Commun. Comput. Phys.
doi: 10.4208/cicp.311009.100310a

Vol. **9**, No. 2, pp. 390-405
February 2011

Stochastic Simulation of the Cell Cycle Model for Budding Yeast

Di Liu*

Department of Mathematics, Michigan State University, East Lansing, MI 48824, USA.

Received 31 October 2009; Accepted (in revised version) 10 March 2010

Communicated by Gang Bao

Available online 27 August 2010

Abstract. We use the recently proposed Nested Stochastic Simulation Algorithm (Nested SSA) to simulate the cell cycle model for budding yeast. The results show that Nested SSA is able to significantly reduce the computational cost while capturing the essential dynamical features of the system.

AMS subject classifications: 65C05, 60G17, 74A25, 92C42

Key words: Stochastic simulation algorithm, stochastic process, bio-chemical reacting network, system biology.

1 Introduction

System biology, which studies integrated cellular reacting networks involving multiple levels of biological activities from gene expression, protein interaction, metabolism to signal transduction, has emerged as a new scientific discipline. Within such functional networks, many types of molecular processes take place on a wide range of time and population scales, under significant influence of random perturbations. From the point of view of modeling, Gene Regulatory Networks (GRNs), unlike protein and metabolic networks, involve fewer number of species and lower population of molecules in a small volume within a cell [1]; therefore stochastic effects have a significant impact on the system and stochastic models are particularly well suited to the study of the functionality of GRNs [2]. The Stochastic Simulation Algorithm (SSA) introduced by Gillespie in [3,4] has been the most successful and promising meso-scale bio-chemical reacting model, as well as an accurate simulation scheme that incorporates stochastic effects. Meanwhile, it is well known that bio-chemical reactions in intracellular networks involving gene expression occur on different time scales, e.g. the fast binding of RNA Polymerase to the DNA

*Corresponding author. *Email address:* richardl@math.msu.edu (D. Liu)

chain versus the relatively slower transcription process, which makes SSAs necessarily inefficient despite its accuracy.

In recent years, the stochastic simulation of intracellular bio-chemical reacting networks with multiple time scales has received a great deal of attention and important progress has been made. The main idea, pursued in different forms by many people, is to capture the effective dynamics on the slow time scale, by assuming the fast processes to be in a quasi-equilibrium distribution [5–11]. In [5], a scheme based on the quasi-equilibrium assumption was proposed supposing that the probability density of the fast species is known exactly as a function of the slow species or can be approximated, e.g. by a Gaussian. The same quasi-equilibrium assumption was used in [6,7], where the effective slow rates are obtained by solving a system of approximate algebraic equations, which are based on extra assumptions on both the reaction rates and the equilibrium distributions of the fast reactions. These limitations are removed in the recent work [8,9], in which stochastic simulation algorithms with nested structures are proposed to deal with the time scale issue. The Nested Stochastic Simulation Algorithm (Nested SSA, or NSSA) proposed in [8,9] relies only on the disparity of the rates, and makes no a priori assumption on the form of the slow and fast variables, nor upon the analytic form of the rate functions. Similar schemes are also proposed in [10,11], with different implementations on sampling the quasi-equilibrium of the fast reactions and time advancing of the slow reactions.

The purpose of the current paper is to test the Nested SSA on the cell cycle model for budding yeast [12]. The cell-division cycle is the sequence of events that take place in a eukaryotic cell leading to its replication. A growing cell replicates all its components and divides them into two daughter cells, so that each daughter has the information and machinery necessary to repeat the process. To account for random fluctuations in the molecular numbers of some major regulatory proteins, it is imperative to incorporate stochastic effects in the dynamics. A stochastic version of the budding yeast cell cycle model has been proposed in the framework of SSA [13,14], which consists of 55 reacting species involved in 82 reactions. Using Nested SSA, we are able to significantly speed up the simulation of the model without losing much accuracy in the key dynamical features of the system, such as the period of the cell cycle. In the following, we will first briefly introduce the Nested SSA and the stochastic cell cycle model for budding yeast. Then we will discuss in detail how NSSA can be applied to improve the efficiency of the stochastic simulation.

2 The nested stochastic simulation algorithm

2.1 Direct SSA

The Stochastic Simulation Algorithm [3,4] describes the time evolution of a spatially homogeneous mixture of chemically reacting molecules contained in a fixed volume V. The solution is assumed to be well mixed and iso-thermal so the details of the diffusion and

transport processes can be neglected and the reaction rates only depend on the populations of reacting molecules. We take N_S species of molecules $S_{i=1,\cdots,N_S}$ involved with M_R reactions $R_{j=1,\cdots,M_R}$, with $x_i \in \mathbb{N}$ being the number of molecules of species S_i. The state of the system is defined by

$$x = (x_1, \cdots, x_{N_S}) \in \mathbb{N}^{N_S}. \tag{2.1}$$

Each reaction R_j is characterized by a rate function $a_j(x)$ and a state change vector $\nu_j \in \mathbb{N}^{N_S}$. We write

$$R_j = (a_j, \nu_j). \tag{2.2}$$

Given state x, the occurrences of the reactions on an infinitesimal time interval dt are independent of each other and the probability of reaction R_j during this time interval is given up to the first order by $a_j(x)dt$. The state of the system after reaction R_j is $x+\nu_j$. The time evolution of the probability distribution of the system $P(x,t)$ is governed by the forward Kolmogorov equation:

$$\frac{\partial P(x,t)}{\partial t} = \sum_j \Big(a_j(x-\nu_j)P(x-\nu_j,t) - a_j(x)P(x,t) \Big). \tag{2.3}$$

The SSA (see also [15] for similar schemes) constructs numerical realizations of the time evolution of the state vector x_t, i.e. simulated trajectories x_t advancing with time t in the state space. To describe the method, we assume that the current time is $t = t_n$, and the state of the system is at $x = x_n$. One version of SSA called the Direct Method performs the following steps:

1. Let $a_0(x) = \sum_j a_j(x)$. Generate independent random numbers r_1 and r_2 with uniform distribution on the unit interval $[0,1]$. Let

$$\delta t_{n+1} = \frac{1}{a_0(x)} \ln\left(\frac{1}{r_1}\right), \tag{2.4}$$

and k_{n+1} be the positive integer such that

$$\sum_{j=1}^{k_{n+1}-1} a_j(x) < r_2 a_0(x) \leq \sum_{j=1}^{k_{n+1}} a_j(x). \tag{2.5}$$

2. Update time and state of the system by

$$t_{n+1} = t_n + \delta t_{n+1}, \qquad x_{n+1} = x_n + \nu_{k_{n+1}}. \tag{2.6}$$

Goto 1. unless certain stopping criterion is met.

