# Exploring dissipative sources of non-Markovian biochemical reaction systems

Xiyan Yang,[1] Yiren Chen,[2] Tianshou Zhou,[3,4] and Jiajun Zhang[3,4,*]

[1]*School of Financial Mathematics and Statistics, Guangdong University of Finance, Guangzhou 510521, People's Republic of China*
[2]*College of Mathematics and Statistics, Shenzhen University, Shenzhen 518060, People's Republic of China*
[3]*School of Mathematics, Sun Yat-Sen University, Guangzhou 510275, People's Republic of China*
[4]*Guangdong Province Key Laboratory of Computational Science, Guangzhou 510275, People's Republic of China*

(Received 11 December 2020; revised 19 March 2021; accepted 29 April 2021; published 18 May 2021)

Many biological processes including important intracellular processes are governed by biochemical reaction networks. Usually, these reaction systems operate far from thermodynamic equilibrium, implying free-energy dissipation. On the other hand, single reaction events happen often in a memory manner, leading to non-Markovian kinetics. A question then arises: how do we calculate free-energy dissipation (defined as the entropy production rate) in this physically real case? We derive an analytical formula for calculating the energy consumption of a general reaction system with molecular memory characterized by nonexponential waiting-time distributions. It shows that this dissipation is composed of two parts: one from broken detailed balance of an equivalent Markovian system with the same topology and substrates, and the other from the direction-time dependence of waiting-time distributions. But, if the system is in a detailed balance and the waiting-time distribution is direction-time independent, there is no energy dissipation even in the non-Markovian case. These general results provide insights into the physical mechanisms underlying nonequilibrium processes. A continuous-time random-walk model and a generalized model of stochastic gene expression are chosen to clearly show dissipative sources and the relationship between energy dissipation and molecular memory.

DOI: 10.1103/PhysRevE.103.052411

## I. INTRODUCTION

Many biological processes including important intracellular processes are governed by biochemical reaction networks. Often, these systems operate far from equilibrium. Nonequilibrium processes are essential for cells to perform internal transport [1,2], drive directional motion [3,4], and generate spatial organization [5,6]. In general, the achievement of biological functions is at the cost of free energy, e.g., cells consume their free energy to achieve their particular functions such as high-fidelity transcription and replication [7,8], accurately sensing and adaption [9,10]. The quantitative or qualitative analysis of free-energy consumption is thus significant for understanding fundamental cell processes and the underlying biophysical mechanisms.

Traditionally, the analysis of dissipative systems and nonequilibrium dynamics is based on the Markovian hypothesis [11–14]; e.g., in order to analyze dynamics of a biochemical reaction network, it is often assumed that reaction rates are constants (this assumption implies that waiting times between individual reaction events follow exponential distributions [15,16]). Under this assumption, the energy dissipation of nonequilibrium systems has been extensively studied [17–19]. For example, the tradeoff between information and dissipation in biological circuits has been discussed using simple models [20]. It is generally believed that free energy is consumed to maintain properties of a functional system, e.g., cells increase, through consuming their free energy, the accuracy, efficiency, and robustness of their functions in noisy environments [21–24]. From the viewpoint of thermodynamics, adenosine triphosphate (ATP) hydrolysis provides energy input to break detailed balance and achieve sensory and adaptive functions at the cellular level [25–27]. Several works showed a close connection between the amount of energy consumed by a Markovian reaction network responsible for performing the cellular computation and the accuracy of the final computed information [8,28–31]. A similar study showed that energy dissipation is needed for regulatory networks, and the amount of dissipated energy limits reliable readout [32]. In addition, other studies showed that Markovian gene expression needs energy dissipation to ensure high-fidelity, specificity, and efficiency of transcription and translation; e.g., fidelity of transcription in general follows energy-dissipative events such as chromatin remodeling, DNA methylation, and histone modification [33]. Recently, it was reported that bursty expression needs to consume more free energy than constitutive expression if both processes are Markovian [34,35].

In the biological world, however, Markovianity is the exception and non-Markovianity is the rule. For example, gene activation in eukaryotic systems is typically a non-Markovian process since transcription begins only when chromatin template accumulates over time until the promoter becomes active. The involved multistep process can create a memory between individual events [36–39], leading to non-Markovian reaction kinetics. More generally, the complex control process of gene expression would involve the recruitment of

*zhjiajun@mail.sysu.edu.cn

 ©2021 American Physical Society

repressors, transcription factors, and mediators, as well as chromatin remodeling or changes in supercoiling, and all these detailed processes can generate nonexponential time intervals between transcription windows [40–42]. Recent studies imply that nontrivial dissipation can be detected even if there are no observed flows in non-Markovian time series [43,44]. These studies indicate that memory plays an important role in biochemical processes. The extensive existence of molecular memory raises an interesting yet fully unsolved question: how do we trace sources of the free-energy dissipation (defined as the entropy production rate in this paper) of a reaction system with molecular memory? Two related issues are, under what conditions of molecular memory there is no free-energy consumption; and does a reaction system consume more free energy in the non-Markovian case than in the Markovian case if mean waiting times between reaction events are the same in the two cases?

There have been some works to address the above or similar questions in terms of continuous-time random walk (CTRW) [45–47]. For example, Martínez *et al.* [48] developed an approach based on the relative entropy to detect time irreversibility and estimate the entropy production of a hidden network with arbitrary waiting-time distributions. They also showed that there is a nontrivial entropy production rate even in the absence of observable currents. Recently, Teza and Stella demonstrated that an exact coarse-graining model with memory can preserve entropy production out of equilibrium [49]. In addition, Bisker *et al.* studied a partitioning of the entropy production related to the observed and hidden variables with integral fluctuation theorems [50]. Despite these efforts, we still lack a general theory for understanding the energy dissipation of general non-Markovian reaction networks and the relationship between energy consumption and molecular memory.

This paper aims to quantify the energy dissipation of a general reaction system with molecular memory characterized by nonexponential waiting-time distributions. Although we have proposed a general framework for modeling this kind of system—-a generalized chemical master equation (gCME), and shown that its steady-state probabilistic behavior is exactly the same as that of an equivalent Markovian reaction network [51], whether both systems consume the same free energy remains unclear. We perform our analysis of entropy production rate based on the CTRW theory [48,52]. Specifically, By considering forward and backward trajectories in the state space of the non-Markovian reaction system if they exist, we analyze and calculate the entropy production rate between these two trajectories, which is used to quantify the energy dissipation of the underlying system. As a result, we successfully derive an analytical formula, which shows that the entropy production rate is composed of two parts: one from the broken detailed balance (BDB) of the topology-equivalent Markovian system, and the other from the direction-time dependence (DTD) of the waiting-time distributions. We also show that if the system is at the detailed balance (DB) and the waiting-time distributions are direction-time independent (DTI), then there is no energy dissipation even in the non-Markovian case. Using our general theory, we analyze two examples: a CTRW model with nonexponential waiting-time distributions, and a generalized ON-OFF model of gene expression with molecular memory. For the former example, we show how molecular memory affects the entropy production rate. For the latter example, we find that different memory mechanisms can induce different patterns of energy dissipation, implying that molecular memory has an important influence on energy dissipation in gene expression. Further study on the energy dissipation of non-Markovian reaction systems can help us understand the underlying nonequilibrium mechanisms.

## II. METHODS

### A. Chemical master equations for modeling biochemical networks with molecular memory

Chemical reactions are the consequence of the interaction between different system components. Consider a general reaction network consisting of $N$ different reactive species (denoted by $X_j$, $j = 1, 2, \ldots, N$) that participate in $L$ reactions. Denote by $\boldsymbol{n} = (n_1, n_2, \ldots n_N)^T$ the microscopic state of the system, where $n_j$ represents the number of molecules for the $j$th reactive species ($j = 1, 2, \ldots, N$). A single event of any reaction, or equivalently a state transition $\boldsymbol{n} \to \boldsymbol{n'}$ of the system (in the following, we also use $\boldsymbol{n} \to \boldsymbol{n'}$ to represent a reaction unless confusions arises), is characterized by the reaction waiting time $t$ whose probability density function (PDF) is denoted by $\psi_{\boldsymbol{nn'}}(t)$, depends, in general, on the system state $\boldsymbol{n}$, where $\boldsymbol{n'} = (n'_1, n'_2, \ldots n'_N)^T$. Thus, a reaction of the underlying system may be, in general, described as

$$\boldsymbol{n} \xrightarrow{\psi_{\boldsymbol{nn'}}(t)} \boldsymbol{n'}. \tag{1}$$

If the cumulative distribution function of $\psi_{\boldsymbol{nn'}}(t)$ is defined as $\Psi_{\boldsymbol{nn'}}(t) = \int_0^t \psi_{\boldsymbol{nn'}}(t')dt'$, then the joint PDF of reaction $\boldsymbol{n} \to \boldsymbol{n'}$ happening with the waiting time $t$ is given by $\phi_{\boldsymbol{nn'}}(t) = \psi_{\boldsymbol{nn'}}(t) \prod_{\boldsymbol{n''} \neq \boldsymbol{n'}} [1 - \Psi_{\boldsymbol{nn''}}(t)]$. Correspondingly, the probability that reaction $\boldsymbol{n} \to \boldsymbol{n'}$ happens is given by $\Phi_{\boldsymbol{nn'}} = \int_0^\infty \phi_{\boldsymbol{nn'}}(t)dt$, where $\Phi_{\boldsymbol{nn'}}$ satisfies the conservation condition $\sum_{\boldsymbol{n'}} \Phi_{\boldsymbol{nn'}} = 1$. The PDF of the residence time at state $\boldsymbol{n}$ is calculated according to $\phi_{\boldsymbol{n}}(t) = \sum_{\boldsymbol{n'}} \phi_{\boldsymbol{nn'}}(t)$. Note that the probability that no reaction occurs at state $\boldsymbol{n}$ in time $[0, t)$ is given by $\Phi_{\boldsymbol{n}}(t) = \sum_{\boldsymbol{n'}} \int_t^\infty \phi_{\boldsymbol{nn'}}(t')dt' = \prod_{\boldsymbol{n'}} [1 - \Psi_{\boldsymbol{nn'}}(t)]$ whereas the mean residence time at state $\boldsymbol{n}$ by $\tau_{\boldsymbol{n}} = \int_0^\infty \Phi_{\boldsymbol{n}}(t)dt = \int_0^\infty t\phi_{\boldsymbol{n}}(t)dt = \int_0^\infty \prod_{\boldsymbol{n'}} [1 - \Psi_{\boldsymbol{nn'}}(t)]\, dt$.

Now, we define

$$\phi_{\boldsymbol{n'|n}}(t) = \frac{\phi_{\boldsymbol{nn'}}(t)}{\Phi_{\boldsymbol{nn'}}}, \tag{2}$$

which represents the conditional waiting-time distribution of reaction $\boldsymbol{n} \to \boldsymbol{n'}$, given state $\boldsymbol{n}$. If waiting-time distributions $\phi_{\boldsymbol{nn'}}(t)$ at state $\boldsymbol{n}$ for different $\boldsymbol{n'}$ satisfy

$$\phi_{\boldsymbol{nn'}}(t) = \Phi_{\boldsymbol{nn'}}\phi_{\boldsymbol{n}}(t), \tag{3}$$

then they are called distributions of direction-time independence [45,46], i.e., the waiting-time distributions can be represented as the product of the waiting-time part that depends only on the originating state, and a transition probability that connects the initial and final states. In the following, DTI is an abbreviation for direction-time independent, whereas DTD is an abbreviation for direction-time dependent. The

combination of Eqs. (2) and (3) implies the relation

$$\phi_{\boldsymbol{n}'|\boldsymbol{n}}(t) = \phi_{\boldsymbol{n}}(t), \tag{4}$$

which indicates that the conditional waiting-time distribution $\phi_{\boldsymbol{n}'|\boldsymbol{n}}(t)$ for next reaction $\boldsymbol{n} \to \boldsymbol{n}'$ does not depend on state $\boldsymbol{n}'$. Obviously, if all reaction processes from state $\boldsymbol{n}$ to state $\boldsymbol{n}'$ are Markovian or memoryless (i.e., the waiting times follow exponential distributions), then the waiting-time distributions of state $\boldsymbol{n}$ are DTI.

Next, we establish a chemical master equation (CME) for the entire reaction network. Let $P_{\boldsymbol{n}}(t)$ denote the probability that the system is in state $\boldsymbol{n}$ at time $t$, and $M_{\boldsymbol{n}\boldsymbol{n}'}(t)$ be the memory function for state transition $\boldsymbol{n} \to \boldsymbol{n}'$. Note that the Laplace transform of $M_{\boldsymbol{n}\boldsymbol{n}'}(t)$ is given by $\tilde{M}_{\boldsymbol{n}\boldsymbol{n}'}(s) = \frac{\tilde{\phi}_{\boldsymbol{n}\boldsymbol{n}'}(s)}{\tilde{\Phi}_{\boldsymbol{n}}(s)}$ [53,54]. Then, according to the probability principle, the change rate (i.e., the derivative with regard to time) of $P_{\boldsymbol{n}}(t)$ is equal to the difference that all possible probability inflows of state $\boldsymbol{n}$ subtract all possible probability outflows of state $\boldsymbol{n}$. Mathematically, this can be expressed as

$$\frac{\partial P_{\boldsymbol{n}}(t)}{\partial t} = \sum_{\boldsymbol{n}'} \int_0^t [M_{\boldsymbol{n}'\boldsymbol{n}}(t-t')P_{\boldsymbol{n}'}(t') - M_{\boldsymbol{n}\boldsymbol{n}'}(t-t')P_{\boldsymbol{n}}(t')]dt'. \tag{5}$$

This master equation is called a generalized CME. It is worth pointing out that this gCME is an extension of the common CME. In fact, we can show that if $\tilde{M}_{\boldsymbol{n}\boldsymbol{n}'}(s)$ is independent of $s$, then $M_{\boldsymbol{n}\boldsymbol{n}'}(t)$ will reduce to a reaction propensity function in the Markovian case.

Note that

$$\int_0^\infty M_{\boldsymbol{n}\boldsymbol{n}'}(t)dt = \tilde{M}_{\boldsymbol{n}\boldsymbol{n}'}(0) = \frac{\tilde{\phi}_{\boldsymbol{n}\boldsymbol{n}'}(0)}{\tilde{\Phi}_{\boldsymbol{n}}(0)} = \frac{\Phi_{\boldsymbol{n}\boldsymbol{n}'}}{\tau_{\boldsymbol{n}}} = \frac{\int_0^\infty \psi_{\boldsymbol{n}\boldsymbol{n}'}(t)\prod_{\boldsymbol{n}''\neq\boldsymbol{n}'}[1-\Psi_{\boldsymbol{n}\boldsymbol{n}''}(t)]\,dt}{\int_0^\infty \prod_{\boldsymbol{n}'}[1-\Psi_{\boldsymbol{n}\boldsymbol{n}'}(t)]\,dt}, \tag{6}$$

which is termed as the effective transition rate for reaction $\boldsymbol{n} \to \boldsymbol{n}'$ and denoted by $K_{\boldsymbol{n}\boldsymbol{n}'}$. Apparently, $K_{\boldsymbol{n}\boldsymbol{n}'}$ integrates the effect of molecular memory if nonexponential waiting-time distributions exist. This is an interesting fact.

