Calculation of switching times in the genetic toggle switch and other bistable systems

Baruch Barzel and Ofer Biham
Racah Institute of Physics, The Hebrew University, Jerusalem 91904, Israel
(Received 24 January 2008; revised manuscript received 22 July 2008; published 28 October 2008)

Genetic circuits with feedback such as the toggle switch often exhibit bistability, namely, two stable states with rare spontaneous transitions between them. These systems can be characterized by the average time between such transitions (referred to as the switching time). However, commonly used deterministic models, based on rate equations, do not account for these fluctuation-induced transitions. Stochastic methods, such as the direct integration of the master equation, do account for the transitions. However, they cannot be used to evaluate the switching time. In order to obtain the switching time, one needs to use Monte Carlo simulations. These methods require the accumulation of statistical data, which limits their accuracy. They may become infeasible when the switching time is long. Here we present an accurate and efficient method for the calculation of the switching time. The method consists of coupled recursion equations for the transition times between microscopic states of the system. Using a suitable definition of the two macroscopic bistable states (in terms of the microscopic states) and the probabilities obtained from the master equation, the method provides the switching time between the two states of the system. The method is demonstrated for the genetic toggle switch. It can be used to evaluate the switching times in a broad range of bistable and multistable systems. We also show that it is suitable for the evaluation of the oscillation periods in oscillatory systems such as the repressilator.

DOI: 10.1103/PhysRevE.78.041919 PACS number(s): 87.10.−e, 87.16.−b

I. INTRODUCTION

Multistability is common in nonlinear systems with feedback loops, particularly those in which the overall feedback throughout the loop is positive. A multistable system exhibits several macroscopic states in which it may remain for a long time. It may switch between these states either spontaneously or as a result of an external trigger. The typical time between spontaneous transitions is one of the essential features that characterize a multistable system. Examples of bistable systems appear in the context of genetic regulatory networks, which describe the interactions between genes and their products in living cells [1]. These genes regulate each other's expression using a combination of transcriptional, post-transcriptional and post-translational regulation mechanisms. Here we focus on transcriptional regulation networks, where transcription factor proteins bind to specific promoter sites on the DNA and regulate the transcription of adjacent genes [2]. In the case of positive regulation, the bound transcription factor activates the transcription while in the case of negative regulation the transcription is suppressed. Genetic networks are complex and typically include thousands of genes. In order to understand the functionality of the network in the cell, it is useful to identify functional modules and analyze each one of them separately [3–5]. Such network modules are commonly simulated using rate equations which are simple and efficient [6–9]. They consist of coupled ordinary differential equations, which account for the temporal variations in the concentrations of molecules such as messenger RNA (mRNA) and proteins. For fixed environmental conditions, these equations often approach a single steady state solution. When some control parameter is varied, a bifurcation may take place, where two distinct stable solutions emerge. Beyond the bifurcation point, the system thus exhibits bistability [10–13]. Within the rate equation approach, these two solutions are completely stable. Thus, the rate equations do not account for spontaneous transitions, which result from stochastic fluctuations. Such fluctuations are important in genetic circuits due to the fact that transcription factors and their binding sites may appear in low copy numbers [14–18].

In order to account for the spontaneous switching events one needs to use stochastic methods. These methods can be implemented either by direct integration of the master equation [19–21] or by Monte Carlo (MC) simulations [22–25]. An advantage of the direct integration of the master equation is that it provides the complete probability distribution over the microscopic states of the system. From the moments of this distribution one can obtain the averages and standard deviations of the molecular concentrations as well as the rates of biochemical reactions. However, the solution of the master equation does not provide dynamical features such as the switching times. In order to obtain the switching times one must perform MC simulations and collect a sufficient amount of statistical data. This is a difficult task because the required simulation time increases with the switching time.

The recently proposed forward flux sampling technique provides a dramatic reduction in the simulation time required for sampling these rare switching events [26]. This technique is based on a ratchetlike mechanism. The transition path between the stable states is divided into sections and the simulation is advanced across each of these sections. The statistical information is then analyzed to obtain the average switching time.

In this paper we present a different method for the calculation of the switching times in bistable and multistable systems. The method is based on a set of coupled recursion equations, which includes one equation for each microscopic state of the system. The equation associated with each of the microscopic states provides the average time it takes the system to get from this state to a predefined target state. The

target state is chosen as one of the multistable states, which consists of a suitable set of microscopic states. The switching time from one of the multistable states to another is obtained using a suitable weighted average, with weights obtained from the master equation.

Unlike MC simulations, this method does not require a stochastic implementation which involves the collection and analysis of statistical data. The number of equations is determined by the size of the state space, with a suitable truncation. Thus, unlike MC simulations, the computational resources required by this method do not scale with the switching time. Furthermore, the results are exact. We demonstrate the method for the calculation of the switching time in the genetic toggle switch. This method can be easily adapted to other bistable and multistable systems. We show that it can also be used to calculate the oscillation periods of oscillatory systems.

This paper is organized as follows. In Sec. II we introduce the genetic toggle switch. In Sec. III we show how to calculate the average switching time using the recursion equation method and compare the results to those obtained from MC simulations. In Sec. IV we apply the method to the calculation of the oscillation period of the repressilator circuit. The results are summarized and discussed in Sec. V.