A slightly different implementation known as the First Reaction Method, was also introduced in [3, 4]. In both versions, the SSA skips time intervals on which there is no reaction event, going directly to the occurrence of the next reaction. Because more random numbers are generated at each time step, the First Reaction Method is usually less efficient than the Direct Method. This was improved in [16] by reusing the random times for reactions that are not affected by the chosen reaction event at each time step. The SSA is exact in the sense that the process generated by SSA has the same probability distribution as the chemical reacting network being simulated.

2.2 Nested SSA

To present the Nested SSA, we focus on the case when there are only two times scales. The general cases is treated in [8, 9]. We assume that the rates $\{a_j(x)\}$'s can be divided into two groups: One group corresponding to the fast processes with rates of order $1/\epsilon$ and the other group corresponding to the slow processes with rates of order 1, with $\epsilon \ll 1$:

$$a(x) = \big(a^s(x), a^f(x)\big), \tag{2.7}$$

where

$$\begin{aligned} a^s(x) &= \big(a_1^s(x), \cdots, a_{M_s}^s(x)\big) = \mathcal{O}(1), \\ a^f(x) &= \big(a_1^f(x), \cdots, a_{M_f}^f(x)\big) = \mathcal{O}\Big(\frac{1}{\epsilon}\Big), \end{aligned} \tag{2.8}$$

in dimensionless units. The corresponding reactions and the associated state change vectors can be grouped accordingly:

$$R^s = (a^s, \nu^s), \qquad R^f = (a^f, \nu^f). \tag{2.9}$$

The Nested SSA consists of two SSAs organized in a hierarchical fashion: an outer SSA on the slow processes only, which uses modified slow rates, and an inner SSA on the fast processes only, which uses the original fast rates and serves to give the modified slow rates. Let $t = t_n$, $x = x_n$ be the current time and state of the system. The steps of the Nested SSA are the following:

1. **Inner SSA**
   Run N independent replicas of SSA with the fast reactions $R^f = (a^f, \nu^f)$ only, for a time interval of T_f. During this calculation, compute the modified slow rates: For $j = 1, \cdots, M_s$, these are

   $$\tilde{a}_j^s = \frac{1}{N} \sum_{k=1}^{N} \frac{1}{T_f} \int_{T_0}^{T_0+T_f} a_j^s(x_k(\tau)) d\tau, \tag{2.10}$$

   where $x_k(\tau)$ is the k-th replica of the auxiliary fast processes at virtual time τ whose initial value is $x_k(0) = x_n$. T_0 is a parameter we choose in order to minimize the effect of the transients in the auxiliary fast processes.

2. **Outer SSA**
   Run one step of SSA for the modified slow reactions

   $$\tilde{R}^s = (\tilde{a}^s, \nu^s), \tag{2.11}$$

   to generate (t_{n+1}, x_{n+1}) from (t_n, x_n). Then goto 1. until a certain stopping criterion is met.

The justification of the Nested SSA is the following. The slow-fast chemical reacting network can be viewed as a singular perturbation problem [17–19]. It is can be proved [9] that the effective dynamics on the slow time scale can be given, up to order $\mathcal{O}(\epsilon)$, by:

$$\bar{R} = (\bar{a}(y), \nu^s), \tag{2.12}$$

where y is the effective slow variables defined to be linearly independent functions conserved in the fast reactions and

$$\bar{a}_j(y) = \langle a_j(x) \rangle_y \equiv \sum_{x \in X} a_j(x) \mu_y(x), \tag{2.13}$$

with $\mu_y(x)$ being the quasi-equilibrium distribution of the fast reactions. Moreover, a rigorous and optimal error estimate for the Nested SSA, as well as generalizations to dynamically partitioning of slow-fast reaction sets and to simulating systems with more than 2 time scales, can be found in [9]. As applications, the virus infection model [20] and the Heat Shock Response of E. Coli. [21] were studied. The same scheme has also been applied to stochastic differential equations with multiple time scales [22], where a rigorous error analysis is also provided.

3 The cell cycle model of budding yeast

Cell cycle is the essential mechanism by which all living things reproduce themselves [23]. It consists of the succession of events whereby one cell grows and divides into two daughter cells that each contains the information and machinery necessary to repeat the process. Between one cell division and the next, all essential components of the cell must be duplicated. The most important component is the genetic material (DNA molecules present in chromosomes), which must be accurately replicated and the two copies carefully segregated to the two daughter cells. The eukaryotic cell cycle consists of 4 phases. The 2 most dramatic events that constitutes the M phase are called mitosis and cytokinesis, in which the nucleus divides and the cell splits into two. During the S (synthesis) phase, the cell replicates its nuclear DNA. The G_1 is the interval between the completion of M phase and the beginning of S phase, while the G_2 phase is the interval between the end of S phase and the beginning of M phase. During the G_1-S-G_2 phases, the cell continues to transcribe genes, synthesize proteins and grow in mass. 2 checkpoints are imposed before the cell enters the S and M phases to make sure that there is no damage to the DNA and its replication.

A lot of knowledge has been accumulated on the molecular mechanism of eukaryotic cell cycle control for budding yeast, which makes Saccharomyces cerevisiae, the unicellular budding yeast, an excellent example to study cell cycle regulation. Molecular biologists have dissected and characterized individual cell cycle components and their interactions that regulate the cell cycle. Read from bottom left toward top right, Fig. 1 shows a consensus picture of the regulatory network of the budding yeast [12]. More information on the model can be found at `http://jigcell.biol.vt.edu`. The original cell cycle model was only deterministic in the form of a system of ODEs. Although being able to reproduce the time scales of the dynamical interactions between different modules of the network, it does not take into account the population scales of the reacting species, which is influenced significantly by random fluctuations. To understand the impact of

