Stochastic Modeling of Protein Field with a Delayed Feedback

D A Bratsun and M D Buzmakov

Department of Applied Physics, Perm National Research Polytechnic University,
11B Professor Pozdeev Street, Perm, 614990, Russia

E-mail: DABracun@pstu.ru

Abstract. Protein fields synthesized by genes play a principal role in the functioning of living systems. The processes of gene regulation determine the properties of these fields. Since the number of nucleotides usually is not large, a deterministic description of the field dynamics is insufficient. In this work, we consider a special kind of protein field, the dynamic behavior of which is described by the non-Markov process. Generally, the dynamics of complex organic compounds is time-dependent and spatially extended, and it may depend on all the previous evolution of the system. We consider a time-delayed repressilator as a model system. We study this system numerically using a modified Gillespie algorithm. New dynamic phenomena, which are visible only within a stochastic description, are reported. We show that synchronization in a gene expression occurs much faster due to the non-linear interaction of noise and delay. It results in almost regular oscillations even below the neutral curve derived within the deterministic analysis. We apply a hybrid approach to study the spatial dynamics of the repressilator proteins. This approach includes a deterministic calculation of the diffusion fluxes between cells and the stochastic simulation of gene regulation processes. We found that the combined action of time-delay, noise, and spatial signaling can lead to pattern formation even when the deterministic description predicts the absolute stability of the system.

1. Introduction

Experimental technologies that help to observe and accurately measure nucleotide fields in cells create the preconditions for the formation of molecular biology as an exact science. This means the widespread use of mathematical tools to describe and model the observed effects, as well as to predict unknown phenomena arising from genetic regulation. Thus, biology undergoes a new stage of its development and turns out to be an exact science rather than a descriptive subject. The paper [1] has triggered the formation of synthetic biology, which is one of the main consequences of the turn. The authors came up with a plasmid of three genes that do not occur together in such a combination in vivo. The elements of the synthetic chain were selected so that the protein of each gene suppressed the expression of the next gene in the chain. A simple mathematical model that was proposed by the authors showed that oscillations should definitely occur in the designed circuit. The experimental implementation of the regulatory chain confirmed this hypothesis. This experiment proved that this could be the way to design artificial schemes with a set of dynamic gene regulations, which became the subject of synthetic biology [2]. Further on, different effects observed in the synthetic genetic circuits were experimentally and theoretically studied in [3-6].
Figure 1. Schematic presentation of a genetic circuit of the repressilator. The repressilator is a genetic regulatory network artificially designed in 2000 [1]. It is a closed chain including three genetic elements lacI, λcI, and tetR, where each of which of natural origin, but all they never occur together in nature. The promoter of each gene controls the following cistron by suppressing the expression of the neighboring gene.

Recently, delays in gene regulation attracted a lot of attention. Notice that the processes of gene regulation generally are multistep biochemical reactions that sequentially generate ensembles of complex organic compounds. For this reason alone, these reactionary processes do not occur instantly but are extended in time. If we take into account that the tissue generates spatially distributed protein fields, then we will inevitably conclude that the processes are performed with a certain time lag. Besides, the very nature of the transcription/translation processes implies a certain time interval between the onset of polymerase binding and protein expression. If this delay is small compared to other characteristic times in the system, these effects can be neglected [7]. The fact that biochemical reactions are greatly delayed in time during a gene expression has been firmly established in research into natural circadian rhythms at different microorganisms, for example, N. crassa and Drosophila [8-10]. Also, genetic engineering technologies allow artificially introducing delays into regulatory gene circuits. The paper [11] studied the behavior of a group of E. coli with implanted retardation. It was shown that the ensemble of bacteria exhibits traveling waves of spatial activity.

This work is devoted to the stochastic analysis of the repressilator with a time-delayed feedback proposed in [12]. We focus on those results where a qualitative difference in the behavior of the repessor was found in the framework of deterministic and stochastic analysis.