II. THE GENETIC SWITCH

The genetic toggle switch consists of two genes x_1 and x_2, which negatively regulate each other's expression by transcriptional regulation. Gene x_i transcribes mRNA molecules m_i, which are translated into repressor proteins X_i, $i=1,2$. For simplicity, in the model used here, the rates of transcription and translation are combined together into a protein synthesis rate g_i (s^{-1}). The degradation rates of these proteins are denoted by d_i (s^{-1}). The binding rate of X_1 (X_2) proteins to the promoter site that regulates the expression of x_2 (x_1) is b_1 (b_2). When a repressor protein X_1 is bound to the promoter site of x_2, the synthesis of X_2 proteins is suppressed (and vice versa). Finally, a bound X_i protein may unbind at a rate u_i (s^{-1}). For simplicity we focus on the symmetric case in which the parameters of the two genes are identical, namely $g_i=g$, $d_i=d$, $b_i=b$, and $u_i=u$, where $i=1,2$.

For a certain range of parameters such genetic systems feature two distinct stable states. One state is reached when the X_1 proteins become dominant. As a result, the promoter site regulating the gene x_2 is more likely to be occupied by an X_1 repressor, causing the suppression of the X_2 proteins. The other possible state emerges when X_2 proteins are dominant and X_1 is suppressed. In this paper we use the following parameter values: $g=0.05$, $d=0.005$, $b=0.1$, and $u=0.005$ (s^{-1}). These are typical values for bacteria such as *Escherichia coli* [3]. The protein formation rate represents typical synthesis times of proteins which are of the order of 10 to 20 s. The degradation rate is consistent with typical half-life times of proteins, which are of the order of several minutes. The binding rate represents the time scale of diffusion across the cell and a specific binding to the promoter site on the DNA. The unbinding rate represents the residence time of the transcription factors on the promoter site, of the order of several minutes.

An example of a genetic switch appears in the λ phage, which infects *E. coli* and other bacteria [1]. It can exist in two exclusive states, one called lysogeny and the other called lysis. When the phage enters its host, it integrates itself into the host's DNA and is duplicated by cell division. It codes for proteins that can identify stress in the host cell. In case of stress, the phage transforms into the lysis state. In this state, it kills the host cell, using its DNA to produce many copies of the phage, which are released and later infect other cells. Other switch circuits exist in the metabolic systems of cells and determine, for example, which type of sugar the cell will digest [27]. The genetic switch may also serve as a memory unit of the cell, and help determine its fate during cell differentiation [28].

Recent advances in synthetic biology enable the construction of genetic circuits with desired properties, that are determined by the network architecture. These networks are constructed from available components, namely, genes and promoters. They do not require the manipulation of the structure of proteins and other regulatory elements at the molecular level. These genes and promoters are often inserted into plasmids rather than on the chromosome. A synthetic toggle switch, that consists of two repressible promoters with mutual negative regulation, was constructed in *E. coli* and the conditions for bistability were examined [6]. The switching between its two states was demonstrated using chemical and thermal induction.

Several variants of the genetic switch have been studied theoretically in recent years using deterministic and stochastic methods. The basic circuit described above is called the general switch [10,11]. A closely related circuit is the exclusive switch in which there is some overlap between the promoter sites of the two genes. As a result, the X_1 and X_2 repressors cannot be bound simultaneously. In yet another variant, the regulation is performed when several proteins bind simultaneously to the promoter. This property is termed cooperative binding [1]. Deterministic analysis of the general switch and the exclusive switch, using rate equations, showed that they exhibit bistability only in presence of cooperative binding [6,7,10–12]. However, stochastic analysis of the exclusive switch showed that it exhibits bistability even without cooperative binding [29,30]. Here we focus on the case of the exclusive switch without cooperative binding, namely, the regulation is performed by single repressor proteins.

A graphic representation of the exclusive switch appears in Fig. 1. We denote the average copy number of X_i proteins per cell in a population of genetically identical cells by $\overline{N_i}$. The average number of bound X_i proteins is denoted by $\bar{r}_i$. Clearly, in this formulation $\bar{N}_i$ and $\bar{r}_i$ are real, positive numbers and $0<\bar{r}_i<1$. The rate equations that describe the exclusive switch are

$$\frac{d\overline{N_1}}{dt} = g_1(1-\bar{r}_2) - d_1\overline{N_1} - b_1(1-\bar{r}_1-\bar{r}_2)\overline{N_1} + u_1\bar{r}_1, \quad (1a)$$

$$\frac{d\overline{N_2}}{dt} = g_2(1-\bar{r}_1) - d_2\overline{N_2} - b_2(1-\bar{r}_1-\bar{r}_2)\overline{N_2} + u_2\bar{r}_2, \quad (1b)$$

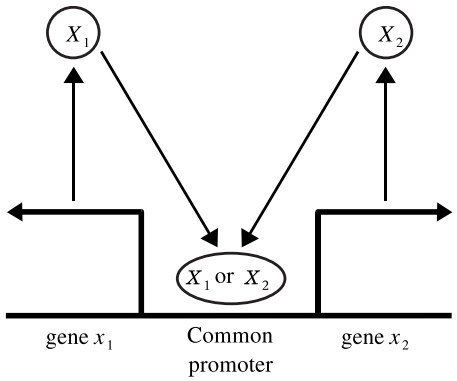

FIG. 1. The exclusive switch consists of two genes, from which the repressor proteins X_1 and X_2 are synthesized. The binding of X_1 to the promoter site represses the production of X_2 and vice versa. The binding is exclusive, namely only one of the repressors may be bound at any given time.

$$\frac{d\bar{r}_1}{dt} = b_1(1-\bar{r}_1-\bar{r}_2)\overline{N_1} - u_1\bar{r}_1, \tag{1c}$$

$$\frac{d\bar{r}_2}{dt} = b_2(1-\bar{r}_1-\bar{r}_2)\overline{N_2} - u_2\bar{r}_2. \tag{1d}$$

In Eqs. (1a) and (1b) the first terms describe the synthesis of proteins, which takes place only when the corresponding promoter site is vacant. The second terms describe protein degradation. The third terms describe the binding of free proteins to the promoter sites. These sites may hold at most one protein each, and only one of them may be occupied at any given time. Thus the binding process takes place only when both sites are vacant. The fourth terms describe the unbinding of proteins from the promoters. In Eqs. (1c) and (1d) the first and second terms describe the effect of binding and unbinding, respectively, on the number of bound repressors. Under steady state conditions, these equations turn out to exhibit a single steady state solution [30]. Thus, within the rate equation formulation the exclusive switch does not exhibit bistability.

