Non-Markovian Processes in Gene Regulation

D.A.Bratsun [a], D.N.Volfson [a,b], J.Hasty[a], and L.S. Tsimring [b]

[a]Department of Bioengineering University of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0412;
[b]Institute for Nonlinear Science, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0402

ABSTRACT

We study the stochastic properties of gene regulation taking into account the non-Markovian character of gene transcription and translation. We show that time delay in protein production or degradation may change the behavior of the system from stationary to oscillatory even when a deterministic counterpart of the stochastic system exhibits no oscillations. Assuming significant decorrelation on the time scale of gene transcription, we deduce a truncated master equation of the reactive system and derive an analytical expression for the autocorrelation function of the protein concentration. For weak feedback the theory agrees well with with numerical simulations based on the modified direct Gillespie method.

Keywords: Gene regulatory networks, intrinsic noise, non-Markovian processes, time delay, Gillespie algorithm

1. INTRODUCTION

In the recent years, there has been a significant interest in the stochastic modeling of gene regulatory networks.[1–3] There is considerable experimental evidence that stochasticity plays a major role in such networks,[4,5] both due to intrinsic (small numbers of molecules involved in biochemical reactions) and extrinsic (complex extracellular signaling, cell cycle irregularities, etc.) factors. There are several different approaches to the modeling of stochastic chemical reactions: numerous flavors of Gillespie algorithm,[6–9] exact master equation analysis, as well as various simplified descriptions based on the Fokker-Planck or Langevin equations (see[3] for a review). One major difficulty that naturally arises in the analysis of gene networks is the vast separation of time scales between the fast reactions (dimerization, binding/unbinding) and slow reactions (transcription, degradation). There have been a number of papers devoted to the development of reduced descriptions of these systems using the idea of quasi-equilibrium of fast processes compared with slow dynamics ([10] and references therein). However, all of these approaches implicitly assume that all of the reactions (fast and slow) are Markovian processes which obey Poissonian statistics. In this regard, it is important to realize that transcription and translation are not only slow but are compound multi-stage reactions involving the sequential assembly of long molecules comprised of many elementary reactions. Thus, by virtue of the central limit theorem, such processes obey Gaussian statistics with a characteristic mean delay time. Both analytical and numerical modeling of such processes is needed in order to account for their non-Markovian nature.

The behavior of stochastic delay-differential equations (SDDEs) has been studied in.[11–13] In these papers, various approximate methods of treating SDDEs were developed. The moment equations for the solutions of linear SDDEs and their stability have been studied in.[11] The limit of small delays, where a univariate non-delayed stochastic differential equation can approximate the initial SDDE, has been considered in[12] in detail. In,[13] a model noise-driven bistable system with delayed feedback in the limit of small noise and small magnitude of the feedback has been reduced to a two-state model with delayed transition rates for which analytical description could be developed. The deterministic dynamics of gene regulatory circuits with delayed transcription/translation have been recently addressed in.[14,15] Lewis[14] also presented the results of approximate stochastic simulations of a simple autorepressor model for Zebrafish somitogenesis using Euler finite difference method.

For further information send correspondence to L.T. at ltsimring@ucsd.edu

Noise in Complex Systems and Stochastic Dynamics III, edited by Laszlo B. Kish, Katja Lindenberg, Zoltán Gingl, Proc. of SPIE Vol. 5845 (SPIE, Bellingham, WA, 2005) · 0277-786X/05/$15 · doi: 10.1117/12.609707

The goal of the present work is to apply and extend these methods to the models of biochemical reactions comprising gene regulation. We will present a stochastic algorithm for simulating non-Markovian reactions. Unlike Ref.,[14] our algorithm generalizes the *exact* direct Gillespie algorithm. We also describe a theoretical approach to formulate the simplified master equation for the biochemical kinetics which involves delayed reactions. It is then used to calculate the correlation function of the delayed stochastic system analytically. We study two specific examples: single-gene transcription with delayed degradation, and a single-gene auto-repressive network with delayed negative feedback. In agreement with earlier numerical work[14] we find that intrinsic stochastic fluctuations lead to the occurrence of oscillations in these systems even when their deterministic analogs do not oscillate.