Furthermore, if the stationary distribution of the reaction system, denoted by $P_{\boldsymbol{n}}$, exists (mathematically proving this point seems difficult, but numerical calculations of examples verify that $P_{\boldsymbol{n}}$ indeed exists), then using $K_{\boldsymbol{n}\boldsymbol{n}'}$, it follows from Eq. (5) that the following the stationary gCME (sgCME) holds:

$$\sum_{\boldsymbol{n}'} (K_{\boldsymbol{n}'\boldsymbol{n}}P_{\boldsymbol{n}'} - K_{\boldsymbol{n}\boldsymbol{n}'}P_{\boldsymbol{n}}) = 0. \tag{7}$$

Interestingly, this equation is actually the CME for a Markovian reaction network with the same topology but with $K_{\boldsymbol{n}\boldsymbol{n}'}$ taken as reaction propensity functions. In other words, using $K_{\boldsymbol{n}\boldsymbol{n}'}$, we can construct an equivalent Markovian reaction network whose steady-state CME takes the form of Eq. (7).

In addition, using the conservation condition that the sum of all the transition probabilities satisfies $\sum_{\boldsymbol{n}'} \Phi_{\boldsymbol{n}\boldsymbol{n}'} = 1$, we can obtain

$$\tau_{\boldsymbol{n}} = \frac{1}{\sum_{\boldsymbol{n}'} K_{\boldsymbol{n}\boldsymbol{n}'}}, \quad \Phi_{\boldsymbol{n}\boldsymbol{n}'} = \frac{K_{\boldsymbol{n}\boldsymbol{n}'}}{\sum_{\boldsymbol{n}'} K_{\boldsymbol{n}\boldsymbol{n}'}}. \tag{8}$$

Note that $\Phi_{\boldsymbol{n}\boldsymbol{n}'}$ is actually an element of the transition probability matrix for the equivalent Markovian reaction network. If we let $\Pi_{\boldsymbol{n}}$ represent the distribution of visiting state $\boldsymbol{n}$, which denotes the fraction of visiting state $\boldsymbol{n}$ in a stationary regime, then $\Pi_{\boldsymbol{n}}$ can be expressed as $\Pi_{\boldsymbol{n}} = \sum_{\boldsymbol{n}'} \Pi_{\boldsymbol{n}'}\Phi_{\boldsymbol{n}'\boldsymbol{n}}$. Since the mean residence time for each reaction at state $\boldsymbol{n}$ is $\tau_{\boldsymbol{n}}$ (we will call the normalization constant $\tau_R = \sum_{\boldsymbol{n}} \Pi_{\boldsymbol{n}}\tau_{\boldsymbol{n}}$ the mean waiting time per reaction in the reaction network), the stationary distribution $P_{\boldsymbol{n}}$ of the whole system can be expressed as

$$P_{\boldsymbol{n}} = \frac{\tau_{\boldsymbol{n}}}{\tau_R}\Pi_{\boldsymbol{n}}. \tag{9}$$

The above analysis gives a framework for modeling a general reaction network. We emphasize that this analysis framework is suitable to not only Markovian (i.e., all the waiting times for reaction events follow exponential distributions) reaction networks but also non-Markovian (i.e., the waiting times for reaction events follow nonexponential distributions) reaction networks, thus having broad applications.

### B. A decomposition principle for entropy production rates

Recall that the relative entropy is defined as

$$E(P\|Q) = \sum_x P(x)\ln\frac{P(x)}{Q(x)}, \tag{10}$$

where $P(x)$ and $Q(x)$ are two probability distributions. The entropy defined in such a manner is an information-theoretic measure of distinguishability [55], and can quantify irreversibility and dissipation in nonequilibrium systems [43,56].

For the above reaction network, let us first consider a trajectory in the state space

$$\gamma = \{\boldsymbol{n}_1 \overset{t_1}{\to} \boldsymbol{n}_2 \overset{t_2}{\to} \cdots \boldsymbol{n}_{l-1} \overset{t_{l-1}}{\longrightarrow} \boldsymbol{n}_l \overset{t_l}{\to} \boldsymbol{n}_{l+1}\}, \tag{11}$$

and its reversal trajectory

$$\tilde{\gamma} = \{\boldsymbol{n}_l \overset{t_l}{\to} \boldsymbol{n}_{l-1} \overset{t_{l-1}}{\longrightarrow} \cdots \boldsymbol{n}_2 \overset{t_2}{\to} \boldsymbol{n}_1 \overset{t_1}{\to} \boldsymbol{n}_0\}, \tag{12}$$

if they exist, where $\sum_i t_i = t$. Equations (11) and (12) imply that the trajectory $\gamma$ and its reversal trajectory $\tilde{\gamma}$ have the same waiting time $t_i$ at state $\boldsymbol{n}_i$. Then, the occurring probabilities of the two trajectories are given, respectively, by

$$P(\gamma) = \phi_{\boldsymbol{n}_1\boldsymbol{n}_2}(t_1)\phi_{\boldsymbol{n}_2\boldsymbol{n}_3}(t_2)\cdots\phi_{\boldsymbol{n}_l\boldsymbol{n}_{l+1}}(t_l), \tag{13a}$$

$$P(\tilde{\gamma}) = \phi_{\boldsymbol{n}_1\boldsymbol{n}_0}(t_1)\phi_{\boldsymbol{n}_2\boldsymbol{n}_1}(t_2)\cdots\phi_{\boldsymbol{n}_l\boldsymbol{n}_{l-1}}(t_l). \tag{13b}$$

Similarly, if we consider two stationary trajectories of the equivalent Markovian reaction network as constructed above: a forward trajectory $\sigma = \{\boldsymbol{n}_1, \boldsymbol{n}_2, \boldsymbol{n}_3, \ldots \boldsymbol{n}_l, \boldsymbol{n}_{l+1}\}$, and a backward trajectory $\tilde{\sigma} = \{\boldsymbol{n}_l, \boldsymbol{n}_{l-1}, \ldots \boldsymbol{n}_2, \boldsymbol{n}_1, \boldsymbol{n}_0\}$, then the occurring probabilities of these two trajectories are given by

$$P(\sigma) = \Phi_{\boldsymbol{n}_1\boldsymbol{n}_2}\Phi_{\boldsymbol{n}_2\boldsymbol{n}_3}\cdots\Phi_{\boldsymbol{n}_l\boldsymbol{n}_{l+1}}, \tag{14a}$$

$$P(\tilde{\sigma}) = \Phi_{\boldsymbol{n}_1\boldsymbol{n}_0}\Phi_{\boldsymbol{n}_2\boldsymbol{n}_1}\cdots\Phi_{\boldsymbol{n}_l\boldsymbol{n}_{l-1}}. \tag{14b}$$

In order to calculate the relative entropy per state transition, we first define memory difference by considering

two state transitions $\boldsymbol{n}' \to \boldsymbol{n}''$ and $\boldsymbol{n}' \to \boldsymbol{n}$, denoted by $D[\phi_{\boldsymbol{n}''|\boldsymbol{n}'}(t)\|\phi_{\boldsymbol{n}|\boldsymbol{n}'}(t)]$, as

$$D[\phi_{\boldsymbol{n}''|\boldsymbol{n}'}(t)\|\phi_{\boldsymbol{n}|\boldsymbol{n}'}(t)] = \frac{1}{\Phi_{\boldsymbol{n}'\boldsymbol{n}''}}\int_0^\infty \phi_{\boldsymbol{n}'\boldsymbol{n}''}(t)\ln\frac{\phi_{\boldsymbol{n}'\boldsymbol{n}''}(t)}{\phi_{\boldsymbol{n}'\boldsymbol{n}}(t)}dt + \ln\frac{\Phi_{\boldsymbol{n}'\boldsymbol{n}}}{\Phi_{\boldsymbol{n}'\boldsymbol{n}''}}, \quad (15)$$

where states $\boldsymbol{n}, \boldsymbol{n}', \boldsymbol{n}''$ constitute a substring of the trajectory $\sigma$, that is, they take the forms:$\boldsymbol{n} = \boldsymbol{n}_m$, $\boldsymbol{n}' = \boldsymbol{n}_{m+1}$, $\boldsymbol{n}'' = \boldsymbol{n}_{m+2}$. Obviously, if two conditional waiting-time distributions $\phi_{\boldsymbol{n}''|\boldsymbol{n}'}(t)$ and $\phi_{\boldsymbol{n}|\boldsymbol{n}'}(t)$ at state $\boldsymbol{n}'$ are DTI and DTD, respectively, then the memory difference is $D[\phi_{\boldsymbol{n}''|\boldsymbol{n}'}(t)\|\phi_{\boldsymbol{n}|\boldsymbol{n}'}(t)] = 0$ and $D[\phi_{\boldsymbol{n}''|\boldsymbol{n}'}(t)\|\phi_{\boldsymbol{n}|\boldsymbol{n}'}(t)] > 0$, respectively. In other words, DTI distributions at state $\boldsymbol{n}'$ cannot produce memory difference whereas DTD distributions at state $\boldsymbol{n}'$ produces a positive memory difference.

Then, the relative entropy per state transition for all trajectories in the reaction network is calculated according to [43,48]

$$\begin{aligned}\delta_G &= \lim_{l\to\infty}\frac{1}{l}\sum_\gamma P(\gamma)\ln\frac{P(\gamma)}{P(\tilde{\gamma})}\\ &= \lim_{l\to\infty}\frac{1}{l}\sum_\sigma\int_0^\infty\cdots\int_0^\infty \phi_{\boldsymbol{n}_1\boldsymbol{n}_2}(t_1)\phi_{\boldsymbol{n}_2\boldsymbol{n}_3}(t_2)\cdots\phi_{\boldsymbol{n}_l\boldsymbol{n}_{l+1}}(t_l)\left[\ln\frac{\phi_{\boldsymbol{n}_1\boldsymbol{n}_2}(t_1)}{\phi_{\boldsymbol{n}_1\boldsymbol{n}_0}(t_1)} + \ln\frac{\phi_{\boldsymbol{n}_2\boldsymbol{n}_3}(t_2)}{\phi_{\boldsymbol{n}_2\boldsymbol{n}_1}(t_2)} + \cdots \ln\frac{\phi_{\boldsymbol{n}_l\boldsymbol{n}_{l+1}}(t_l)}{\phi_{\boldsymbol{n}_l\boldsymbol{n}_{l-1}}(t_l)}\right]dt_1dt_2\cdots dt_l.\end{aligned} \quad (16)$$

Substituting Eq. (15) into Eq. (16) and by simple calculation, we can obtain the following expression for the relative entropy (see Appendix A for derivation):

$$\delta_G = \sum_{\boldsymbol{n},\boldsymbol{n}'}\Pi_{\boldsymbol{n}}\Phi_{\boldsymbol{n}\boldsymbol{n}'}\ln\frac{\Phi_{\boldsymbol{n}\boldsymbol{n}'}}{\Phi_{\boldsymbol{n}'\boldsymbol{n}}} + \sum_{\boldsymbol{n},\boldsymbol{n}',\boldsymbol{n}''}\Pi_{\boldsymbol{n}}\Phi_{\boldsymbol{n}\boldsymbol{n}'}\Phi_{\boldsymbol{n}'\boldsymbol{n}''}D[\phi_{\boldsymbol{n}''|\boldsymbol{n}'}(t)\|\phi_{\boldsymbol{n}|\boldsymbol{n}'}(t)], \quad (17)$$

where $\Pi_{\boldsymbol{n}}\Phi_{\boldsymbol{n}\boldsymbol{n}'}$ is the probability for state transition $\boldsymbol{n} \to \boldsymbol{n}'$ whereas $\Pi_{\boldsymbol{n}}\Phi_{\boldsymbol{n}\boldsymbol{n}'}\Phi_{\boldsymbol{n}'\boldsymbol{n}''}$ is the probability for state transitions $\boldsymbol{n} \to \boldsymbol{n}' \to \boldsymbol{n}''$ along a forward trajectory. Note that the first sum on the right-hand side of Eq. (17) is just the relative entropy for the constructed Markovian reaction network, whereas the second sum represents the relative entropy generated due to the DTD of waiting-time distributions. In other words, the relative entropy for the entire reaction network can be split into two parts: the one for the equivalent Markovian reaction network and the other from the DTD of waiting-time distributions.

On the other hand, for a nonequilibrium system, energy dissipation can be quantified by the entropy production rate [13]. Here, we try to derive a formula for calculating the entropy production rate for the above reaction network. Note that the entropy production per unit time, denoted by $\dot{S}_G$, is calculated according to $\dot{S}_G = \frac{\delta_G}{\tau_R}$, where $\tau_R$ is the mean waiting time for every reaction in the reaction network. Thus, according to Eq. (17), $\dot{S}_G$ can be decomposed as

$$\dot{S}_G = \dot{S}_E + \dot{S}_M, \quad (18)$$

where

$$\dot{S}_E = \sum_{\boldsymbol{n},\boldsymbol{n}'}P_{\boldsymbol{n}}K_{\boldsymbol{n}\boldsymbol{n}'}\ln\frac{K_{\boldsymbol{n}\boldsymbol{n}'}}{K_{\boldsymbol{n}'\boldsymbol{n}}}, \quad (18a)$$

which is the entropy production rate for the equivalent Markovian system, and

$$\dot{S}_M = \sum_{\boldsymbol{n},\boldsymbol{n}',\boldsymbol{n}''}P_{\boldsymbol{n}}K_{\boldsymbol{n}\boldsymbol{n}'}\frac{K_{\boldsymbol{n}'\boldsymbol{n}''}}{\sum_{\boldsymbol{n}''}K_{\boldsymbol{n}'\boldsymbol{n}''}}D[\phi_{\boldsymbol{n}''|\boldsymbol{n}'}(t)\|\phi_{\boldsymbol{n}|\boldsymbol{n}'}(t)], \quad (18b)$$

which represents the entropy production due to the property of the waiting-time distributions (more precisely, it is zero if all the waiting-time distributions are DTI, and positive otherwise). In particular, if the waiting times for all the states follow exponential distributions, the above energy dissipation decomposition reduces to a previous result obtained in the Markovian case [11].

Note that $\dot{S}_E = 0$ if the equivalent Markovian system is in a detailed balance, and $\dot{S}_E > 0$ if the detailed balance is broken; $\dot{S}_M = 0$ if all the waiting-time distributions are DTI, and $\dot{S}_M > 0$ if waiting-time distributions are DTD. Thus, we obtain three possible decomposition patterns for the entropy production rate of an nonequilibrium reaction system: (1) $\dot{S}_E > 0$ and $\dot{S}_M > 0$; (2) $\dot{S}_E > 0$ and $\dot{S}_M = 0$; and (3) $\dot{S}_E = 0$ and $\dot{S}_M > 0$, referring to Fig. 1. The decomposition given by Eq. (18) shows an important difference in energy dissipation (or entropy production rate) between the non-Markovian system and its equivalent Markovian system, that is, in contrast to the latter, the former can have extra dissipation if waiting-time distributions are DTD.