In order to show that the exclusive switch without cooperative binding does exhibit bistability, stochastic methods based on the master equation are required. The instantaneous state of the switch at time t is given by $\vec{N}(t) = (N_1, N_2, r_1, r_2)$, where $N_i = 0, 1, 2, \ldots$ $(i=1,2)$ is the copy number of free proteins of type X_i, and $r_i = 0, 1$ is the copy number of bound proteins of this type. In the exclusive switch, either $r_1 = 1$ or $r_2 = 1$, but not both of them at the same time. The master equation is expressed in terms of the probabilities $P(N_1, N_2, r_1, r_2)$ for the system to be in each one of the states $\vec{N} = (N_1, N_2, r_1, r_2)$. It takes the form

$$\begin{aligned}\dot{P}(\vec{N}) = \sum_{i=1}^{2} \{ & g_i[(1-r_j)P(\ldots, N_i - 1, \ldots) \\ & - (1-r_j)P(N_1, N_2, r_1, r_2)] \\ & + d_i[(N_i+1)P(\ldots, N_i+1, \ldots) - N_i P(N_1, N_2, r_1, r_2)] \\ & + b_i[(N_i+1)r_i(1-r_j)P(\ldots, N_i+1, \ldots, r_i-1, \ldots) \\ & - N_i(1-r_i)(1-r_j)P(N_1, N_2, r_1, r_2)] \\ & + u_i[(1-r_i)(1-r_j)P(\ldots, N_i-1, \ldots, r_i+1, \ldots) \\ & - r_i(1-r_j)P(N_1, N_2, r_1, r_2)]\},\end{aligned} \tag{2}$$

where $j = 3-i$. The first term accounts for the transcription and translation processes resulting in the formation of X_1 and X_2 proteins. The second term accounts for protein degradation. The last two terms describe the processes of binding and unbinding of the repressors to/from the promoter site. Note that the formation of X_1 (X_2) proteins takes place only when $r_2 = 0$ $(r_1 = 0)$. In numerical simulations the master equation is truncated such that $N_i = 0, 1, 2, \ldots, N_i^{\max}$. The cutoffs $N_i^{\max}$, $i=1,2$, are chosen such that the probability of the system to occupy states beyond the cutoffs is negligible. Direct integration of the master equation provides the complete probability distribution at steady state. We also define the marginal probability distribution

$$P(N_1, N_2) = \sum_{r_1=0}^{1} \sum_{r_2=0}^{1} P(N_1, N_2, r_1, r_2). \tag{3}$$

This distribution, under steady state conditions, is shown in Fig. 2. It exhibits two peaks, one dominated by X_1 proteins and the other dominated by X_2 proteins. In the domain between these peaks the probability nearly vanishes. This indicates that under these conditions the system is bistable. However, the switching time between the two states cannot be extracted directly from the solution of the master equation.

In the master equation the bistable states are not uniquely defined. Each bistable state is a macroscopic state which consists of a bunch of microscopic states of high probability. The two bistable states are separated from each other by a corridor in which all states are of low probability. In a symmetric system such as the genetic switch, it is possible to divide the state space symmetrically. Using such division, we define the macroscopic states $S_1 = \{\vec{N} = (N_1, N_2, r_1, r_2) | N_1 > N_2, r_1 = 1\}$ and $S_2 = \{\vec{N} = (N_1, N_2, r_1, r_2) | N_2 > N_1, r_2 = 1\}$ as the two bistable states of the switch, dominated by X_1 and X_2 proteins, respectively.

III. CALCULATION OF SWITCHING TIMES

The commonly used method to obtain the switching times is MC simulations. In MC simulations of the switch system, the instantaneous state of the system is represented by a site on the four-dimensional lattice $\vec{N} = (N_1, N_2, r_1, r_2)$. In each MC step the state of the system is updated from $\vec{N}$ to some $\vec{N}'$, where one of the following processes takes place: formation of a new protein, degradation of an existing protein, binding of a free protein or unbinding of a bound protein. Each process is characterized by a suitable rate. The synthesis rate of protein X_1 is $R(\vec{N} \to \vec{N}') = g_1(1-r_2)$ where $\vec{N}' = (N_1+1, N_2, r_1, r_2)$. The degradation rate of protein X_1 is $R(\vec{N} \to \vec{N}') = d_1 N_1$, where $\vec{N}' = (N_1-1, N_2, r_1, r_2)$. The binding rate of protein X_1 is $R(\vec{N} \to \vec{N}') = b_1 N_1 (1-r_1-r_2)$, where $\vec{N}'$