2. SINGLE-GENE PROTEIN PRODUCTION-DEGRADATION MODEL

We begin with an extremely idealized model for unregulated single-gene protein production and delayed protein degradation. Here we ignore the distinction between transcription and translation and call the combined transcription/translation process *production*. This simple model is considered here to illustrate the methods and techniques we use to analyze non-Markovian effects in more realistic models of gene regulation.

The reaction of production of protein X is written as

$$\emptyset^t \xrightarrow{A} X^t, \tag{1}$$

where A is the protein production rate. Here and below the superscript indicates time at which the particular reacting component is taken. The degradation of protein with rate B is written as

$$X^t \xrightarrow{B} \emptyset^t. \tag{2}$$

Both reactions (1) and (2) are assumed to happen without any delay and so they obey the Poissonian statistics. Let us suppose that there is another "delayed" mechanism of protein degradation

$$X^t \xrightarrow{C} \emptyset^{t+\tau}. \tag{3}$$

which leads to the decay of one protein molecule after a certain time delay τ after this reaction is initiated. This type of process may occur if the protein degradation is mediated by production of a certain enzyme (a protease) which takes a finite time to accomplish.

Delayed degradation provides a simple form of a delayed negative feedback. It is well known that sufficiently strong delayed negative feedback may lead to periodic oscillations.

Let us first briefly outline the deterministic dynamics of this system in the rate approximation which is described by the following linear delay-differential equation:

$$\frac{dx}{dt} = A - Bx(t) - Cx(t-\tau). \tag{4}$$

This system has one fixed point $x^* = A/(B+C)$ whose stability determines the transition to oscillations. Looking for a solution $x(t)$ in the form $x(t) \sim e^{\lambda t}$, we find the eigenvalues $lambda = \mu + i\omega$ in the following form

$$\mu = \frac{1}{\tau} Re(W(-\tau C e^{\tau B})) - B, \tag{5}$$

$$\omega = \frac{1}{\tau} Im(W(-\tau C e^{\tau B})). \tag{6}$$

where $W(z)$ is the Lambert function defined via equation $W(z)e^{W(z)} = z$. By solving (5), (6) for $\mu = 0$ and $\omega \neq 0$ we find the condition for Hopf bifurcation. The neutral curve is shown in Fig.1, where the instability domain is located above the curve.

Evidently, in this linear model, the amplitude of the oscillations grows indefinitely without saturation. In a real system, nonlinearity saturates this exponential growth.

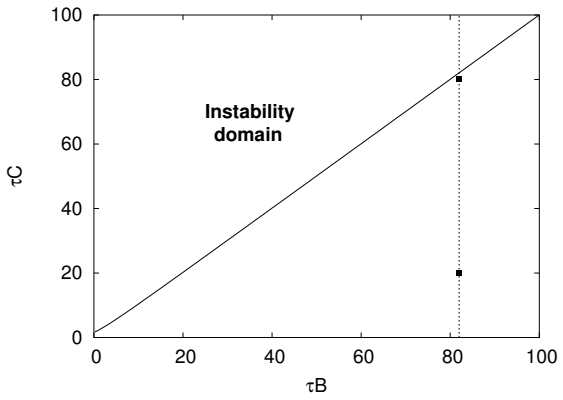

Figure 1. Neutral curve of the Hopf bifurcation (solid line). The upper and lower squares indicate the fixed parameters for which correlation functions are plotted in Fig.3a,b respectively. The dashed line shows the cross-section of parameter space which corresponds to Fig.4

Stochastic description

Now we take into account the fact that chemical reactions (1-3) occur randomly in time according to their respective rates. Since the number of molecules involved is often small, random fluctuations of copy numbers are important and a stochastic approach should be used to describe the behavior of such a system.

Let us denote $P(n,t)$ the probability of having n monomers at time t. Then the master equation for the time evolution of the probability $P(n,t)$ can be written as

$$\begin{aligned}\frac{dP(n,t)}{dt} &= A\left(P(n-1,t)-P(n,t)\right)+B\left((n+1)P(n+1,t)-nP(n.t)\right)+\\ &+ C\sum_{m=0}^{\infty} m\left(P(n,t;m,t-\tau)-\Theta_n P(n,t;m,t-\tau)\right), \quad n=0..\infty \end{aligned} \tag{7}$$

where $P(n,t;m,t-\tau)$ is the joint probability of having n molecules at time t and m molecules at time $t-\tau$ and the multiplier

$$\Theta_n = \begin{cases} 0, & n=0 \\ 1, & n>0 \end{cases} \tag{8}$$

is added to account for the fact that $P(n,t), P(n,t;m,t-\tau)$ should remain zero for negative n.

