Computation of stationary distributions in stochastic models of cellular processes with molecular memory

Jiajun Zhang^{1, 2} and Tianshou Zhou^{1, 2, *}

1 Key Laboratory of Computational Mathematics, Guangdong Province

2 School of Mathematics, Sun Yat-Sen University, Guangzhou, 510275, P. R. China

* mcszhtsh@mail.sysu.edu.cn

Abstract

Modeling stochastic dynamics of intracellular processes has long rested on Markovian (i.e., memoryless) hypothesis. However, many of these processes are non-Markovian (i.e., memorial) due to, e.g., small reaction steps involved in synthesis or degradation of a macroscopic molecule. When interrogating aspects of a cellular network by experimental measurements (e.g., by single-molecule and single-cell measurement technologies) of network components, a key need is to develop efficient approaches to simulate and compute joint distributions of these components. To cope with this computational challenge, we develop two efficient algorithms: stationary generalized Gillespie algorithm and stationary generalized finite state projection, both being established based on a stationary generalized chemical master equation. We show how these algorithms can be combined in a streamlined procedure for evaluation of non-Markovian effects in a general cellular network. Stationary distributions are evaluated in two models of constitutive and bursty gene expressions as well as a model of genetic toggle switch, each considering molecular memory. Our approach significantly expands the capability of stochastic simulation to investigate gene regulatory network dynamics, which has the potential to advance both understanding of molecular systems biology and design of synthetic circuits.

Author summary

Cellular systems are driven by interactions between subsystems via time-stamped discrete events, involving numerous components and reaction steps and spanning several time scales. Such biochemical reactions are subject to inherent noise due to the small numbers of molecules. Also, they could involve several small steps, creating a memory between individual events. Delineating

these molecular stochasticity and memory of biomolecular networks are continuing challenges for molecular systems biology. We present a novel approach to compute the probability distribution in stochastic models of cellular processes with molecular memory based on stationary generalized chemical master equation. We map a stochastic system with memory onto a Markovian model with effective reaction propensity functions. This formulation enables us to efficiently develop algorithms under the Markovian framework, and thus systematically analyze how molecular memories regulate stochastic behaviors of biomolecular networks. Here we propose two representative algorithms: stationary generalized Gillespie algorithm and stationary generalized finite state projection algorithm. The former generate realizations with Monte Carlo simulation, but the later compute approximations of the probability distribution by solving a truncated version of stochastic process. Our approach is demonstrated by applying it to three different examples from systems biology: generalized birth-death process, a stochastic toggle switch model, and a 3-stage gene expression model.

1. Introduction

Cellular processes are biochemical and hence noisy: in each cell, the numbers or concentrations of species molecules are subject to random fluctuations due to the small numbers of these molecules or environmental perturbations or both. This stochasticity, which has been verified by single-molecule experiments and can have a significant impact on cellular functioning, has been intensively studied in the past few decades [1,2]. For a biochemical reaction system, if the stochastic motion of the reactants is assumed to be uninfluenced by previous states and only by the current state [3,4], then the stochastic effect can be well revealed by analyzing the chemical master equation (CME) [3-5]. This Markovian assumption has also led to important successes in the description of many stochastic processes on networks [3-8].

However, as a general rule, the dynamics of a given reactive species results from its interactions with the environment and cannot be described as a Markov process. Indeed, although the evolution of the set of all microscopic degrees of freedom of a biochemical system is Markovian, the dynamics restricted to the species only is not [9,10] (Fig.1 (A)). This is typically the case for a promoter cycle, whose non-Markovian dynamics result from the fact that the promoter gradually becomes mature

through a series of inactive states (with several hidden Poisson steps) before being activated, leading to narrowly distributed gestation periods between transcription windows [11]. More generally, the expressions of most genes are controlled in a complex manner involving several repressors, transcription factors and mediators, as well as chromatin remodeling or changes in supercoiling. These unspecified regulations or other processes can be modelled by non-exponential waiting time distributions [11-15]. Other experimentally observed examples of non-Markovian dynamics include the human activity patterns [16], natural phenomena [17], biological processes [18], biochemical reactions [19, 20], etc. Examples listed in Table 1 indicate that waiting-time distributions between biochemical events are not limited to exponential distributions but may be diverse. Indeed, the time scales of ultrafast biochemical processes are of the same order as the time scale of the bath they interact with, leading to failure of the Markov approximation.

Table 1. Literature survey on waiting-time distributions.