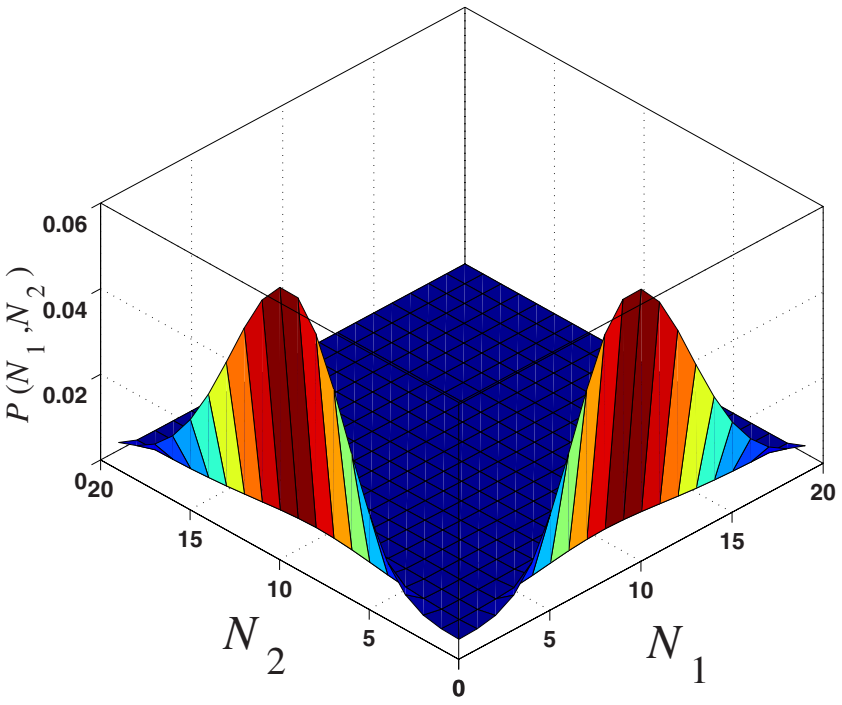

FIG. 2. (Color online) The probability distribution $P(N_1,N_2)$ as obtained from the master equation for the exclusive switch. Two distinct peaks are observed, one dominated by X_1 proteins and the other is dominated by X_2 proteins. The probability density in the range between these peaks is very low, indicating that the system is bistable.

$=(N_1-1,N_2,r_1+1,r_2)$. The unbinding rate of protein X_1 is $R(\vec{N}\to\vec{N}')=u_1r_1(1-r_2)$, where $\vec{N}'=(N_1+1,N_2,r_1-1,r_2)$. The rates of processes involving X_2 proteins can be expressed in a similar way. When the MC step is executed, one of these processes is chosen, with a probability proportional to its rate. The probability that the move $\vec{N}\to\vec{N}'$ will be chosen is

$$P(\vec{N}\to\vec{N}')=\frac{R(\vec{N}\to\vec{N}')}{\sum_{\vec{N}''}R(\vec{N}\to\vec{N}'')}. \tag{4}$$

The rate of exiting the state $\vec{N}$ is equal to the sum of the rates of the individual processes. Therefore, the average time it takes the system to exit the state $\vec{N}$ is given by

$$\tau(\vec{N})=\left[\sum_{\vec{N}''}R(\vec{N}\to\vec{N}'')\right]^{-1}. \tag{5}$$

However, the process of exiting the state $\vec{N}$ is stochastic and Markovian. Therefore, the elapsed time t should be advanced by Δt, drawn from the distribution [22,23]

$$P(\Delta t)=\frac{1}{\tau(\vec{N})}e^{-\Delta t/\tau(\vec{N})}. \tag{6}$$

In Fig. 3 we present the copy numbers of the free and bound proteins vs time as obtained from a MC simulation. Clearly, the system behaves as a switch with two distinct stable states. One state is dominated by X_1 proteins (and $r_1=1$), while the other state is dominated by X_2 proteins (and $r_2=1$). Once in a while, a spontaneous and abrupt transition between the two states takes place. To understand the transition process, consider the case in which the switch is in the state S_1. In this case the X_1 proteins are abundant ($N_1\gg 1$), one of them is bound to the x_2 promoter ($r_1=1$) and the X_2 proteins are suppressed ($N_2\ll N_1$). When the bound X_1 protein unbinds, it is most likely to be replaced by another X_1 protein rather than an X_2 protein. However, in some rare cases (with probability $\sim N_2/N_1$) an X_2 protein may bind. As a result, the synthesis of X_2 proteins will be enabled while the production of X_1 proteins will be suppressed. On average, the X_2 protein stays bound for a period of $1/u_2$ s. The expectation value for the number of X_2 proteins synthesized during this time is g_2/u_2. In order for the transition to take place, after the bound X_2 protein will unbind, it must be replaced by

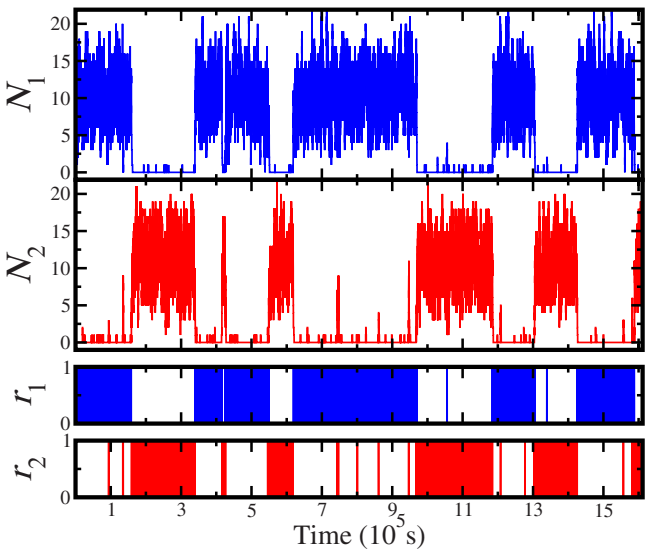

FIG. 3. (Color online) The numbers of free proteins X_1 and X_2 and bound proteins r_1 and r_2 vs time for the exclusive switch. Two distinct states are observed, with spontaneous transitions between them. In each state one protein is dominant and the other is suppressed. Here the switching time is of the order of 10^5 (s).