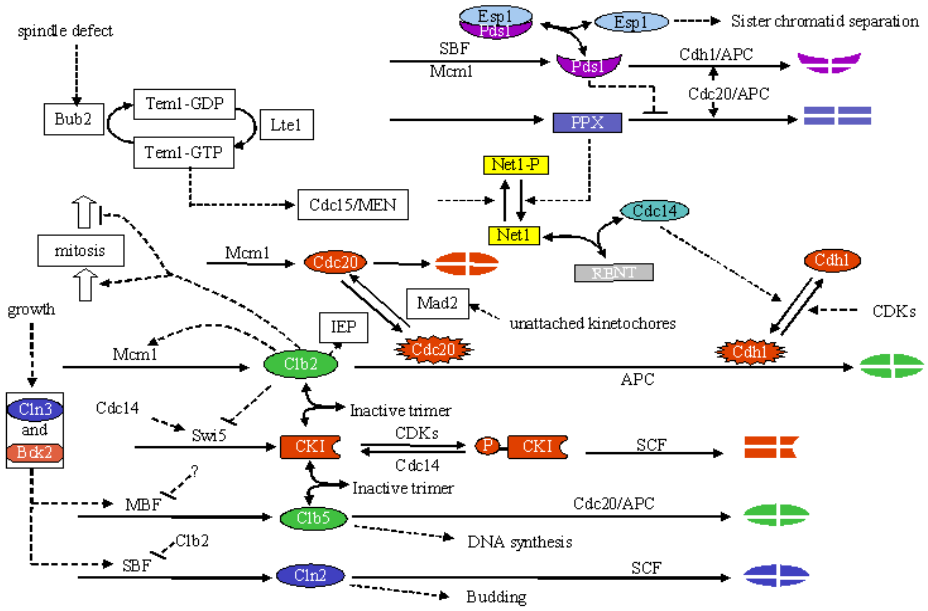

Figure 1: Reaction diagram for the cell cycle regulation of budding yeast.

stochastic effects on the dynamics, a stochastic version of the model has been recently proposed in [13, 14] in the framework of SSA using experimentally collected data on the populations of the reacting species [24] (also available at `http://www.yeastgenome.org`). In Table 1, we list the total 82 reaction channels and the corresponding reaction rates of the model. We use *struct* to represent the structural proteins making up the reacting species. We also denote it by *degraded* when a reacting species is degraded or dissociated. There are also 25 algebraic relations between the reaction species populations and the reaction rates that are listed in Table 1. The parameters in the reaction rates and the initial conditions for the simulations are given in Table 2.

In Table 1, $Swi5T$ denotes the total population of Swi5, while $Swi5$ denotes the active population of Swi5. Reaction a_{43} implies that they share the same decay rate. At each reaction event of the degradation of Swi5, a random number is generated according to the proportion of active Swi5 among total Swi5 to determine whether an active Swi5 should be degraded. The same notation is also used for the reacting species $Cdc20$, $Cdh1$, $Cdc14$, $Net1$ in reaction a_{50}, a_{53}, a_{61}, a_{69}, respectively.

Besides the reacting species that are in discrete quantities, there are also continuous variables in the system. The mass of the cell is supposed to grow exponentially until the cell division:

$$\frac{d}{dt}mass = k_g \cdot mass. \tag{3.1}$$

Notice that *mass* enters the dynamical system as a multiplier of the rates of synthesis of cyclins Cln2, $Clb2$ and $Clb5$. At division of the cell, the *mass* is divided between the

Table 1: The cell cycle model of the budding yeast.

I. Reaction channels and reaction rates

$$\text{Cln2} \xrightleftharpoons[a_2=(k'_{s,n2}+k''_{s,n2}\cdot SBF)\cdot mass]{a_1=k_{d,n2}\cdot Cln2} \textit{struct/degraded} \qquad (*)$$

$$\text{Clb5} \xrightleftharpoons[a_4=(k'_{s,b5}+k''_{s,b5}\cdot MBF)\cdot mass]{a_3=V_{d,b5}\cdot Clb5} \textit{struct/degraded}$$

$$\text{C5P} \xrightarrow{a_5=k_{d3,c1}\cdot C5P} \text{Clb5}$$

$$\text{Clb5 + Sic1} \xrightleftharpoons[a_7=k_{di,b5}\cdot \text{C5}]{a_6=k_{as,b5}\cdot Sic1\cdot Clb5} \text{C5}$$

$$\text{F5P} \xrightarrow{a_8=k_{d3,f6}\cdot F5P} \text{Clb5}$$

$$\text{Clb5 + Cdc6} \xrightleftharpoons[a_{10}=k_{di,f5}\cdot F5]{a_9=k_{as,f5}\cdot Cdc6\cdot Clb5} \text{F5}$$

$$\text{Clb2} \xrightleftharpoons[a_{12}=(k'_{s,b2}+k''_{s,b2}\cdot Mcm1)\cdot mass]{a_{11}=V_{d,b2}\cdot Clb2} \textit{struct/degraded}$$

$$\text{C2P} \xrightarrow{a_{13}=k_{d3,c1}\cdot C2P} \text{Clb2}$$

$$\text{Clb2 + Sic1} \xrightleftharpoons[a_{15}=k_{di,b2}\cdot \text{C2}]{a_{14}=k_{as,b2}\cdot Sic1\cdot Clb2} \text{C2}$$

$$\text{F2P} \xrightarrow{a_{16}=k_{d3,f6}\cdot F2P} \text{Clb2}$$

$$\text{Clb2 + Cdc6} \xrightleftharpoons[a_{18}=k_{di,f2}\cdot F2]{a_{17}=k_{as,f2}\cdot Cdc6\cdot Clb2} \text{F2}$$

$$\text{Sic1} \xleftarrow{a_{19}=k'_{s,c1}+k''_{s,c1}\cdot Swi5} \textit{struct/degraded}$$

$$\text{C2} \xrightarrow{a_{20}=V_{d,b2}\cdot \text{C2}} \text{Sic1}$$

$$\text{C5} \xrightarrow{a_{21}=V_{d,b5}\cdot \text{C5}} \text{Sic1}$$

$$\text{Sic1} \xrightleftharpoons[a_{23}=k_{pp,c1}\cdot Cdc14\cdot Sic1P]{a_{22}=V_{kp,c1}\cdot Sic1} \text{Sic1P}$$

$$\text{Sic1P} \xrightarrow{a_{24}=k_{d3,c1}\cdot Sic1P} \textit{degraded}$$

$$\text{C2P} \xrightarrow{a_{25}=V_{d,b2}\cdot C2P} \text{Sic1P}$$

$$\text{C5P} \xrightarrow{a_{26}=V_{d,b5}\cdot C5P} \text{Sic1P}$$

Continued on next page

Table 1 – continued from previous page

$$\text{C2} \underset{a_{28}=k_{pp,c1}\cdot Cdc14\cdot C2P}{\overset{a_{27}=PV_{kp,c1}\cdot C2}{\rightleftharpoons}} \text{C2P}$$

$$\text{C5} \underset{a_{30}=k_{pp,c1}\cdot Cdc14\cdot C5P}{\overset{a_{29}=PV_{kp,c1}\cdot C5}{\rightleftharpoons}} \text{C5P}$$