Description	Timescale	Distribution	References
Transcription activation time	minute	exponential	19,20
Transcription inactivation time	hour	phase-type	19,20
Enzymatic reaction time	millisecond	exponential/phase-type	21
Lambda induction in bacteria (lysis)	minute~hour	Gamma/phase-type	22
Transcription initiation time	minute	Gamma/phase-type	23
Cytokine secretion onset time	hour	Gamma/phase-type	24
ESC state residence time	hour	Gamma/phase-type	18
Cell-type specific waiting time	day	Fréchet type	25

An important step towards a general understanding of non-Markovian dynamics was recently made based on the continuous time random walk (CTRW) theory. Namely, a generalized chemical master equation (gCME) was derived, which can account for non-exponential inter-reaction times and the resulting non-Markovian character of reaction dynamics in time [26,27]. However, this equation is rather formal, and the “memory functions” involved are implicitly expressed by waiting time distributions. This leads to notorious difficulties in analysis and simulation, greatly limiting applications of the gCME. An alternative method is the kinetic Monte Carlo approach, which instead seeks either exact or approximate realizations of the underlying non-Markov process. Earlier, along the line of CTRW, Gillespie developed a stochastic simulation algorithm of random walk with residence time depending on transition probability rates [28], which is actually a prototype of the famous Gillespie algorithm (GA) [29]. Based on the renewal theory, Boguñá and colleagues

extended the GA to general renewal processes [30]. Subsequently, Vestergaard and Géniois developed a temporal GA, which can simulate non-Markovian dynamics [31]. Recently, Masuda and Rocha proposed a Laplace GA, which can be applied to renewal processes where survival functions are monotonic [32]. In a word, these numerical methods generate realizations enough to obtain statistics for pivotal events. Unfortunately, in many cases, biologically important events occur rarely when reaction waiting times span multiple time scales or/and follow non-exponential distributions, implying that the previous methods are unfeasible in this case. It is an urgent need but is still a significant challenge to develop effective methods to simulate and compute the stochastic dynamics of non-Markovian reaction processes.

As a fact of matter, the emergence of measurement techniques such as single-cell fluorescent microscopy [33], flow and mass cytometry [34], single-cell qPCR [35] and single-cell RNA-seq [36] currently allow for access to increasingly rich data on approximately steady-state distributions of interested molecular species and waiting time distributions of key reaction events. Indeed, steady-state distribution is an important quantity needed to characterize the stationary behavior of many biochemical network systems. To assess how experimental data can be informative, it is also needed to calculate or simulate aspects of steady-state distribution [37-41].

Fig 1. A framework for computing steady-state probability distributions of stochastic reaction processes with molecular memory. (A) Map a multi-step biochemical process onto a single-step reaction process with non-exponential waiting time. (B) Convert a non-Markovian problem into a Markovian problem.

Here we propose two efficient stochastic simulation algorithms: stationary generalized GA (sgGA) that generates realizations with Monte Carlo simulation, and stationary generalized finite state projection (sgFSP) that computes approximate probability distributions. These algorithms are established based on a stationary gCME (sgCME) for a general biomolecular network with arbitrary waiting time distributions. While a key point of the sgCME is that it converts a computationally challenging non-Markovian problem into a computable Markovian one (Fig.1 (B)) by introducing an effective transition rate for every reaction in the network [27], our algorithms are designed to sufficiently exploit this characteristic and hence can be used not only in simulation and analysis of non-Markovian dynamics but also in discovery of biological knowledge underlying non-Markovian effects in complex cellular network systems.

The paper is organized as follows. First, we introduce a general cellular network model with any waiting-time distributions between biochemical reactions. Then, we model this network with a gCME expressed in Laplace transforms and derive a sgCME where effective transition rates decoding non-Markovian effects are explicitly given. Next, we propose two efficient algorithms: sgGA and sgFSP. As an illustration, we have applied these methods to the analysis of a well-characterized birth-death process and investigate the memory effect on stochastic gene expression. Furthermore, we also demonstrated the applicability of these methods to two more complex biochemical network: a stochastic toggle switch model, and a 3-stage stochastic gene expression model. Finally, we revisit derivation of the most common waiting time distributions in reaction systems and give their physical explanations, with the aim to help the reader's understanding.

2. A general theory for reaction networks

Consider a reaction network that models a general cellular process. Assume that this network contains N different species (denoted by X_j , $j=1,2,\dots,N$), which participate in M different reactions of the form

where parameters r_{ij} and s_{ij} are stoichiometric coefficients, vector $\mathbf{n} = (n_1, \dots, n_N)^T$ represents the microscopic state of the underlying system with n_j being the molecule number of reactive

species X_j and T being transpose, and $\psi_i(t; \mathbf{n})$ be the probability density function (PDF) of the inter-reaction waiting time of the i th reaction (see interpretation below). For convenience, let differences $\nu_{ij} = s_{ij} - r_{ij}$ be stored in the $M \times N$ stoichiometric matrix S . Note that if the waiting time of all reactions follow geometric distributions, the generalized biochemical reaction network is reduced to the common biochemical reaction network.