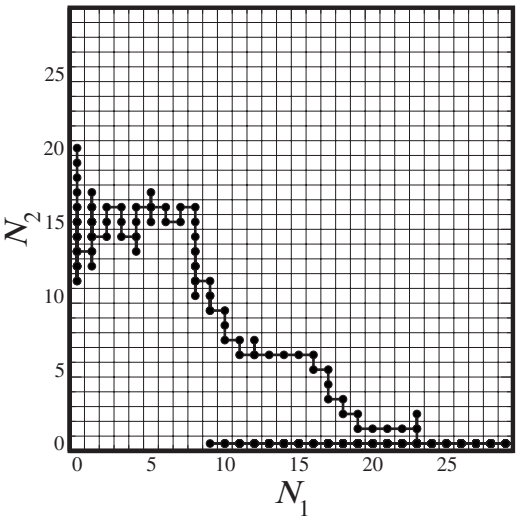

FIG. 4. The transition path from an initial state dominated by X_1 proteins to a final state dominated by X_2 proteins. This path can be viewed as a generalized random walk in the (N_1,N_2) plane.

another X_2 rather than an X_1 protein. The likelihood of this course of events depends on the ratio between the copy numbers of the two proteins at that time. Using this argument, it was shown that the switching time of the exclusive switch, namely, the average time between spontaneous transitions, can be approximated by $bg/(ud^2)$ [30].

The time evolution of the switch can be described by a generalized random walk in the four-dimensional space spanned by $\vec{N}=(N_1,N_2,r_1,r_2)$. In Fig. 4 we show an example of a single transition of the exclusive switch projected on the (N_1,N_2) plane. In each step, the random walk may hop from its current state $\vec{N}$ to one of several adjacent states $\vec{N}'$ which can be reached by the synthesis, degradation, binding, or unbinding of a single X_1 or X_2 protein. The probability of each one of these moves to be chosen is given by Eq. (4).

For each microscopic state $\vec{N}$, we define by $T(\vec{N})=T(\vec{N}\to S_2)$ the average time it takes the system, starting from the state $\vec{N}$, to reach for the first time, one of the microscopic states included in S_2. Clearly, for states that satisfy $\vec{N}\in S_2$, the time $T(\vec{N})=0$. We construct a set of coupled recursion equations which relate the times $T(\vec{N})$ between adjacent microscopic states. For all the states that satisfy $\vec{N}\notin S_2$ the equations take the form

$$T(\vec{N}) = \tau(\vec{N}) + \sum_{\vec{N}'} P(\vec{N}\to\vec{N}')T(\vec{N}'), \tag{7}$$

where $P(\vec{N}\to\vec{N}')$ is given by Eq. (4) and $\tau(\vec{N})$ is given by Eq. (5). For the states that satisfy $\vec{N}\in S_2$ the equations are simply

$$T(\vec{N}) = 0. \tag{8}$$

The recursion equations are linear algebraic equations—one equation for each microscopic state of the system. These equations can also be expressed in a matrix form

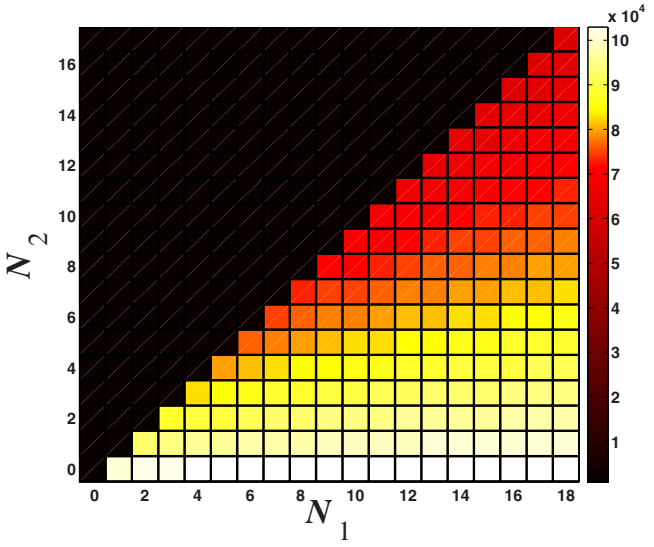

FIG. 5. (Color online) The average transition time $T(\vec{N}\to S_2)$ from the microscopic state $\vec{N}=(N_1,N_2)$ to the bistable state S_2. The results plotted in the lower-right side are for states which belong to S_1, for which $r_1=1$ and $r_2=0$. The results plotted in the upper-left side, are for states which belong to S_2, for which $r_1=0$, $r_2=1$, and $T(\vec{N}\to S_2)=0$.

$$M\vec{T} = \vec{\tau}, \tag{9}$$

where $\vec{T}$ is a vector which consists of the elements $T(\vec{N})$ for all the microscopic states of the system. The vector $\vec{\tau}$ is a vector that consists of all the values $\tau(\vec{N})$ for $\vec{N}\notin S_2$, and has a value of zero for states that satisfy $\vec{N}\in S_2$. The elements of the matrix M for $\vec{N}\notin S_2$, are $M_{\vec{N},\vec{N}'}=\delta_{\vec{N},\vec{N}'}-P(\vec{N}\to\vec{N}')$, and for $\vec{N}\in S_2$, $M_{\vec{N},\vec{N}'}=\delta_{\vec{N},\vec{N}'}$, where $\delta_{\vec{N},\vec{N}'}=1$ if $\vec{N}=\vec{N}'$ and 0, otherwise.