2. Genetic circuit of the repressilator

Figure 1 shows a diagram of a gene regular circuit of the repressilator [1,6]. One can see that a plasmid of the repressilator has three cistrons. The first cistron contains a gene tetR-lite (a gene and its promoter are blue) which is controlled by a promoter lac of gene lacI-lite (both are yellow) serving to be a repessor. Expression of tetR-lite is typically accompanied by the expression of the corresponding protein which will be called X for short. X protein dimerizes and can serve to be a repressor for λ phage through tetR promoter (the corresponding gene λcI-lite is red in the picture). The scheme shows that a promoter of λ phage serves to be a repessor of gene lacI-lite expression. The protein coded by the latter gene will be labeled Z. Thus, the circle closes. The main repressilator’s core determining its time dynamics consists of three genes which are interconnected via promoters with negative feedback. Let us enumerate the main assumptions which are used in developing a model [12]:

- a repressilator is symmetric, i.e. all genes have the same characteristics;
- a repressilator is regulated by a dimer form of protein;
- a promoter of each gene can regulate the expression of the next gene in the circuit;
- a negative feedback is delayed for each couple of genes with the time delay τ;
- any delay is much longer than the other typical characteristic times determining a non-linear dynamics of a repressilator.
Table 1. A list of reactions at the repressilator gene regulatory network shown in Figure 1.

| Reaction Type               | Reaction                                      |
|-----------------------------|------------------------------------------------|
| Dimerization of proteins    | $X + X \xrightarrow{k_1} X_2$                |
| Dedimerization of proteins  | $Y + Y \xrightarrow{k_1} Y_2$                |
|                            | $Z + Z \xrightarrow{k_1} Z_2$                |
| Binding of dimers           | $X_2 \xrightarrow{k_{-1}} X + X$             |
|                            | $Y_2 \xrightarrow{k_{-1}} Y + Y$             |
|                            | $Z_2 \xrightarrow{k_{-1}} Z + Z$             |
| Unbinding of dimers         | $D_0^X + Z_2 \xrightarrow{k_2} D_1^X$        |
|                            | $D_0^Y + X_2 \xrightarrow{k_2} D_1^Y$        |
|                            | $D_0^Z + Y_2 \xrightarrow{k_2} D_1^Z$        |
| Protein synthesis delayed in time | $D_0^X(t) \xrightarrow{A} D_0^X + X(t + \tau)$ |
|                            | $D_0^Y(t) \xrightarrow{A} D_0^Y + Y(t + \tau)$ |
|                            | $D_0^Z(t) \xrightarrow{A} D_0^Z + Z(t + \tau)$ |
| Degradation of proteins     | $X \xrightarrow{B} \emptyset$                |
|                            | $Y \xrightarrow{B} \emptyset$                |
|                            | $Z \xrightarrow{B} \emptyset$                |

Considering the above-mentioned comments, Table 1 gives a complete list of biochemical reactions running once the repressilator is switched on. The respective reaction rates are indicated above the arrows. Operators-sites in Table 1 are taken into account by introducing a special discrete function $DE\{D_0, D_1\}$ which can take the value $D_0$ when an operator-site is open and the value $D_1$ when the operator-site is closed. Formally, an operator-site condition could be regarded as an additional reagent in the system. If a repressor dimer is bound, an operator-site is closed, and no gene transcription can occur. Thus, the system has a set of negative feedbacks.

In [12], we have derived the following system of delay differential equations governing the evolution of a symmetric repressilator within the deterministic description:

\[
\frac{dx}{dt} = \frac{1}{1 + 4\epsilon x} \left( \frac{A}{1 + \epsilon \delta z^2(t - \tau)} - Bx \right),
\]

\[
\frac{dy}{dt} = \frac{1}{1 + 4\epsilon y} \left( \frac{A}{1 + \epsilon \delta x^2(t - \tau)} - By \right),
\]

\[
\frac{dz}{dt} = \frac{1}{1 + 4\epsilon z} \left( \frac{A}{1 + \epsilon \delta y^2(t - \tau)} - Bz \right),
\]

where

\[
\epsilon = \frac{k_1}{k_{-1}}, \quad \delta = \frac{k_2}{k_{-2}}.
\]