This set of equations is not closed because the one-point probability distribution is determined by the two-point probability distributions on the r.h.s. In order to make progress, we assume that the time delay τ is large compared with other characteristic times of the system so that events at time t and $t-\tau$ are effectively decoupled. Under this approximation, we can write

$$P(n,t;m,t-\tau) = P(n,t)P(m,t-\tau) \tag{9}$$

In fact, large τ is a necessary but not a sufficient condition for this to hold, because as we will see below, strong delayed feedback leads to significant correlations over long periods of time. Therefore, an additional condition

for the applicability of (9) is to have relatively weak feedback. Adopting this approximation, we obtain

$$\begin{aligned}\frac{dP(n,t)}{dt} &= A\left(P(n-1,t)-P(n,t)\right)+B\left((n+1)P(n+1,t)-nP(n,t)\right)+\\ &\quad +C\sum_{m=0}^{\infty} mP(m,t-\tau)\left(P(n+1,t)-\Theta_n P(n,t)\right),\\ &= A\left(P(n-1,t)-P(n,t)\right)+B\left((n+1)P(n+1,t)-nP(n,t)\right)+\\ &\quad +C\langle n(t-\tau)\rangle\left(P(n+1,t)-\Theta_n P(n,t)\right), \quad n=0..\infty \end{aligned} \tag{10}$$

Autocorrelation function

For a stationary random process $n(t)$, the autocorrelation function, $K(T)$, is given by

$$K(T) = \langle n(t)n(t+T)\rangle - \langle n(t)\rangle^2 \tag{11}$$

$$= \sum_{n=0}^{\infty} nP_s(n)\langle n', T|n, 0\rangle - \langle n\rangle^2. \tag{12}$$

where P_s is the stationary probability and $\langle n', T|n, 0\rangle$ is the mean number of proteins at time T given that it was equal to n at time 0. In order to calculate the correlation function we make use of the generating function

$$G(s,t) = \sum_{n=0}^{\infty} s^n P(n,t) \tag{13}$$

One can easily see the relation between $G(s,t)$ and various moments (only the first two moments are shown):

$$\frac{\partial G}{\partial s}|_{s=1} = \sum_{n=0}^{\infty} ns^{n-1}P(n,t)|_{s=1} = \sum_{n=0}^{\infty} nP(n,t) = \langle n(t)\rangle, \tag{14}$$

$$\frac{\partial^2 G}{\partial s^2}|_{s=1} = \sum_{n=0}^{\infty} n(n-1)s^{n-2}P(n,t)|_{s=1} = \sum_{n=0}^{\infty} n(n-1)P(n,t) = \langle n^2(t)\rangle - \langle n(t)\rangle. \tag{15}$$

Using the generating function, we can convert the infinite set of ordinary differential equations (7), to a single partial differential equation for $G(s,t)$:

$$\frac{\partial G}{\partial t} = (s-1)\left(AG(t) - B\frac{\partial G(t)}{\partial s} + C\frac{\langle n(t-\tau)\rangle}{s}(P_0(t) - G(t))\right). \tag{16}$$

If the mean number of protein molecules is large, the states with $n = 0$ are rare, and $P_0(t)$ can be neglected in (16). Then the stationary solution for G is readily derived

$$G_s(s) = s^{-\frac{C}{B}<n>_s} e^{\frac{A}{B}(s-1)}, \tag{17}$$

where $\langle n\rangle_s$ stands for the stationary mean: $\langle n\rangle_s = \langle n(t\to\infty)\rangle$ and the normalization $G_s(1) = 1$ has been taken into account.

