Optimal transition paths of phenotypic switching in a non-Markovian self-regulation gene expression

Hongwei Yin [1,2,*] Shuqin Liu,[2] and Xiaoqing Wen[1,2,†]

[1]*School of Mathematics and Statistics, Xuzhou University of Technology, Xuzhou 221111, People's Republic of China*
[2]*School of Science, Nanchang University, Nanchang 330031, People's Republic of China*

Gene expression is a complex biochemical process involving multiple reaction steps, creating molecular memory because the probability of waiting time between consecutive reaction steps no longer follows exponential distributions. What effect the molecular memory has on metastable states in gene expression remains not fully understood. Here, we study transition paths of switching between bistable states for a non-Markovian model of gene expression equipped with a self-regulation. Employing the large deviation theory for this model, we analyze the optimal transition paths of switching between bistable states in gene expression, interestingly finding that dynamic behaviors in gene expression along the optimal transition paths significantly depend on the molecular memory. Moreover, we discover that the molecular memory can prolong the time of switching between bistable states in gene expression along the optimal transition paths. Our results imply that the molecular memory may be an unneglectable factor to affect switching between metastable states in gene expression.

I. INTRODUCTION

Phenotypic switching in gene expression, which allows cells to switch between distinct gene-product states in response to internal or external stochastic signals, has been found in some biological entities, such as bacteria, yeast, and cancer cells [1–4]. In fact, it has been confirmed that different states in living systems, such as stem-cell decision, disease states, and cancer subtypes, are associated with phenotypic switching, whose probability distribution is multimodality [5–7]. Thus, understanding how multiple gene-expression states switch is of significance.

In gene regulatory networks, general mechanisms underlying switching between metastable states of gene expression are positive-feedback gene circuits perturbed by intracellular and extracellular stochastic signals [8–14], for example, signal transduction of λ phage lysis [15], T-cell receptor signaling pathways [16], and the P53 regulatory network [17]. In previous research, the process of gene expression can be mainly described by chemical master equations (CMEs), and for linear biochemical reaction models their analytical solutions can be analytically given out [18,19]. However, analytical results for nonlinear gene regulatory models are impossible to be found since positive feedback in gene regulations are mathematically defined as nonlinear functions (for example, Hill functions). But, this kind of model can be well approximated by Wentzel-Kramers-Brillouin (WKB) method [20], which can efficiently find optimal transition paths of switching between metastable states for multiple-state systems [21].

How metastable states in gene expression dynamics switch each other always attracts researchers' attention [22,23]. The WKB method is very effective to explore this issue because it can capture a path switching between two states in the sense of maximum probability, meaning that this path is the most likelihood. At present, the WKB method has been successfully applied in gene regulatory networks. In detail, Assaf *et al.* [23] employed the WKB method to study the dynamical behaviors of a generic-feedback on-off switch and to obtain accurate results for probability distributions of messenger RNA (mRNA) and protein copy numbers and for the mean switching time. In Ref. [24], they further applied this method to study a model of a Boolean-regulated genetic switch and to obtain the result that extrinsic noise can significantly alter the lifetimes of the phenotypic states. Ruben *et al.* in Ref. [25] used the WKB approximation to research the effect of noise in gene expression on pattering time and boundary position. Lv *et al.* in Ref. [26] made use of the WKB method to study a simplified budding yeast cell cycle model, showing that this cycle is robust.

It is worth noting that the above examples and models always assume that biochemical reactions involved in gene expression to be Markovian processes, that is, next biochemical reactions depend on only current states for a system, and are uninfluenced by previous states. This memoryless assumption on biochemical reactions is based on the fact that the probability distributions of the waiting time between successive biochemical reaction events follow exponential distributions [27,28]. However, gene expression is, in reality, a complex stochastic process with multistep biochemical reaction events [29], such as switching of promoter activity, chromatin remodeling, histone modifications, transcription initiation, alternative splicing, recruitment of transcription factors and polymerases elongating, protein translation,

_____

*hongwei-yin@hotmail.com
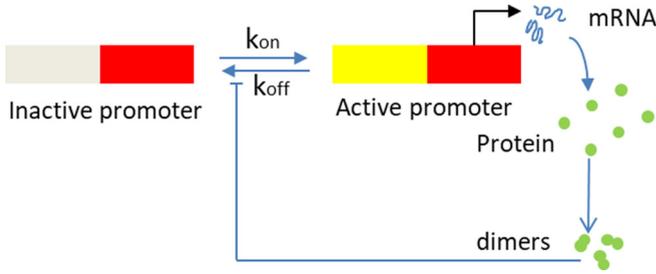†Corresponding author: wendyxiaoqing@sina.com

FIG. 1. A sketch map of gene expression with a self-regulation circuit.

protein dimers, and so on, see Fig. 1. These multistep biochemical reactions can add up to form a nonexponential distribution of waiting time between successive protein production events, implying that the process of gene expression possesses the characteristic of the molecular memory. Moreover, the recent literature [30] showed that the phases of the inactive promoter in the prolactin gene have different probability distributions, implying the existence of the molecular memory during gene transcription. Besides, mRNA degradation may also have the molecular memory since the degradation process consists of multiple small biochemical reactions [31], for example, in *Escherichia coli (E. coli)*, the degradation of mRNAs is mediated by the combined action of endo- and exo-ribonuclease [31]. As a matter of fact, the molecular memory can obviously affect stochastic dynamical behaviors of gene expression. For instance, there exists an optimal strength of the molecular memory that minimizes the Fano factor of gene bursty expression [32]. Also, the molecular memory can induce bimodality in the probability sense for a genetic toggle switch without cooperative binding [33]. Although the molecular memory of gene expression has been investigated, some questions remain open, for example, how would the molecular memory affect optimal transition paths of switching between metastable states in gene expression? To answer this question briefly, in this paper we will introduce a simple gene expression model, which involves a self-regulation feedback and the molecular memory.

So far, there have been a lot of gene regulatory models, which are usually described by the CMEs because they are supposed to be memoryless. Also, there are some works [34–37] to model the molecular memory by using an implicit method, that is, all biochemical reaction events and reaction components involved in gene expression are fully modeled by the CMEs. But, this straightforward method will be difficult to identify parameters' values in the reaction events, and result in a high-dimension system with more expensive computational complexity, whose analytic solution is rarely given out. An alternative approach is to simplify the biochemical reactions in gene regulatory networks into a low-dimension system by explicitly introducing the molecular memory. For example, nonexponential probability distributions are used to describe stochastic waiting times of multistep reaction processes [35,38]. The mathematical form of this approach usually involves complex convolutions and is completely different from the classical CMEs [32,33,39,40].

In this paper, we apply the continuous time random walk (CTRW) theory to model a non-Markovian gene expression

with a positive self-regulation circuit. First, we introduce a generalized waiting-time probability distribution for a process of gene expression, and build a generalized CME model. Second, we obtain a stationary generalized CME model by some complex calculation. Third, for this stationary model we adopt the WKB framework to study how the molecular memory affects the optimal transition paths of switching between metastable states of this model. We find that the molecular memory can significantly affect the optimal transition paths and the switching time.