From Eq. (18b), we know that nonzero entropy production rate $\dot{S}_M$ is contributed by memory differences, and it is positive if the conditional waiting-time distributions $\phi_{\boldsymbol{n}''|\boldsymbol{n}'}(t)$ and $\phi_{\boldsymbol{n}|\boldsymbol{n}'}(t)$ at state $\boldsymbol{n}'$ are DTD. To quantitatively show how molecular memory impacts dissipation in the above non-Markovian reaction system, we define mean memory difference as

$$\langle D\rangle = \sum_{\boldsymbol{n},\boldsymbol{n}',\boldsymbol{n}''}P_{\boldsymbol{n}'}D[\phi_{\boldsymbol{n}''|\boldsymbol{n}'}(t)\|\phi_{\boldsymbol{n}|\boldsymbol{n}'}(t)], \quad (19)$$

which characterizes the mean effect of molecular memory.

## III. APPLICATIONS

Using the above general theory, here we analyze two examples: a CTRW model and a model of stochastic gene

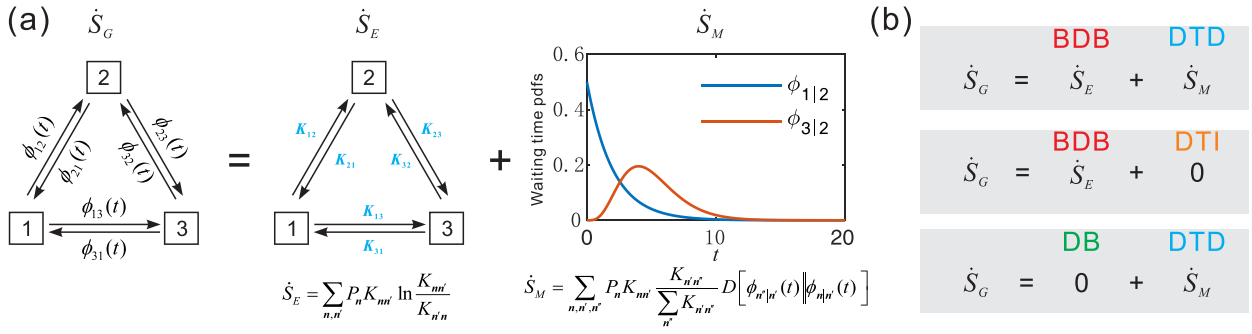


FIG. 1. Shown is an example for energy dissipation decomposition. (a) Schematic description of the decomposition, showing that the total entropy production rate is split into two parts: one for the equivalent Markovian system, the other due to molecular memory characterized by nonexponential waiting-time distributions. (b) Three possible decomposition patterns for nontrivial entropy production rate, where BDB is an abbreviation for broken detailed balance; DTD for direction-time dependence; DTI for direction-time independence; and DB for detailed balance.

expression. For the former, we show that DB and DTI together correspond to no energy dissipation even though the random-walk process is non-Markovian. For the latter example, we find that different memory mechanisms can lead to different patterns of dissipation. These results indicate that molecular memory has an important influence on energy dissipation of biochemical reaction systems.

### A. Analysis of a CTRW model

#### 1. Model setting

Consider a CTRW model, referring to Fig. 1(a) where three states 1, 2, and 3 form a loop. Here we use mixed Gamma distributions to model the memory effect at every state, and assume the waiting-time distributions for jump from state $i$ to state $j$, $\phi_{ij}(t)$ $(i, j = 1, 2, 3)$, take the forms: $\phi_{12}(t) = p_1 \frac{\lambda_{12}^{k_{12}} t^{k_{12}-1}}{\Gamma(k_{12})} e^{-\lambda_{12} t}$, $\phi_{13}(t) = q_1 \frac{\lambda_{13}^{k_{13}} t^{k_{13}-1}}{\Gamma(k_{13})} e^{-\lambda_{13} t}$, $\phi_{23}(t) = p_2 \frac{\lambda_{23}^{k_{23}} t^{k_{23}-1}}{\Gamma(k_{23})} e^{-\lambda_{23} t}$, $\phi_{21}(t) = q_2 \frac{\lambda_{21}^{k_{21}} t^{k_{21}-1}}{\Gamma(k_{21})} e^{-\lambda_{21} t}$, $\phi_{31}(t) = p_3 \frac{\lambda_{31}^{k_{31}} t^{k_{31}-1}}{\Gamma(k_{31})} e^{-\lambda_{31} t}$, and $\phi_{32}(t) = q_3 \frac{\lambda_{32}^{k_{32}} t^{k_{32}-1}}{\Gamma(k_{32})} e^{-\lambda_{32} t}$, where $p_j + q_j = 1$ $(j = 1, 2, 3)$. Symbols $\lambda_{ij}$ $(i, j = 1, 2, 3)$ are positive constants and will be called memory parameters throughout this paper, and symbols $k_{ij}$ $(i, j = 1, 2, 3)$ are positive integers and will be called memory indices (remark: a larger value of memory index means that the memory is stronger). We point out that all $k_{ij} = 1$ and $\lambda_{jk} = \lambda_{ji}$ correspond to the Markovian case whereas other parameter values to the non-Markovian case. Note that if $\lambda_{jk} \neq \lambda_{ji}$ $(\lambda_{jk} = \lambda_{ji})$, then the waiting-time distributions at state $j$ will be DTD (DTI).

In the following, for analysis convenience but without loss of generality, we set $k_{ij} = 2$ $(i, j = 1, 2, 3)$. Then, the effective transition rates, $K_{ij}$, take the forms $K_{12} = \frac{p_1\lambda_{12}\lambda_{13}}{2(p_1\lambda_{13}+q_1\lambda_{12})}$, $K_{13} = \frac{q_1\lambda_{12}\lambda_{13}}{2(p_1\lambda_{13}+q_1\lambda_{12})}$, $K_{23} = \frac{p_2\lambda_{21}\lambda_{23}}{2(p_2\lambda_{21}+q_2\lambda_{23})}$, $K_{21} = \frac{q_2\lambda_{21}\lambda_{23}}{2(p_2\lambda_{21}+q_2\lambda_{23})}$, $K_{31} = \frac{p_3\lambda_{31}\lambda_{32}}{2(p_3\lambda_{32}+q_3\lambda_{31})}$, and $K_{32} = \frac{q_3\lambda_{31}\lambda_{32}}{2(p_3\lambda_{32}+q_3\lambda_{31})}$. In addition, the stationary distribution $P_i$ $(i = 1, 2, 3)$ can be easily obtained by solving the following sgCME:

$$\begin{aligned} -(K_{12} + K_{13})P_1 + K_{21}P_2 + K_{31}P_3 &= 0, \\ K_{12}P_1 - (K_{21} + K_{23})P_2 + K_{32}P_3 &= 0, \\ K_{13}P_1 + K_{23}P_2 - (K_{32} + K_{31})P_3 &= 0, \end{aligned} \tag{20}$$

with the conservation condition $\sum_{i=1}^{3} P_i = 1$.

#### 2. Calculation of energy dissipation

First, we can show that the dissipation due to broken detailed balance is given by

$$\dot{S}_E = \frac{1}{U}(K_{12}K_{23}K_{31} - K_{21}K_{32}K_{13}) \ln \frac{K_{12}K_{23}K_{31}}{K_{21}K_{32}K_{13}}, \tag{21}$$

where $U = K_{12}K_{23} + K_{23}K_{31} + K_{31}K_{12} + K_{21}K_{13} + K_{32}K_{13} + K_{21}K_{32} + K_{12}K_{32} + K_{23}K_{13} + K_{31}K_{21}$. Note that $\ln(K_{12}K_{23}K_{31})/(K_{21}K_{32}K_{13}) = \ln(p_1p_2p_3)/(q_1q_2q_3)$. Therefore, the condition for DB reduces to $p_1p_2p_3 = q_1q_2q_3$, which is independent of the memory parameters.

Second, the dissipation due to molecular memory takes the form

$$\begin{aligned} \dot{S}_M = {} & P_1K_{12}p_2D[\phi_{3|2}(t)\|\phi_{1|2}(t)] + P_1K_{13}q_3D[\phi_{2|3}(t)\|\phi_{1|3}(t)] + P_2K_{23}p_3D[\phi_{1|3}(t)\|\phi_{2|3}(t)] + P_2K_{21}q_1D[\phi_{3|1}(t)\|\phi_{2|1}(t)] \\ & + P_3K_{31}p_1D[\phi_{2|1}(t)\|\phi_{3|1}(t)] + P_3K_{32}q_2D[\phi_{1|2}(t)\|\phi_{3|2}(t)], \end{aligned} \tag{22}$$

where the memory difference is given by

$$D[\phi_{k|j}(t)\|\phi_{i|j}(t)] = 2\left(\ln\frac{\lambda_{jk}}{\lambda_{ji}} + \frac{\lambda_{ji}}{\lambda_{jk}} - 1\right), \tag{22a}$$

with $\phi_{k|j}(t)$ and $\phi_{i|j}(t)$ being conditional waiting-time distributions of jumps $j \to k$ and $j \to i$, respectively. From Eqs. (22) and (22a), we find that if there is a state $j$ whose waiting-time distribution is DTD, i.e., if the ratio of memory parameter satisfies $\lambda_{jk}/\lambda_{ji} \neq 1$, then we have $\dot{S}_M > 0$.

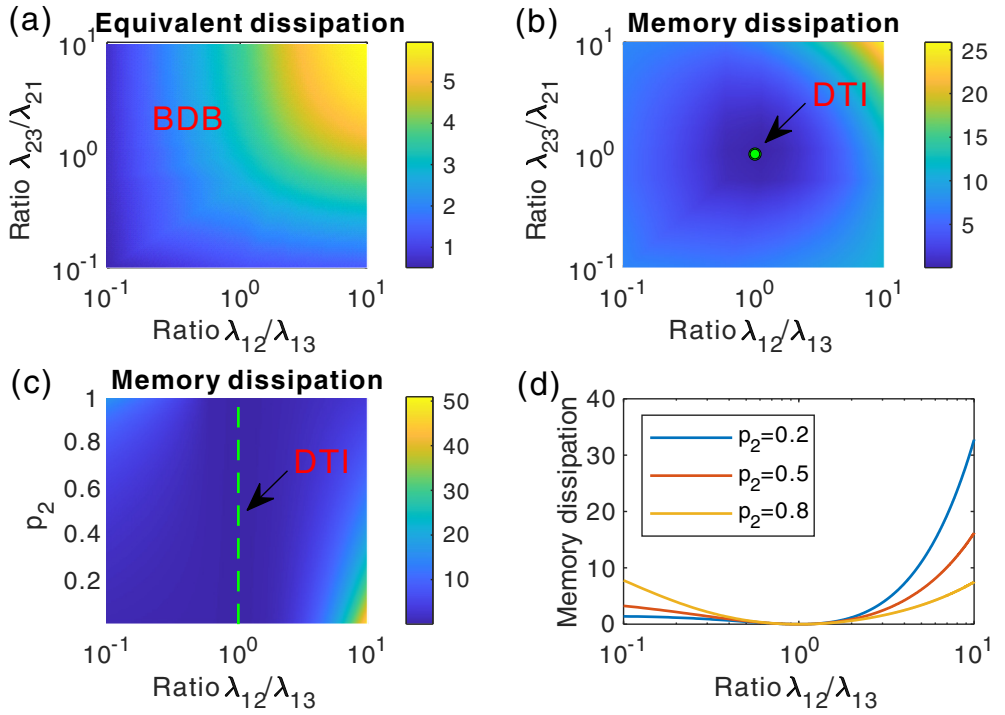


FIG. 2. Energy dissipation in a CTRW model. (a) The global scenario for dependence of the equivalent dissipation $\dot{S}_E$ on both ratios $\lambda_{12}/\lambda_{13}$ and $\lambda_{23}/\lambda_{21}$. (b) The global scenario for dependence of the memory dissipation $\dot{S}_M$ on both ratios $\lambda_{12}/\lambda_{13}$ and $\lambda_{23}/\lambda_{21}$. (c) the global scenario for dependence of the memory dissipation $\dot{S}_M$ on both ratio $\lambda_{12}/\lambda_{13}$ and transition probability $p_2$ if the system is in DB. (d) Special case of (c), which shows dependence of the memory dissipation $\dot{S}_M$ on ratio $\lambda_{12}/\lambda_{13}$ for three fixed transition probabilities $p_2$. The transition probabilities are set as: $p_1 = 0.6$, $p_2 = 0.7$, $p_3 = 0.8$ for (a) and (b), and $p_1 = 1 - p_2$, $p_3 = 0.5$ for (c) and (d), and other parameters are set as: $\lambda_{13} = 10$, $\lambda_{21} = 20$, $\lambda_{31} = \lambda_{32} = 50$. In addition, $\lambda_{23} = 20$ is set for (c) and (d).

On the other hand, if the waiting-time distributions of all states $j(j = 1, 2, 3)$ are DTI, i.e., all the ratios of memory parameters equal the unit, i.e., if $\lambda_{jk}/\lambda_{ji} = 1$, then we can get $\dot{S}_M = 0$. Moreover, we find that the memory difference at state $j$ as defined by Eq. (22a) is a monotonically decreasing function of ratio $\lambda_{jk}/\lambda_{ji}$ if $\lambda_{jk}/\lambda_{ji} < 1$, but a monotonically increasing function of ratio $\lambda_{jk}/\lambda_{ji}$ if $\lambda_{jk}/\lambda_{ji} > 1$. These analyses imply that both the memory difference and the energy dissipation have a direct correlation with the ratios of memory parameters, $\lambda_{jk}/\lambda_{ji}$. In the following, we will analyze how the ratios of memory parameters quantitatively affect the energy dissipation of the underlying system.

### 3. *Numerical results*

The numerical results are shown in Fig. 2. For convenience, we will call the entropy production rate corresponding to the equivalent Markovian system "equivalent dissipation," whereas the entropy production rate corresponding to waiting-time distributions "memory dissipation."

Figure 2(a) shows how the ratios of memory parameters affect the equivalent dissipation whereas Fig. 2(b) shows how these ratios affect the memory dissipation in the case of BDB. We observe that when one of the memory parameter ratios is fixed, the equivalent dissipation monotonically increases with the other ratio of memory parameters. This is because increasing the ratio of memory parameter can lead to a faster switching rate, which in turn can result in more energy consumption (this is a known fact in the Markovian case [34]). On the other hand, we observe that the memory dissipation has a minimal value 0 when the two ratios of memory parameter simultaneously equal 1, i.e., the waiting-time distributions for all states are DTI. In addition, the memory dissipation increases as the ratio of memory parameter is far from 1. In particular, it can become larger as this ratio further increases, implying that a stronger memory leads to more memory dissipation. Figure 2(c) shows how both the ratio of memory parameter and transition probability $p_2$ affect the memory dissipation in the case of DB (i.e., the equivalent dissipation $\dot{S}_E = 0$). For this, we change transition probabilities $p_1$ or $p_2$ while keeping the property of DB. From Fig. 2(d), we observe that there is trivial (nontrivial) dissipation if the ratio of memory parameter equals (does not equal) 1. These results imply that DB and DTI together are equivalent to the absence of energy dissipation in the non-Markovian process.