$$\text{Cdc6} \xleftarrow{a_{31}=k'_{s,f6}+k''_{s,f6}\cdot Swi5+k'''_{s,f6}\cdot SBF} \textit{struct}$$

$$\text{F2} \xrightarrow{a_{32}=V_{d,b2}\cdot F2} \text{Cdc6}$$

$$\text{F5} \xrightarrow{a_{33}=V_{d,b5}\cdot F5} \text{Cdc6}$$

$$\text{Cdc6} \underset{a_{35}=k_{pp,f6}\cdot Cdc14\cdot Cdc6P}{\overset{a_{34}=V_{kp,f6}\cdot Cdc6}{\rightleftharpoons}} \text{Cdc6P}$$

$$\text{Cdc6P} \xrightarrow{a_{36}=k_{d3,f6}\cdot Cdc6P} \textit{degraded}$$

$$\text{F2P} \xrightarrow{a_{37}=V_{d,b2}\cdot F2P} \text{Cdc6}$$

$$\text{F5P} \xrightarrow{a_{38}=V_{d,b5}\cdot F5P} \text{Cdc6}$$

$$\text{F2} \underset{a_{40}=k_{pp,f6}\cdot Cdc14\cdot F2P}{\overset{a_{39}=V_{kp,f6}\cdot F2}{\rightleftharpoons}} \text{F2P}$$

$$\text{F5} \underset{a_{42}=k_{pp,f6}\cdot Cdc14\cdot F5P}{\overset{a_{41}=V_{kp,f6}\cdot F5}{\rightleftharpoons}} \text{F5P}$$

$$\text{Swi5T}\backslash\text{Swi5} \underset{a_{44}=k'_{s,swi}+k''_{s,swi}\cdot Mcm1}{\overset{a_{43}=k_{d,swi}\cdot(Swi5T\backslash Swi5)}{\rightleftharpoons}} \textit{struct/degraded}$$

$$\text{Swi5} \underset{a_{46}=k_{a,swi}\cdot Cdc14\cdot(Swi5T-Swi5)}{\overset{a_{45}=k_{i,swi}\cdot Clb2\cdot Swi5}{\rightleftharpoons}} \textit{struct/degraded}$$

$$\text{APC\_P} \underset{a_{48}=k_{a,apc}\cdot Clb2\cdot(1.15e3-APC\_P)/(J_{a,apc}+1.15e3-APC\_P)}{\overset{a_{47}=k_{i,apc}\cdot APC\_P/(J_{i,apc}+APC\_P)}{\rightleftharpoons}} \textit{struct/degraded}$$

$$\text{Cdc20T} \xleftarrow{a_{49}=k'_{s,20}+k''_{s,20}\cdot Mcm1} \textit{struct}$$

$$\text{Cdc20T}\backslash\text{Cdc20A} \xrightarrow{a_{50}=k_d20\cdot Cdc20T\backslash Cdc20A} \textit{degraded}$$

$$\text{Cdc20A} \underset{a_{52}=(k'_{a,20}+k''_{a,20}\cdot APC\_P)\cdot(Cdc20T-Cdc20A)}{\overset{a_{51}=k_{mad2}\cdot Cdc20A}{\rightleftharpoons}} \textit{struct/degraded}$$

$$\text{Cdh1T}\backslash\text{Cdh1} \underset{a_{54}=k_{s,cdh}}{\overset{a_{53}=k_{d,cdh}\cdot(Cdh1T\backslash Cdh1)}{\rightleftharpoons}} \textit{struct/degraded}$$

Continued on next page

Table 1 – continued from previous page

$$\text{Cdh1} \xrightleftharpoons[a_{56}=V_{a,cdh}\cdot(Cdh1T-Cdh1)/(J_{a,cdh}+Cdh1T-Cdh1)]{a_{55}=V_{i,cdh}\cdot Cdh1/(J_{i,cdh}+Cdh1)} \textit{struct/degraded}$$

$$\text{Tem1} \xrightleftharpoons[a_{58}=k_{lte1}\cdot(Tem1T-Tem1)/(J_{a,tem}+Tem1T-Tem1)]{a_{57}=k_{bub2}\cdot Tem1/(J_{i,tem}+Tem1)} \textit{struct/degraded}$$

$$\text{Cdc15} \xrightleftharpoons[a_{60}=(k'_{a,15}\cdot(Tem1T-Tem1)+k''_{a,15}\cdot Tem1+k'''_{a,15}\cdot Cdc14)\cdot(Cdc15T-Cdc15)]{a_{59}=k_{i,15}\cdot Cdc15} \textit{struct/degraded}$$

$$\text{Cdc14T}\backslash\text{Cdc14} \xrightleftharpoons[a_{62}=k_{s,14}]{a_{61}=k_{d,14}\cdot(Cdc14T\backslash Cdc14)} \textit{struct/degraded} \quad (**)$$

$$\text{RENT} \xrightarrow{a_{63}=k_{d,net}\cdot RENT} \text{Cdc14} \quad (**)$$

$$\text{Cdc14} \xrightleftharpoons[a_{65}=k_{d,net}\cdot RENTP,\quad a_{66}=k_{di,rentp}\cdot RENTP]{a_{64}=k_{as,rentp}\cdot Net1P\cdot Cdc14} \textit{struct/degraded} \quad (**)$$

$$\text{Cdc14 + Net1} \xrightleftharpoons[a_{68}=k_{di,rent}\cdot RENT]{a_{67}=k_{as,rentp}\cdot Net1\cdot Cdc14} \text{RENT} \quad (**)$$

$$\text{Net1T}\backslash\text{Net1} \xrightleftharpoons[a_{70}=k_{s,net}]{a_{69}=k_{d,net}\cdot(Net1T\backslash Net1)} \textit{struct/degraded} \quad (**)$$

$$\text{RENT} \xrightarrow{a_{71}=k_{d,14}\cdot RENT} \text{Net1} \quad (**)$$

$$\text{Net1} \xrightleftharpoons[a_{73}=V_{pp,net}\cdot Net1P]{a_{72}=V_{kp,net}\cdot Net1} \textit{struct/degraded} \quad (**)$$

$$\text{RENT} \xrightleftharpoons[a_{75}=V_{pp,net}\cdot RENTP]{a_{74}=V_{kp,net}\cdot RENT} \textit{struct/degraded} \quad (**)$$

$$\text{PPX} \xrightleftharpoons[a_{77}=k_{s,ppx}]{a_{76}=V_{d,ppx}\cdot PPX} \textit{struct/degraded}$$

$$\text{Pds1} \xrightleftharpoons[a_{79}=k'_{s,pds}+k''_{s1,pds}\cdot SBF+k''_{s2,pds}\cdot Mcm1]{a_{78}=V_{d,pds}\cdot Pds1} \textit{struct/degraded}$$

$$\text{Pds1 + Esp1} \xrightleftharpoons[a_{81}=k_{di,esp}\cdot PE]{a_{80}=k_{as,esp}\cdot Esp1\cdot Pds1} \textit{struct/degraded}$$

$$\text{Esp1} \xleftarrow{a_{82}=V_{d,pds}\cdot PE} \textit{degraded}$$

II. Algebraic relations

$$GK(V_a,V_i,J_a,J_i)=\frac{2.\cdot J_i\cdot V_a}{V_i-V_a+J_a\cdot V_i+J_i\cdot V_a+\sqrt{(V_i-V_a+J_a\cdot V_i+J_i\cdot V_a)^2-4.\cdot(V_i-V_a)\cdot J_i\cdot V_a}}$$