In order to obtain the equations for the mean and higher moments, we expand the generating function into a power series of $s-1$:

$$G(s-1,t) = 1 + (s-1)\alpha(t) + \frac{1}{2}(s-1)^2\beta(t) + ... \tag{18}$$

where the functions $\alpha(t)$ and $\beta(t)$ are related to the lowest two moments, $\alpha(t) = \langle n(t)\rangle,\ \beta(t) = \langle n^2(t)\rangle - \langle n(t)\rangle$. Substituting (18) into (16) we obtain

$$\frac{d\alpha}{dt} = A - B\alpha(t) - C\alpha(t-\tau), \tag{19}$$

$$\frac{1}{2}\frac{d\beta}{dt} = A\alpha(t) - C\alpha(t-\tau) - B\beta(t) + C\alpha(t)\alpha(t-\tau). \tag{20}$$

As expected, the equation for the mean (19) coincides with Eq.(4) obtained above with the rate-equation approximation. It can be readily solved with the symmetry condition $\alpha(t) = \alpha(-t)$ and initial condition $\alpha(0) = n$:

$$\alpha(t) = (n-1)\frac{\sigma(t)}{1-\zeta e^{-\lambda\tau}} + \left(1 - \frac{A}{B+C}\right)\frac{\sigma(t)}{1-\zeta e^{-\lambda\tau}} + \frac{A}{B+C}, \tag{21}$$

where

$$\begin{aligned}
&\sigma(t) = e^{-\lambda t} - \zeta e^{\lambda(t-\tau)}, \quad 0 < t < \tau \\
&\sigma(N\tau + t) = e^{-Bt}(\sigma(N\tau) - C\int_0^t \sigma((N-1)\tau + t')e^{Bt'}dt'), \\
&\lambda = \sqrt{B^2 - C^2}, \quad \zeta = \frac{1}{C}(B - \lambda).
\end{aligned} \tag{22}$$

The solution (21),(22) exists if $B > C$, i.e. below the deterministic Hopf bifurcation. To complete the calculation we find

$$\alpha_s = \frac{A}{B+C}, \tag{23}$$

$$\beta_s = \frac{A(AB + BC + C^2)}{B(B+C)^2} \tag{24}$$

and insert (21)-(24) into (12):

$$\begin{aligned}
K(T) &= \frac{\sigma(T)}{1-\zeta e^{-\lambda\tau}}\sum_{n=0}^{\infty} n(n-1)P_s(n) + \left(1 - \frac{A}{B+C}\right)\frac{\sigma(T)}{1-\zeta e^{-\lambda\tau}}\sum_{n=0}^{\infty} nP_s(n) = \\
&= \frac{\sigma(T)}{1-\zeta e^{-\lambda\tau}}\frac{d^2G_s(s)}{ds^2}|_{s=1} + \left(1 - \frac{A}{B+C}\right)\frac{\sigma(T)}{1-\zeta e^{-\lambda\tau}}\frac{dG_s(s)}{ds}|_{s=1} = \frac{A}{B}\frac{\sigma(T)}{(1-\zeta e^{-\lambda\tau})},
\end{aligned} \tag{25}$$

where $T = N\tau + t$.

Stochastic simulations based on Gillespie method

In order to test the validity of our approximations and analytical results, we performed numerical stochastic simulations of the original system of chemical kinetic reactions (1-3). Here we introduce modifications to the direct Gillespie (DG) algorithm[6] which would allow us to treat delayed reactions.

Suppose the system consists of N components X_i which react through M elementary reaction channels R_μ. According to the DG scheme, time is advanced from one elementary reaction to the next. At every "stop" one has to determine the time to the next reaction and which reaction it will be. For Markovian processes, the distribution of times until the next reaction is exponential

$$P(\tau) = \sum_\mu a_\mu \exp\left(-\triangle t \sum_\mu a_\mu\right) \tag{26}$$

where $a_\mu = c_\mu h_\mu$ is propensity of the channel R_μ. The choice of the next reaction is made based upon the following discrete distribution,

$$P(\mu = \mu') = a_{\mu'} / \sum_\mu a_\mu. \tag{27}$$

In the case when some of the channels are non-Markovian, we modify the Gillespie algorithm as follows. At every "stop" we perform the same selection of the next reaction time according to the distributions (26), (27). If the next reaction time is chosen to be t_* but the selected reaction is delayed, it is placed in a stack, so it will actually be completed at time $t_* + \tau$. If however the chosen reaction is Markovian, the time to the next reaction t_m is compared to the times of previously scheduled delayed reactions. If none of those scheduled reactions are