Recall that BDB can lead to energy dissipation in a Markovian process [57]. Here we have shown that BDB or DTD can also lead to energy dissipation in the non-Markovian process. It is worth mentioning that the previous studies implied several same characteristics such as the steady-state flux, exit probabilities, and mean trapping times between a non-Markovian CTRW process and its equivalent Markovian process [45]. However, we can distinguish these two processes by the fact that the non-Markovian CTRW has extra dissipation in contrast to the Markovian CTRW.

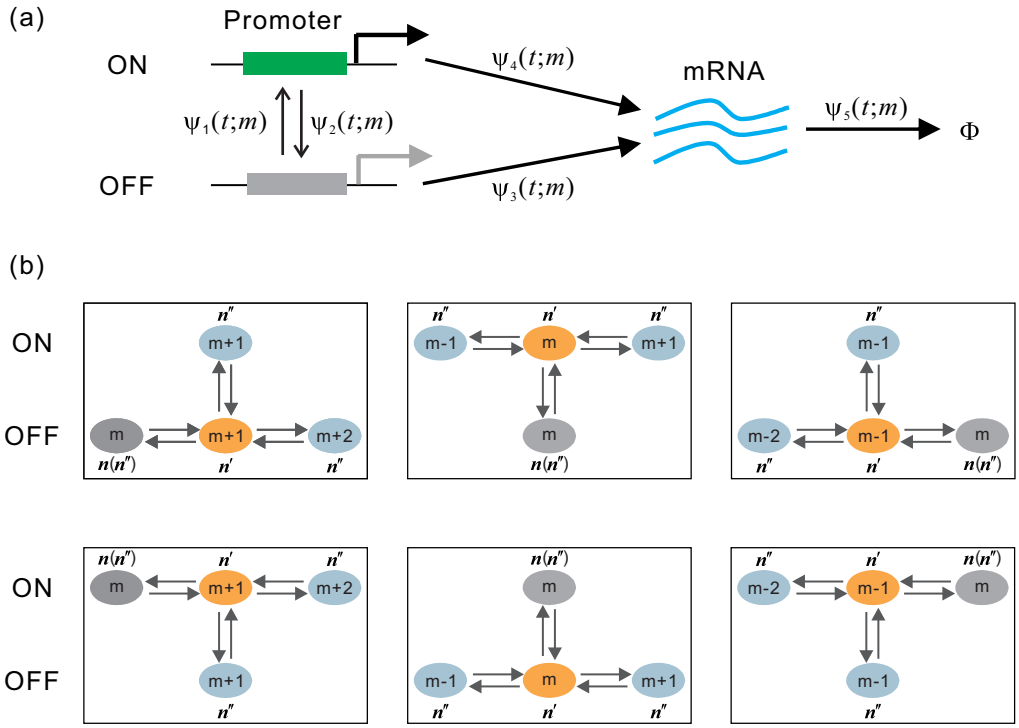


FIG. 3. (a) Schematic description of a generalized gene expression model, where the reaction processes are characterized by waiting-time distributions as indicated. (b) Schematic description of state transitions with six possible patterns. The top row figures show that the current state is at OFF state whereas the bottom row figures show that the current state is at ON state. In all cases, $\boldsymbol{n}$, $\boldsymbol{n'}$, and $\boldsymbol{n''}$ denote the current state, next state, and last sate, respectively. Note that the current state $\boldsymbol{n}$ may be the last state $\boldsymbol{n''}$.

### B. Analysis of a generalized gene expression model

#### 1. Model description

Regulation of gene expression often involves binding and unbinding of many transcription factors (TFs), leading to complex promoter structure [58,59]. For example, although the promoter for repressor maintenance (PRM) promoter of phage lambada in *Escherichia coli* is regulated by two different TFs binding to two sets of three operators, this can lead to the PRM promoter states' number up to 128 [60]. Compared with prokaryote promoter, eukaryotic promoters would be more complex since they would involve nucleosomes that compete with or are removed by TFs [61]. In addition, epigenetic regulation via histone modifications can also lead to complex promoter structures [62]. A previous study showed that the process of gene activation can create narrowly distributed gestation periods between transcription windows [37]. In addition to gene activation, gene expression also involves transcription initiation, elongation, and release [63–65]. For example, Xu *et al.* presented a coarse-grained model that the delay from initiation to elimination of nascent RNA is modeled as a fixed quantity, and predicted discontinuous distribution of nascent RNA [66]. Filatova *et al.* considered a stochastic model in which the elongation time is modeled as a random variable, and uncovered the explicit dependence of the statistics of both types of RNA on transcriptional parameters [67]. Jiang *et al.* used neural networks to approximate the time-dependent distributions and infer the parameters of gene expression models [68]. These studies indicate that molecular memory exists extensively in gene expression. Given the complexity, we introduce waiting-time distributions to simplify the modeling of complex processes occurring in gene expression. As such, we can assume that the promoter of a gene has one active (ON) state and one inactive (OFF) state, although it would have many activity states in a realistic case.

Let $m$ be the molecular number of mRNAs. For the gene expression model depicted in Fig. 3(a), biochemical reactions are listed below:

$$\begin{aligned}
&\mathrm{OFF} \xrightarrow{\psi_1(t;m)} \mathrm{ON}, \quad \mathrm{ON} \xrightarrow{\psi_2(t;m)} \mathrm{OFF},\\
&\mathrm{OFF} \xrightarrow{\psi_3(t;m)} \mathrm{OFF} + \mathrm{mRNA},\\
&\mathrm{ON} \xrightarrow{\psi_4(t;m)} \mathrm{ON} + \mathrm{mRNA},\\
&\mathrm{mRNA} \xrightarrow{\psi_5(t;m)} \phi,
\end{aligned} \tag{23}$$

where $\psi_1(t;m)$ and $\psi_2(t;m)$ are waiting-time distributions for promoter switches from OFF to ON and vice versa, respectively, $\psi_3(t;m)$ and $\psi_4(t;m)$ for mRNA production at OFF state and ON state, respectively, and $\psi_5(t;m)$ for mRNA degradation. Since Gamma distributions can model multistep processes, we can set $\psi_1(t;m) = [\Gamma(L_1)]^{-1}(\lambda_1)^{L_1} t^{L_1-1} e^{-\lambda_1 t}$, $\psi_2(t;m) = [\Gamma(L_2)]^{-1}(\lambda_0)^{L_2} t^{L_2-1} e^{-\lambda_0 t}$, where $L_1$ and $L_2$ are positive integers with each called memory index, $\lambda_1$ and $\lambda_0$ are the mean rates of gene activation and inactivation, respectively. In addition, we set $\psi_3(t;m) = \mu_0 e^{-\mu_0 t}$, $\psi_4(t;m) = \mu e^{-\mu t}$, and $\psi_5(t;m) = \delta m e^{-\delta m t}$ for simplicity, where $\mu$ is the transcription rate whereas $\mu_0$ is the leakage rate, and $\delta$ is the degradation rate.

According to the definition of $\Psi_i(t;m)$ and the setting of $\psi_i(t;m)$, simple calculations can show $\Psi_1(t;m) = 1-\sum_{i=0}^{L_1-1}\frac{(\lambda_1 t)^i}{i!}e^{-\lambda_1 t}$, $\Psi_2(t;m) = 1-\sum_{i=0}^{L_2-1}\frac{(\lambda_0 t)^i}{i!}e^{-\lambda_0 t}$, $\Psi_3(t;m) = 1-e^{-\mu_0 t}$, $\Psi_4(t;m) = 1-e^{-\mu t}$, and $\Psi_5(t;m) = 1-e^{-\delta m t}$. In addition, we have $\phi_1(t;m) = \frac{(\lambda_1)^{L_1}}{\Gamma(L_1)}t^{L_1-1}e^{-(\lambda_1+\mu_0+\delta m)t}$,

$$\phi_2(t;m) = \frac{(\lambda_0)^{L_2}}{\Gamma(L_2)}t^{L_2-1}e^{-(\lambda_0+\mu+\delta m)t},$$

$$\phi_3(t;m) = \mu_0 e^{-(\lambda_1+\mu_0+\delta m)t}\left[\sum_{i=0}^{L_1-1}\frac{(\lambda_1 t)^i}{i!}\right],$$

$$\phi_4(t;m) = \mu e^{-(\lambda_0+\mu+\delta m)t}\left[\sum_{i=0}^{L_2-1}\frac{(\lambda_0 t)^i}{i!}\right],$$

$$\phi_5(t;m) = \delta m e^{-[\lambda_1(1-I_{\rm on})+\lambda_0 I_{\rm on}+\mu_0(1-I_{\rm on})+\mu I_{\rm on}+\delta m]t}\left[\sum_{i=0}^{L_1-1}\frac{(\lambda_1 t-\lambda_1 I_{\rm on}t)^i}{i!}\right]\left[\sum_{j=0}^{L_2-1}\frac{(\lambda_0 I_{\rm on}t)^j}{j!}\right],$$

where the indicator function $I_{\rm on}=1$ if the gene is at the ON state, but $I_{\rm on}=0$ if the gene is at the OFF state (we will always make this understanding throughout this paper). Thus,$\Phi_1(m) = (\frac{\lambda_1}{\lambda_1+\mu_0+\delta {\rm m}})^{L_1}$, $\Phi_2(m) = (\frac{\lambda_0}{\lambda_0+\mu+\delta {\rm m}})^{L_2}$, $\Phi_3(m) = \mu_0\sum_{i=0}^{L_1-1}\frac{\lambda_1{}^i}{(\lambda_1+\mu_0+\delta {\rm m})^{i+1}}$, $\Phi_4(m) = \mu\sum_{i=0}^{L_2-1}\frac{\lambda_0{}^i}{(\lambda_0+\mu+\delta m)^{i+1}}$, and

$$\Phi_5(m) = \delta m\sum_{i=0}^{L_1-1}\sum_{j=0}^{L_2-1}\binom{i+j}{j}\frac{(\lambda_1-\lambda_1 I_{\rm on})^i(\lambda_0 I_{\rm on})^j}{[\lambda_1(1-I_{\rm on})+\lambda_0 I_{\rm on}+\mu_0(1-I_{\rm on})+\mu I_{\rm on}+\delta m]^{i+j+1}}.$$

According to definition, the mean residence time at state $\boldsymbol{n} = ({\rm off}, {\rm on}, m)$ is given by

$$\tau_{\boldsymbol{n}} = \sum_{i=0}^{L_1-1}\sum_{j=0}^{L_2-1}\binom{i+j}{j}\frac{(\lambda_1-\lambda_1 I_{\rm on})^i(\lambda_0 I_{\rm on})^j}{[\lambda_1(1-I_{\rm on})+\lambda_0 I_{\rm on}+\mu_0(1-I_{\rm on})+\mu I_{\rm on}+\delta m]^{i+j+1}}. \tag{24}$$

The effective transition rates are given by

$$\begin{aligned} K_1(m) &= (\mu_0+\delta m)\frac{(\lambda_1)^{L_1}/(\lambda_1+\mu_0+\delta m)^{L_1}}{1-(\lambda_1)^{L_1}/(\lambda_1+\mu_0+\delta m)^{L_1}},\\ K_2(m) &= (\mu+\delta m)\frac{(\lambda_0)^{L_2}/(\lambda_0+\mu+\delta m)^{L_2}}{1-(\lambda_0)^{L_2}/(\lambda_0+\mu+\delta m)^{L_2}},\\ K_3(m) &= \mu_0, \quad K_4(m) = \mu, \quad K_5(m) = \delta m. \end{aligned} \tag{25}$$

Let $P_0(m)$ and $P_1(m)$ be the stationary probabilities that mRNA has $m$ molecules at states OFF and ON, respectively. Then we can obtain the following sgCME:

$$\begin{aligned} -K_1(m)P_0(m)+K_2(m)P_1(m)+K_3(m-1)P_0(m-1)-K_3(m)P_0(m)+K_5(m+1)P_0(m+1)-K_5(m)P_0(m) = 0,\\ K_1(m)P_0(m)-K_2(m)P_1(m)+K_4(m-1)P_1(m-1)-K_4(m)P_1(m)+K_5(m+1)P_1(m+1)-K_5(m)P_1(m) = 0. \end{aligned} \tag{26}$$

In general, Eq. (26) has no analytical solution but we can solve it using a numerical method (see Appendix B for details). Furthermore, the mean mRNA is calculated according to

$$\langle m\rangle = \sum_m mP(m), \tag{27}$$

where $P(m) = P_0(m)+P_1(m)$ is the total stationary probability. Note that if we take $K_i(m)$ as the reaction-propensity function for the $i$th reaction in the topologically equivalent gene expression model without molecular memory, then the stationary CME of the non-Markovian system is exactly the same as that of the equivalent Markovian system.

### *2. Decomposition of energy dissipation*

For convenience, we call the energy dissipation corresponding to the equivalent Markovian system "equivalent dissipation," and that corresponding to the memory part "memory dissipation." First, the equivalent dissipation shown in Eq. (18a) becomes

$$\begin{aligned} \dot{S}_E = &\sum_m P_0(m)\left[K_1(m)\log\frac{K_1(m)}{K_2(m)}+K_3(m)\log\frac{K_3(m)}{K_5(m+1)}+K_5(m)\log\frac{K_5(m)}{K_3(m-1)}\right]\\ &+\sum_m P_1(m)\left[K_2(m)\log\frac{K_2(m)}{K_1(m)}+K_4(m)\log\frac{K_4(m)}{K_5(m+1)}+K_5(m)\log\frac{K_5(m)}{K_4(m-1)}\right]. \end{aligned} \tag{28}$$

Note that if all the reaction processes are Markovian, Eq. (28) further reduces to a previous case [8]. In the following, we denote $R_i(m) = \frac{\phi_i(t;m)}{\Phi_i(m)}$ $(1 \leqslant i \leqslant 4)$ and $R_5^X(m) = \frac{\phi_5^X(t;m)}{\Phi_5^X(m)}$ (where $X =$ on or off) for convenience.