Continued on next page

Table 1 – continued from previous page

$$V_{d,b5}=k'_{d,b5}+kp''_{d,b5}\cdot Cdc20A$$

$$V_{d,b2}=k'_{d,b2}+k''_{d,b2}\cdot Cdh1+k_{d,b2p}\cdot Cdc20A$$

$$Cln3=[C_0\cdot D_{n3}\cdot mass/(J_{n3}+D_{n3}\cdot mass)]$$

$$Bck2=[B_0\cdot mass]$$

$$V_{a,sbf}=k_{a,sbf}\cdot(\epsilon_{sbf,n2}\cdot Cln2+\epsilon_{sbf,n3a}\cdot Cln3+\epsilon_{sbf,n3b}\cdot Bck2+\epsilon_{sbf,b5}\cdot Clb5)$$

$$V_{i,sbf}=k'_{i,sbf}+k''_{i,sbf}\cdot Clb2$$

$$SBF=long(4.55e4\cdot GK(V_{a,sbf},V_{i,sbf},J_{a,sbf},J_{i,sbf}))$$

$$MBF=SBF$$

$$Mcm1=[2.1e4\cdot GK(k_{a,mcm}\cdot Clb2,k_{i,mcm},J_{a,mcm},J_{i,mcm})]$$

$$Clb5T=Clb5+C5+C5P+F5+F5P$$

$$Clb2T=Clb2+C2+C2P+F2+F2P$$

$$Sic1T=Sic1+Sic1P+C2+C2P+C5+C5P$$

$$Cdc6T=Cdc6+Cdc6P+F2+F2P+F5+F5P$$

$$V_{kp,c1}=k_{d1,c1}+k_{d2,c1}\cdot(\epsilon_{c1,n3}\cdot Cln3+\epsilon_{c1,k2}\cdot Bck2+\epsilon_{c1,n2}\cdot Cln2+\epsilon_{c1,b5}\cdot Clb5+\epsilon_{c1,b2}\cdot Clb2)/(J_{d2,c1}+Sic1T)$$

$$V_{kp,f6}=k_{d1,f6}+k_{d2,f6}\cdot(\epsilon_{f6,n3}\cdot Cln3+\epsilon_{f6,k2}\cdot Bck2+\epsilon_{f6,n2}\cdot Cln2+\epsilon_{f6,b5}\cdot Clb5+\epsilon_{f6,b2}\cdot Clb2)/(J_{d2,f6}+Cdc6T)$$

$$CKIT=Sic1T+Cdc6T$$

$$RENTP=Cdc14T-RENT-Cdc14$$

$$Net1P=Net1T-Net1-Cdc14T+Cdc14$$

$$PE=Esp1T-Esp1$$

$$V_{a,cdh}=k'_{a,cdh}+k''_{a,cdh}\cdot Cdc14$$

$$V_{i,cdh}=k'_{i,cdh}+k''_{i,cdh}\cdot(\epsilon_{cdh,n3}\cdot Cln3+\epsilon_{cdh,n2}\cdot Cln2+\epsilon_{cdh,b2}\cdot Clb2+\epsilon_{cdh,b5}\cdot Clb5)$$

$$V_{pp,net}=k'_{pp,net}+k''_{pp,net}\cdot PPX$$

$$V_{kp,net}=(k'_{kp,net}+k''_{kp,net}\cdot Cdc15)\cdot mass$$

$$V_{d,ppx}=k'_{d,ppx}+k''_{d,ppx}\cdot(J_{20,ppx}+Cdc20A)\cdot J_{pds}/(J_{pds}+Pds1)$$

$$V_{d,pds}=k'_{d1,pds}+k''_{d2,pds}\cdot Cdc20A+k''_{d3,pds}\cdot Cdh1$$

Table 2: Parameters and initial conditions.

A. Parameters for the cell cycle model

$k_g=0.007702$	$k'_{s,n2}=0.$	$k''_{s,n2}=4.12e-3$	$k_{d,n2}=0.12$	$k'_{s,b5}=1.8$	$k''_{s,b5}=2.47e-4$
$k'_{d,b5}=0.01$	$k''_{d,b5}=1.4e-3$	$k'_{s,b2}=2.25$	$k''_{s,b2}=4.29e-3$	$k'_{d,b2}=0.003$	$k''_{d,b2}=0.004$
$k_{d,b2p}=1.31e-3$	$k'_{s,c1}=27.$	$k''_{s,c1}=0.191$	$k_{d1,c1}=0.01$	$k_{d2,c1}=1.$	$k_{d3,c1}=1.$
$k_{pp,c1}=1.65e-3$	$k'_{s,f6}=54.$	$k''_{s,f6}=0.191$	$k'''_{s,f6}=1.98e-4$	$k_{d1,f6}=0.01$	$k_{d2,f6}=1.$
$k_{d3,f6}=1.$	$k_{pp,f6}=1.65e-3$	$k_{as,b5}=0.0222$	$k_{di,b5}=0.06$	$k_{as,f5}=4.44e-6$	$k_{di,f5}=0.01$
$k_{as,b2}=0.0222$	$k_{di,b2}=0.05$	$k_{as,f2}=6.67e-3$	$k_{di,f2}=0.5$	$k'_{s,swi}=7.05$	$k''_{s,swi}=5.37e-3$
$k_{d,swi}=0.08$	$k_{a,swi}=8.26e-4$	$k_{i,swi}=2.22e-5$	$k_{a,apc}=5.11e-2$	$k_{i,apc}=1.72e2$	$k'_{s,20}=0.687$
$k''_{s,20}=3.27e-3$	$k_{d,20}=0.3$	$k'_{a,20}=.05$	$k''_{a,20}=1.74e-4$	$k_{s,cdh}=1.$	$k_{d,cdh}=0.01$
$k'_{a,cdh}=1.$	$k''_{a,cdh}=0.0331$	$k'_{i,cdh}=0.1$	$k''_{i,cdh}=8.$	$k_{s,14}=484$	$k_{d,14}=0.1$
$k_{s,net}=203$	$k_{d,net}=0.03$	$k'_{a,15}=3.5e-6$	$k''_{a,15}=1.75e-3$	$k'3_{a,15}=4.13e-7$	$k_{i,15}=0.5$
$k'_{pp,net}=0.05$	$k''_{pp,net}=3.e-4$	$k'_{kp,net}=0.01$	$k''_{kp,net}=2.52e-3$	$k_{asrent}=8.26e-2$	$k_{as,rentp}=4.13e-4$
$k_{di,rent}=1.$	$k_{di,rentp}=2.$	$k_{s,ppx}=100.$	$k'_{d,ppx}=0.17$	$k''_{d,ppx}=1.75e-2$	$k'_{s,pds}=0.$
$k''_{s1,pds}=6.6e-5$	$k''_{s2,pds}=2.62e-4$	$k'_{d1,pds}=0.01$	$k''_{d2,pds}=1.75e-3$	$k''_{d3,pds}=4.e-4$	$k_{as,esp}=0.5$
$k_{di,esp}=0.5$	$k_{s,ori}=2.$	$k_{d,ori}=0.06$	$k_{s,bud}=0.2$	$k_{d,bud}=0.06$	$k_{s,spn}=0.1$
$k_{d,spn}=0.06$	$k_{a,sbf}=0.38$	$k'_{i,sbf}=0.6$	$k''_{i,sbf}=3.56e-3$	$k_{a,mcm}=4.45e-5$	$k_{i,mcm}=0.15$