II. THE NON-MARKOVIAN GENE EXPRESSION WITH A SELF-REGULATION LOOP

Here, we will build a non-Markovian model of gene expression with a positive feedback, as shown in Fig. 1. During the process of gene expression, a gene has two states: the active state (ON) and the inactive state (OFF), which can switch each other [41]. Here, we consider the simplest model of gene expression, in which the gene is always in the active state. Also, introducing generalized probability distributions of waiting time between successive protein copy events can simplify the processes of gene transcription and protein translation. This simplification is from the viewpoint of queuing theory, which has been successfully used in Ref. [38]. For clarity, we list the simplified biochemical reaction events in the process of gene expression as follows:

$$\text{DNA} \xrightarrow{p_1(t)} \text{DNA} + X, \quad X \xrightarrow{p_2(t)} \emptyset, \tag{1}$$
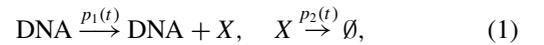
where X represents the gene product; $p_1(t)$ and $p_2(t)$ are the waiting-time probability density functions (PDFs) of producing and degrading proteins, respectively. The mean waiting time of producing proteins is given by $E_1 = \int_0^\infty t p_1(t) dt$, and then $1/E_1$ is the mean rate of producing proteins. $\int_t^\infty p_1(t') dt'$ should be read as the probability that protein's production will not occur in the time interval $[0, t]$, $\int_t^\infty p_2(t') dt'$ being the probability that the degradation of one protein will not occur in $[0, t]$. Then $(\int_t^\infty p_2(t') dt')^n$ is the probability that none of n proteins will degrade in $[0, t]$. Therefore,

$$\psi(t, n) = n p_2(t) \left(\int_t^\infty p_2(t') dt' \right)^{n-1}, \tag{2a}$$

is the waiting-time PDF that one of n proteins will degrade at time t. Without loss of generality, in the microcosmic scale our model always assumes that at an instantaneous moment one protein is either produced or degraded, meaning that protein's production and degradation cannot occur at the same time. Let

$$\phi_1(t, n) = p_1(t) \int_t^\infty \psi(t', n) dt', \tag{2b}$$

represent the waiting-time PDF of the event occurring in which at time t protein's production occurs but its degradation does not when the system has n proteins, and let

$$\phi_2(t, n) = \psi(t, n) \int_t^\infty p_1(t') dt', \tag{2c}$$

represent the waiting-time PDF of the event occurring in which protein's production does not occur but its degradation does at time t for n proteins.

Here, our model considers that the positive self-regulation of gene expression is the Hill function as

$$f(X) = \alpha + \frac{\beta X^m}{X^m + K}, \qquad (3)$$

where α is a basic rate of producing proteins, β is the strength of the feedback, m is the Hill coefficient, K is the parameter of X needed for half-maximum. Here, we define a function

$$F(n) = \Omega f\left(\frac{n}{\Omega}\right), \qquad (4)$$

where n is the protein's copy number, Ω is the system's volume. According to the CTRW theory [42], we can obtain a generalized chemical master equation (seeing Appendix A for detail),

$$\frac{\partial P_n(t)}{\partial t} = \int_0^t \left(\mathbf{E}^{-1} - \mathbf{I}\right) M_1(t - t', n) P_n(t') dt'$$
$$+ \int_0^t (\mathbf{E} - \mathbf{I}) M_2(t - t', n) P_n(t') dt', \qquad (5)$$

where $P_n(t)$ is the probability of having n proteins at time t; $\mathbf{E}$ and $\mathbf{E}^{-1}$ are one-step operators with respect to the state of protein's copy number n. For example, for a state function $\varphi(n)$, we define $E[\varphi(n)] = \varphi(n+1)$ an $E^{-1}[\varphi(n)] = \varphi(n-1)d$. $\mathbf{I}$ is a unit operator. $M_1(t, n)$ and $M_2(t, n)$, respectively, represent the memory functions for protein's production and degradation [42], whose Laplace forms (referring to Appendix A) are given by

$$\widetilde{M}_i(s, n) = \frac{s\widetilde{\phi}_i(s, n)}{1 - \widetilde{\phi}_1(s, n) - \widetilde{\phi}_2(s, n)}, \quad i = 1, 2, \qquad (6)$$

where the notation $\tilde{f}$ denotes the Laplace transform of a function f with respect to time t. For the model (5), we often focus on its steady state. Since the model (5) involves a convolution, we apply the Laplace transform with respect to time t to Eq. (5). Then, one can obtain

$$s\widetilde{P}_n(s) - P_n(0) = \left(\mathbf{E}^{-1} - \mathbf{I}\right)\widetilde{M}_1(s, n)\widetilde{P}_n(s)$$
$$+ (\mathbf{E} - \mathbf{I})\widetilde{M}_2(n, s)\widetilde{P}_n(s). \qquad (7)$$

Multiplying by s on both sides of Eq. (7) and applying the final value theorem of the Laplace transform, $\lim_{s \to 0} s\widetilde{P}_n(s) = P_n$, yield

$$\left(\mathbf{E}^{-1} - \mathbf{I}\right) m_1(n) P_n + (\mathbf{E} - \mathbf{I}) m_2(n) P_n = 0, \qquad (8)$$

where P_n is the solution to the corresponding stationary equation of (5). Let $m_1(n) = \lim_{s \to 0} \widetilde{M}_1(s, n)$ and $m_2(n) = \lim_{s \to 0} \widetilde{M}_2(s, n)$, and

$$m_1(n) = \frac{\int_0^\infty p_1(t)\left[\int_t^\infty \psi(n, t') dt'\right] dt}{\int_0^\infty \left[\int_t^\infty p_1(t') dt' \int_t^\infty \psi(n, t') dt'\right] dt}, \qquad (9a)$$

$$m_2(n) = \frac{\int_0^\infty \psi(n, t)\left[\int_t^\infty p_1(t') dt'\right] dt}{\int_0^\infty \left[\int_t^\infty p_1(t') dt' \int_t^\infty \psi(n, t') dt'\right] dt}, \qquad (9b)$$

referring to Appendix B.

Some biological experiments have implied that the Erlang distribution can better fit with some experimental data than the

exponential distribution [36], where the Erlang distribution is defined as

$$Er(t) = \frac{\mu^k}{\Gamma(k)} t^{k-1} e^{-\mu t}, \quad 0 < t < \infty, \qquad (10)$$

k is an arbitrary positive integer and μ is an arbitrary positive constant. Its mean and variance are k/μ and k/μ^2, respectively. In fact, gene expression following the Erlang distribution has a reasonable biological explanation: the process of producing proteins essentially consists of k stages, through each of which the time to progress is assumed to be the exponential distribution with the mean $1/\mu$ [36], and consequently the process of gene expression follows the convolution of k identical exponential distributions. Moreover, we have the following formula:

$$\overbrace{\mu e^{-\mu t} * \mu e^{-\mu t} * \cdots * \mu e^{-\mu t}}^{k \text{ stages}} = \frac{\mu^k}{\Gamma(k)} t^{k-1} e^{-\mu t} \qquad (11)$$

where $*$ denotes the convolution. When $k = 1$, the Erlang distribution reduces to the exponential distribution with the mean $1/\mu$. As k increases, the Erlang distribution becomes more symmetrical and more closely centered around its mean. To investigate the effect of the molecular memory on optimal transition paths of switching between metastable states in gene expression, we further simplify the waiting-time PDFs of protein's production and degradation, that is, $p_1(t)$ and $p_2(t)$, respectively, follow an Erlang probability distribution and an exponential probability distribution, that is,

$$p_1(t) = \frac{\mu^k}{\Gamma(k)} t^{k-1} e^{-\mu t}, \quad p_2(t) = \delta e^{-\delta t}. \qquad (12)$$