Here, $x(t), y(t), z(t)$ are continuous functions of time, denoting, respectively, the concentration of proteins $X, Y, Z$ in the monomer form. The introduction of a delay, which implies the occurrence of multi-stage reactions, during which ensembles of complex connections are sequentially formed, allows the order of the system to be halved in comparison with the model suggested in [1].
3. Gillespie algorithm

One of the most famous methods for the numerical study of a small-sized stochastic reactive system is the Gillespie algorithm [Gillespie, 1977], which has become truly classical due to its simplicity and reliability. Mathematically, the Gillespie algorithm is a kind of Monte Carlo method for numerical stochastic analysis and reproduces the solution for the probability distribution given by the master equation. Generally, it is believed that averaging over realizations should converge to the solution obtained within the framework of the deterministic description.

Let us briefly describe the main idea of the Gillespie method developed for Markov processes. Suppose there are $K$ chemical compounds $X_i$ entering into multi-stage reactions characterized by a set of chemical channels $R_{\mu}$. Let there be $M$ such channels. According to the algorithm, the multi-component reactions can be broken down into elementary acts. When performing each act, it is necessary to calculate the time after which the next elementary reaction will occur, and also to determine the channel through which the connections come into contact. For Markov processes, the distribution of time intervals between reactions obeys the Poisson distribution:

$$P(\Delta t) = \sum_{\mu=1}^{M} a_\mu \exp \left( -\Delta t \sum_{\mu=1}^{M} a_\mu \right) ,$$

where

$$a_\mu = c_\mu h_\mu$$

is the probability (propensity) that the $\mu$ chemical channel is currently being implemented. As can be seen from definition (6), the value of $a_\mu$ depends on the reaction rate and the kinetics of elementary reaction events. The distribution of probabilities is discrete and is normalized to the value of a reliable event, which is determined by the sum of the probabilities of all reactive events (some reaction from the set will inevitably occur anyway):

$$P(\mu = \mu^\ast) = \frac{a_{\mu^\ast}}{\sum_{\mu=1}^{M} a_\mu} .$$

Thus, the algorithm includes a preliminary calculation of propensities (6) based on the known kinetics of chemical reactions. After that, at each time step, two random numbers are generated from the segment $[0,1]$. The first number, according to distribution (5), determines the time of the onset of the nearest reactionary act $\Delta t$, and the second number, according to distribution (7), determines the specific channel $\mu^\ast$ through which the system realizes its dynamics. After completing the time step, the system state vector $X_i$ is updated, and the whole procedure is repeated.
A modification of the algorithm for the case of time-delayed processes has been proposed by one of the authors in [7]. In fact, it can be easily generalized to the case of arbitrary non-Markovian processes. The main idea of improving the algorithm is illustrated in Figure 2. Suppose one of the channels is time-delayed. If a delayed reaction queue drops out, then it is not executed, and a record about it is pushed onto the memory stack to be executed at time $\Delta t + \tau$. This step looks simple and logical. It is more difficult to determine what to do if the next reaction is to be performed, but in the time interval $[t, t + \Delta t]$ the delayed reaction is scheduled to be executed. It was shown in [7] that it is correct to proceed as follows: the last value for the time step should be ignored, and the system forcibly jumps to a point in time that was previously planned but delayed (see diagram in Figure 2).

It should be noted that any kind of a non-Markov behavior seriously complicates the calculation procedure since during the calculation it is necessary to keep a huge amount of data in memory. Even for the simplest non-Markov type, dynamics with a fixed lag time, hundreds of thousands of deferred reactions quickly accumulate in the memory stack. In addition, in gene regulation, the reactions with long lag times, which lead to nontrivial effects, are of particular interest.