Then, the memory dissipation can be decomposed as

$$\dot{S}_M = \dot{S}_M^{(\mathrm{off})} + \dot{S}_M^{(\mathrm{on})}, \tag{29}$$

where

$$\begin{aligned}
\dot{S}_M^{(\mathrm{off})} = & \sum_m P_0(m)K_1(m)\frac{K_4(m)D[R_4(m)\|R_2(m)] + K_5(m)D\big[R_5^{\mathrm{on}}(m)\big\|R_2(m)\big]}{K_2(m)+K_4(m)+K_5(m)} \\
& + \sum_m P_0(m)K_3(m)\frac{K_1(m+1)D\big[R_1(m+1)\big\|R_5^{\mathrm{off}}(m+1)\big] + K_3(m+1)D\big[R_3(m+1)\big\|R_5^{\mathrm{off}}(m+1)\big]}{K_1(m+1)+K_3(m+1)+K_5(m+1)} \\
& + \sum_m P_0(m)K_5(m)\frac{K_1(m-1)D[R_1(m-1)\|R_3(m-1)] + K_5(m-1)[R_5^{\mathrm{off}}(m-1)\|R_3(m-1)]}{K_1(m-1)+K_3(m-1)+K_5(m-1)},
\end{aligned} \tag{29a}$$

$$\begin{aligned}
\dot{S}_M^{(\mathrm{on})} = & \sum_m P_1(m)K_2(m)\frac{K_3(m)D[R_3(m)\|R_1(m)] + K_5(m)D[R_5^{\mathrm{off}}(m)\|R_1(m)]}{K_1(m)+K_3(m)+K_5(m)} \\
& + \sum_m P_1(m)K_4(m)\frac{K_4(m+1)D\big[R_4(m+1)\big\|R_5^{\mathrm{on}}(m+1)\big] + K_2(m+1)D\big[R_2(m+1)\big\|R_5^{\mathrm{on}}(m+1)\big]}{K_2(m+1)+K_4(m+1)+K_5(m+1)} \\
& + \sum_m P_1(m)K_5(m)\frac{K_2(m-1)D[R_2(m-1)\|R_4(m-1)] + K_5(m-1)D\big[R_5^{\mathrm{on}}(m-1)\big\|R_4(m-1)\big]}{K_2(m-1)+K_4(m-1)+K_5(m-1)}.
\end{aligned} \tag{29b}$$

In the following, we consider two special kinds of cases, which correspond to two different kinds of memory mechanisms:

*Case 1:* $L_1 \geqslant 1, \quad L_2 = 1$

In this case, the reaction processes are multistep activation with memory and single-step deactivation without memory. So Eq. (29a) reduces to

$$\dot{S}_M^{(\mathrm{off})} = \sum_m P_0(m)K_3(m)\frac{K_1(m+1)D\big[R_1(m+1)\big\|R_5^{\mathrm{off}}(m+1)\big]}{K_1(m+1)+K_3(m+1)+K_5(m+1)} + \sum_m P_0(m)K_5(m)\frac{K_1(m-1)D[R_1(m-1)\|R_3(m-1)]}{K_1(m-1)+K_3(m-1)+K_5(m-1)}, \tag{30}$$

where the memory difference is given by

$$\begin{aligned}
D\big[R_1(m)\big\|R_5^{\mathrm{off}}(m)\big] &= D[R_1(m)\|R_3(m)] \\
&= -\ln\Gamma(L_1) + (L_1-1)\psi(L_1) + \ln(\lambda_1+\mu_0+\delta m) + \ln\left(\sum_{i=0}^{L_1-1}\frac{\lambda_1^i}{(\lambda_1+\mu_0+\delta m)^{i+1}}\right) \\
&\quad - \int_0^\infty \frac{(\lambda_1+\mu_0+\delta m)^{L_1}}{\Gamma(L_1)} t^{L_1-1}e^{-(\lambda_1+\mu_0+\delta m)t}\ln\left(\sum_{i=0}^{L_1-1}\frac{(\lambda_1 t)^i}{i!}\right)dt.
\end{aligned} \tag{30a}$$

Equation (29b) reduces to

$$\dot{S}_M^{(\mathrm{on})} = \sum_m P_1(m)K_2(m)\frac{K_3(m)D[R_3(m)\|R_1(m)] + K_5(m)D[R_5^{\mathrm{off}}(m)\|R_1(m)]}{K_1(m)+K_3(m)+K_5(m)}, \tag{31}$$

where the memory difference is given by

$$D[R_3(m)\|R_1(m)] = D\big[R_5^{\rm off}(m)\big\|R_1(m)\big] = \ln\Gamma(L_1) - (L_1-1)\frac{\sum_{i=0}^{L_1-1}\frac{(\lambda_1)^i\psi(i+1)}{(\lambda_1+\mu_0+\delta{\rm m})^{i+1}}}{\sum_{i=0}^{L_1-1}\frac{(\lambda_1)^i}{(\lambda_1+\mu_0+\delta{\rm m})^{i+1}}} - \ln(\lambda_1+\mu_0+\delta m)$$
$$- \ln\left(\sum_{i=0}^{L_1-1}\frac{\lambda_1^i}{(\lambda_1+\mu_0+\delta m)^{i+1}}\right)$$
$$+ \left(\sum_{i=0}^{L_1-1}\frac{(\lambda_1)^i}{(\lambda_1+\mu_0+\delta m)^{i+1}}\right)^{-1}\int_0^\infty e^{-(\lambda_1+\mu_0+\delta m)t}\left(\sum_{i=0}^{L_1-1}\frac{(\lambda_1 t)^i}{i!}\right)\ln\left(\sum_{i=0}^{L_1-1}\frac{(\lambda_1 t)^i}{i!}\right)dt. \tag{31a}$$

In Eq. (31a), $\psi(x)$ is the Digamma function, and satisfies the iteration relationship: $\psi(x+1)=\psi(x)+1/x$, where $\psi(1)=-\gamma$ with $\gamma$ being the Euler's constant.

*Case 2:* $L_1 = 1, \quad L_2 \geqslant 1$

In this case, the reaction processes are single-step activation without memory and multistep deactivation with memory. So Eq. (29a) reduces to

$$\dot{S}_M^{(\rm off)} = \sum_m P_0(m)K_1(m)\frac{K_4(m)D[R_4(m)\|R_2(m)] + K_5(m)D\big[R_5^{\rm on}(m)\big\|R_2(m)\big]}{K_2(m)+K_4(m)+K_5(m)}, \tag{32}$$

where the memory difference is given by

$$D[R_4(m)\|R_2(m)] = D\big[R_5^{\rm on}(m)\big\|R_2(m)\big]$$
$$= \ln\Gamma(L_2) - (L_2-1)\frac{\sum_{i=0}^{L_2-1}\frac{(\lambda_0)^i\psi(i+1)}{(\lambda_0+\mu+\delta{\rm m})^{i+1}}}{\sum_{i=0}^{L_2-1}\frac{(\lambda_0)^i}{(\lambda_0+\mu+\delta{\rm m})^{i+1}}} - \ln(\lambda_0+\mu+\delta m) - \ln\left(\sum_{i=0}^{L_2-1}\frac{\lambda_0^i}{(\lambda_0+\mu+\delta m)^{i+1}}\right)$$
$$+ \left(\sum_{i=0}^{L_2-1}\frac{(\lambda_0)^i}{(\lambda_0+\mu+\delta m)^{i+1}}\right)^{-1}\int_0^\infty e^{-(\lambda_0+\mu+\delta m)t}\left(\sum_{i=0}^{L_2-1}\frac{(\lambda_0 t)^i}{i!}\right)\ln\left(\sum_{i=0}^{L_2-1}\frac{(\lambda_0 t)^i}{i!}\right)dt. \tag{32a}$$

Equation (29b) reduces to

$$\dot{S}_M^{(\rm on)} = \sum_m P_1(m)\frac{K_4(m)K_2(m+1)D\big[R_2(m+1)\big\|R_5^{\rm on}(m+1)\big]}{K_2(m+1)+K_4(m+1)+K_5(m+1)} + \sum_m P_1(m)\frac{K_5(m)K_2(m-1)D[R_2(m-1)\|R_4(m-1)]}{K_2(m-1)+K_4(m-1)+K_5(m-1)}, \tag{33}$$

where the memory difference is given by

$$D\big[R_2(m)\big\|R_5^{\rm on}(m)\big] = D[R_2(m)\|R_4(m)]$$
$$= -\ln\Gamma(L_2) + (L_2-1)\psi(L_2) + \ln(\lambda_0+\mu+\delta m) + \ln\left(\sum_{i=0}^{L_2-1}\frac{(\lambda_0)^i}{(\lambda_0+\mu+\delta m)^{i+1}}\right)$$
$$- \int_0^\infty \frac{(\lambda_0+\mu+\delta m)^{L_2}}{\Gamma(L_2)}t^{L_2-1}e^{-(\lambda_0+\mu+\delta m)t}\ln\left(\sum_{i=0}^{L_2-1}\frac{(\lambda_0 t)^i}{i!}\right)dt. \tag{33a}$$

From the above analysis, we can see that there is a nontrivial entropy production rate corresponding to the memory if $L_1 > 1$ ($L_2 > 1$ ) in both cases. Furthermore, we find that the nontrivial effect results from the difference between the waiting-time distributions of state $\boldsymbol{n} = ({\rm off}, {\rm on}, m)$, i.e., the memory difference. So it is reasonable to speculate that memory dissipation is used for computing the memory difference (or for processing the memory information).

To quantitatively analyze how molecular memory affects the energy dissipation, we compute the mean memory difference according to

$$\langle D\rangle = \sum_m [P_0(m)D_{\rm off}(m) + P_1(m)D_{\rm on}(m)], \tag{34}$$

where $D_{\rm off}(m)$ and $D_{\rm on}(m)$ denote the memory difference at OFF state and ON state, respectively, and can be expressed as

$$D_{\rm off}(m) = D\big[R_1(m)\big\|R_5^{\rm off}(m)\big] + D\big[R_5^{\rm off}(m)\big\|R_1(m)\big] + D[R_3(m)\|R_1(m)] + D[R_1(m)\|R_3(m)], \quad D_{\rm on}(m) = 0,$$

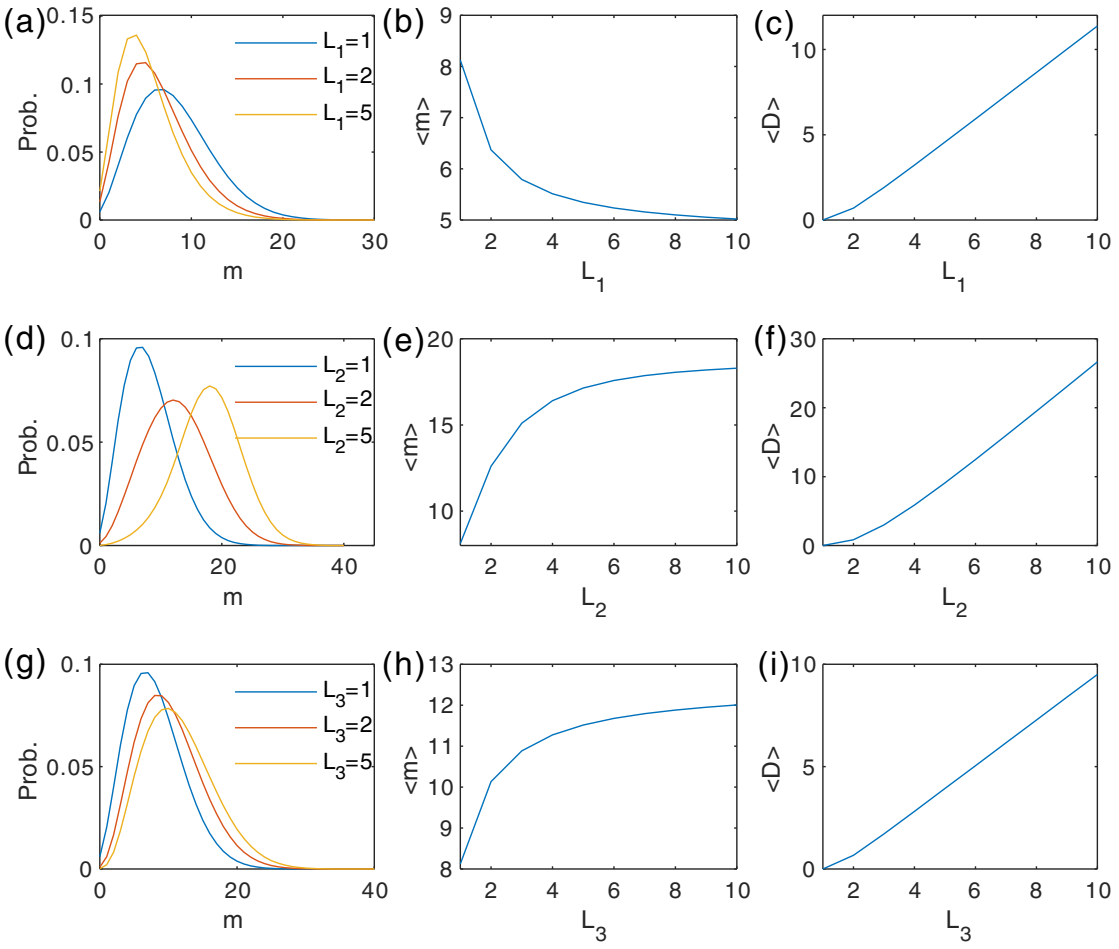


FIG. 4. Influence of molecular memory on mRNA expression and memory difference. (a) Stationary mRNA distributions in Case 1; (b) Influence of memory index $L_1$ on mean mRNA level in Case 1; (c) Influence of $L_1$ on mean memory difference in Case 1; (d) Stationary mRNA distributions in Case 2; (e) Influence of memory index $L_2$ on mean mRNA level in Case 2; (f) Influence of $L_2$ on mean memory difference in Case 2. (g) Stationary mRNA distributions in Case 3; (h) Influence of memory index $L_3$ on mean mRNA level in Case 3; (i) Influence of $L_3$ on mean memory difference in Case 3. Parameters are set as $\mu_0 = 1$, $\mu = 20$, $\lambda_1 = 3L_1$, $\lambda_0 = 5L_2$, $\delta = L_3$.

for Case 1, and

$$D_{\text{off}}(m) = 0$$
$$D_{\text{on}}(m) = D[R_4(m)\|R_2(m)] + D[R_2(m)\|R_4(m)] + D\big[R_2(m)\big\|R_5^{\text{on}}(m)\big] + D\big[R_5^{\text{on}}(m)\big\|R_2(m)\big],$$

for Case 2.

Recent experimental data analysis indicated that the waiting-time distribution describing nascent RNA removal is nonexponential [63], implying that the mRNA degradation is a multistep process with memory. So we also consider a nonexponential waiting-time distribution for degradation with the memory effect modeled by Gamma distribution. For convenience, this memory mechanism is called Case 3, and its analysis is referred to as Appendix C.

In addition, we are more interested in the mean dissipation rate, which can be used to show characteristics of energy dissipation. Therefore, we define the "mean equivalent dissipation" as the ratio of the equivalent dissipation over the mean mRNA, and the "mean memory dissipation" as the ratio of the memory dissipation over the mean memory difference, that is,

$$\dot{W}_E = \frac{\dot{S}_E}{\langle m\rangle}, \quad \dot{W}_M = \frac{\dot{S}_M}{\langle D\rangle}. \tag{35}$$

Equation (35) indicates that the mean equivalent dissipation is proportional to the energy dissipation required per mRNA molecule, whereas the mean memory dissipation is proportional to the energy dissipation required for each unit of memory difference.