Let $P(t; \mathbf{n})$ represent the probability that the system is at state $\mathbf{n}$ at time t , and $\phi_i(t; \mathbf{n})$ be the joint PDF of the i th reaction happening and the reaction waiting time being t , which can be expressed as $\phi_i(t; \mathbf{n}) = \psi_i(t; \mathbf{n}) \prod_{j \neq i} [1 - \Psi_j(t; \mathbf{n})]$. Introduce memory function $M_i(t; \mathbf{n})$ for the i th reaction, where the memory is created if inter-reaction waiting time distribution $\psi_i(t; \mathbf{n})$ is non-exponential. Following the standard derivation of the Montroll-Weiss equation [42-44] in combination with the idea of chemical continuous time random walks (CTRW) [26,27], we can derive the following gCME in the sense of Laplace transform

$$s\tilde{P}(s; \mathbf{n}) = \tilde{P}(0; \mathbf{n}) + \sum_{i=1}^M [\tilde{M}_i(s; \mathbf{n} - \mathbf{v}_i) \tilde{P}(s; \mathbf{n} - \mathbf{v}_i) - \tilde{M}_i(s; \mathbf{n}) \tilde{P}(s; \mathbf{n})] \quad (2)$$

where $\mathbf{v}_i = (\nu_{i1}, \dots, \nu_{iN})^T$, and $\tilde{P}(s; \mathbf{n})$ and $\tilde{M}_i(s; \mathbf{n})$ are Laplace transforms of $P(t; \mathbf{n})$ and $M_i(t; \mathbf{n})$, respectively.

Interestingly, if we define $K_i(\mathbf{n}) = \lim_{s \rightarrow 0} \tilde{M}_i(s; \mathbf{n})$ and , then we can show [27]

$$K_i(\mathbf{n}) = \frac{\int_0^{+\infty} \psi_i(t; \mathbf{n}) \prod_{j \neq i} [1 - \Psi_j(t; \mathbf{n})] dt}{\int_0^{+\infty} \prod_{j=1}^M [1 - \Psi_j(t; \mathbf{n})] dt}, \quad i = 1, \dots, M \quad (3)$$

where $\Psi_i(t; \mathbf{n}) = \int_0^t \psi_i(t'; \mathbf{n}) dt'$ is the cumulative distribution function. Function $K_i(\mathbf{n})$ will be called the effective transition rate of the i th reaction. Different from reaction propensity function $a_i(\mathbf{n})$ in the Markovian case, which is a polynomial of $\mathbf{n}$, function $K_i(\mathbf{n})$ may be a rational function of $\mathbf{n}$. Note that effective transition rate $K_i(\mathbf{n})$ decodes the effect of non-Markovity in the original reaction system.

Now, we assume that the stationary distribution of the underlying process exists, denoted by $P(\mathbf{n})$. Then, based on the chemical CRTW theory and effective reaction propensity functions

$K_i(\mathbf{n})$, we can derive the following sgCME in compact form as [27]

$$\sum_{i=1}^M \left(\prod_{j=1}^N \mathbb{E}^{-V_{ij}} - 1 \right) K_i(\mathbf{n}) P(\mathbf{n}) = 0 \quad (4)$$

where $\mathbb{E}^{-V_{ij}}$ is a step operator with the operation rule below: $\mathbb{E}^{-V_{ij}} f(\mathbf{n}) = f(n_1, \dots, n_j - s_{ij}, \dots, n_N)$ for any function $f(\mathbf{n})$. If we construct an equivalent reaction network with $K_i(\mathbf{n})$ taken as reaction propensity functions, then Eq. (4) has two direct and important implications: (i) the original non-Markovian issue is transformed to an equivalent Markovian one; (ii) although the temporal dynamics of individual reactants would be non-Markovian in the original system, the set of all microscopic degrees of freedom of a reaction system finally evolves to be Markovian, regardless of both the topology of the underlying network and the form of inter-reaction time or/and delay distributions.

3. Two efficient algorithms

Based on the above formulation, i.e., Eq. (4) with Eq. (3), we now easily establish numerical algorithms. Here we present two efficient algorithms: sgGA and sgFSP. The former generates realizations with Monte Carlo simulation whereas the later computes an approximate probability distribution. In spite of natural extensions of the corresponding classical algorithms, sgGA and sgFSP are the first algorithms which can efficiently solve a mathematically intractable non-Markovian issue.

3.1 Generalized Gillespie algorithm

Recall that traditional GA introduces two random numbers: one for modeling which reaction will happen and the other for modeling when the next reaction will happen. The GA can provide statistically exact methods for simulating stochastic dynamics as interacting sequences of discrete events [29]. However, this algorithm exploits the hypothesis that inter-event times follow exponentially distributions. In the present study, we relax this hypothesis and propose a so-called sgGA based on Eq. (4), which is an extended version of the GA but can be applied to stochastic analysis of a general biochemical network with arbitrary waiting-time distributions. It is worth pointing out that many other extensions of the GA can also be established in cases of non-Markovian

reactions.