The solution of this set of equations provides the average time required to reach the macroscopic state S_2 from any microscopic state $\vec{N}$ of the system. These average transition times, obtained from Eq. (9), are shown in Fig. 5. The calculation of the switching time from state S_1 to state S_2 requires a combination of the results obtained from Eq. (9) with the steady state solution of the master equation. This switching time is obtained as the weighted average

$$T(S_1\to S_2) = \sum_{\vec{N}\in S_1} P(\vec{N}|\vec{N}\in S_1)T(\vec{N}), \tag{10}$$

where $P(\vec{N}|\vec{N}\in S_1)$ is the probability of the system to be in the microscopic state $\vec{N}$ given that it is in the macroscopic state S_1.

The switching times obtained from the recursion equations (9) and (10) are shown in Fig. 6. The switching times are presented as a function of the binding strength $k=b/u$ [Fig. 6(a)] and of the parameter $q=g/d^2$ [Fig. 6(b)]. The results obtained from the recursion equation method (circles) are in perfect agreement with those obtained from MC simu-

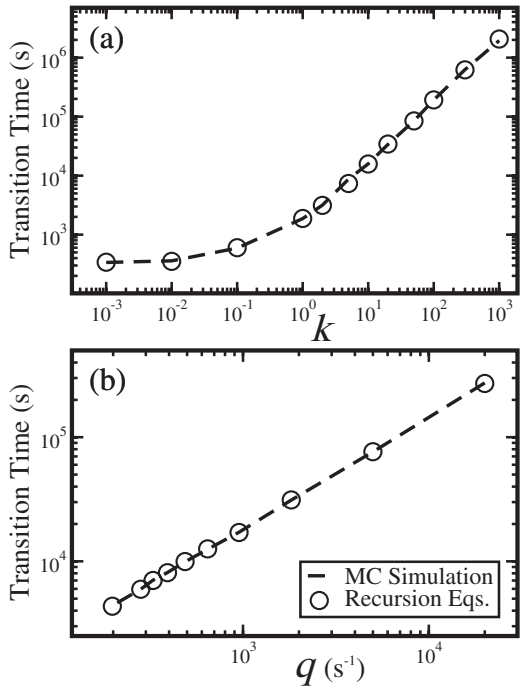

FIG. 6. The switching time for the exclusive switch vs the binding strength k (a) and vs the parameter $q=g/d^2$ (b), as obtained from the recursion equations (circles). The results are in perfect agreement with those obtained from MC simulations (dashed lines).

lations (dashed line). In Fig. 6(a) the parameter u is varied keeping $b=0.1$ (s^{-1}). In Fig. 6(b) the parameters g and d are varied keeping the ratio $g/d=10$ and $k=20$.

The recursion equations method is not limited to bistable systems. It is also applicable to multistable systems with any number of macroscopic states. Consider, for example, a tristable system with macroscopic states S_1, S_2, and S_3. To apply the method, one should first define each macroscopic state in terms of the microscopic states of the system. The next step is to evaluate the average residence time of the system in each one of these states using the recursion equations (9). For example, in order to calculate the residence time in the state S_1, we define the union of S_2 and S_3 as the target state. The residence time is then obtained as the weighted average over all microscopic states in S_1, where the weights are obtained from the master equation.

IV. CALCULATION OF CYCLE PERIODS

Another application of the recursion equations method is for the evaluation of the average cycle periods of stochastic oscillatory systems. In the context of genetic regulatory networks, oscillators are known to be of great importance. A notable example is the circadian clocks, which exhibit oscillations with an average period of about 24 h. To demonstrate how the recursion equations are applied to oscillatory systems we consider the repressilator circuit [9]. The repressilator is a genetic regulatory module consisting of three proteins X_1, X_2, and X_3, which negatively regulate each other's synthesis by transcriptional regulation in a cyclic fashion. More specifically, X_1 regulates X_2, X_2 regulates X_3, and X_3 regulates X_1 (Fig. 7). The production rate of the X_i protein is denoted by g_i (s^{-1}) and its degradation rate is d_i (s^{-1}). The binding and unbinding rates are b_i and u_i (s^{-1}), respectively. Here we assume that bound proteins degrade at a rate d_{r_i} (s^{-1}). For simplicity, we consider the symmetric case in which the parameters of all the genes are identical, namely, $g_i=g$, $d_i=d$, $b_i=b$, $u_i=u$, and $d_{r_i}=d_r$ for $i=1$, 2, and 3. More specifically, the values of the parameter are $g=0.03$, $d=d_r=0.003$, $b=0.5$, and $u=0.01$ (s^{-1}).

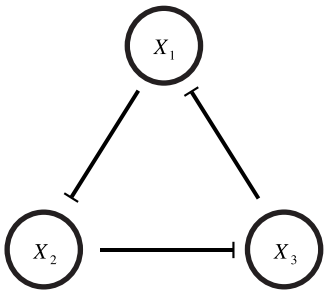

FIG. 7. Schematic plot of the repressilator circuit. It consists of three proteins X_1, X_2, and X_3, which negatively regulate each other in a cyclic manner. The flat-headed arrows denote negative transcriptional regulation.