Our assumption means that producing proteins is non-Markovian, and degrading proteins is Markovian. For this case, after performing some calculation, Eq. (9) is rewritten into

$$\psi(t, n) = n\delta e^{-\delta t}\left[\int_t^\infty \delta e^{-\delta t'} dt'\right]^{n-1} = n\delta e^{-n\delta t}, \qquad (13a)$$

$$m_1(n) = \frac{\int_0^\infty \frac{\mu^k}{\Gamma(k)} t^{k-1} e^{-\mu t}\left[\int_t^\infty n\delta e^{-n\delta t'} dt'\right] dt}{\int_0^\infty \left[\int_t^\infty \frac{\mu^k}{\Gamma(k)} t'^{k-1} e^{-\mu t'} dt'\right]\left[\int_t^\infty n\delta e^{-n\delta t'} dt'\right] dt}$$
$$= \frac{\mu^k(n\delta + \mu)^{-k}}{\sum_{i=0}^{k-1}\left[\mu^i(n\delta + \mu)^{-i-1}\right]} = \frac{\mu^k n\delta}{(n\delta + \mu)^k - \mu^k}, \qquad (13b)$$

$$m_2(n) = \frac{\int_0^\infty n\delta e^{-n\delta t}\left[\int_t^\infty \frac{\mu^k}{\Gamma(k)} t'^{k-1} e^{-\mu t'} dt'\right] dt}{\int_0^\infty \left[\int_t^\infty \frac{\mu^k}{\Gamma(k)} t'^{k-1} e^{-\mu t'} dt'\right]\left[\int_t^\infty n\delta e^{-n\delta t'} dt'\right] dt}$$
$$= \frac{\sum_{i=0}^{k-1}\left[\mu^i \delta n(n\delta + \mu)^{-i-1}\right]}{\sum_{i=0}^{k-1}\left[\mu^i(n\delta + \mu)^{-i-1}\right]} = \delta n. \qquad (13c)$$

We want to uncover what effect the molecular memory has on optimal transition paths of switching between metastable states in gene expression. According to the mathematical meaning of Eq. (4), we take $\mu = F(n)$, and then

$$m_1(n) = \frac{[F(n)]^k \delta n}{[n\delta + F(n)]^k - [F(n)]^k}. \qquad (14)$$
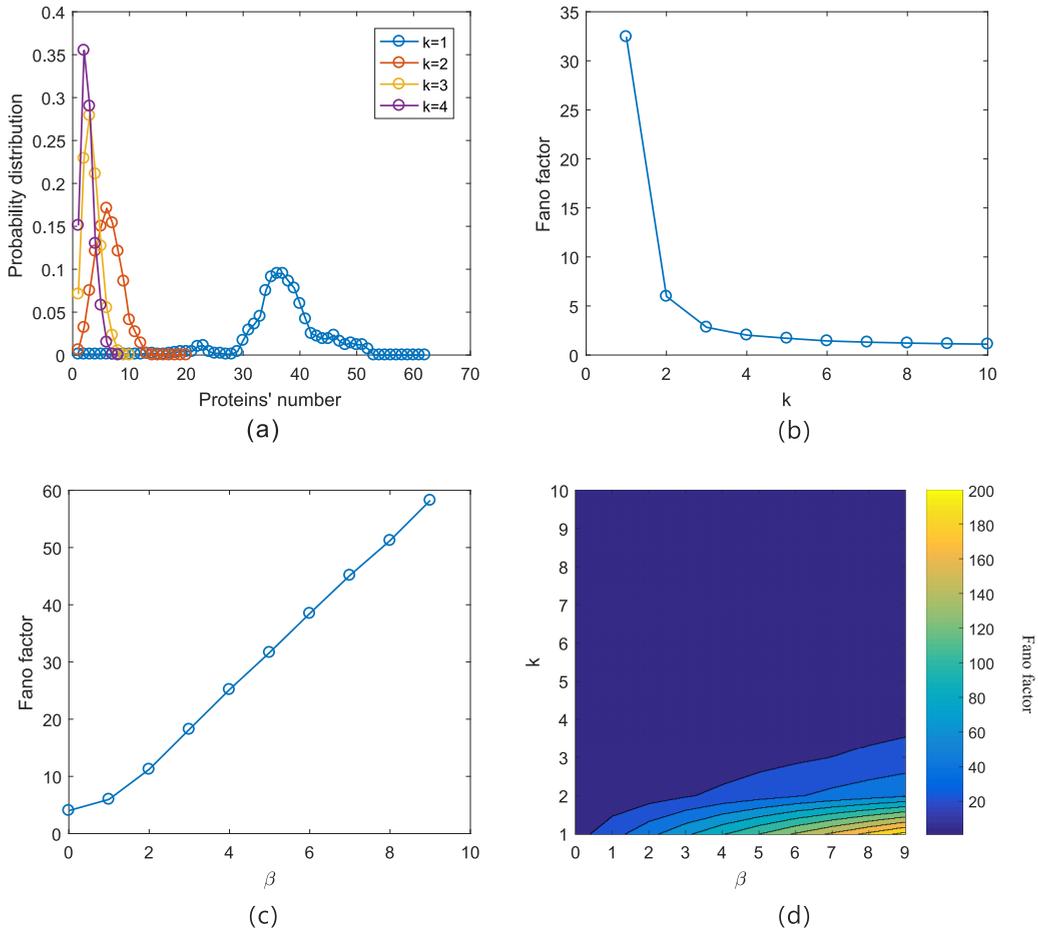
FIG. 2. The stationary probability distributions and the Fano factors of gene expression. (a) The probability distributions for the different molecular memories k. $\alpha = 0.5$, $\beta = 1$, $K = 5$, $m = 2$, $\delta = 0.2$, and $\Omega = 5$. (b) The Fano factors in the protein copy number for the different molecular memory parameter k. $\alpha = 0.5$, $\beta = 1$, $K = 5$, $m = 2$, $k = 2$, and $\Omega = 5$. (c) The Fano factors for the different feedback strength β. (d) The Fano factor versus k and β. $\alpha = 0.5$, $K = 5$, $m = 2$, $\delta = 0.2$, and $\Omega = 5$.

The literature [32,40] has uncovered that the molecular memory is equivalent to a negative feedback. In Fig. 2(a), using a generalized Gillespie stochastic simulation method [43], which can efficiently simulate non-Markovian reaction processes, we plot the probability distributions of the protein copy number for the different strength of the molecular memory k. We observe that under the fixed parameters, the shape of the stationary probability distribution of the protein copy number turns sharp with the increase of k in the waiting-time probability distribution, implying that the molecular memory can repress stochastic fluctuation in protein copy number. Here, we use the Fano factor (F), a natural quantity to measure deviation of a given probability distribution from Poisson distribution with $F = 1$ and defined as the variance divided by the mean, to characterize variability in the gene copy number [5]. We plot the Fano factors of protein's distributions as the function of the molecular memory in Fig. 2(b), showing that the Fano factor decreases with the molecular memory. For $k = 1$ (that is, the case of Markov), the probability distribution of the protein copy number is obviously super-Poisson. Also, for the weaker molecular memory (such as $k = 2$), the probability distributions of the protein copy number are

still super-Poisson. For the stronger molecular memory, the distributions, however, can be sub-Poisson. Also, for the fixed parameter k, we show how the strength of the positive feedback, β, can affect the Fano factor in gene expression in Fig. 2(c). We find that the positive feedback can magnify the Fano factor, as shown in Refs. [44,45]. Then, a natural question is what effect the combination of the molecular memory and the positive feedback has on the probability distribution in protein's copy number. Figure 2(d) demonstrates how the Fano factor in protein's copy number is affected by the parameters k and β. We observe that the larger parameter k can more significantly reduce the Fano factor even when the feedback strength is stronger.