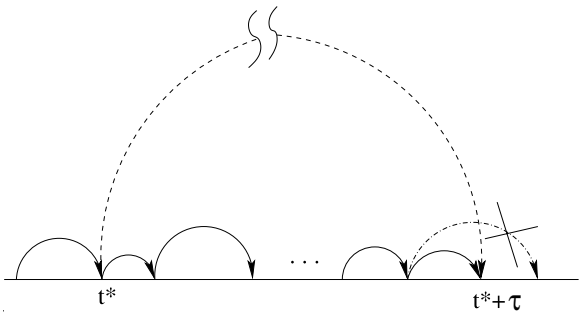

Figure 2. Scheme of numerical simulation

to occur before t_*, the time is advanced to t_*, and the process repeats. If however, there is a delayed reaction scheduled for completion for $t_d < t_m$, the last selection is ignored, the time advances to t_d, the scheduled reaction is performed, and the selection process repeats (see Fig.2, where $t_d = t^* + \tau$).

Let us formulate modifications to the Direct Gillespie algorithm which are necessary to correctly simulate the stochastic system with some delayed reactions. Our algorithm consists of the following steps (see also Fig. 2):

1. Input values for initial state $(x_1, \ldots, x_N)$, set $t = 0$.
2. Compute propensities $a_\mu, \mu = 1..M$.
3. Generate uniform random numbers u_1, u_2.
4. Compute the time interval $\triangle t$ until the next reaction according to distribution (26), viz. $\triangle t = -(\ln u_1)/\sum_\mu a_\mu$.
5. Check if there has been a delayed reaction scheduled to occur at time t_d within the range between t and $t + \triangle t$.

   a) if yes, then the results of steps 2,3 and 4 are ignored, time is advanced to $t = t_d$, X_i are updated according to the delayed reaction. Return to step 2.

   b) If no, proceed to the step 6.
6. Find the channel of the next reaction μ from the distribution (27), viz. take μ to be the integer for which $\sum_{\nu=1}^{\mu-1} < u_2 a_0 \leq \sum_{\nu=1}^{\mu}$, where a_0 is the total propensity, $a_0 = \sum_{\nu=1}^{M} a_\mu$.
7. Update time $t \to t + \triangle t$. If the selected reaction μ is Markovian, update X_i in accordance with $R(\mu)$. If the reaction is delayed, X_i update is postponed for time $t + \tau$. Return to step 2.

We calculated the correlation function numerically according to the standard formula implying ergodicity of the underlying stochastic process

$$K(T) = \lim_{t\to\infty} \frac{1}{t} \int_0^t n(t')n(t'+T)dt' - \left[\lim_{t\to\infty} \frac{1}{t} \int_0^t n(t')dt'\right]^2 . \tag{28}$$

The comparison between correlation functions derived analytically and numerically is shown in Fig.3 for two different sets of parameters indicated by black squares in Fig.1. We use $A = 100, B = 4.1$ and $\tau = 20$. So the delay time is large compared with the characteristic equilibration time $B^{-1} \approx 0.25$, which is necessary to justify the master equation (7). One can see that when the parameters are such that the system is near the Hopf bifurcation (Fig.3a), the agreement between our analytical method and the direct Gillespie simulation is not as good. This is to be expected since the processes at times t and t_τ are strongly correlated (the secondary peak of the correlation function is large). However, if the system is far from the bifurcation and the influence of the delay term is weaker, then the curves virtually coincide (Fig.3b). In Figure 4 we plot the height of the secondary

