochastic differential equations with time delay, Math. Comput. Simulation 54 (2000) 189–205.

[15] J. Lewis, Autoinhibition with transcriptional delay: a simple mechanism for the zebrafish somitogenesis oscillator, Curr. Biol. 13 (2003) 1398–1408.

[16] X. Mao, S. Sabanis, Numerical solutions of stochastic differential delay equations under local Lipschitz condition, J. Comput. Appl. Math. 151 (2003) 215–227.

[17] N.A.M. Monk, Oscillatory expression of Hes1, p53, and NF-kappaB driven by transcriptional time delay, Curr. Biol. 33 (2003) 1409–1413.

[18] P. Smolen, D.A. Baxter, J.H. Byrne, Effects of macromolecular transport and stochastic fluctuations on dynamics of genetic regulatory systems, Amer. J. Physiol.—Cell PH 277 (1999) C777–C790.

[19] M. Thattai, A. van Oudenaarden, Intrinsic noise in gene regulatory networks, Proc. Nat. Acad. Sci. 98 (2001) 8614–8619.

[20] T. Tian, K. Burrage, Binomial leap methods for simulating chemical kinetics, J. Chem. Phys. 121 (2004) 10356–10364.

[21] T. Tian, K. Burrage, Bistability and switching in the lysis/lysogeny genetic regulatory network of Bacteriophage lambda, J. Theor. Biol. 227 (2004) 229–237.