Similar to the traditional GA, we draw two random numbers r_1 and r_2 from the uniform distribution in the unit interval, and set

$$\tau = \frac{1}{\sum_{i=1}^M K_i(\mathbf{n})} \ln\left(\frac{1}{r_1}\right) \quad (5a)$$

$$j = \text{the smallest integer satisfying } \sum_{i=1}^j K_i(\mathbf{n}) > r_2 \sum_{i=1}^M K_i(\mathbf{n}) \quad (5b)$$

where $K_i(\mathbf{n})$ is calculated according to Eq. (3). Then, the sgGA, which constructs an exact numerical realization of the non-Markovian process $\mathbf{n}(t)$, is stated below

0. Initialize the time $t = t_0$ and the system's state $\mathbf{n} = \mathbf{n}_0$.

1. With the system in state $\mathbf{n}$ at time t_0 , evaluate all the $K_i(\mathbf{n})$ and their sum $\sum_{i=1}^M K_i(\mathbf{n})$.

2. Generate values of τ and j using Eq. 5(a) and 5(b).

3. Update the system's state based on the selected reaction and by replacing $t \leftarrow t + \tau$ and

$$\mathbf{n} \leftarrow \mathbf{n} + (v_{1i}, v_{2i}, \dots, v_{Ni})^T.$$

4. Record trajectories $\mathbf{n}(t)$ as desired. Return to Step 1, or else, end simulation.

After discarding the early part of the trajectories (the burn-in phase), the remaining realization values are used to estimate the steady-state probability distribution and related statistics.

3.2 Generalized finite state projection algorithm

The FSP method developed initially by Munsky and Khammash has been very successful in solving the CME even for relatively large systems [45]. Based on the idea of FSP, here we propose a sgFSP, which enables to compute stationary probability distributions in systems of non-Markovian biochemical networks.

First, the sgCME (i.e., Eq. 4) can be rewritten in the following form:

$$\mathbf{A}\mathbf{P} = 0 \quad (6)$$

where all possible states space $\mathcal{S}$ are enumerated as a vector $(\mathbf{n}_1, \mathbf{n}_2, \dots)^T$, and $\mathbf{P}$ is a vector of

the stationary probability distributions, i.e., $\mathbf{P}(t) = (P(t; \mathbf{n}_1), P(t; \mathbf{n}_2), \dots)^T$. The i th diagonal

element of matrix $\mathbf{A}$ is negative with magnitude equal to the sum of the effective transition rates

of reactions that leave the i th state, i.e., $A_{ii} = -\sum_{l \neq i} K_l(\mathbf{n}_i)$, and off-diagonal element A_{ij} is

positive with magnitude being $K_l(\mathbf{n}_j)$ if there is a reaction $l \in \{1, \dots, L\}$ such that $\mathbf{n}_j = \mathbf{n}_i + (v_{i1}, v_{i2}, \dots, v_{iN})^T$ and zero otherwise. In other words, the state reaction matrix $\mathbf{A}$ comprises of elements:

$$A_{ij} = \begin{cases} -\sum_{l \neq i} K_l(\mathbf{n}_i) & \forall i = j \\ K_l(\mathbf{n}_i) & \forall j \neq i \text{ such that } \mathbf{n}_j = \mathbf{n}_i + (v_{i1}, v_{i2}, \dots, v_{iN})^T \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

We point out that if the number of the states is infinite or extremely large, the sgCME is very difficult to solve, but one can get a good approximation of the solution by making use of truncation methods such as finite state projection [45], sliding window [46], finite buffer method [47,48] and quantized tensor trains methods [49]. In these methods, truncation is made to guarantee both that the number of states retained is small enough and that the truncated equation is efficiently solved. On contrary, if the number of states retained is very large, the truncated equation is still difficult to solve.

Then, we adopt finite state projection as a foundation to construct a sgFSP. Let $\mathcal{M}$ denote a Markov chain on the state set $\mathcal{S}$, whose master equation is given by $\partial_t \mathbf{P}(t) = \mathbf{A} \mathbf{P}(t)$. Let $\mathcal{M}_j$ denote a reduced Markov chain, which comprises of the state set $\mathcal{S}_j$ plus a single absorbing state $\mathcal{S}_{j'}$ (referring to Fig. 2(B)). The stationary probability distribution satisfying the master equation of $\mathcal{M}_j$ is computed according to

$$\begin{bmatrix} \mathbf{A}_j \\ -\mathbf{1}^T \mathbf{A}_j \end{bmatrix} \begin{bmatrix} \mathbf{P}_j \\ \mathbf{P}_{j'} \end{bmatrix} = 0 \quad (8)$$

where $\mathbf{1}^T = (1, 1, \dots, 1)$ is a row vector with the same dimension as that of $\mathbf{A}_j$, $\mathbf{P}_j$ and $\mathbf{P}_{j'}$ are the stationary probability distributions corresponding to state sets $\mathcal{S}_j$ and $\mathcal{S}_{j'}$, respectively.

4. Implementations

In order to show the effectiveness of sgGA and sgFSP, here we analyze three representative examples: a generalized birth-death model, a generalized toggle switch model, and an on-off model of gene expression, each considering molecular memory.