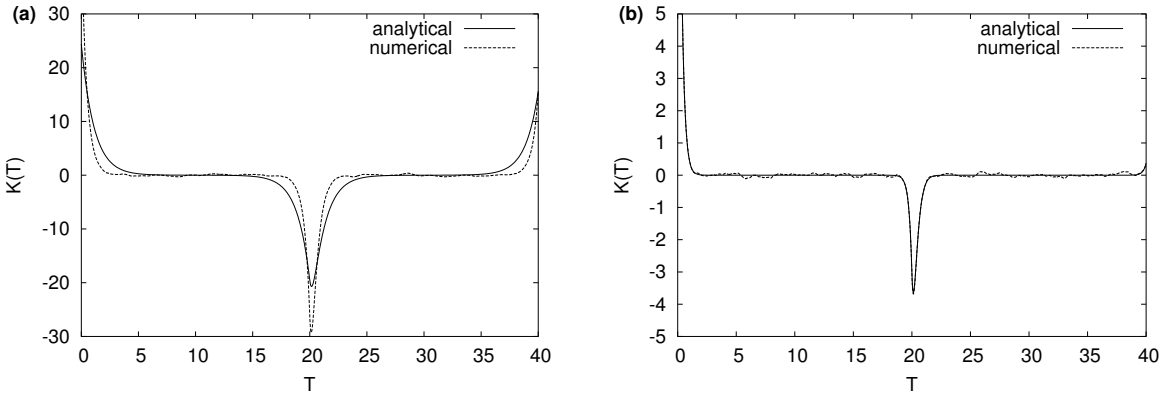

Figure 3. Comparison of correlation functions obtained analytically (solid line) and numerically (dashed line) for $\tau = 20$, $A = 100$, $B = 4.1$. (a) $C = 4$, (b) $C = 1$ (indicated by black squares in Fig.1)

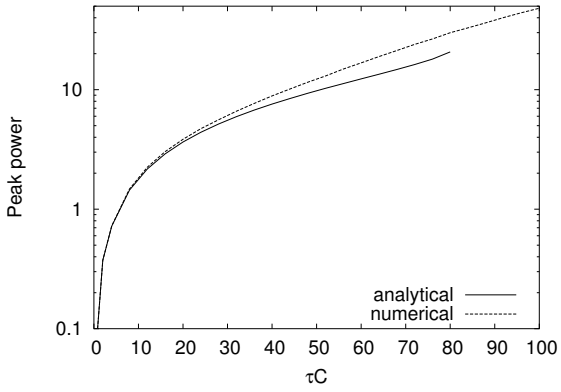

Figure 4. Comparison of height of correlation function's second peaks as a function the parameter C. Fixed parameters are $\tau = 20$, $A = 100$, $B = 4.1$. This picture corresponds to range of parameters indicated in Fig.1 by dashed vertical line. Analytical curve is shown only in the range $0 < C < 4.1$ where the solution (21-22) exists.

peak located near $\tau = 20$ as a function of the parameter τC. One can see that the ratio of peaks approaches unity as $C \to 0$.

As we can see from these results, the system exhibits oscillatory properties even below the Hopf bifurcation: the correlation function has peaks approximately at multiples of the time delay τ.

3. SINGLE-GENE AUTO-REPRESSOR MODEL

The purpose of the very simple model considered in the previous section was to illustrate the methods of analyzing the stochastic non-Markovian biochemical reactions. In this section we analyze a more complicated model which is actually relevant for genetic regulation. Namely, we consider a single gene protein synthesis with negative auto-regulation. This is a popular motif in genetic regulatory circuits, and its dynamics has been analyzed within both deterministic and stochastic frameworks.[1,3] Here we generalize this system by taking into account that transcription of auto-repressor protein takes a finite amount of time τ. We postulate that the chemical state of the operator site $D^t \in \{D_0^t, D_1^t\}$ determines the production of protein at time $t + \tau$ If the operator at time t is unoccupied (D_0^t) then the protein may be produced at time $t + \tau$ with a certain probability A per unit time. Otherwise if the operator is occupied (D_1^t), the production at time $t + \tau$ is blocked. The transitions between operator states, which we denote as D_0 (unoccupied) and D_1 (occupied) occur with rates k_1, k_{-1}, are written as

$$D_0^t + X^t \xrightarrow{k_1} D_1^t, \tag{29}$$

$$D_1^t \xrightarrow{k_{-1}} D_0^t + X^t. \tag{30}$$

Protein production-degradation reactions can be written in the following form:

$$\emptyset \xrightarrow{AS(t)} X^{t+\tau}, \tag{31}$$

$$X^t \xrightarrow{B} \emptyset. \tag{32}$$

Here $S(t) = 0$ for unoccupied operator state D_0^t and 1 for occupied state D_1^t. Thus, the reactions (31-32) have negative feedback with (29-30) through the reaction rate in (31). It is important to note that the protein production occurs with time lag τ can happen at time $t + \tau$ only if the operator is unoccupied at the time t.

Without going into the detailed description of this system, let us summarize the main results. It can be shown that in the deterministic limit, the fixed point corresponding to the constant protein production, is globally stable and does not exhibit Hopf bifurcation.