$k_{mad2}=8.$ (for $ORI>1.$ and $SPN<1.$) or =.1 (otherwise).

$k_{bub2}=573.$ (for $ORI>1.$ and $SPN<1.$) or =114.6 (otherwise).

$k_{lte1}=573.$ (for $SPN>1.$ and $Clb2<K_{ez}$) or =57.3 (otherwise).

$\epsilon_{sbf,n2}=1.6e-3$	$\epsilon_{sbf,n3a}=0.0102$	$\epsilon_{sbf,n3b}=0.0125$	$\epsilon_{sbf,b5}=8.89e-4$	$\epsilon_{c1,n3}=0.689$	$\epsilon_{c1,n2}=0.108$
$\epsilon_{c1,k2}=0.0847$	$\epsilon_{c1,b5}=0.1$	$\epsilon_{c1,b2}=0.45$	$\epsilon_{f6,n3}=0.689$	$\epsilon_{f6,n2}=0.108$	$\epsilon_{f6,k2}=0.0844$
$\epsilon_{f6,b5}=0.1$	$\epsilon_{f6,b2}=0.55$	$\epsilon_{cdh,n3}=2.55e-4$	$\epsilon_{cdh,n2}=3.2e-4$	$\epsilon_{cdh,b5}=3.56e-3$	$\epsilon_{cdh,b2}=5.33e-4$
$\epsilon_{ori,b5}=4.e-4$	$\epsilon_{ori,b2}=2.e-4$	$\epsilon_{bud,n3}=5.1e-5$	$\epsilon_{bud,n2}=2.e-4$	$\epsilon_{bud,b5}=4.44e-4$	
$C_0=392.$	$D_{n3}=1.$	$B_0=43.2$	$Tem1T=573$	$Cdc15T=238$	$Esp1T=100$
$J_{d2,c1}=112.5$	$J_{d2,f6}=112.5$	$J_{a,apc}=115.$	$J_{i,apc}=115.$	$J_{a,cdh}=3.$	$J_{i,cdh}=3.$
$J_{a,tem}=57.3$	$J_{i,tem}=57.3$	$J_{a,sbf}=0.01$	$J_{i,sbf}=0.01$	$J_{a,mcm}=0.1$	$J_{i,mcm}=0.1$
$J_{spn}=315.$	$J_{n3}=6.$	$J_{20,ppx}=17.2$	$J_{pds}=4.$	$K_{ez}=2700.$	$K_{ez2}=1800.$

$f=exp(-1.026+32.\cdot k_g)$

B. Initial conditions for the daughter cell

$mass=1.2060$	$F5=1$	$Cdc14=1133$	$Cln2=82$	$F2P=62$	$Net1T=6776$
$Clb5=117$	$F5P=1$	$Net1=45$	$Clb2=331$	$Swi5T=1377$	$RENT=2540$
$Sic1=52$	$Swi5=1348$	$PPX=1232$	$Sic1P=14$	$APC_P=117$	$Pds1=3$
$C2=536$	$Cdc20T=220$	$Esp1=30$	$C5=158$	$Cdc20A=51$	$C2P=54$
$Cdh1T=100$	$C5P=16$	$Cdh1=93$	$Cdc6=242$	$Tem1=518$	$Cdc6P=35$
$Cdc15=156$	$F2=531$	$Cdc14T=4840$			

mother cell and the daughter cell such that $(1-f)\cdot mass$ is given to the new born daughter and $f\cdot mass$ remains with the mother cell.

Governed by the following equations, flag variables are also introduced to continuously monitor and regulate the progress of critical events of the cell cycle:

$$\begin{aligned}
\frac{dBUD}{dt} &= k_{s,bud}\cdot(\epsilon_{bud,n2}\cdot \mathrm{Cln2}+\epsilon_{bud,n3}\cdot \mathrm{Cln3}+\epsilon_{bud,b5}\cdot Clb5)-k_{d,bud}\cdot BUD,\\
\frac{dORI}{dt} &= k_{s,ori}\cdot(\epsilon_{ori,b5}\cdot Clb5+\epsilon_{ori,b2}\cdot Clb2)-k_{d,ori}\cdot ORI,\\
\frac{dSPN}{dt} &= k_{s,spn}\cdot Clb2/(J_{spn}+Clb2)-k_{d,spn}\cdot SPN.
\end{aligned} \tag{3.2}$$

BUD represents proteins that are phosphorylated and subsequently initiate a new bud for the new daughter cell when the phosphorylation state reaches a threshold, BUD=1. In a similar manner, $ORI=1$ signals the onset of DNA synthesis. ORI is reset to zero only if $Clb2+Clb5$ drops below another threshold K_{ez2}. The checkpoint is lifted when SPN=1, which represents alignment of all chromosomes on the metaphase plate. When $Clb2$ drops below K_{ez}, we reset BUD and SPN to zero and the division of the cell is initiated.