The repressilator was found to exhibit oscillations between three different states, where each state is dominated by one of the proteins. Consider a situation in which the number of X_1 proteins is large. The negative regulation leads to a reduction in the number of X_2 proteins. As the number of X_2 proteins is reduced, the synthesis of X_3 proteins is enhanced. As a result, the synthesis of X_1 proteins will be repressed, resulting in a cyclic behavior. Deterministic simulations of the repressilator using rate equations show periodic oscillations [9]. However, comparison of the deterministic analysis with the results of stochastic simulations shows that the former does not accurately predict the average period of these oscillations [31]. In order to obtain accurate results for the average oscillation period, one needs to perform MC simulations and collect a sufficient amount of statistical data. The microscopic states of the repressilator system are given by $\vec{N}(t)=(N_1,N_2,N_3,r_1,r_2,r_3)$, where $N_i=1,\ldots,N_i^{\max}$ is the number of free X_i proteins and $r_i=0,1$ is the number of bound X_i proteins ($i=1,2,3$). The repressilator cycles between the three macroscopic states, dominated by X_1, X_3, X_2, and X_1 again. In the transition between the X_1 and X_3 dominated states, the repressilator passes through the transition state $S_{1\to3}=\{\vec{N}=(N_1,N_2,N_3,r_1,r_2,r_3)|N_1=N_3,N_1>N_2\}$. Similarly, in the transition between the X_3 and X_2 dominated states, it passes via $S_{3\to2}=\{\vec{N}=(N_1,N_2,N_3,r_1,r_2,r_3)|N_3=N_2,N_3>N_1\}$, and in the transition between the X_2 and X_1 dominated states, it passes via $S_{2\to1}=\{\vec{N}=(N_1,N_2,N_3,r_1,r_2,r_3)|N_2=N_1,N_2>N_3\}$.

Consider the average transition time Δt needed for the repressilator to advance from $S_{1\to3}$ to the successive transition state $S_{3\to2}$. In the symmetric case considered here, the average period of a complete cycle of the repressilator will be $T=3\Delta t$. In the MC simulations, the transition probabilities $P(\vec{N}\to\vec{N}')$ are defined as in Eq. (4), and the exit time from

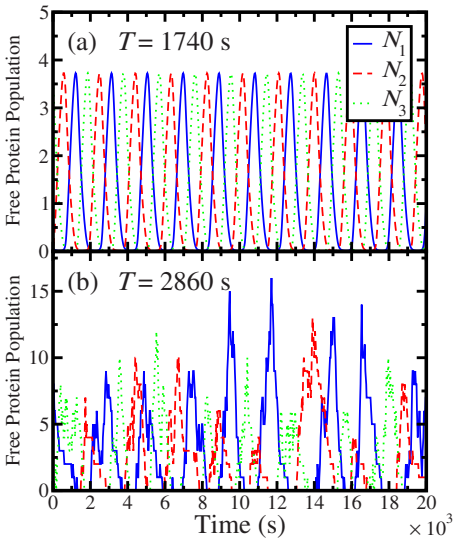

FIG. 8. (Color online) (a) The populations of free proteins from X_1 (solid line), X_2 (dashed line), and X_3 (dotted line) vs time, as obtained from the rate equations, for the repressilator circuit. The system performs regular oscillations with a period of $T \simeq 1740$ (s). (b) The results of a Monte Carlo simulation for the reressillator circuit. Here, stochasticity is taken into account and the oscillations are not regular. The average period of oscillation is $T \simeq 2860$ (s), which is significantly larger than the period predicted by the deterministic approach (rate equations).

the state $\vec{N}$ is defined as in Eq. (5). The period T is obtained by averaging over a large number of cycles.

In Fig. 8(a) we show the copy numbers of the X_1 (solid line), X_2 (dashed line), and X_3 (dotted line) proteins vs time as obtained from the rate equations for the repressilator circuit. Periodic oscillation are observed. The period of these oscillations is $T \simeq 1740$ s. In Fig. 8(b) we show the copy numbers of the X_1, X_2, and X_3 proteins vs time as obtained from MC simulations. Here the average oscillation period is higher, and equals to $T \simeq 2860$ s.

To apply the recursion equation method we solve Eq. (9), where the target state is given by the set of microscopic states that satisfy $\vec{N} \in S_{3\to 2}$. Here we assigned the cutoff for the free proteins to be $N_i^{\max}=19$, leading to a total of $(20 \times 2)^3 = 64\,000$ linear algebraic equations. The result of Eq. (9) is $T(\vec{N} \to S_{3\to 2})$, which is the average time it takes the repressilator to reach $S_{3\to 2}$ starting from an initial state $\vec{N}$. The average time the repressilator spends in a state dominated by the X_3 protein is $\Delta t = \Sigma_{\vec{N} \in S_{1\to 3}} T(\vec{N} \to S_{3\to 2}) P(\vec{N} | \vec{N} \in S_{1\to 3})$. Here, $P(\vec{N})$ is the probability distribution obtained from the master equation and $P(\vec{N} | \vec{N} \in S_{1\to 3})$ is the probability for the repressilator to be at the state $\vec{N}$, given that it is at the macroscopic state $S_{1\to 3}$. For the parameters used here, the result obtained from the recursion equations is $\Delta t = 912$ (s), leading to a cycle period of $T=2736$ (s). This result is in good agreement (within 5%) with the cycle period of $T \simeq 2860$ (s) obtained from the MC simulations.

V. SUMMARY AND DISCUSSION

We have presented a method for the calculation of the switching times in bistable and multistable systems. The method is based on a set of coupled recursion equations, which includes one equation for each microscopic state of the system. These equations provide the average time it takes the system to get from each microscopic state to the target state, chosen as one of the multistable states of the system. The switching time from one multistable state to another is obtained using a suitable weighted average, with weights obtained from the master equation. The method is demonstrated for the calculation of the switching time in the genetic toggle switch. The results are found to be in perfect agreement with those obtained from MC simulations. The method is not limited to multistable systems and can be applied to oscillatory systems as well. To demonstrate its applicability, we have also used the method to obtain the oscillation period of the repressilator. The results were found to be in good agreement with those obtained from MC simulations.