Nevertheless, stochastic fluctuation lead to persistent (albeit irregular) oscillations with characteristic period close to two time delay times. In order to describe this process quantitatively, we introduce two probabilities, $P_n^0(t)$ and $P_n^1(t)$, for the number of proteins to be equal to n at time t and for the state of the operator at time $t - \tau$ to be D_0 or D_1, respectively. Then the master equations for the reactions (29-32) have the form

$$\begin{aligned}
\frac{dP^0(n,t)}{dt} &= A(P^0(n-1,t) - P^0(n,t)) + B((n+1)P^0(n+1,t) - nP^0(n,t)) - \\
&- k_1 \sum_{n=0}^{\infty} n[(P^0(n,t-\tau) + P^1(n,t-\tau)]P^0(n,t) + k_{-1}P^1(n,t), \\
\frac{dP^1(n,t)}{dt} &= B((n+1)P^1(n+1,t) - nP^1(n,t)) + \\
&+ k_1 \sum_{n=0}^{\infty} n[(P^0(n,t-\tau) + P^1(n,t-\tau)]P^0(n,t) - k_{-1}P^1(n,t).
\end{aligned} \tag{33}$$

Here again we made the implicit assumption that the processes at times t and $t - \tau$ are weakly correlated, and to the first approximation the two-point probability distribution function may be factorized, $P(n,t;m,t-\tau) \approx$

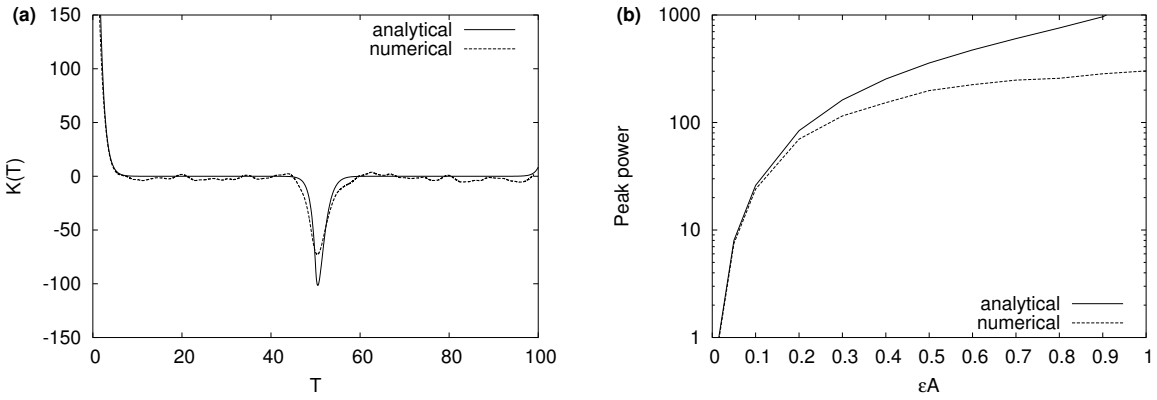

Figure 5. (a) Comparison of correlation functions obtained analytically (solid line) and numerically (dashed line) for $A = 100$, $\varepsilon = 0.002$. (b) Comparison of height of correlation function's second peaks as a function the parameter εA. Fixed parameters are $\tau = 50$, $B = 1$.

$P(n,t)P(m,t-\tau)$. In order to calculate the correlation function we again use the method of generating functions. Here we introduce two generating functions

$$G_i(s,t) = \sum_{n=0}^{\infty} s^n P_n^i(t), \quad i = 1,2 \tag{34}$$

which correspond to the two operator states and their sum gives the full generating function $G(s,t) = G_0(s,t) + G_1(s,t)$. The equations for these functions read

$$\begin{aligned} \frac{\partial G_0}{\partial t} &= (s-1)\left(AG_0 - B\frac{\partial G_0}{\partial s}\right) - k_1\langle n(t-\tau)\rangle G_0 + k_{-1}G_1, \\ \frac{\partial G_1}{\partial t} &= -(s-1)B\frac{\partial G_1}{\partial s} + k_1\langle n(t-\tau)\rangle G_0 - k_{-1}G_1. \end{aligned} \tag{35}$$

Finding the solutions of these equations near $s = 1$ is analogous to the previous Section, and we present here only the final result. The correlation function reads

$$K(T) = \left(\frac{A^2(B+k_{-1})}{B((B+k_{-1})(B+\varepsilon A)+k_1 A)} + \frac{A(B+\varepsilon A - A)}{(B+\varepsilon A)^2}\right)\frac{\sigma(T)}{1-\zeta e^{-\lambda\tau}}.$$

where σ, λ, and ζ are given by (22).