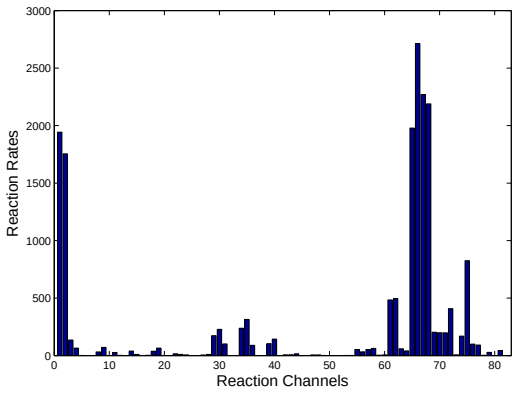

Figure 2: Reaction rates for the cell cycle model.

4 Simulation of cell cycle model with NSSA

We used both Direct SSA and Nested SSA to simulate the cell cycle model. The time scale separation can be illustrated by Fig. 2, which gives the magnitude of the reaction rates at time $t=2000$. It can be seen that the time scale separation is of the order of $\mathcal{O}(10^2)$, with the fastest rates to be of order $\mathcal{O}(10^3)$, while most of other rates are of $\mathcal{O}(10)$. The corresponding fast reactions are marked with $(*)$ and $(**)$ in Table 1. Notice that the fast reactions belong to the modules that synthesis and degrade $\mathrm{Cln2}$ and $Cdc14$,

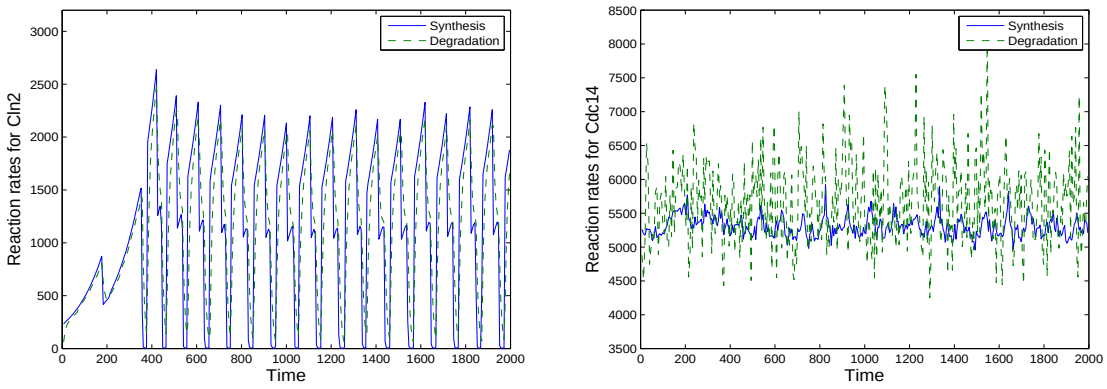

Figure 3: Reaction rates for $C\ln 2$ and $Cdc14$.

corresponding to reactions marked with $(*)$ and $(**)$, respectively. The cyclin $C\ln 2$ is primarily responsible for bud emergence while $Cdc14$ is the phosphatase required the cell for exit of M phase and return to $G1$ phase. The net effect of these reactions are for $C\ln 2$ and $Cdc14$ to reach an quasi-equilibrium while the other species in these modules do not influence the rest of the network. Therefore even the reaction rates $a_{63}-a_{64}$ and $a_{69}-a_{71}$ are only $\mathcal{O}(10^1)-\mathcal{O}(10^2)$, they are still counted in the fast module for $Cdc14$ since they are dominated by faster reactions in the module. Notice that the fast modules are affected by the slow modules through populations of species like $Clb5$ and $Cdc15$. To justify the validity of Nested SSA, we need to verify the quasi-equilibrium assumption, which requires the time scale separation to hold. For this purpose, we evaluate the rates of the synthesis and degradation of $C\ln 2$ and $Cdc14$. It can be seen from Fig. 3 that the reactions for $Cdc14$ remain fast all the time and the corresponding rates remain balanced most of time. But the reactions for $C\ln 2$ constantly switch between slow and fast sets, with the magnitude of the rates spanning a wide range of $\mathcal{O}(1)-\mathcal{O}(10^3)$. This means we have to dynamically partition between the fast and slow reactions sets in Nested SSA. We make the following rule for the reactions marked with $(*)$ in Table 1 such that if $a_1+a_2>1000$, then the synthesis and degradation of $C\ln 2$ is counted as fast reactions, otherwise if $a_1+a_2<=1000$, they are treated as slow reactions. Although the convergence for the adaptive NSSA is still not well addressed, the numerical result here shows that this simple mechanism still achieve efficiency while keeping accuracy of NSSA.

To handle the continuous variable $mass$, we simply solve (3.1) such that

$$mass(t+\tau)=mass(t)\exp(k_g\tau), \tag{4.1}$$

where τ is generated at each time step of Outer SSA. The flag variables BUD, ORI and SPN are solved using the Duhamel's principle:

$$\begin{aligned} BUD(t+\tau)=&\frac{1-\exp(-k_{d,bud}\tau)}{k_{d,bud}}\cdot k_{s,bud}\cdot\left(\epsilon_{bud,n2}\cdot C\ln 2+\epsilon_{bud,n3}\cdot C\ln 3+\epsilon_{bud,b5}\cdot Clb5\right)\\ &-\exp(-k_{d,bud}\tau)\cdot BUD, \end{aligned}$$

$$ORI(t+\tau)=\frac{1-\exp(-k_{d,ori}\tau)}{k_{d,ori}}\cdot k_{s,ori}\cdot\left(\epsilon_{ori,b5}\cdot Clb5+\epsilon_{ori,b2}\cdot Clb2\right)-\exp(-k_{d,ori}\tau)\cdot ORI,$$
$$SPN(t+\tau)=\frac{1-\exp(-k_{d,spn}\tau)}{k_{d,spn}}\cdot k_{s,spn}\cdot Clb2/\left(J_{spn}+Clb2\right)-\exp(-k_{d,spn}\tau)\cdot SPN, \quad (4.2)$$

where all the variables on the right hand side are evaluated at time t.