### *3. Numerical results*

Here, we use numerical results to demonstrate how molecular memory (characterized by one of three different memory indexes) qualitatively affects mRNA expression level, memory difference, and energy dissipation. From the above analysis, we know that the equivalent dissipation is caused by the BDB of the equivalent Markovian system and is responsible for mRNA production, whereas the memory dissipation is generated due to the DTD of the waiting-time distributions and contributes to computing memory difference. For

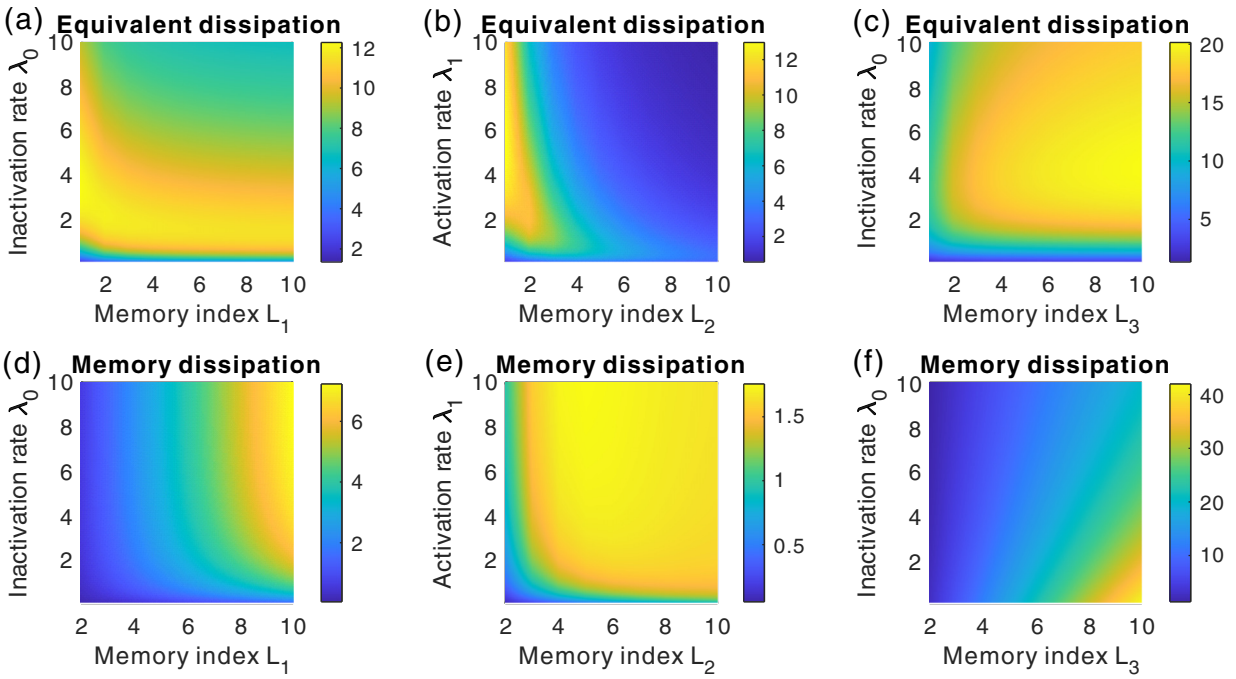


FIG. 5. Effects of molecular memory on energy dissipation. Heatmaps (a), (d), respectively, show the equivalent dissipation and the memory dissipation as a function of both memory index $L_1$ and inactivation rate $\lambda_0$. Heatmaps (b), (e), respectively, show the equivalent dissipation and the memory dissipation as a function of both memory index $L_2$ and activation rate $\lambda_1$. Heatmaps (c), (f), respectively, show the equivalent dissipation and the memory dissipation as a function of both memory index $L_3$ and inactivation rate $\lambda_0$. The setting of parameter values is the same as in Fig. 4.

clarity, we keep a constant average time $\lambda_1/L_1 = 3$ for Case 1, $\lambda_0/L_2 = 5$ for Case 2, and $\delta/L_3 = 1$ for Case 3.

*a. Molecular memory can tune gene expression but always increases memory difference.* Numerical results are shown in Fig. 4, where the top row, the middle row, and the bottom row correspond to Case 1, Case 2, and Case 3, respectively. From the first column, we observe that the increase of memory index $L_1$ can cause stationary mRNA distributions moving to the left-hand side whereas the increase of memory indexes $L_2$ and $L_3$ can cause stationary mRNA distributions moving to the right-hand side, implying that molecular memory can tune the stationary distributions. From the second column, we can see that the mean mRNA is monotonically decreasing in memory index $L_1$, whereas monotonically increasing in memory indexes $L_2$ and $L_3$, implying that molecular memory can tune the expression level. From the last column, we observe that the mean memory difference is always a monotonically increasing function of the memory index, implying that a stronger memory can lead to a larger mean memory difference.

In a word, molecular memory can tune mRNA expression level and increase memory difference, indicating that molecular memory has a non-negligible effect on gene expression. Based on the above analysis, we speculate that the memory difference may be related with the memory information. The increase in mean memory difference would mean more memory information being processed in a non-Markovian gene-expression system, and computing memory difference (used in processing the memory information) requires consumption of energy. In the following, we further analyze the influence of molecular memory on energy dissipation.

*b. Effects of molecular memory on energy dissipation.* Numerical results are shown in Fig. 5, where the first, second, and last columns correspond to Case 1, Case 2, and Case 3, respectively. From the top row of this figure, we observe that the dependence of equivalent dissipation on different memory indexes has different modes, referring to Figs. 5(a)–5(c). For example, with the increase of memory index $L_1(L_2)$, the equivalent dissipation first increases and then decreases for a smaller inactivation rate $\lambda_0$ (activation rate $\lambda_1$), whereas it always decreases for a larger $\lambda_0$ ($\lambda_1$). However, the equivalent dissipation always increases with the increase of memory index $L_3$ for arbitrarily fixed inactivation rate $\lambda_0$. In addition, there are some differences: for Case 1 and Case 3, larger equivalent dissipation exists in a wide range of memory index, but for Case 2, it exists in a narrow range of memory index. These results imply that the modes of energy dissipation through breaking the DB in a non-Markovian gene-expression system would be complex.

The bottom row of Fig. 5 shows the dependence of memory dissipation on both memory index and the switching rate of the promoter. For a given switching rate, we observe that the memory dissipation is almost a monotonically increasing function of the three memory indexes. Note that there are some differences: In contrast to the equivalent dissipation, the larger memory dissipation in Case 1 and Case 3 exists in a narrow range of memory index, whereas that in Case 2 exists in a wide range of memory index. In short, molecular memory can induce different modes of the equivalent dissipation but almost increases the memory dissipation.

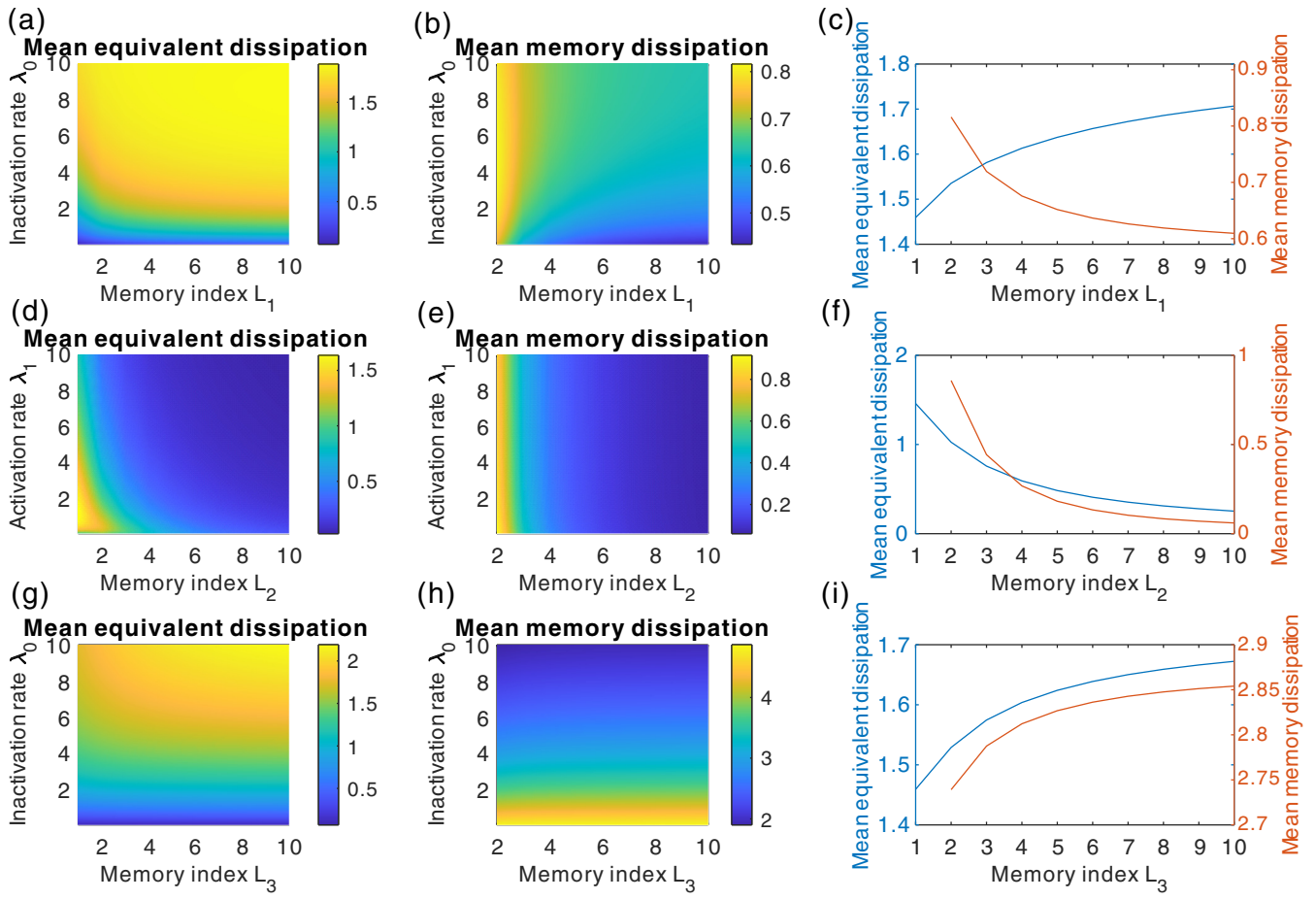


FIG. 6. Effects of molecular memory on mean dissipation. Heatmaps (a), (b), respectively, show the mean equivalent dissipation and the mean memory dissipation as a function of both memory index $L_1$ and inactivation rate $\lambda_0$. Heatmaps (d), (e), respectively, show the mean equivalent dissipation and the mean memory dissipation as a function of both memory index $L_2$ and activation rate $\lambda_1$. Heatmaps (g), (h), respectively, show the mean equivalent dissipation and the mean memory dissipation as a function of both memory index $L_3$ and inactivation rate $\lambda_0$. (c), (f), (i) Dependence of the mean equivalent dissipation and mean memory dissipation on $L_1$ for a fixed $\lambda_0 = 5$, $L_2$ for a fixed $\lambda_1 = 3$, and $L_3$ for a fixed $\lambda_0 = 5$, respectively. The setting of other parameter values is the same as in Fig. 4.

*c. Effects of molecular memory on mean dissipation.* Figure 6 illustrates how molecular memory affects mean dissipation; here the top, middle, and bottom rows correspond to Case 1, Case 2, and Case 3, respectively. First, we analyze the mean equivalent dissipation, referring to the first column of Fig. 6. We observe that if the switching rate of the promoter is fixed, the mean equivalent dissipation is a monotonically increasing function of memory index in Case 1 and Case 3, implying that molecular memory can lead to more energy dissipation needed for the production of one mRNA molecule. However, the mean equivalent dissipation monotonically decreases with the increase of memory index in Case 2, implying that molecular memory can reduce energy dissipation required for the production of one mRNA molecule. Similar to the equivalent dissipation, larger mean equivalent dissipation in Case 1 and Case 3 exists in a wider range of memory index with a larger inactivation rate, referring to Figs. 6(a) and 6(g), and larger mean equivalent dissipation in Case 2 exists in a narrower range of memory index with a lower activation rate, referring to Fig. 6(d).

Next, we analyze the mean memory dissipation, referring to the second column of Fig. 6. If the switching rate of the promoter is fixed, we observe that with the increase of memory index, the mean memory dissipation in Case 1 and Case 2 significantly reduces, whereas the one in Case 3 slightly increases (comparing the subfigures in the last column of Fig. 6), implying that different memory mechanisms can lead to different dissipation required for each unit of memory difference. In addition, we observe that larger mean memory dissipation in Case 1 and Case 2 emerges in a smaller range of memory index with a wider range of switching rate of the promoter, whereas in Case 3 it emerges in a larger range of memory index with a smaller range of switching rate of the promoter, implying that the state transitions of the promoter also influence the mean memory dissipation.

Finally, we simply state possible implications of the above findings. The above analysis shows that molecular memory can tune two kinds of energy dissipations, and different memory mechanisms can lead to different modes of energy dissipation. A previous study showed that molecular memory is in effect equivalent to a feedback [51], and different feedback regulations can induce different and even counterintuitive effects on gene expression [69]. Here, different memory mechanisms are equivalent to different regulation

mechanisms, so they may lead to different effects on dissipation. Recent experimental evidence of time-resolved data showed that the inactive phases of the promoter are a multistep process, showing strong memory [40], and the analysis of multistep activation memory mechanism implies that more strong memory can lead to more memory dissipation. On the other hand, the dissipation of biological systems usually follows certain design principles for optimal evolutionary fitness [70]. Thus, we speculate that there would be a tradeoff between energy dissipation and the achievement of biological functions.

## IV. SUMMARY AND CONCLUSION

Nonequilibrium mechanisms play important roles in many biological processes such as high-fidelity DNA transcription [7], internal transport [2], and spatial organization [6]. Energy dissipation is a characteristic of these nonequilibrium processes. For example, the dissipation of the chemical energy from the hydrolysis of ATP can drive unidirectional transitions between states of a molecular motor [71], and such unbalanced transitions break DB, resulting in directional motion of the individual motor. From the perspective of stochastic thermodynamics, the entropy production rate is a good way to quantify the energy dissipation [13]. Previous studies on energy dissipation in biochemical reaction systems are mainly based on Markovian hypothesis, i.e., the entropy production rate is derived in the framework of the Markovian process [11]. However, intracellular reaction processes are not necessarily Markovian but may be non-Markovian. For example, the complex control process of gene expression can generate nonexponential time intervals between transcription windows [40], and in elongation process [66]. How energy dissipation is quantified in the non-Markovian process is a key for understanding nonequilibrium mechanisms. Here, based on the CTRW theory we have derived and decomposed the total entropy production rate for a general biochemical-reaction network with arbitrary waiting-time distributions. We found that the energy dissipation can be split into two parts: one from BDB of the topology-equivalent network system, and the other from DTD of the waiting-time distributions. In particular, the total energy dissipation rate can reduce to that in the Markovian case if the waiting-time distributions at all states are exponential. Our decomposition of the total energy dissipation rate can help us to distinguish the non-Markovian system and its equivalent Markovian system, although the two systems have the same stationary dynamical behaviors.

We have applied the above method to two examples: a CTRW model and a generalized ON-OFF model of gene expression. Analysis of CTRW model indicated that DB and DTI together are equivalent to no energy dissipation in the non-Markovian process. In other words, even if the condition of DB is satisfied, there is still nontrivial dissipation in the non-Markovian process. Analysis of the gene-expression model indicated that different memory mechanisms can induce different modes of energy dissipation, implying that molecular memory has an important influence on gene expression. Based on these obtained results, we speculate that the energy dissipation due to the BDB of the equivalent Markovian system would be used mainly in achieving particular functions such as fidelity of transcription [57], whereas the dissipation from the DTD of the waiting-time distributions would be used mainly in processing memory information [72].