4.1 A model of birth and death with molecular memory

The first example to be analyzed is a birth-death model, which is a natural and formal framework for modeling a vast variety of biological processes such as population dynamics, speciation, genome evolution [27,50]. Consider a birth-death network consisting of the following two non-Markovian reactions (Fig. 2(A))

where two waiting time distributions are set as $\psi_1(t;n) = [\Gamma(k_b)]^{-1} [\lambda_b(n)]^{k_b} t^{k_b-1} e^{-\lambda_b(n)t} \mathbb{I}_{(0,\infty)}(t)$ and $\psi_2(t;n) = [\Gamma(k_d)]^{-1} [\lambda_d(n)]^{k_d} t^{k_d-1} e^{-\lambda_d(n)t} \mathbb{I}_{(0,\infty)}(t)$. Symbol $\mathbb{I}_A(x)$ is an indicator function of set A , i.e., $\mathbb{I}_A(x) = 1$ if $x \in A$ and $\mathbb{I}_A(x) = 0$ otherwise. Symbol $\Gamma(\cdot)$ is the common Gamma function. In numerical simulation, we limit our consideration to the case of Gamma distributions, although our approach may be applied to other kinds of distributions suggested in the literature [51] (additional explanations can be found in the next section).

It is worth pointing out that the above formulation contains the stochastic model of a gene's constitutive expression as its special case, where protein synthesis corresponds to birth while protein degradation to death [52,53]. Here, we consider the following two cases of gene expression (referring to Figure 2(C) and (F)):

Case 1 Gene expression without feedback

$$\lambda_b(n) = \lambda_b, \quad \lambda_d(n) = \lambda_d n \quad (10)$$

Case 2 Gene expression with feedback

$$\lambda_b(n) = \frac{\lambda_{\max} K^H}{K^H + n^H} + \varepsilon, \quad \lambda_d(n) = \lambda_d n \quad (11)$$

where $\lambda_{\max}$ is the maximum transcription rate, K is the equilibrium constant of the transcription factor binding to its DNA binding site, and ε represents a small leakiness of the promoter. Hill coefficient $H > 0$ corresponds to negative feedback and $H < 0$ to positive feedback.

In order to show how sgFSP is performed, we set $\mathcal{S}_j = \{0, 1, \dots, N\}$ (Fig. 2(B)). In addition, denote $\mathbf{P} = (P(0), \dots, P(N))^T$, where the conservative condition $\sum_{i=1}^N P(i) = 1$ must hold. Introduce a matrix of the form

$$\mathbf{A} = \begin{pmatrix} -K_1(0) & K_2(1) & & & \\ K_1(0) & -(K_1(1)+K_2(1)) & K_2(2) & & \\ & \ddots & \ddots & \ddots & \\ & & K_1(N-2) & -(K_1(N-1)+K_2(N-1)) & K_2(N) \\ & & & K_1(N-1) & -K_1(N) \end{pmatrix} \quad (12)$$

By solving the algebraic equation $\mathbf{A}\mathbf{P}=0$ with $\sum_{i=1}^N P(i)=1$, we can obtain a numerical $\mathbf{P}$.

For clarity, we distinguish the following four cases to demonstrate numerical results: (1) $\psi_1(t;n)$ is a Gamma distribution whereas $\psi_2(t;n)$ is an exponential distribution (i.e., $k_d=1$) in **case 1** (Fig. 2(D)); (2) $\psi_2(t;n)$ is a Gamma distribution whereas $\psi_1(t;n)$ is an exponential distribution (i.e., $k_b=1$) in **case 1** (Fig. 2(E)); (3) $\psi_1(t;n)$ is a Gamma distribution for positive feedback regulation whereas $\psi_2(t;n)$ is an exponential distribution in **case 2** (Fig. 2(G)); (4) $\psi_1(t;n)$ is a Gamma distribution for negative feedback regulation whereas $\psi_2(t;n)$ is an exponential distribution in **case 2** (Fig. 2(H)). In all the cases, we set k_b/λ_b , k_d/λ_d and $k_b/\lambda_{\max}$ as constants, so that the average waiting times remain unchanged. Recall that a Gamma distribution can model a multistep process. Therefore, integer parameter k_b (or k_d) represents the step number in our case.

Fig 2. (A) Schematic of generalized birth-death process. (B) A reduced Markov chain with a single absorbing state $\mathcal{S}_f$. (C) A constitutive gene expression model, where (D) shows protein distributions for different k_b , and (E) show protein distributions for different k_d . (F) A gene self-regulatory model, where (G) shows protein distributions for different k_b and (H) shows protein distributions for different k_d . Some parameter values are set as: (D) $k_d = 1, \lambda_d = 1$, (E) $k_b = 1, \lambda_b = 20$, (F) $K = 10, H = -2, \varepsilon = 0.1, k_d = 1, \lambda_d = 1$ and (G) $K = 10, H = 2, \varepsilon = 0.1, k_d = 1, \lambda_d = 1$.