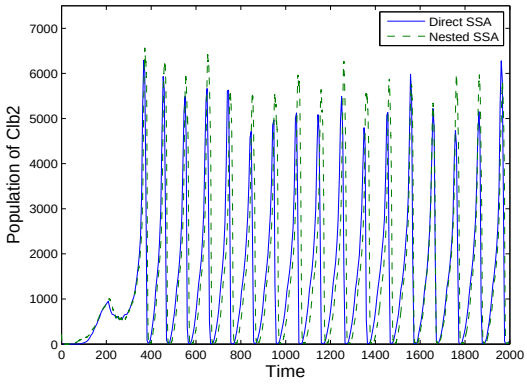

Figure 4: Cell cycles obtained by SSAs.

In Fig. 4, we show the numerical results obtained using both Direct SSA and Nested SSA. We have chosen the numerical parameters in Nested SSA such that

$$N=1, \quad T_0=0, \quad T_f=10^{-6}. \quad (4.3)$$

To further test the efficiency of Nested SSA quantitatively, we run the simulation till about 10^4 periods of cycles and measure the average and variance of the effective period. The direct SSA will require 2.42×10^4 CPU seconds to finish the computation, and give an average mean cycle period of 101.2246 seconds with a variance of 9.4288. In Nested SSA, the parameter T_f is increased gradually each time, which will require a longer CPU time but increase accuracy. The results are given in Table 3. It can be seen that the Nested SSA is about 4 times faster than Direct SSA while the relative error on the key dynamical feature of the effective period is only .0014%.

Table 3: Efficiency of Nested SSA for the period of cell cycle.

T_f	10^{-9}	10^{-8}	10^{-7}	10^{-6}
CPU	5900	5820	6074	6032
$\overline{Period}$	101.2257	101.2261	101.2260	101.2260
$\mathtt{var}(Period)$	11.4156	9.8154	11.1637	10.3799

5 Conclusion

We tested the Nested Stochastic Simulation Algorithm on the cell cycle model of budding yeast. The results show Nested SSA as a very efficient and accurate methods for simulating large scale reacting networks with multiple time scales. Future investigations will be focused on analyzing convergence of adaptive schemes and applying the method to the study of more biological problems.

Acknowledgments

The author would like to thank Weinan E and Eric Vanden-Eijnden for stimulating discussions. The work is supported by grants NSF-DMS 0845061 and NSF-DMS 0829515.

References

[1] N. Fedoroff and W. Fontana, Small numbers of big molecules, Science, 297, 1129-1131, 2002.
[2] H. McAdams and A. Arkin, It's a noisy business! Genetic regulation at the nanomolar scale, Trends. Genet., 15, 65-69, 1999.
[3] D. Gillespie, A general method for numerically simulating the stochastic time evolution of coupled chemical reactions, J. Comp. Phys., 22, 403-434, 1976.
[4] D. Gillespie, Exact stochastic simulation of coupled chemical reactions, J. Phys. Chem., 81, 2340-2361, 1977.
[5] C. Rao and A. Akin, Stochastic chemical kinetics and the quasi-steady-state assumption: application to the Gillespie algorithm, J. Chem. Phys., 118, 4999-5010, 2003.
[6] Y. Cao, D. Gillespie, L. Petzold, The slow scale stochastic simulation algorithm, J. Chem. Phys., 122, 014116, 2005.
[7] Y. Cao, D. Gillespie, L. Petzold, Multiscale stochastic simulation algorithm with stochastic partial equilibrium assumption for chemically reacting systems, J. Comp. Phys., 206, 395-411, 2005.
[8] W. E, D. Liu and E. Vanden-Eijnden, Nested stochastic simulation algorithm for chemical kinetic systems with disparate rates, J. Chem. Phys., 123, 194107, 2005.
[9] W. E, D. Liu and E. Vanden-Eijnden, Nested stochastic simulation algorithms for chemical kinetic systems with multiple time scales, J. Comp. Phys., 221, 158-180, 2007.
[10] H. Salis and Y. Kaznessisa, An equation-free probabilistic steady-state approximation: Dynamic application to the stochastic simulation of biochemical reaction networks, J. Chem. Phys., 123, 214106, 2005.
[11] A. Samant and D. G. Vlachos, Overcoming stiffness in stochastic simulation stemming from partial equilibrium: A multiscale Monte Carlo algorithm, J. Chem. Phys., 123, 144114, 2005.
[12] K. C. Chen, L. Calzone, A. Csikasz-Nagy, F. R. Cross, B. Novak and J. J. Tyson, Integrative analysis of cell cycle control in budding yeast, Mol. Biol. Cell, 15, 3841-3862, 2004.
[13] P. Wang, R. Randhawa, C. Shaffer, Y Cao, and W Baumann, Converting Macromolecular Regulatory Models from Deterministic to Stochastic Formulation, High Performance Computing and Simulation Symposium (HPCS 2008).

[14] T.-H. Ahn, L. T. Watson, Y. Cao, C. A. Shaffer and W. T. Baumann, Cell cycle modeling for budding yeast with stochastic simulation algorithms, Computer Modeling in Engineering & Sciences, 51, 27-52, 2009. Also available at `http://eprints.cs.vt.edu`.
[15] A. Bortz, M. Kalos and J. Lebowitz, A new algorithm for Monte Carlo simulation of Ising spin systems, J. Comp. Phys., 17, 10-18, 1975.
[16] M. A. Gibson and J. Bruck, Efficient exact stochastic simulation of chemical systems with many species and many channels, J. Phys. Chem. A, 104, 1876-1889, 2000.
[17] R. Z. Khasminskii, Principle of averaging for parabolic and elliptic differential equations and for Markov processes with small diffusion, Theory Probab. Appl., 8, 1-21, 1963.
[18] T. Kurtz, A limit theorem for perturbed operator semigroups with applications for random evolutions, J. Funct. Ana., 12, 55-67, 1973.
[19] G. Papanicolaou, Introduction to the asymptotic analysis of stochastic differential equations, Lectures in Applied Mathematics, R. C. Diprima Edt., 16, 1977.
[20] R. Srivastava, L. You, J. Summers and J. Yin, Stochastic vs. deterministic modeling of intracellular viral kinetics, J. Theor. Biol., 218, 309-321, 2002.
[21] R. Srivastava, M. Peterson and W. Bently, Stochastic kinetic analysis of the Escherichia coli stress circuit using σ^{32}-targeted antisens, Biotechnol. Bioeng., 75, 120-129, 2001.
[22] W. E, D. Liu and E. Vanden-Eijnden, Analysis of multiscale methods for stochastic differential equations, Comm. on Pure and Appl. Math., 58, 1544-1585, 2005.
[23] B. Alberts, et al., Essential Cell Biology, 2nd. ed., Garland Science, 2003.
[24] S. Ghaemmaghami, et al., Global analysis of protein expression in yeast, Nature, 425, 737-41, 2003.