In summary, we have developed a method for decomposing the energy dissipation of a general reaction network. This method has a good perspective for understanding the nonequilibrium dynamics of non-Markovian biochemical reaction systems.

## ACKNOWLEDGMENTS

This work was supported by the National Nature Science Foundation of People's Republic of China (Grants No. 11601094, No. 11475273, No. 11631005, No. 11775314, No. 11931019, and No. 12001377), Guangdong Key Research and Development Project (Grant No. 2019B0233002), and Guangdong Province Key Laboratory of Computational Science at the Sun Yat-sen University (Grant No. 2020B1212060032).

## APPENDIX A: DERIVATION OF EQ. (17)

Here, we derive the relative entropy per state in Eq. (17). Substituting Eq. (15) into Eq. (16), we have

$$
\begin{aligned}
\delta_G &= \lim_{l\to\infty}\frac{1}{l}\sum_{\gamma}P(\gamma)\ln\frac{P(\gamma)}{P(\tilde{\gamma})}\\
&= \lim_{l\to\infty}\frac{1}{l}\sum_{\sigma}\int_0^\infty\cdots\int_0^\infty \phi_{\boldsymbol{n}_1\boldsymbol{n}_2}(t_1)\phi_{\boldsymbol{n}_2\boldsymbol{n}_3}(t_2)\cdots\phi_{\boldsymbol{n}_l\boldsymbol{n}_{l+1}}(t_l)\left[\ln\frac{\phi_{\boldsymbol{n}_1\boldsymbol{n}_2}(t_1)}{\phi_{\boldsymbol{n}_1\boldsymbol{n}_0}(t_1)}+\ln\frac{\phi_{\boldsymbol{n}_2\boldsymbol{n}_3}(t_2)}{\phi_{\boldsymbol{n}_2\boldsymbol{n}_1}(t_2)}+\cdots\ln\frac{\phi_{\boldsymbol{n}_l\boldsymbol{n}_{l+1}}(t_l)}{\phi_{\boldsymbol{n}_l\boldsymbol{n}_{l-1}}(t_l)}\right]dt_1dt_2\cdots dt_l\\
&= \lim_{l\to\infty}\frac{1}{l}\sum_{i=1}^{l}\left[\sum_{\sigma}\int_0^\infty\cdots\int_0^\infty \phi_{\boldsymbol{n}_1\boldsymbol{n}_2}(t_1)\phi_{\boldsymbol{n}_2\boldsymbol{n}_3}(t_2)\cdots\phi_{\boldsymbol{n}_{i-1}\boldsymbol{n}_i}(t_{i-1})\phi_{\boldsymbol{n}_{i+1}\boldsymbol{n}_{i+2}}(t_{i+1})\cdots\phi_{\boldsymbol{n}_l\boldsymbol{n}_{l+1}}(t_l)dt_1\cdots dt_l\right]\\
&\quad\times\left[\int_0^\infty\phi_{\boldsymbol{n}_i\boldsymbol{n}_{i+1}}(t_i)\ln\frac{\phi_{\boldsymbol{n}_i\boldsymbol{n}_{i+1}}(t_i)}{\phi_{\boldsymbol{n}_i\boldsymbol{n}_{i-1}}(t_i)}dt_i\right]\\
&= \lim_{l\to\infty}\frac{1}{l}\sum_{i=1}^{l}\left(\sum_{\sigma}\Phi_{\boldsymbol{n}_1\boldsymbol{n}_2}\Phi_{\boldsymbol{n}_2\boldsymbol{n}_3}\cdots\Phi_{\boldsymbol{n}_{i-1}\boldsymbol{n}_i}\Phi_{\boldsymbol{n}_{i+1}\boldsymbol{n}_{i+2}}\cdots\Phi_{\boldsymbol{n}_l\boldsymbol{n}_{l+1}}\right)\left\{\Phi_{\boldsymbol{n}_i\boldsymbol{n}_{i+1}}\ln\frac{\Phi_{\boldsymbol{n}_i\boldsymbol{n}_{i+1}}}{\Phi_{\boldsymbol{n}_i\boldsymbol{n}_{i-1}}}+\Phi_{\boldsymbol{n}_i\boldsymbol{n}_{i+1}}D\big[\phi_{\boldsymbol{n}_{i+1}|\boldsymbol{n}_i}(t)\big\|\phi_{\boldsymbol{n}_{i-1}|\boldsymbol{n}_i}(t)\big]\right\}
\end{aligned}
$$

$$= \lim_{l\to\infty}\frac{1}{l}\sum_{\sigma}\Phi_{\boldsymbol{n}_1\boldsymbol{n}_2}\Phi_{\boldsymbol{n}_2\boldsymbol{n}_3}\cdots\Phi_{\boldsymbol{n}_l\boldsymbol{n}_{l+1}}\left[\sum_{i=1}^{l}\ln\frac{\Phi_{\boldsymbol{n}_i\boldsymbol{n}_{i+1}}}{\Phi_{\boldsymbol{n}_i\boldsymbol{n}_{i-1}}}\right] + \lim_{l\to\infty}\frac{1}{l}\sum_{\sigma}\Phi_{\boldsymbol{n}_1\boldsymbol{n}_2}\Phi_{\boldsymbol{n}_2\boldsymbol{n}_3}\cdots\Phi_{\boldsymbol{n}_l\boldsymbol{n}_{l+1}}\left\{\sum_{i=1}^{l}D\big[\phi_{\boldsymbol{n}_{i+1}|\boldsymbol{n}_i}(t)\big\|\phi_{\boldsymbol{n}_{i-1}|\boldsymbol{n}_i}(t)\big]\right\}$$

$$= \lim_{l\to\infty}\frac{1}{l}D[P(\sigma)\|P(\tilde{\sigma})] + \sum_{\boldsymbol{n},\boldsymbol{n}',\boldsymbol{n}''}\Pi_{\boldsymbol{n}}\Phi_{\boldsymbol{n}\boldsymbol{n}'}\Phi_{\boldsymbol{n}'\boldsymbol{n}''}D[\phi_{\boldsymbol{n}''|\boldsymbol{n}'}(t)\|\phi_{\boldsymbol{n}|\boldsymbol{n}'}(t)]$$

$$= \sum_{\boldsymbol{n},\boldsymbol{n}'}\Pi_{\boldsymbol{n}}\Phi_{\boldsymbol{n}\boldsymbol{n}'}\ln\frac{\Phi_{\boldsymbol{n}\boldsymbol{n}'}}{\Phi_{\boldsymbol{n}'\boldsymbol{n}}} + \sum_{\boldsymbol{n},\boldsymbol{n}',\boldsymbol{n}''}\Pi_{\boldsymbol{n}}\Phi_{\boldsymbol{n}\boldsymbol{n}'}\Phi_{\boldsymbol{n}'\boldsymbol{n}''}D[\phi_{\boldsymbol{n}''|\boldsymbol{n}'}(t)\|\phi_{\boldsymbol{n}|\boldsymbol{n}'}(t)]. \tag{A1}$$

## APPENDIX B: NUMERICAL ALGORITHM FOR SOLVING EQ. (26)

We give a numerical algorithm to solve Eq. (26) in the main text, which is based on the truncation of sgCME. Set $0 \leqslant n \leqslant N$, and introduce a column vector

$$\boldsymbol{P}(m) = [[P_0(0)\quad P_1(0)]^T\ \ [P_0(1)\quad P_1(1)]^T\quad\cdots\quad[P_0(N)\quad P_1(N)]^T]^T, \tag{B1}$$

and a matrix

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{T}_0-\boldsymbol{G}_0-\boldsymbol{D}_0 & \boldsymbol{D}_1 & & & & & \\ \boldsymbol{G}_0 & \boldsymbol{T}_1-\boldsymbol{G}_1-\boldsymbol{D}_1 & \boldsymbol{D}_2 & & & & \\ & \boldsymbol{G}_1 & \boldsymbol{T}_2-\boldsymbol{G}_2-\boldsymbol{D}_2 & \boldsymbol{D}_3 & & & \\ & & \ddots & \ddots & & & \\ & & & \boldsymbol{G}_{N-2} & \boldsymbol{T}_{N-1}-\boldsymbol{G}_{N-1}-\boldsymbol{D}_{N-1} & \boldsymbol{D}_N \\ & & & & \boldsymbol{G}_{N-1} & \boldsymbol{T}_N-\boldsymbol{G}_N-\boldsymbol{D}_N \end{pmatrix}, \tag{B2}$$

where

$$\boldsymbol{T}_i = \begin{pmatrix} -K_1(i) & K_2(i) \\ K_1(i) & -K_2(i) \end{pmatrix},\qquad \boldsymbol{G}_i = \begin{pmatrix} K_3(i) & 0 \\ 0 & K_4(i) \end{pmatrix},\qquad \boldsymbol{D}_i = \begin{pmatrix} K_5(i) & 0 \\ 0 & K_5(i) \end{pmatrix}. \tag{B3}$$

By solving the algebraic equation $\boldsymbol{AP}=\boldsymbol{0}$ with the conservation condition: $\sum_{i=0}^{n}P(i)=1$, we can obtain a numerical $\boldsymbol{P}$.

## APPENDIX C: ANALYSIS FOR THE NON-MARKOVIAN DEGRADATION

In Case 3, we assume that the waiting-time distribution of the degradation reaction is Gamma distribution whereas the waiting-time distributions of other reactions are exponential distributions. Then, we set $\psi_1(t;m)=\lambda_1 e^{-\lambda_1 t}$, $\psi_2(t;m)=\lambda_0 e^{-\lambda_0 t}$, where $\lambda_1$ and $\lambda_0$ are the rates of gene activation and inactivation, respectively; $\psi_3(t;m)=\mu_0 e^{-\mu_0 t}$, $\psi_4(t;m)=\mu e^{-\mu t}$, where $\mu$ is the transcription rate whereas $\mu_0$ is the leakage rate. $\psi_5(t;m)=[\Gamma(L_3)]^{-1}(\delta m)^{L_3}t^{L_3-1}e^{-\delta m t}$, where $L_3$ is positive integer and is called memory index, $\delta$ is the mean degradation rate. Simple calculation can show $\Psi_1(t;m)=1-e^{-\lambda_1 t}$, $\Psi_2(t;m)=1-e^{-\lambda_0 t}$, $\Psi_3(t;m)=1-e^{-\mu_0 t}$, $\Psi_4(t;m)=1-e^{-\mu t}$, and $\Psi_5(t;m)=1-\sum_{i=0}^{L_3-1}\frac{(\delta m t)^i}{i!}e^{-\delta m t}$. In addition, we have $\phi_1(t;m)=\lambda_1 e^{-(\lambda_1+\mu_0+\delta m)t}[\sum_{i=0}^{L_3-1}\frac{(\delta m t)^i}{i!}]$, $\phi_2(t;m)=\lambda_0 e^{-(\lambda_0+\mu+\delta m)t}[\sum_{i=0}^{L_3-1}\frac{(\delta m t)^i}{i!}]$, $\phi_3(t;m)=\mu_0 e^{-(\lambda_1+\mu_0+\delta m)t}[\sum_{i=0}^{L_3-1}\frac{(\delta m t)^i}{i!}]$, $\phi_4(t;m)=\mu e^{-(\lambda_0+\mu+\delta m)t}[\sum_{i=0}^{L_3-1}\frac{(\delta m t)^i}{i!}]$, $\phi_5(t;m)=[\Gamma(L_3)]^{-1}(\delta m)^{L_3}t^{L_3-1}e^{-[\lambda_1(1-I_{\rm on})+\lambda_0 I_{\rm on}+\mu_0(1-I_{\rm on})+\mu I_{\rm on}+\delta m]t}$. Thus, $\Phi_1(m)=\lambda_1\sum_{i=0}^{L_3-1}\frac{(\delta m)^i}{(\lambda_1+\mu_0+\delta m)^{i+1}}$, $\Phi_2(m)=\lambda_0\sum_{i=0}^{L_3-1}\frac{(\delta \mathrm{m})^i}{(\lambda_0+\mu+\delta \mathrm{m})^{i+1}}$ $\Phi_3(m)=\mu_0\sum_{i=0}^{L_3-1}\frac{(\delta m)^i}{(\lambda_1+\mu_0+\delta m)^{i+1}}$, $\Phi_4(m)=\mu\sum_{i=0}^{L_3-1}\frac{(\delta \mathrm{m})^i}{(\lambda_0+\mu+\delta \mathrm{m})^{i+1}}$ and $\Phi_5(m)=(\frac{\delta m}{\lambda_1(1-I_{\rm on})+\lambda_0 I_{\rm on}+\mu_0(1-I_{\rm on})+\mu I_{\rm on}+\delta m})^{L_3}$. The mean residence time at state $\boldsymbol{n}=(\mathrm{off},\mathrm{on},m)$ is given by

$$\tau_{\boldsymbol{n}} = \sum_{i=0}^{L_3-1}\frac{(\delta m)^i}{[\lambda_1(1-I_{\rm on})+\lambda_0 I_{\rm on}+\mu_0(1-I_{\rm on})+\mu I_{\rm on}+\delta m]^{i+1}}. \tag{C1}$$

The effective transition rates are given by

$$K_1(m)=\lambda_1,\quad K_2(m)=\lambda_0,\quad K_3(m)=\mu_0,\quad K_4(m)=\mu,$$
$$K_5(m)=[\lambda_1(1-I_{\rm on})+\lambda_0 I_{\rm on}+\mu_0(1-I_{\rm on})+\mu I_{\rm on}]\frac{(\delta m)^{L_3}/[\lambda_1(1-I_{\rm on})+\lambda_0 I_{\rm on}+\mu_0(1-I_{\rm on})+\mu I_{\rm on}+\delta m]^{L_3}}{1-(\delta m)^{L_3}/[\lambda_1(1-I_{\rm on})+\lambda_0 I_{\rm on}+\mu_0(1-I_{\rm on})+\mu I_{\rm on}+\delta m]^{L_3}} \tag{C2}$$

Substituting Eq. (C2) into Eqs. (26)–(28), respectively, we can obtain the stationary probability, the mean mRNA, and the equivalent dissipation.