III. WKB APPROXIMATION

Define the protein density x in the system's volume Ω by $x = n/\Omega$. Here, we use the WKB approximation to deal with the stationary system (8). The WKB ansatz reads

$$P_n \approx \exp[-\Omega S(x)], \qquad (15)$$

where $S(x)$ is a mechanical quantity known as the action [46–48], which is defined by

$$S(x) = \int_{t_1}^{t_2} L(x, x') dt, \qquad (16)$$

where $L(x, x')$ is a Lagrangian function and will be given out in Eq. (18). Substituting Eq. (15) into Eq. (8), doing some algebra calculation, and keeping the leading order of Ω, yield the Hamilton equation in the form $H(x, S') = 0$ (see Appendix C), where H is called the Hamiltonian function and given by

$$H(x, \lambda) = M(x)(e^\lambda - 1) + \delta x(e^{-\lambda} - 1), \qquad (17a)$$

where $\lambda = S'$ is a conjugate momentum, and

$$M(x) = m_1(x\Omega)/\Omega = \frac{[F(x\Omega)]^k \delta x}{[x\Omega\delta + F(x\Omega)]^k - [F(x\Omega)]^k}$$

$$= \frac{[f(x)]^k \delta x}{[x\delta + f(x)]^k - [f(x)]^k}. \qquad (17b)$$

Then, the Lagrangian $L(x, x')$ in (16) is given by

$$L(x, x') = \sup_{\lambda \in R} \{x'\lambda - H(x, \lambda)\}. \qquad (18)$$

If $k = 1$ (that is, the process of producing proteins is Markovian), then $M(x)$ reduces into $f(x)$, which is a classical feedback function. But, for $k \geqslant 2$, $M(x)$ always depends on the parameter k, implying that the process of producing proteins has the characteristic of the molecular memory.

In previous studies, which assumed that producing proteins is Markovian, the corresponding Hamiltonian function is composed of a feedback function of producing proteins and a function of degrading proteins [23]. Besides these two functions, in our model (17), the Hamilton function H includes the parameter of the molecular memory k. This is the most significant difference from the previous studies. Next, we can write the canonical equations of motion of motion of Eq. (17a) as follows:

$$\dot{x} = \frac{\partial H}{\partial \lambda} = M(x)e^\lambda - \delta x e^{-\lambda}, \qquad (19a)$$

$$\dot{\lambda} = -\frac{\partial H}{\partial \lambda} = -M'(x)(e^\lambda - 1) - \delta(e^{-\lambda} - 1), \qquad (19b)$$

where

$$M'(x) = \frac{f^k(x)k\delta^2 x^2 (x\delta + f(x))^k}{(-(x\delta + f(x))^k + f^k(x))^2 (x\delta + f(x)) f(x)} \frac{\beta x^m m K}{(x^m + K)^2 x}$$

$$- \frac{\delta f^k(x)((-f(x) + \delta(k-1)x)(x\delta + f(x))^k + f^k(x)(x\delta + f(x)))}{(x\delta + f(x))(-(x\delta + f(x))^k + f^k(x))^2}.$$

The Hamiltonian equations (19) are a two-dimensional dynamical system. For the case without the molecular memory (that is, $k = 1$), its corresponding Hamilton equations (19) are given by

$$\dot{x} = \frac{\partial H}{\partial \lambda} = f(x)e^\lambda - \delta x e^{-\lambda}, \qquad (20a)$$

$$\dot{\lambda} = -\frac{\partial H}{\partial \lambda} = -f'(x)(e^\lambda - 1) - \delta(e^{-\lambda} - 1), \qquad (20b)$$

where

$$f'(x) = \frac{\beta x^m m K}{(x^m + K)^2 x}.$$

The present literature mostly focused on finding optimal transition paths for Hamilton systems without the molecular memory as Eq. (20), and no literature, as far as we know, studies it for the Hamilton systems with the molecular memory like Eq. (19). Equations (19) along the deterministic line of $\lambda = 0$ are described by

$$\dot{x} = \frac{\partial H(x, \lambda)}{\partial \lambda}\bigg|_{\lambda=0} = M(x) - \delta x, \qquad (21)$$

which is a classical rate equation for a deterministic model for gene expression with the molecular memory. For the parameters $\alpha = 0.2$, $\beta = 5$, $K = 10$, $m = 10$, and $\delta = 0.2$, the system (21) with $k = 1$ (i.e, a Markovian process) has an unique positive equilibrium point, but the system (21) with

the molecular memory (such as $k = 2, 3, \ldots, 10$) has three positive equilibrium points (one is unstable, and the other two are all stable), which is shown in Fig. 3(a). This implies that the molecular memory can induce bistable states [33,40]. For the different strength of the molecular memory, the mean-field (21) can have three different positive equilibrium points, denoted, respectively, by x_0, x_1 and x_2 with $x_0 < x_1 < x_2$. Wherein, for the system (21), x_0 and x_2 are stable points, but x_1 is an unstable point. They are sometimes called the low (x_0), intermediate (x_1), and high (x_2) states of gene expression, respectively. To uncover the difference of the positive equilibrium points for the different molecular memories, Fig. 3(b) with $\delta = 0.35$ shows the equilibrium points for the different parameter k, implying that the strength of the molecular memory can reduce the value of the high state (x_2) in gene expression. Furthermore, the Hamiltonian equations (19) also have three fixed states $(x_0, 0)$, $(x_1, 0)$, and $(x_2, 0)$ in the two-dimensional phase plane (x, λ), which, for convenience, are still called the low, intermediate, and high states, respectively. The three fixed states, however, are all unstable saddle points, referring to Fig. 4. In detail, x_0 and x_2 are stable on the local manifold of $\lambda = 0$, and unstable on the local manifold of $\lambda \neq 0$, but the stability of x_1 is the opposite. Here, for the model (19) we are mainly interested in the optimal transition paths of switching between the low and high states of gene expression. According to Eq. (15), the optimal transition path means that the action S is minimum along it, implying that the probability of switching along this path is maximal among
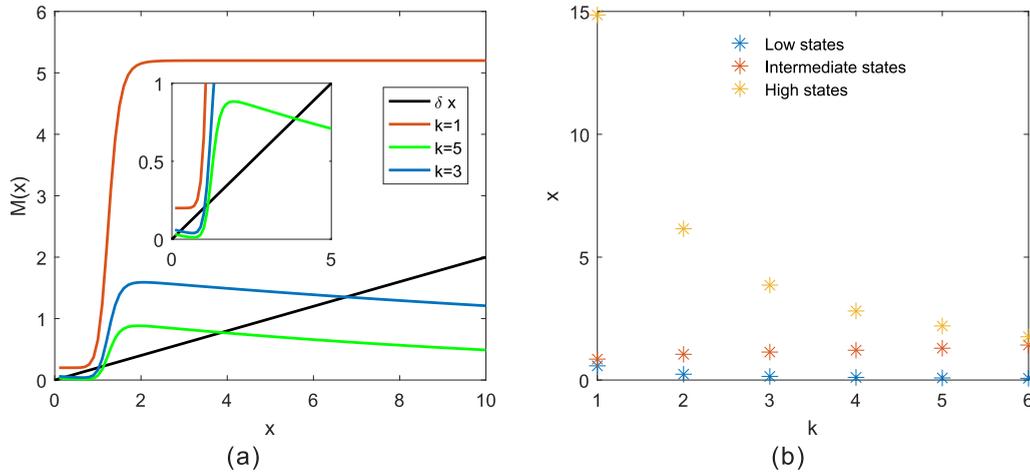
FIG. 3. (a) The curves of the functions in the system (21). The black line is δx. The red curve is the feedback function $f(x)$ without the molecular memory, the green and blue curves corresponding to the feedback functions with $k = 5$ and $k = 3$, respectively. (b) Equilibrium points versus the strength of the molecular memory k, where the low, intermediate and high states correspond to the points x_0 (stable), x_1 (unstable) and x_2 (stable), respectively. The parameters are taken as $\alpha = 0.2$, $\beta = 5$, $K = 10$, $m = 10$, $\delta = 0.2$ in (a), and $\delta = 0.35$ in (b).