Numerical results are demonstrated in Fig. 2. In Fig. 2(D)(E), lines correspond to distributions obtained by sgFSP and empty circles to distributions obtained by sgGA. From these two panels, we observe that the two obtained kinds of distributions match remarkably well, implying that the two methods are equivalent. Therefore, we focus on the performance of sgFSP in the following. We also observe that protein distributions shifts toward the direction that k_b increases (Fig. 2(D)(G)(H)), but toward the opposite direction (Fig. 2 (E)). Interestingly, protein populations demonstrate bimodal distribution when $k_b > 1$ (Fig. 2 (E)), implying that molecular memory can induce population diversity.

4.2 Gene model of toggle switch with molecular memory

Recall that a toggle switch network (Fig. 3(A)) can model the cross-repression between the determinants of different cellular states, which can result in a definite choice between two outcomes [54-56]. Conventional models of genetic toggle switch consider exponential waiting time distributions. However, the expression of a gene in general involves a multi-steps process. Indeed, transcriptional repressor monomer (A or B) binds first to dimers and then to specific DNA sequences near the promoter, repressing the production of transcriptional repressor monomer (B or A). This multi-step process can result in non-exponential waiting times and hence create molecular memories between reaction events. Here, we consider a stochastic model of genetic toggle switch with memories, which consists of the following four reactions (or referring to Fig. 3(B)):

where $\psi_g^A(t;n), \psi_{\text{deg}}^A(t;n), \psi_g^B(t;n)$ and $\psi_{\text{deg}}^B(t;n)$ are waiting time distributions. We assume that

these distributions are:

(1) Transcription process for gene A: $\psi_g^A(t; \mathbf{n}) = \left[\Gamma(k_g^A) \right]^{-1} \left[\lambda_g^A(\mathbf{n}) \right]^{k_g^A} t^{k_g^A-1} e^{-\lambda_g^A(\mathbf{n})t} \mathbb{I}_{(0,\infty)}(t)$ with

$$\lambda_g^A(\mathbf{n}) = \frac{\alpha_A}{1 + \beta_A n_B^{H_A}};$$

(2) Degradation process for gene A: $\psi_{\text{deg}}^A(t; \mathbf{n}) = \lambda_{\text{deg}}^A n_A e^{-\lambda_{\text{deg}}^A n_A t} \mathbb{I}_{(0,\infty)}(t);$

(3) Transcription process for gene B: $\psi_g^B(t; \mathbf{n}) = \left[\Gamma(k_g^B) \right]^{-1} \left[\lambda_g^B(\mathbf{n}) \right]^{k_g^B} t^{k_g^B-1} e^{-\lambda_g^B(\mathbf{n})t} \mathbb{I}_{(0,\infty)}(t)$ with

$$\lambda_g^B(\mathbf{n}) = \frac{\alpha_B}{1 + \beta_B n_A^{H_B}};$$

(4) Degradation process for gene B: $\psi_{\text{deg}}^B(t; \mathbf{n}) = \lambda_{\text{deg}}^B n_B e^{-\lambda_{\text{deg}}^B n_B t} \mathbb{I}_{(0,\infty)}(t).$

Note that $k_g^A = 1$ and $k_g^B = 1$ correspond to the Markovian case whereas $k_g^A > 1$ or $k_g^B > 1$

corresponds to the non-Markovian case. In numerical simulations, we always set $k_g^B = 1$.

Fig 3. (A) Schematic for a model of toggle switch with memory, where two genes are repressed to each other. (B) Four reactions corresponding to (A), where waiting times for synthesis and degradation of each protein follow distributions, respectively. Default parameter values are taken as: $\beta_A = \beta_B = 1$, $H_A = H_B = 1$, $\lambda_{\text{deg}}^A = \lambda_{\text{deg}}^B = 1$. (C) and (F) Joint distributions of proteins A and B, obtained by sgFSP; (D) and (G) Heat maps in the plane of protein A and B; (E) and (H) Time series of the difference between the levels of protein A and B, obtained by sgGA. (C)-(E) correspond to

exponential waiting times, where parameter values are set as $k_g^A = 1, \alpha_A = 20$; (F)-(H) correspond to non-exponential waiting times, where parameter values are set as $k_g^A = 2, \alpha_A = 40$.

Numerical results are demonstrated in Fig. 3, where Fig. 3(C)-(E) correspond to the case of exponential waiting times whereas Fig. 3(F)-(H) to the case of non-exponential waiting times. We observe that if the waiting time for synthesis of protein A follows an exponential distribution (i.e., if $k_g^A = 1$ is set), the steady-state joint distribution of protein A and B is unimodal, referring to Fig. 3(C)(D). However, if the waiting time for synthesis of protein A follows a non-exponential distribution (i.e., if $k_g^A = 2$ is set), the steady-state joint distribution of protein A and B is bimodal, referring to Fig. 3(F)(G). To examine the time dependence of the populations of two proteins in a single cell, we first perform stochastic simulations with sgGA and then calculate the difference between the levels of proteins A and B. The numerical results are shown in Fig. 3(E)(H). Comparing Fig. 3(E) with Fig. 3(H), we find that two switching states occur only in the case of non-exponential waiting times or memory. In a word, we conclude from Fig. 3 that molecular memory can induce bimodal distributions in the toggle switch model described in Fig. 3(A) or Fig. 3(B).