In addition, the memory dissipation can be decomposed as

$$\dot{S}_M = \dot{S}_M^{(\mathrm{off})} + \dot{S}_M^{(\mathrm{on})}, \tag{C3}$$

where

$$\dot{S}_M^{(\mathrm{off})} = \sum_m P_0(m)K_1(m)\frac{K_5(m)D\big[R_5^{\mathrm{on}}(m)\big\|R_2(m)\big]}{K_2(m)+K_4(m)+K_5(m)} + \sum_m P_0(m)K_5(m)\frac{K_5(m-1)\big[R_5^{\mathrm{off}}(m-1)\big\|R_3(m-1)\big]}{K_1(m-1)+K_3(m-1)+K_5(m-1)}$$
$$+\sum_m P_0(m)K_3(m)\frac{K_1(m+1)D\big[R_1(m+1)\big\|R_5^{\mathrm{off}}(m+1)\big]+K_3(m+1)D\big[R_3(m+1)\big\|R_5^{\mathrm{off}}(m+1)\big]}{K_1(m+1)+K_3(m+1)+K_5(m+1)}, \tag{C3a}$$

$$\dot{S}_M^{(\mathrm{on})} = \sum_m P_1(m)K_2(m)\frac{K_5(m)D\big[R_5^{\mathrm{off}}(m)\big\|R_1(m)\big]}{K_1(m)+K_3(m)+K_5(m)} + \sum_m P_1(m)K_5(m)\frac{K_5(m-1)D\big[R_5^{\mathrm{on}}(m-1)\big\|R_4(m-1)\big]}{K_2(m-1)+K_4(m-1)+K_5(m-1)}$$
$$+\sum_m P_1(m)K_4(m)\frac{K_4(m+1)D\big[R_4(m+1)\big\|R_5^{\mathrm{on}}(m+1)\big]+K_2(m+1)D\big[R_2(m+1)\big\|R_5^{\mathrm{on}}(m+1)\big]}{K_2(m+1)+K_4(m+1)+K_5(m+1)}, \tag{C3b}$$

with the memory differences given by

$$D\big[R_5^{\mathrm{on}}(m)\big\|R_2(m)\big] = D\big[R_5^{\mathrm{on}}(m)\big\|R_4(m)\big] = -\ln\Gamma(L_3)+(L_3-1)\psi(L_3)+\ln(\lambda_0+\mu+\delta m)+\ln\left(\sum_{i=0}^{L_3-1}\frac{(\delta m)^i}{(\lambda_0+\mu+\delta m)^{i+1}}\right)$$
$$-\int_0^\infty \frac{(\lambda_0+\mu+\delta m)^{L_3}}{\Gamma(L_3)}t^{L_3-1}e^{-(\lambda_0+\mu+\delta m)t}\ln\left(\sum_{i=0}^{L_3-1}\frac{(\delta mt)^i}{i!}\right)dt. \tag{C3c}$$

$$D\big[R_5^{\mathrm{off}}(m)\big\|R_3(m)\big] = D\big[R_5^{\mathrm{off}}(m)\big\|R_1(m)\big] = -\ln\Gamma(L_3)+(L_3-1)\psi(L_3)+\ln(\lambda_1+\mu_0+\delta m)+\ln\left(\sum_{i=0}^{L_3-1}\frac{(\delta m)^i}{(\lambda_1+\mu_0+\delta m)^{i+1}}\right)$$
$$-\int_0^\infty \frac{(\lambda_1+\mu_0+\delta m)^{L_3}}{\Gamma(L_3)}t^{L_3-1}e^{-(\lambda_1+\mu_0+\delta m)t}\ln\left(\sum_{i=0}^{L_3-1}\frac{(\delta mt)^i}{i!}\right)dt. \tag{C3d}$$

$$D\big[R_1(m)\big\|R_5^{\mathrm{off}}(m)\big] = D\big[R_3(m)\big\|R_5^{\mathrm{off}}(m)\big]$$
$$= \ln\Gamma(L_3)-(L_3-1)\frac{\sum_{i=0}^{L_3-1}\frac{(\delta m)^i\psi(i+1)}{(\lambda_1+\mu_0+\delta m)^{i+1}}}{\sum_{i=0}^{L_3-1}\frac{(\delta m)^i}{(\lambda_1+\mu_0+\delta m)^{i+1}}}-\ln(\lambda_1+\mu_0+\delta m)-\ln\left(\sum_{i=0}^{L_3-1}\frac{(\delta m)^i}{(\lambda_1+\mu_0+\delta m)^{i+1}}\right)$$
$$+\left(\sum_{i=0}^{L_3-1}\frac{(\delta m)^i}{(\lambda_1+\mu_0+\delta m)^{i+1}}\right)^{-1}\int_0^\infty e^{-(\lambda_1+\mu_0+\delta m)t}\left(\sum_{i=0}^{L_3-1}\frac{(\delta mt)^i}{i!}\right)\ln\left(\sum_{i=0}^{L_3-1}\frac{(\delta mt)^i}{i!}\right)dt, \tag{C3e}$$

$$D\big[R_4(m)\big\|R_5^{\mathrm{on}}(m)\big] = D\big[R_2(m)\big\|R_5^{\mathrm{on}}(m)\big]$$
$$= \ln\Gamma(L_3)-(L_3-1)\frac{\sum_{i=0}^{L_3-1}\frac{(\delta m)^i\psi(i+1)}{(\lambda_0+\mu+\delta m)^{i+1}}}{\sum_{i=0}^{L_3-1}\frac{(\delta m)^i}{(\lambda_0+\mu+\delta m)^{i+1}}}-\ln(\lambda_0+\mu+\delta m)-\ln\left(\sum_{i=0}^{L_3-1}\frac{(\delta m)^i}{(\lambda_0+\mu+\delta m)^{i+1}}\right)$$
$$+\left(\sum_{i=0}^{L_3-1}\frac{(\delta m)^i}{(\lambda_0+\mu+\delta m)^{i+1}}\right)^{-1}\int_0^\infty e^{-(\lambda_0+\mu+\delta m)t}\left(\sum_{i=0}^{L_3-1}\frac{(\delta mt)^i}{i!}\right)\ln\left(\sum_{i=0}^{L_3-1}\frac{(\delta mt)^i}{i!}\right)dt. \tag{C3f}$$

Finally, we compute the mean memory difference according to Eq. (34), i.e., the following expression:

$$\langle D\rangle = \sum_m [P_0(m)D_{\mathrm{off}}(m)+P_1(m)D_{\mathrm{on}}(m)],$$

with

$$D_{\mathrm{off}}(m) = D\big[R_1(m)\big\|R_5^{\mathrm{off}}(m)\big]+D\big[R_5^{\mathrm{off}}(m)\big\|R_1(m)\big]+D\big[R_3(m)\big\|R_5^{\mathrm{off}}(m)\big]+D\big[R_5^{\mathrm{off}}(m)\big\|R_3(m)\big],$$
$$D_{\mathrm{on}}(m) = D\big[R_4(m)\big\|R_5^{\mathrm{on}}(m)\big]+D\big[R_5^{\mathrm{on}}(m)\big\|R_4(m)\big]+D\big[R_2(m)\big\|R_5^{\mathrm{on}}(m)\big]+D\big[R_5^{\mathrm{on}}(m)\big\|R_2(m)\big].$$

[1] S. Datta, *Electronic Transport in Mesoscopic Systems* (Cambridge University Press, Cambridge, 1997).
[2] M. E. Cates, Rep. Prog. Phys. **75**, 042601(2012).
[3] C. P. Brangwynne, G. H. Koenderink, F. C. MacKintosh, and D. A. Weitz, J. Cell Biol. **183**, 583 (2008).
[4] R. D. Astumian and M. Bier, Phys. Rev. Lett. **72**, 1766 (1994).
[5] K. C. Huang, Y. Meir, and N. S. Wingreen, Proc. Natl. Acad. Sci., USA **100**, 12724 (2003).
[6] J. Halatek and E. Frey, Cell Rep. **1**, 741 (2012).
[7] J. J. Hopfield, Proc. Natl. Acad. Sci, USA **71**, 4135 (1974).
[8] A. Murugan, D. A. Huse, and S. Leibler, Proc. Natl. Acad. Sci, USA **109**, 12034 (2012).
[9] G. Lan, P. Sartori, S. Neumann, V. Sourjik, and Y. Tu, Nat. Phys. **8**, 422 (2012).
[10] P. Mehta and D. J. Schwab, Proc. Natl. Acad. Sci. USA **109**, 17978 (2012).
[11] J. L. Lebowitz and H. Spohn, J. Stat. Phys. **95**, 333 (1999).
[12] T. Harada and S-I. Sasa, Phys. Rev. E **73**, 026131 (2006).
[13] H. Ge and H. Qian, Phys. Rev. E **81**, 051133 (2010).
[14] Y. Yang and H. Qian, Phys. Rev. E **101**, 022129 (2020).
[15] C. Gardiner, *Stochastic Methods: A Handbook for the Natural and Social Sciences* (Springer, New York, 2009).
[16] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, The Netherlands, 2007).
[17] A. C. Barato and U. Seifert, Phys. Rev. Lett. **114**, 158101 (2015).
[18] T. R. Gingrich, J. M. Horowitz, N. Perunov, and J. L. England, Phys. Rev. Lett. **116**, 120601 (2016).
[19] J. M. Horowitz, K. Zhou, and J. L. England, Phys. Rev. E **95**, 042102 (2017).
[20] P. Szymańska-Rożek, D. Villamaina, J. Miekisz, and A. M. Walczak, Entropy **21**, 1212 (2019).
[21] H. Qian, J. Mol. Biol. **362**, 387 (2006).
[22] N. Barkai and S. Leibler, Nature (London) **387**, 913 (1997).
[23] U. Alon, M. G. Surette, N. Barkai, and S. Leibler, Nature (London) **397**, 168 (1999).
[24] C. C. Govern and P. R. ten Wolde, Phys. Rev. Lett. **109**, 218103 (2012).
[25] P. Sartori and Y.Tu, Phys. Rev. Lett. **115**, 118102 (2015).
[26] G. Lan and Y. Tu, Rep. Prog. Phys. **79**, 052601(2016).
[27] H. Qian, Biophys. Chem. **105**, 585 (2003).
[28] A. H. Lang, C. K. Fisher, T. Mora, and P. Mehta, Phys. Rev. Lett. **113**, 148103 (2014).
[29] D. Hartich, A. C. Barato, and U. Seifert, New J. Phys. **17**, 055026 (2015).
[30] C. C. Govern and P. R. ten Wolde, Proc. Natl. Acad. Sci. USA **111**, 17486 (2014).
[31] C. C. Govern and P. R. ten Wolde, Phys. Rev. Lett. **113**, 258102 (2014).
[32] T. E. Ouldridge, C. C. Govern, and P. R. ten Wolde, Phys. Rev. X **7**, 021004 (2017).
[33] D. Pincus, BMC Biol. **13**, 9 (2015).
[34] L. Huang, Z. Yuan, P. Liu, J. Yu, and T. Zhou, Chaos **25**, 123101 (2015).
[35] H. Wang, P. Liu, Q. Li, and T. Zhou, FEBS Lett. **592**, 1135 (2018).
[36] M. R. Green, Mol. Cell **18**, 399 (2005).
[37] J. M. Pedraza and J. Paulsson, Science **319**, 339 (2008).
[38] T. Jia and R. V. Kulkarni, Phys. Rev. Lett. **106**, 058102 (2011).
[39] B. Qiu, J. Zhang, and T. Zhou, Phys. Rev. E **101**, 022409 (2020).
[40] 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 *et al.*, PLoS Biol. **9**, e1000607 (2011).
[41] D. M. Suter, N. Molina, D. Gatfield, K. Schneider, U. Schibler, and F. Naef, Science **332**, 472 (2011).
[42] P. S. Stumpf, R. C. G. Smith, M. Lenz, A. Schuppert, F. J. Müller, A. Babtie, T. E. Chan, M. P. H. Stumpf, C. P. Please, S. D. Howison *et al.*, Cell Syst. **5**, 268 (2017).
[43] É. Roldán and J. M. R. Parrondo, Phys. Rev. Lett. **105**, 150607 (2010).
[44] É. Roldán and J. M. R. Parrondo, Phys. Rev. E **85**, 031129 (2012).
[45] H. Qian and H. Wang, Europhys. Lett. **76**, 15 (2006).
[46] H. Wang and H. Qian, J. Math. Phys. **48**, 013303 (2007).
[47] C. Maes, K. Netočný, and B. Wynants, J. Phys. A: Math. Theor. **42**, 365002 (2009).
[48] I. A. Martínez, G. Bisker, J. M. Horowitz, and J. M. R. Parrondo, Nat. Commun. **10**, 3542 (2019).
[49] G. Teza and A. L. Stella, Phys. Rev. Lett. **125**, 110601 (2020).
[50] G. Bisker, M. Polettini, T. R. Gingrich, and J. M. Horowitz, J. Stat. Mech. (2017) 093210.
[51] J. Zhang and T. Zhou, Proc. Natl. Acad. Sci. USA **116**, 23542 (2019).
[52] T. Aquino and M. Dentz, Phys. Rev. Lett. **119**, 230601 (2017).
[53] U. Landman, E. W. Montroll, and M. F. Shlesinger, Proc. Natl. Acad. Sci. USA **74**, 430 (1977).
[54] H. Scher and C. H. Wu, Proc. Natl. Acad. Sci. USA **78**, 22 (1981).
[55] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (John Wiley & Sons, Hoboken, NJ, 2012).
[56] C. Maes and K. Netočný, J. Stat. Phys. **110**, 269 (2003).
[57] F. S. Gnesotto, F. Mura, J. Gladrow, and C. P. Broedersz, Rep. Prog. Phys. **81**, 066601 (2018).
[58] J. Zhang and T. Zhou, Biophys. J. **106**, 479 (2014).
[59] T. Liu, J. Zhang, and T. Zhou, PLoS Comput. Biol. **12**, e1004917 (2016).
[60] J. M. G. Vilar and L. Saiz, Bioinformatics **26**, 2060 (2010).
[61] G. Hornung, R. Bar-Ziv, D. Rosin, N. Tokuriki, D. S. Tawfik, M. Oren, and N. Barkai, Genome Res. **22**, 2409 (2012).
[62] L. Weinberger, Y. Voichek, I. Tirosh, G. Hornung, I. Amit, and N. Barkai, Mol. Cell **47**, 193(2012).
[63] H. Xu, L. A. Sepúlveda, L. Figard, A. M. Sokac, and I. Golding, Nat. Methods **12**, 739 (2015).
[64] Z. Cao and R. Grima, Proc. Natl. Acad. Sci. USA **117**, 4682 (2020).
[65] J. Liu, D. Hansen, E. Eck, Y. J. Kim, M. Turner, S. Alamos, and H. G. Garcia, bioRxiv, doi:10.1101/2020.08.29.273474.
[66] H. Xu, S. O. Skinner, A. M. Sokac, and I. Golding, Phys. Rev. Lett. **117**, 128101 (2016).
[67] T. Filatova, N. Popovic, and R. Grima, Math. Biol. **83**, 3 (2021).
[68] Q. Jiang, X. Fu, S. Yan, R. Li, W. Du, Z. Cao, F. Qian, and R. Grima, bioRxiv, doi:10.1101/2020.12.15.422883.
[69] L. Huang, Z. Yuan, and T. Zhou, Phys. Rev. E **90**, 052702 (2014).
[70] B. Schwanhäusser, D. Busse, N. Li, G. Dittmar, J. Schuchhardt, J. Wolf, W. Chen, and M. Selbach, Nature (London) **473**, 337 (2011).
[71] A. Ajdari, J. Prost, F. Ju, and F. Jülicher, Rev. Mod. Phys. **69**, 1269 (1997).
[72] J. E. Purvis and G. Lahav, Cell **152**, 945 (2013).