all possible paths. The optimal transition path of switching between the states $(x_0, \ 0)$ and $(x_2, 0)$ corresponds to the zero energy ($H = 0$) in the phase space (x, λ). In this phase space, $(x_2, \ 0)$ and $(x_0, \ 0)$ are saddle points, but locally stable nodes on the line along $\lambda = 0$ (the zero conjugate momentum). Thus, the optimal transition path of switching from $(x_0, \ 0)$ to $(x_2, \ 0)$ is composed of two segments: the first segment is the heteroclinic trajectory with nonzero momentum that connects the point $(x_0, \ 0)$ with the intermediate point $(x_1, \ 0)$; the second segment is the line of $\lambda = 0$ from $(x_1, \ 0)$ to $(x_2, \ 0)$, referring to Fig. 4. Similarly, the optimal transition path from $(x_2, \ 0)$ to $(x_0, \ 0)$ is composed of the heteroclinic trajectory with nonzero momentum connecting $(x_2, \ 0)$ and $(x_1, \ 0)$ and the line of $\lambda = 0$ from $(x_1, \ 0)$ to $(x_0, \ 0)$. Then, the minimum action from

x_0 (the low state) to x_2 (the high state) can be given by

$$S_{\text{opt}}(x_0 \rightarrow x_2) = \int_{x_0}^{x_2} \lambda_{\text{opt}}(x)dx = \int_{x_0}^{x_1} \lambda_{\text{opt}}(x)dx. \quad (22)$$

The minimum action from x_2 to x_0 is similarly given by

$$S_{\text{opt}}(x_2 \rightarrow x_0) = \int_{x_2}^{x_1} \lambda_{\text{opt}}(x)dx. \quad (23)$$

Next, we will uncover how the molecular memory affects the optimal transition paths and the actions S_{opt} along them. We have known that for the same parameters in Eq. (19), the different molecular memories can produce different equilibrium points [referring to Fig. 3(b)]. To compare how the molecular memory affects the optimal transition paths and the actions along them, these paths should have the same starting and ending points. To this end, we introduce a continuous rescaled map Γ, defined by

$$\Gamma(s) = \begin{cases} x_0 + s(x_1 - x_0), & 0 \leqslant s < 1, \\ x_1 + (s-1)(x_2 - x_1), & 1 \leqslant s \leqslant 2, \end{cases} \quad (24)$$

with $\Gamma(0) = x_0$, $\Gamma(1) = x_1$ and $\Gamma(2) = x_2$. Figure 5, which is performed by the method of the geometric minimum-action method [49], shows the optimal transition paths of switching between the points $(x_0, \ 0)$ and $(x_2, \ 0)$ and their corresponding actions. Wherein, the rescaled points $(2, 0)$, $(0, 0)$, and $(1, 0)$ correspond to the high state $(x_2, \ 0)$, the low state $(x_0, \ 0)$, and the intermediate state $(x_1, \ 0)$, respectively. From Fig. 5(a), we find that the optimal transition path of switching from the low state to the high state turns long with the strength of the molecular memory, the corresponding action also increasing [referring to Fig. 5(c)]. However, the optimal transition path of switching from the high state to the low state turns short with the strength of the molecular memory, and the corresponding action decreases, as seen in Figs. 5(b) and 5(d). In a word, for the stronger molecular memory, the switching from the low state to the high state along the optimal transition path needs more action than the switching from the high state to the
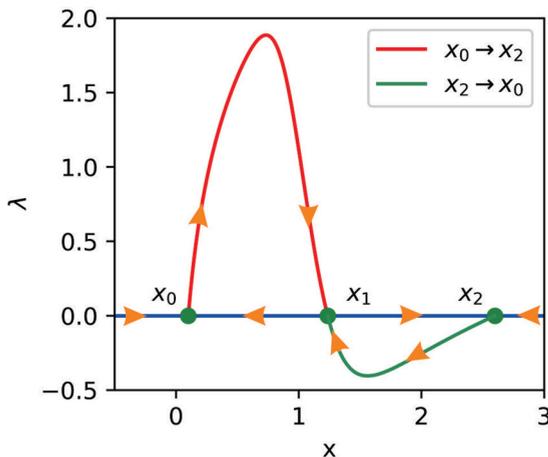

FIG. 4. The phase diagram $(x, \ \lambda)$ of the Hamilton system (19). The points $(x_0, \ 0)$, $(x_1, \ 0)$ and $(x_2, \ 0)$ are all unstable saddle points. The parameters are taken as $\alpha = 0.2$, $\beta = 5$, $K = 10$, $m = 10$, $k = 1$, and $\delta = 2$.
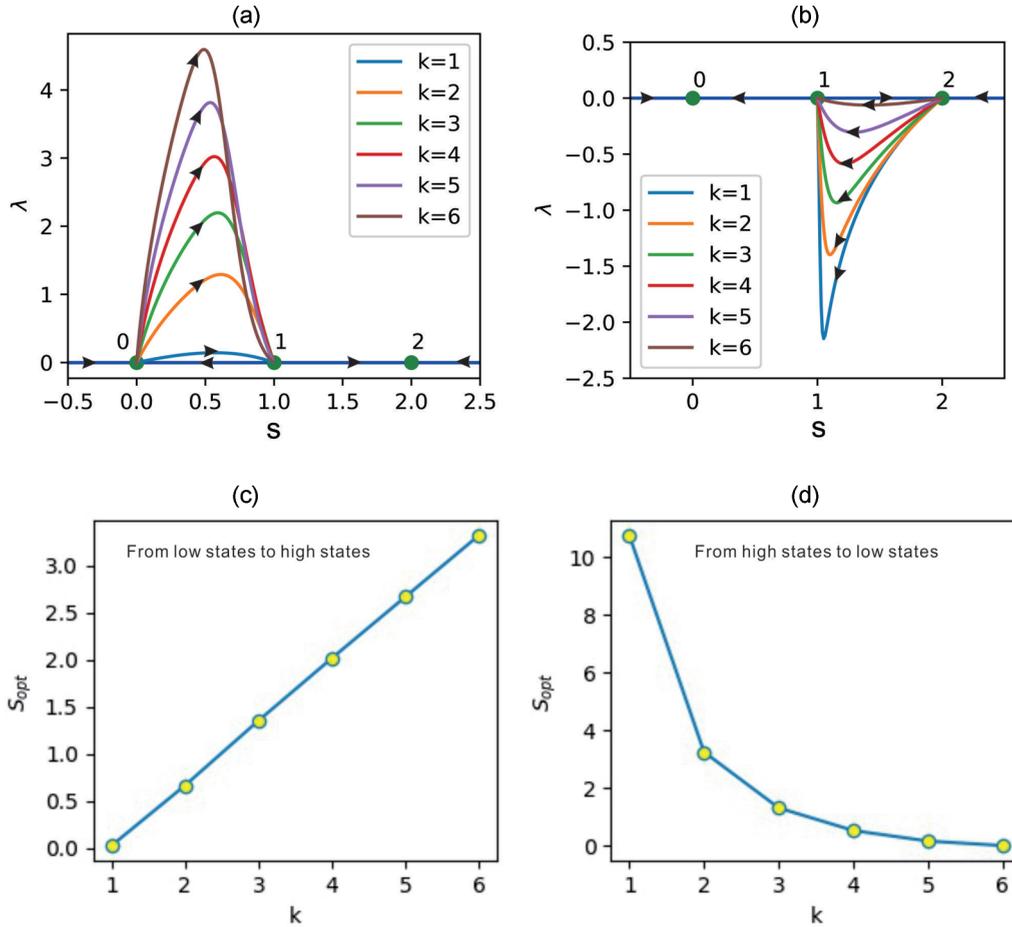
FIG. 5. The optimal transition paths and the minimal actions versus the strength of the molecular memory k. (a) The optimal transition paths of switching from the point $(0,\ 0)$ to the point $(2,\ 0)$. (b) The optimal transition paths of switching from the point $(2,\ 0)$ to the point $(0,\ 0)$. (c) The minimal actions along the optimal transition paths in the subplot A. (d) The minimal actions along the optimal transition paths in the subplot B. The parameters are set as $\alpha = 0.2$, $\beta = 5$, $K = 10$, $m = 10$, and $\delta = 0.35$.

low state, but for the weaker molecular memory, the opposite result occurs. The reason is that the stronger molecular memory in the gene regulatory network, which implies a longer mean time to produce proteins, can repress gene expression essentially. This is one of the main results of our paper.