Bistable and multistable systems typically exhibit very long switching times. As a result, Monte Carlo methods in which the running time scales with the switching time of the system may become infeasible. The recursion equations method is advantageous because it does not scale with the switching time. Instead, the complexity of the recursion equations method scales with the number of microscopic states of the system. The method involves the solution of two sets of equations. One set consists of the master equation and the other is the set of recursion equations for the transition times from each microscopic state to the target state. The number of equations in each of the two sets is equal to the number of microscopic states in the system. This number quickly increases with the complexity of the circuit which is analyzed, limiting the applicability of the method for large circuits. However, desktop computers nowadays are capable of solving up to several millions of such coupled linear equations easily, making the method applicable to a broad range of systems. In summary, the recursion equation method and MC simulations complement each other. For the analysis of large circuits with short switching times, MC simulations are more efficient, while for small circuits which exhibit long switching times the recursion equations method is preferable.

An interesting question is whether the method can be generalized to the calculation of the entire probability distribution of switching times. In discrete time systems, such as random walkers, the distributions of first passage times can be calculated using methods that resemble the approach we applied here. This is done by writing down recursion equations for the probabilities $P_{\vec{N}}(t)$, $t=1,2,\ldots$, to reach the target state after exactly t steps, starting from state $\vec{N}$. In this approach, each microscopic state of the system requires to evaluate a vector $P_{\vec{N}}(t)$ instead of a single number $\langle t \rangle_{\vec{N}}$ when only the average is calculated. However, the problem considered here is of continuous time. In this case, the recursion equations become integral equations, which are extremely difficult to solve.

[1] M. Ptashne, *A Genetic Switch: Phage λ and Higher Organisms*, 2nd ed. (Cell Press and Blackwell Scientific Publications, Cambridge, MA, 1992).
[2] M. Ptashne and A. Gann, *Genes and Signals* (Cold Spring Harbor Laboratory Press, New York, 2002).
[3] U. Alon, *An Introduction to Systems Biology: Design Principles of Biological Circuits* (Chapman & Hall/CRC, London, 2006).
[4] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, Science **298**, 824 (2002).
[5] R. Milo, S. Itzkovitz, N. Kashtan, R. Levitt, S. Shen-Orr, I. Ayzenshtat, M. Sheffer, and U. Alon, Science **303**, 1538 (2004).
[6] T. S. Gardner, C. R. Cantor, and J. J. Collins, Nature (London) **403**, 339 (2000).
[7] J. L. Cherry and F. R. Adler, J. Theor. Biol. **203**, 117 (2000).
[8] N. Rosenfeld, M. B. Elowitz, and U. Alon, J. Mol. Biol. **323**, 785 (2002).
[9] M. B. Elowitz and S. Leibler, Nature (London) **403**, 335 (2000).
[10] P. B. Warren and P. R. ten Wolde, Phys. Rev. Lett. **92**, 128101 (2004).
[11] P. B. Warren and P. R. ten Wolde, J. Phys. Chem. B **109**, 6812 (2005).
[12] A. M. Walczak, M. Sasai, and P. Wolynes, Biophys. J. **88**, 828 (2005).
[13] T. B. Kepler and T. C. Elston, Biophys. J. **81**, 3116 (2001).
[14] H. H. McAdams and A. Arkin, Proc. Natl. Acad. Sci. U.S.A. **94**, 814 (1997).
[15] A. Arkin, J. Ross, and H. H. McAdams, Genetics **149**, 1633 (1998).
[16] A. Becskei and L. Serrano, Nature (London) **405**, 590 (2000).
[17] J. Paulsson and M. Ehrenberg, Phys. Rev. Lett. **84**, 5447 (2000).
[18] J. Paulsson, Nature (London) **427**, 415 (2004).
[19] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, North-Holland, 1992).
[20] C. W. Gardiner, *Handbook of Stochastic Methods* (Springer, Berlin, 2004).
[21] B. Munsky and M. Khammash, J. Chem. Phys. **124**, 044104 (2006).
[22] D. T. Gillespie, J. Phys. Chem. **81**, 2340 (1977).
[23] M. E. J. Newman and G. T. Barkema, *Monte Carlo Methods in Statistical Physics* (Clarendon Press, Oxford, 1999).
[24] D. T. Gillespie, J. Chem. Phys. **115**, 1716 (2001).
[25] C. V. Rao and A. P. Arkin, J. Chem. Phys. **118**, 4999 (2003).
[26] R. J. Allen, P. B. Warren, and P. R. ten Wolde, Phys. Rev. Lett. **94**, 018104 (2005).
[27] M. Lewis, C. R. Biol. **328**, 521 (2005).
[28] P. Laslo, C. J. Spooner, A. Warmflash, D. W. Lancki, H.-J. Lee, R. Sciammas, B. N. Gantner, A. R. Dinner, and H. Singh, Cell **126**, 755 (2006).
[29] A. Lipshtat, A. Loinger, N. Q. Balaban, and O. Biham, Phys. Rev. Lett. **96**, 188101 (2006).
[30] A. Loinger, A. Lipshtat, N. Q. Balaban, and O. Biham, Phys. Rev. E **75**, 021904 (2007).
[31] A. Loinger and O. Biham, Phys. Rev. E **76**, 051917 (2007).