We performed direct Gillespie simulations of the stochastic model of single-gene auto-repressor. The comparison between correlation functions derived analytically and numerically is shown in Figs.5(a,b). One can notice that the structure of the correlation function is well described by the theory for small A, but the heights of analytical and numerical peaks (Fig.5b) diverge quickly as εA grows. This is to be expected, because the validity of our approximate theory is limited to small ϵ.

4. CONCLUSIONS

We have developed deterministic and stochastic models of the transcriptional regulation with delayed feedback and have studied them both analytically and numerically. The stochastic problems with delayed kinetics are generally difficult because of the non-Markovian nature of the dynamics. Such non-Markovian processes are ubiquitous in transcriptional gene regulation, because transcription step is comprised of many elementary binding/unbinding reactions. However, the main features of such systems can be deduced within simplified models. In this paper, we studied two such models for a single-gene transcriptional gene regulations. We were able to

derive the analytical formulas for autocorrelation functions and compare it with those obtained numerically using suitably modified direct Gillespie algorithm. We have shown that delayed transcription leads to characteristic oscillations of the protein concentration even below the Hopf bifurcation of the deterministic analog of this stochastic system.

ACKNOWLEDGMENTS

This work was supported by the the NSF Division of Molecular and Cellular Bioscience and the National Institutes of Health (NIH R01 GM69811-01).

REFERENCES

1. J. Hasty, F. Isaacs, M. Dolnik, D. McMillen, and J. J. Collins, "Designer gene networks: towards fundamental cellular control," *Chaos* **11**, pp. 207–220, 2001.
2. M. Thattai and A. van Oudenaarden, "Intrinsic noise in gene regulatory networks," *Proc. Natl. Acad. Sci. U.S.A.* **98**, pp. 8614–8619, 2001.
3. T. B. Kepler and C. Elston, "Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations," *Biophys. J.* **81**, pp. 3116–3136, 2001.
4. K. Ahmad and S. Henikoff, "Modulation of a transcription factor counteracts heterochromatic gene silencing in drosophila," *Cell* **81**, pp. 839–847, 2001.
5. M. Wijgerde, F. Grosveld, and P. Fraser, "Transcription complex stability and chromatin dynamics in vivo," *Nature* **377**, pp. 209–213, 1995.
6. D. T. Gillespie, "Exact stochastic simulation of coupled chemical reactions," *J. Phys. Chem.* **81**, pp. 2340–2361, 1977.
7. M. A. Gibson and J. Bruck, "'efficient exact stochastic simulation of chemical systems with many species and many channels," *J. Phys. Chem* **104**, pp. 1876–1889, 2000.
8. D. T. Gillespie, "Approximate accelerated stochastic simulation of chemically reacting systems," *J. Chem. Phys.* **115**, pp. 1716–1733, 2001.
9. D. Adalsteinsson, D. McMillen, and T. C. Elston, "Biochemical network stochastic simulator (bionets): Software for stochastic modeling of biochemical networks," *BMC Bioinformatics* **5**, p. 24, 2004.
10. C. V. Rao and A. P. Arkin, "Stochastic chemical kinetics and the quasi-steady-state assumption: application to the gillespie algorithm," *J. Chem. Phys.* **118**, pp. 4999–5010, 2003.
11. M. C. Mackey and I. G. Nechaeva, "Solution moment stability in stochastic differential delay equations," *Phys. Rev. E* **52**, pp. 3366–3376, 1955.
12. S. Guillouzic, I. L'Heureux, and A. Longtin, "Small delay approximation of stochastic differential delay equations," *Phys. Rev. E* **59**, pp. 3970–3982, 1999.
13. L. Tsimring and A. Pikovsky, "Noise-induced dynamics in bistable systems with delay," *Phys. Rev. Lett.* **87**, p. 250602, 2001.
14. J. Lewis, "Autoinhibition with transcriptional delay: a simple mechanism for the zebrafish somitogenesis oscillator," *Current Biology* **13**, pp. 1398–1408, 2003.
15. N. A. M. Monk, "Oscillatory expression of hes1, p53, and nf-kappab driven by transcriptional time delays," *Current Biology* **13**, pp. 1409–1413, 2003.