IV. SWITCHING TIME

Now, we will study the time of switching between the low state and the high state of gene expression. According to the calculation method of the switching time in multistep reactions in Ref. [50] (see Appendix D), the time of switching from the low state x_0 to the high state x_2 can be plotted in Fig. 6(a) for the different parameters k. By a similar method, we also plot the time of switching from x_2 to x_0 in Fig. 6(b). From Fig. 6, we find that the times of switching from x_0 to x_2 and from x_2 to x_0 increase with the parameter k, implying that the molecular memory can prolong the time of switching between the low high states in gene expression. This is another main result in this paper.

V. CONCLUSION

Bistable models of genetic toggle switch have been extensively studied about the production of bistability, stochastic switching induced by extrinsic or intrinsic signaling, and the optimal transition paths of switching between metastable states and so on. In the present literature, the WKB method is a common method to explore the optimal transition paths of switching between metastable states for a genetic toggle switch. It is worth noting that the literature always assumed that the process of gene expression is Markovian [51], which is usually modeled by the CME. However, gene expression is an extremely complex process, which involves plenty of small biochemical reaction events, such as switching of activity of genetic promoters, accumulation of transcription factors, mRNA transcription, posttranscriptional modification, protein translation, and so on. Completing each reaction event must go through a waiting time. Consequently, the whole process of gene expression is no longer Markovian, but governed by the molecular memory. Although the molecular memory has been studied in gene expression, the researched questions mainly focused on how it affects burst in gene expression [32] and on whether or not it can induce a bimodality in the
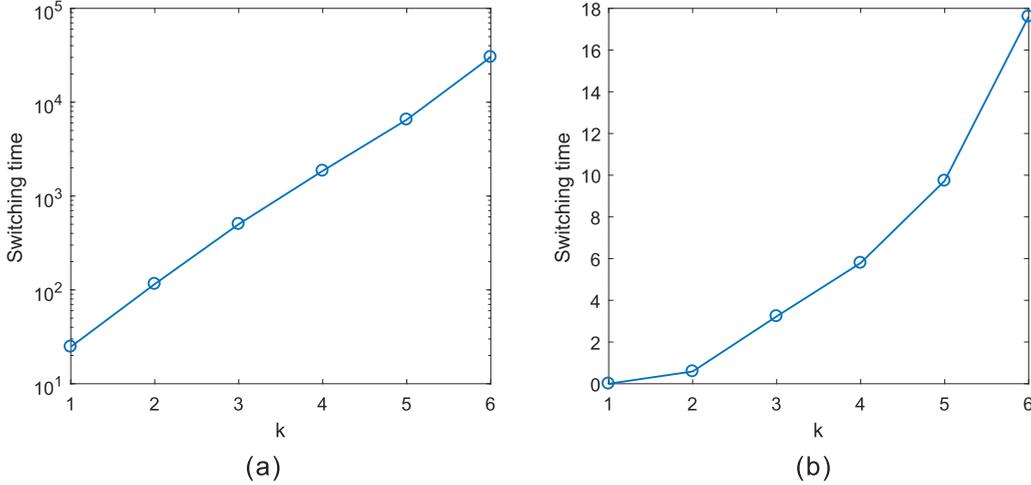
FIG. 6. The switching time versus the molecular memory k. (a) The time of switching from the low state to the high state. (b) The time of switching from the high state to the low state. The parameters are set as $\alpha = 0.2$, $\beta = 5$, $K = 10$, $m = 10$, $\Omega = 50$, and $\delta = 0.35$.

sense of probability distributions [33]. However, no paper, to our knowledge, has studied how the molecular memory affects optimal transition paths of switching between metastable states in gene expression. Here, we have researched the gene regulatory model with the positive feedback, which is described by the generalized CME. We have found that the molecular memory can induce bistable states in gene expression. This result is consistent with that in Ref. [33]. In addition, we have shown that the action along the optimal transition path of switching from the low state to the high state increases with the strength of the molecular memory, and that it, along with the optimal transition path of switching from the high state to the low state decreases with the strength of the molecular memory. Moreover, we have demonstrated that the time of switching between the low and high states increases with the strength of the molecular memory. Our results have implied that the molecular memory can much affect switching between metastable states in gene expression.

In this paper, our model is a classical genetic self-regulation model, and considers neither the promoter architecture of gene transcription nor switching between the active and inactive states of promoters on DNA. In fact, biochemical reactions in gene expression involve multiple transcription factors regulating the activity of the promoters. Our model assumes that a gene is always in an active state, and the positive self-regulatory feedback depends on the quantity of transcription factors, whose mathematical form is a Hill function. This feedback is a canonical and a simple form, but neglects the spatial distribution of regulatory factors and regulatory delays. Moreover, our model is very ideal because it has neglected interaction among regulations of multiple genes. A canonical example is that the genes *lacI* and *tecR* mutually repress each other in the lac operon of *E. coli* [52,53]. Finally, there are some other ways to describe the molecular memory in gene expression, such as queuing models [54] and delay models [55,56]. Therefore, our model needs to be further modified. The above factors will be investigated in our future work.

APPENDIX A: A GENERALIZED CHEMICAL MASTER EQUATION

Our model (5) is a generalized CME, which is non-Markovian in contrast to the conventional gene regulatory model that is Markovian. In order to reduce complexity, we will focus on the biochemical reactions (1). We define $R_k(t, n)$ as a joint PDF of the waiting time that protein's copy number reaches n after undergoing k reaction steps, and then we have

$$R_{k+1}(t, n) = \int_0^t [R_k(t', n-1)\phi_1(t-t', n-1) \\ + R_k(t', n+1)\phi_2(t-t', n+1)]dt', \quad \text{(A1)}$$

and $R_0(t, n) = P_n(t)\delta(t)$, where $P_n(t)$ is a probability that protein's copy number reaches n at time t; $\delta(t)$ is a delta function; $\phi_1(t, n)$ and $\phi_2(t, n)$ have been defined in Eq. (2). Note that the total $R(t, n) = \sum_{k=0}^{\infty} R_k(t, n)$ is the PDF of the waiting time that a protein's copy number reaches n after undergoing any number of reaction steps. Let

$$\Phi(t, n) = 1 - \int_0^t \phi_1(t', n) + \phi_2(t', n)dt' \quad \text{(A2a)}$$

represent the probability that no biochemical reaction occurs during the time interval $(0, t)$ when the system has n proteins. Therefore, we have

$$P_n(t) = \int_0^t R(t', n)\Phi(t-t', n)dt'. \quad \text{(A2b)}$$

Marking the Laplace transform on the both sides of (A1) yields

$$\tilde{R}_{k+1}(s, n) = \tilde{R}_k(s, n-1)\tilde{\phi}_1(s, n-1)$$
$$+ \tilde{R}_k(s, n+1)\tilde{\phi}_2(s, n+1). \tag{A3}$$

Summing both sides of Eq. (A3) over k, we obtain

$$\tilde{R}(s, n) = \tilde{R}_0(s, n) + \tilde{R}(s, n-1)\tilde{\phi}_1(s, n-1)$$
$$+ \tilde{R}(s, n+1)\tilde{\phi}_2(s, n+1). \tag{A4}$$

Making the Laplace transform on both sides of (A2b) yields

$$\tilde{P}_n(s) = \tilde{R}(s, n)[1 - \tilde{\phi}_1(s, n) - \tilde{\phi}_2(s, n)]/s. \tag{A5}$$

Substituting Eq. (A5) into Eq. (A4), we obtain

$$s\tilde{P}_n(s) = P_n(0) + \tilde{R}(s, n-1)\tilde{\phi}_1(s, n-1)$$
$$+ \tilde{R}(s, n+1)\tilde{\phi}_2(s, n+1)$$
$$- \tilde{R}(s)[\phi_1(s, n) + \phi_2(s, n)]. \tag{A6}$$