4.3 On-off model of gene expression with molecular memory

The 3-stage model of gene expression [12-15,57,58] can find its prototypes in prokaryotic and eukaryotic cells. This model, which can capture many essential features of transcriptional biology, has extensively been used and studied. The basic hypothesis of the model is that the gene promoter exists either in an activated state that produces mRNA or in a repressed state that is unproductive, and that an mRNA undergoes a competition between translation and degradation. In traditional studies, it is also assumed that every chemical reaction involved happens probabilistically at a fixed rate, i.e., waiting times between reaction events are assumed to follow exponential distributions. However, biochemical events do not necessarily occur in single steps of individual molecules. For example, a promoter's gradual maturing through a series of inactive states (many of which have been unspecified) before activation can result in non-exponential distributed gestation periods between transcription windows. Also for example, mRNAs or proteins' senescing through a series of states (many of which have also been unspecified) before they are finally degraded can also lead to non-exponential waiting times.

In order to model various possible regulations involved in gene expression, we replace exponential waiting time distributions with non-exponential waiting time distributions for transitions between promoter states, transcription, translation and degradation. Thus, the traditional 3-stage model of gene expression is modified as (also referring to Fig. 4(A)):

where OFF denotes an inactive promoter state, ON denotes an active promoter state, and the first two reactions describe the promoter switch between its OFF and ON states. The ON-state DNA can generate mRNAs that are further translated into proteins and are eventually degraded. Functions $\psi_{\text{on}}(t;\mathbf{n})$, $\psi_{\text{off}}(t;\mathbf{n})$, $\psi_{\text{m}}(t;\mathbf{n})$, $\psi_{\text{mdeg}}(t;\mathbf{n})$, $\psi_{\text{p}}(t;\mathbf{n})$ and $\psi_{\text{pdeg}}(t;\mathbf{n})$ in Eq. (14), which are waiting time distributions, can describe complex regulatory processes. For convenience, we denote by $\mathbf{n} = (\text{on}, \text{off}, m, p)^T$ the state vector of the gene system, where *on*, *off*, *m*, *p* represent the numbers of DNA molecules at off and on states, the numbers of mRNA and protein molecules, respectively. We assume that all these waiting time distributions are Gamma distributions, that is,

- (1) Promoter switch process: $\psi_{\text{on}}(t;\mathbf{n}) = \left((\lambda_{\text{on}} \text{off})^{k_{\text{on}}} / \Gamma(k_{\text{on}}) \right) t^{k_{\text{on}}-1} e^{-\lambda_{\text{on}} \text{off} t} \mathbb{I}_{(0,\infty)}(t)$,
 $\psi_{\text{off}}(t;\mathbf{n}) = \left((\lambda_{\text{off}} \text{on})^{k_{\text{off}}} / \Gamma(k_{\text{off}}) \right) t^{k_{\text{off}}-1} e^{-\lambda_{\text{off}} \text{on} t} \mathbb{I}_{(0,\infty)}(t)$.
- (2) Transcription process: $\psi_{\text{m}}(t;\mathbf{n}) = \left((\lambda_{\text{m}} \text{on})^{k_{\text{m}}} / \Gamma(k_{\text{m}}) \right) t^{k_{\text{m}}-1} e^{-\lambda_{\text{m}} \text{on} t} \mathbb{I}_{(0,\infty)}(t)$.
- (3) mRNA degradation process: $\psi_{\text{mdeg}}(t;\mathbf{n}) = \left((\lambda_{\text{mdeg}} m)^{k_{\text{mdeg}}} / \Gamma(k_{\text{mdeg}}) \right) t^{k_{\text{mdeg}}-1} e^{-\lambda_{\text{mdeg}} m t} \mathbb{I}_{(0,\infty)}(t)$.
- (4) Translation process: $\psi_{\text{p}}(t;\mathbf{n}) = \left((\lambda_{\text{p}} m)^{k_{\text{p}}} / \Gamma(k_{\text{p}}) \right) t^{k_{\text{p}}-1} e^{-\lambda_{\text{p}} m t} \mathbb{I}_{(0,\infty)}(t)$.
- (5) Protein degradation process: $\psi_{\text{pdeg}}(t;\mathbf{n}) = \left((\lambda_{\text{pdeg}} p)^{k_{\text{pdeg}}} / \Gamma(k_{\text{pdeg}}) \right) t^{k_{\text{pdeg}}-1} e^{-\lambda_{\text{pdeg}} p t} \mathbb{I}_{(0,\infty)}(t)$.