4. Temporal dynamics of the repressilator

We have shown in [12] that the system of DDEs (1–3) has a single steady-state $\{x^*, y^*, z^*\}$, the stability of which can only change in an oscillatory manner. In the case of a symmetric repressilator, the neutral curve for the Hopf bifurcation is implicitly given by the following algebraic equation:

$$
\text{Re} \left[ \Lambda \left( \frac{-2\tau \varepsilon \delta X^* B^2}{A(1+4\varepsilon X^*)} e^{1+4\varepsilon X^*} \right) \right] - \frac{\tau B}{1+4\varepsilon X^*} = 0, \tag{8}
$$

where we assumed that $x^* = y^* = z^* = X^*$. Here, $\Lambda$ stands for the Lambert function.

Consider a typical cross-section of the stability map shown in Figure 3. At the protein degradation $B = 5$, the neutral value of the protein production is $A^* = 70.81$. Figure 4a shows a comparison between deterministic and stochastic simulations performed slightly above the neutral curve at $A = 80$ (see Figure 3). Here we found the effect of spontaneous degradation of oscillations resulting in arbitrary changes in the phase of oscillations. One can see that the standard averaging over realizations used in stochastic analysis leads to a distortion of the perception: the graph can be interpreted as damping of oscillations (Figure 4a). In contrast, Figure 4b shows one specific run of stochastic simulation that demonstrates full-scale oscillations in the system. It can be explained by the spontaneous degeneration of oscillations occurring from time to time. After some time, the oscillations are restored, however, their phase changes randomly concerning the initial oscillations. When a signal is averaged over a large ensemble of realizations, we obtain an oscillatory process that differs significantly from deterministic oscillations.

Figure 3. Neutral curve of the oscillatory instability of a symmetric repressilator with a delay in the plane of protein production and degradation, obtained within the deterministic description given by (1–3). The instability is above the curve. The black squares correspond to the parameter values for which the stochastic calculations are performed: $\varepsilon = 0.1$, $\delta = 0.2$, $\tau = 20$, $B = 5$, $A = 70$ (below the curve), $A = 80$ (above the curve).
The random shifts of the oscillation phase resulting in a mean-field decay on the background of growing oscillations of individual stochastic elements have been discussed in [14,15] in the context of a stochastic parametric resonance in detail.

In the subcritical region, we can observe the existence of quasi-regular oscillations, which is illustrated by the autocorrelation function of the signal calculated at $A = 70$ (the point below the neutral curve in Figure 3) and shown in Figure 5. Autocorrelations indicate that the dynamic system is sensitive to synchronous oscillations with a period of $2\tau$. An interesting feature of the process is a gradual increase in autocorrelations for more and more distant points of the time series, which indicates a gradual self-tuning of the repressilator.

5. Spatial dynamics of the repressilator
To get insight into the spatial effects generated by stochasticity, we developed a hybrid model that is constructed as follows. The protein dynamics in a single cell is obtained by performing direct Gillespie simulations of a repressilator model given by a list of the reactions (see Table 1). The signaling between cells is organized as diffusive transport from one cell to the other. For simplicity, we assume that monomers of the basic protein X seep through cell membranes according to the finite-difference formula:
Figure 6. Evolution of the pattern formed by the X protein of the repressilator in two-dimensional tissue consisting of more than 1500 cells obtained in the stochastic simulation. The frames from left to right and from up to down correspond to times representing the dynamics within one oscillatory period. The fixed parameters are $\tau = 20$, $B = 5$, $A = 70$, $k_1 = 100$, $k_{-1} = 1000$, $k_2 = 200$, $k_{-2} = 1000$, $D = 0.05$. The parameter values are taken below deterministic Hopf bifurcation (see Figure 3).

\[
X_j^{t+\Delta t} = X_j^t + D\Delta t \sum_{(j \in \text{adj}(j))} L_{ij} \left( X_i^t - X_j^t \right), \tag{9}
\]

where $\lceil \ldots \rceil$ stands for the ceiling function which maps the smallest integer not less than the function argument, the subscripts refer to $i$th and $j$th cells, $D$ is the diffusion coefficient and $\text{adj}(j)$ stands for “adjacent to $j$-cell”. It is assumed in (9) that the signaling species is transported diffusively from one cell to the other, whereas its flux does not depend on the distance between the two cells $i$ and $j$ but is proportional to the boundary length $L_{ij}$. This implies that the transport is limited by the transfer through cell membranes. The initial configuration of the system is a regular hexagonal lattice comprising 1560 cells. The shape and location of each cell are defined by its nodes. The tissue as a whole has the form of a stripe with two free borders with periodic boundary conditions applied there.