Now, we let

$$\tilde{M}_1(s, n) = \frac{s\tilde{\phi}_1(s, n)}{1 - \tilde{\phi}_1(s, n) - \tilde{\phi}_2(s, n)}, \tag{A7a}$$

$$\tilde{M}_2(s, n) = \frac{s\tilde{\phi}_2(s, n)}{1 - \tilde{\phi}_1(s, n) - \tilde{\phi}_2(s, n)}. \tag{A7b}$$

According to Eq. (A5), we have

$$\tilde{M}_1(s, n)\tilde{P}_n(s) = \tilde{R}(s, n)\tilde{\phi}_1(s, n), \tag{A8a}$$

$$\tilde{M}_2(s, n)\tilde{P}_n(s) = \tilde{R}(s, n)\tilde{\phi}_2(s, n). \tag{A8b}$$

Inserting Eq. (A8) into Eq. (A6) yields the generalized chemical master equation in the form of the Laplace transform,

$$s\tilde{P}_n(s) = P_n(0) + \tilde{M}_1(s, n-1)\tilde{\phi}_1(s, n-1)$$
$$+ \tilde{M}_2(s, n+1)\tilde{\phi}_2(s, n+1)$$
$$- \tilde{M}_1(s, n)\tilde{\phi}_1(s, n) - \tilde{M}_2(s, n)\tilde{\phi}_2(s, n)]. \tag{A9}$$

Taking the inverse Laplace transform on both sides of Eq. (A9), we can obtain the master Eq. (5) in the main text.

APPENDIX B: EXISTENCE OF $\lim_{s\to 0} \tilde{M}_i(n, s)$

In fact, the existence of the limit $\lim_{s\to 0} \tilde{M}_i(n, s)$, $i = 1, 2$ has been shown in Ref. [32]. Here, we restate this proof. In detail,

$$\tilde{M}_1 = \frac{s\int_0^\infty e^{-st}\left[p_1(t)\int_t^\infty \psi(t', n)dt'\right]dt}{1 - \int_0^\infty \left[e^{-st}p_1(t)\int_t^\infty \psi(t', n)dt'\right]dt - \int_0^\infty \left[e^{-st}\psi(t, n)\int_t^\infty p_1(t')dt'\right]dt}$$

$$= \frac{s\int_0^\infty e^{-st}\left[p_1(t)\int_t^\infty \psi(t', n)dt'\right]dt}{1 + \int_0^\infty e^{-st}d\left[\int_t^\infty p_1(t')dt'\int_t^\infty \psi(t', n)dt'\right]}$$

$$= \frac{\int_0^\infty e^{-st}\left[p_1(t)\int_t^\infty \psi(t', n)dt'\right]dt}{\int_0^\infty \left[e^{-st}\int_t^\infty p_1(t')dt'\int_t^\infty \psi(t', n)dt'\right]dt}. \tag{B1}$$

Thus, we obtain

$$\lim_{s\to 0} \tilde{M}_1(n, s) = \frac{\int_0^\infty \left[p_1(t)\int_t^\infty \psi(t', n)dt'\right]dt}{\int_0^\infty \left[\int_t^\infty p_1(t')dt'\int_t^\infty \psi(t', n)dt'\right]dt}. \tag{B2}$$

Similarly, we can obtain

$$\lim_{s\to 0} \tilde{M}_2(n, s) = \frac{\int_0^\infty \left[\psi(t, n)\int_t^\infty p_1(t')dt'\right]dt}{\int_0^\infty \left[\int_t^\infty p_1(t')dt'\int_t^\infty \psi(t', n)dt'\right]dt}. \tag{B3}$$

APPENDIX C: DERIVATION FOR THE HAMILTONIAN FUNCTION

Here, we give out the detailed derivation for the Hamiltonian function (17a) from Eq. (8).
We can rewrite (8) as the following form:

$$m_1(n-1)P_{n-1} + m_2(n+1)P_{n+1} - [m_1(n) + m_2(n)]P_n = 0. \tag{C1}$$

As usual, we assume $\Omega \gg 1$, and let $x = n/\Omega$. The transition rate $m_r(n) = m_r(\Omega x)$, $r = 1, 2$, can be represented as the following expansion in Ω:

$$m_r(n) = m_r(\Omega x) = \Omega M_r(x) + u_r(x) + O\left(\frac{1}{\Omega}\right), \tag{C2}$$

where x and the scaled transition rates M_r and u_r are O(1). Now, we approximate Eq. (C1) by using the WKB method. Assume that the stationary probability P_n has the form of the WKB ansatz,

$$P_n \approx \exp[-\Omega S(x)]. \tag{C3}$$

This ansatz and Eq. (C2) are substituted into the master equation (C1), yielding

$$0 = \left[\Omega M_1 \left(x - \frac{1}{\Omega} \right) + u_1 \left(x - \frac{1}{\Omega} \right) + O\left(\frac{1}{\Omega} \right) \exp\left[-\Omega S \left(x - \frac{1}{\Omega} \right) \right] + \left[\Omega M_2 \left(x + \frac{1}{\Omega} \right) + u_2 \left(x + \frac{1}{\Omega} \right) O\left(\frac{1}{\Omega} \right) \right]$$

$$\times \exp\left[-\Omega S \left(x + \frac{1}{\Omega} \right) \right] - \left[\Omega M_1(x) + u_1(x) + \Omega M_2(x) + u_2(x) + O\left(\frac{1}{\Omega} \right) \right] \exp[-\Omega S(x)]. \quad \text{(C4)}$$

In the above equation, the functions of $x \pm 1/\Omega$ are performed by a Taylor series expansion, that is,

$$M_1(x - 1/\Omega) = M_1(x) - M_1'(x)/\Omega + O(1/\Omega), \quad \text{(C5a)}$$

$$u_1(x - 1/\Omega) = u_1(x) - u_1'(x)/\Omega + O(1/\Omega), \quad \text{(C5b)}$$

$$M_2(x + 1/\Omega) = M_2(x) + M_2'(x)/\Omega + O(1/\Omega), \quad \text{(C5c)}$$

$$u_2(x + 1/\Omega) = u_2(x) + u_2'(x)/\Omega + O(1/\Omega), \quad \text{(C5d)}$$

$$S(x \pm 1/\Omega) = S(x) \pm S'(x)/\Omega + O(1/\Omega). \quad \text{(C5e)}$$

Substituting the above approximations into Eq. (C4), one can obtain the leading order term of Ω, as follows:

$$0 = M_1(x)\exp[-\Omega S(x) + S'(x)] - M_1(x))\exp[-\Omega S(x)]$$
$$+ M_2(x)\exp[-\Omega S(x) - S'(x)] - M_2(x)\exp[-\Omega S(x)]. \quad \text{(C6)}$$

Multiplying both sides of Eq. (C6) by $\exp[-\Omega S(x)]$ yields

$$0 = M_1(x)[\exp[S'(x)] - 1] + M_2(x)[\exp[S'(x)] - 1]. \quad \text{(C7)}$$

Let $M_1(x) = M(x)$, where $M(x)$ has been defined in Eq. (17b), $M_2 = \delta x$ and $S'(x) = \lambda$. Finally, one obtains

$$H(x, \lambda) = M(x)(e^\lambda - 1) + \delta x(e^\lambda - 1), \quad \text{(C8)}$$

which is the Hamiltonian function in the main text.