As an example of our stochastic simulations, consider the case when the deterministic description of the system predicts the stationary behavior (it is indicated by the lower black square in Figure 3). Starting with random initial conditions, the system fairly quickly falls into a fully synchronized mode of oscillations with a common period close to $2\tau$. We found also the effect of clustering when the cells split into two approximately equal communities, which collectively oscillate in anti-phase (Figure 6).
In fact, the clustering in the system with a large number of elements exchanging chemical signals has become at the center of attention of many scientists recently. For example, the group of synthetic genetic oscillators has been studied in [16]. Tissue was found to be divided into two groups of oscillating cells. It is believed that clustering is likely to be the most important characteristic of most communities and could be the reason for further cell differentiation in organs.

6. Conclusion
In this paper, we found that stochastic analysis based on the modified Gillespie algorithm can discover new types of the dynamic behavior of the system under the consideration. It turns out that the interaction of noise and delay can lead to not only quantitative but also qualitative changes. Near the neutral curve, the system demonstrates the process of spontaneous degradation and excitation of periodic oscillations, while each time the phase of the restored oscillations changed unpredictably. We have shown that the standard averaging of this dynamic behavior over the realizations leads to quasi-stationary behavior, which is not true. In the case of higher supercriticalities, we found that noise contributes to a more efficient self-tuning of the symmetric repressilator to the joint work of genes.

The phenomenon is somewhat related to stochastic resonance, in which external noise forces the system to switch to periodic oscillations, i.e. make a qualitative transition in self-organization. With stochastic resonance, a sufficient amplitude of the noise, at which the transition occurs, is important. In our case, the noise is the intrinsic noise of chemical reactions, which cannot be regulated externally. However, the nature of gene regulation is such that fluctuations in protein fields are large. This results in the transition of the system to the teamwork of repressilator genes.

Acknowledgments
This work was supported by the Ministry of Science and Higher Education of the Russian Federation (project No. FSNM-2020-0026).

References
[1] Elowitz M B and Leibler S 2000 Nature 403 335
[2] Hasty J, Dolnik M, McMillen D, Collins J J 2001 Chaos 11 207
[3] Fung E, Wong W W, Suen J K, Bulter T, Lee S and Liao J C 2005 Nature 435 118
[4] Friedland A E, Lu T K, Wang X, Shi D, Church G and Collins J J 2009 Science 324 1199
[5] Bonnet J, Yin P, Ortiz M E, Subsoontorn P and Endy D 2013 Science 340 599
[6] Potvin-Trottier L, Lord N D, Vinnicombe G and Paulsson J 2016 Nature 538 514
[7] Bratsun D, Volfson D, Hasty J and Tsimring L 2005 Proc. Natl. Acad. Sci. U.S.A. 102 14593
[8] Denault D, Loros J and Dunlap J 2001 EMBO J. 20 109
[9] Sriram K and Gopinathan M S 2004 J. Theor. Biol. 231 23
[10] Bratsun D A, Merkuriev D V, Zakharov A P and Pismen L M 2016 J. Biol. Phys. 42 107
[11] Danino T, Mondragón-Palomo O, Tsimring L and Hasty J 2010 Nature 463 326
[12] Bratsun D A, Lorgov E S and Poluyanov A O 2018 Comput. Res. Model. 10 241
[13] Gillespie D T 1977 J. Phys. Chem. 81 2340
[14] Klyatskin V I 2017 Fundamentals of Stochastic Nature Sciences (Berlin: Springer)
[15] Khamsinskii R 2012 Stochastic stability of differential equations (Berlin: Springer)
[16] Koseska A, Ullner E, Volkov E, Kurths J and Garcia-Ojalvo J 2010 J. Theor. Biol. 263 189