APPENDIX D: SWITCHING TIME OF THE GENERALIZED MULTISTEP REACTION SCHEME

We here give out the formula to calculate the time of switching of multistep biochemical reactions. According to Ref. [50], the time of switching from the low state x_0 to the high state x_2 can be given by

$$\tau = \frac{2\pi}{w_{+(x_0)}} \frac{\exp\left[-\int_{x_0}^{x_1} (u_+/w_+ - u_-/w_-)dx \right]}{\sqrt{|S''(x_1)|S''(x_0)}}$$
$$\times \exp[\Omega[S(x_1) - S(x_0)]], \quad \text{(D1)}$$

where

$$S(x_1) - S(x_0) = \int_{x_0}^{x_1} \ln[w_-(x)/w_+(x)]dx.$$

For our model, we have

$$w_+ = \frac{[f(x)]^k \delta x}{[x\delta + f(x)]^k - [f(x)]^k}, \quad u_+ = x\Omega M'(x), \quad \text{(D2a)}$$

$$w_- = \delta x, \quad u_- = \delta x\Omega \quad \text{(D2b)}$$

[1] B. Munsky, G. Neuert, and O. A. Van, Science **336**, 183 (2012).

[2] M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain, Science **297**, 1183 (2002).

[3] J. M. Raser, O, and E. K. Shea, Science **304**, 1811 (2004).

[4] M. Acar, J. T. Mettetal, and A. van Oudenaarden, Nat. Genet. **40**, 471 (2008).

[5] D. L. Jones, R. C. Brewster, and R. Phillips, Science **346**, 1533 (2014).

[6] A. S. Bodrova, A. V. Chechkin, A. G. Cherstvy, H. Safdari, I. M. Sokolov, and R. Metzler, Sci. Rep. **6**, 30520 (2016).

[7] D. A. Potoyan and P. G. Wolynes, J. Chem. Phys. **143**, 195101 (2015).

[8] M. Kheir Gouda, M. Manhart, and G. Balzsi, Proc. Natl. Acad. Sci. **116**, 25162 (2019).

[9] D. A. Charlebois, G. Balázsi, and M. Kærn, Phys. Rev. E **89**, 052708 (2014).

[10] S. Misra, S. Ghatak, and B. P. Toole, J. Biol. Chem. **280**, 20310 (2005).

[11] K. S. Farquhar, D. A. Charlebois, M. Szenk, J. Cohen, D. Nevozhay, and G. Balzsi, Nat. Commun. **10**, 2766 (2019).

[12] G. Balzsi and M. K. Belete, Phys. Rev. E **92**, 062716 (2015).

[13] D. Nevozhay, R. M. Adams, E. Van Itallie, M. R. Bennett, and G. Balzsi, PLoS Comput. Biol. **8**, e1002480 (2012).

[14] M. Scott, B. Ingalls, and M. Kærn, Chaos **16**, 026107 (2006).

[15] T. Tian and K. Burrage, J. Theor. Biol. **227**, 229 (2004).

[16] T. Lipniacki, B. Hat, J. R. Faeder, and W. S. Hlavacek, J. Theor. Biol. **254**, 110 (2008).

[17] K. Puszyski, B. Hat, and T. Lipniacki, J. Theor. Biol. **254**, 452 (2008).

[18] T. Zhou and J. Zhang, SIAM J. Appl. Math. **72**, 789 (2012).

[19] V. Shahrezaei and P. S. Swain, Proc. Natl. Acad. Sci. **105**, 17256 (2008).

[20] O. Ovaskainen and B. Meerson, Trends Ecol. Evol. **25**, 643 (2010).

[21] E. Forgoston and R. Moore, SIAM Rev. **60**, 969 (2018).

[22] E. Vanden-Eijnden and M. Heymann, Phys. Rev. Lett. **100**, 140601 (2008).

[23] M. Assaf, E. Roberts, and Z. Luthey-Schulten, Phys. Rev. Lett. **106**, 24810224 (2011).

[24] E. Roberts, Z. Luthey-Schulten, N. Goldenfeld, and M. Assaf, Phys. Rev. Lett. **111**, 058102 (2013).

[25] R. Perez-Carrasco, P. Guerrero, J. Briscoe, and K. M. Page, PLoS Comput Biol **12**, e1005154 (2016).

[26] C. Lv, X. Li, F. Li, and T. Li, PLoS Comput Biol **11**, e1004156 (2015).

[27] S. Watanabe and N. Ikeda, *Stochastic Differential Equations and Diffusion Processes* (Elsevier, Amsterdam, 1981).

[28] C. W. Gardiner, *Stochastic Methods: A Handbook for the Natural and Social Sciences* (Springer, New York, 2009).

[29] M. K. Prajapat and A. S. Ribeiro, R. Soc. Open Sci. **5**, 181170 (2018).

[30] C. V. Harper, B. Finkenstädt, D. J. Woodcock, S. Friedrichsen, S. Semprini, L. Ashall, D. G. Spiller, J. J. Mullins, D. A. Rand, J. R. E. Davis, and M. R. H. White, PLoS Biol. **9**, e1000607 (2011).

[31] A. J. Carpousis, N. F. Vanzo, and L. C. Raynal, Trends Genet. **15**, 24 (1999).

[32] H. Qiu, B. Zhang, and T. Zhou, Phys. Rev. E **100**, 012128 (2019).

[33] T. Zhou, J. Zhang, and B. Qiu, Phys. Rev. E **101**, 022409 (2020).

[34] F. Nikakhtar, M. R. R. Tabar, S. Rahvar, R. K. Sheth, K. Lehnertz, M. Sahimi, and S. Baghram, Phys. Rev. E **99**, 062101 (2019).

[35] R. Torabi and J. Davidsen, Phys. Rev. E **100**, 052217 (2019).

[36] C. A. Yates, M. J. Ford, and R. L. Mort, Bull. Math. Biol. **79**, 2905 (2017).

[37] T. Galla and T. Brett, Phys. Rev. Lett. **110**, 250601 (2013).

[38] N. Kumar, A. Singh, and R. V. Kulkarni, PLoS Comput. Biol. **11**, e1004292 (2015).

[39] H. Qiu, B. Zhang, and T. Zhou, Phys. Rev. E **101**, 012405 (2020).

[40] J. Zhang and T. Zhou, Proc. Natl. Acad. Sci. USA **116**, 23542 (2019).

[41] N. Kumar, T. Platini, and R. V. Kulkarni, Phys. Rev. Lett. **113**, 268105 (2014).

[42] M. Dentz and T. Aquino, Phys. Rev. Lett. **119**, 230601 (2017).

[43] M. Boguñá, L. F. Lafuerza, R. Toral, and M. A. Serrano, Phys. Rev. E **90**, 042108 (2014).

[44] X. J. Tian, X. P. Zhang, F. Liu, and W. Wang, Phys. Rev. E **80**, 011926 (2009).

[45] R. Hermsen and E. David, PLoS Comput Biol **7**, e1002265 (2011).

[46] K. Proesmans and B. Derrida, J. Stat. Mech. Theory Exp. (2019) 023201.

[47] A. Michael and B. Meerson, J. Phys. A **50**, 263001 (2017).

[48] B. E. Shay and Michael, J. Stat. Mech. Theory Exp. (2016) 113501.

[49] M. Heymann and E. Vanden-Eijnden, Commun. Pure Appl. Math. **61**, 1052 (2008).

[50] C. Escudero and A. Kamenev, Phys. Rev. E **79**, 041149 (2009).

[51] H. Chen, P. Thill, and J. Cao, J. Chem. Phys. **144**, 175104 (2016).

[52] S. N. Semenov, L. J. Kraft, A. Ainla, M. Zhao, M. Baghbanzadeh, V. E. Campbell, K. Kang, J. M. Fox, and G. M. Whitesides, Nature (London) **537**, 656 (2016).

[53] G. M. Suel, J. Garcia-Ojalvo, L. M. Liberman, and M. B. Elowitz, Nature **440**, 545 (2006).

[54] N. Yi, G. Zhuang, L. Da, and Y. Wang, J. Chem. Phys. **136**, 144108 (2012).

[55] D. Bratsun, D. Volfson, L. S. Tsimring, and J. Hasty, PNAS **102**, 14593 (2005).

[56] Q. Wu and T. Tian, Sci. Rep. **6**, 31909 (2016).