As is well known, the relative rate of the constant gene-state transition rates has significant impact on gene expression dynamics. Here, we compute protein distributions in two dynamic regimes: ‘fast’ gene-state dynamics (possibly due to fast binding and unbinding of transcription factors)) and ‘slow’ gene-state dynamics (possibly due to slower changes in chromatin

configuration). Note that parameters L_{on} and L_{off} in the model represent the step numbers of multistep processes, which result in Gamma distributed waiting times. In simulation, the average waiting times is kept constant. Numerical results are shown in Fig.4 (B)-(G). From these panels, we observe that whichever “fast” and “slow” switches, protein distributions shifts to smaller number density as L_{on} is increased (Fig.4 (B)(E)), but shifts higher number density as L_{off} (L_{tot}) is increased (Fig.4 (C)(D)(F)(G)). Interestingly, protein populations tend to limit distributions as L_{off} ($L_{\text{off}}, L_{\text{tot}}$) is increased.

Fig 4. (A) Schematic for a 3-stage model of gene expression with molecular memory. (B)-(G) Protein distributions obtained by sgFSP, where default parameter values are set as following $k_{\text{on}} = 1$, $\lambda_{\text{on}} = 1$ (corresponding to fast switch) or $\lambda_{\text{on}} = 0.1$ (corresponding to slow switch), $k_{\text{off}} = 1$, $\lambda_{\text{off}} = 0.5$ (fast switch) or $\lambda_{\text{off}} = 0.05$ (slow switch), $k_m = 1$, $\lambda_m = 5$, $k_{\text{mdeg}} = 1$, $\lambda_{\text{mdeg}} = 1$, $k_p = 1$, $\lambda_p = 2$, $k_{\text{pdeg}} = 1$, $\lambda_{\text{pdeg}} = 0.2$. In (B) and (E), $k_{\text{on}} = L_{\text{on}}$, $\lambda_{\text{on}} = L_{\text{on}}$ (fast switch) and $\lambda_{\text{on}} = 0.1L_{\text{on}}$ (slow switch). In (C) and (F), $k_{\text{off}} = L_{\text{off}}$, $\lambda_{\text{off}} = 0.5L_{\text{off}}$ (fast switch) and $\lambda_{\text{off}} = 0.05L_{\text{off}}$ (slow switch). In (D) and (G), all the default parameter values are multiplied by L_{tot} .

5. Discussion

A key challenge in stochastic modeling of cellular networks is in raising computational efficiency involving the jointing effects of stochastic fluctuations (or the noise) and molecular memory. We have developed a novel methodological approach, which is based on sgCME, to compute steady-state probability distributions in systems of general cellular reaction processes with molecular memory. While the sgCME would open a new way to address this classically intractable problem from a Markovian viewpoint, the approach can develop a set of efficient simulation methods. Here, we have proposed a Monte Carlo simulation algorithm (i.e., sgGA) to generate realizations of stochastic trajectories, and an approximate algorithm (i.e., sgFSP) to directly compute steady-state probability distributions. Analysis of three representative exemplars has verified the effectiveness of these algorithms.

Several potential weaknesses yet computational challenges with our methods exist both in simulation of transient stochastic dynamics and in choice of an appropriate waiting time distribution. First, our methods rest on sgCME. Although this equation can reveal long time behavior of stochastic process with memory, but lacks moderately transient dynamics. To overcome this difficulty, there are two possible roadmaps: One is that one may take the whole line back to the gCME formulation, numerical integrate the corresponding integro-differential equation, and store the whole historical data; And another is to resort to Monte Carlo simulation. Future extensions of the algorithms are under way. Second, the waiting time distribution selection is essentially a statistic inference problem in systems biology. Large amount of experimental data imply that Gamma distribution and phase-type distribution are good choices. The specifics are outside the scope of this article and will be discussed elsewhere.

We stress that our framework opens a door to analyze and simulate stochastic behaviors of the systems of cellular processes with molecular memory. The presented formulation enables us to efficiently develop algorithms under a Markovian framework with effective transition rates instead of reaction propensity functions. And our theory is general in that it can deal with biochemical networks of arbitrary topology and arbitrary types of waiting time distributions. Therefore, theoretically, successful analysis methods and simulation algorithms developed in the Markovian framework can apply to non-Markovian problems [59]. In addition, it is worth pointing out that our

formalism is not restricted to biochemical networks considered here and can be easily extended to any non-Markovian networks with a finite/infinite set of discrete states, such as epidemic spreading on complex networks [60-63].

Finally, the emergence of measurement techniques currently allow for access to increasingly rich data on approximately steady-state distributions of interested molecular species and waiting time distributions of key reaction events. The presented methods are computationally efficient and scalable. This will facilitate the quantitative mechanistic modelling of complex cellular processes and the exploitation of cell-to-cell variability and molecular memory.

Appendix: Biophysical foundation of waiting-time distributions

The timing of cellular processes cannot be described often by simple Poisson processes, which are traditionally characterized by exponential distributed waiting time distributions and hence can be modelled by classical chemical master equations. However, many single-peaked and even multi-modal waiting-time distributions have been observed empirically, reflecting complexity of cellular processes modelled by biochemical reactions networks. Indeed, non-exponential waiting-time distributions are signatures of relevant intracellular processes without a priori knowledge. A main advantage of gCME capturing effects of non-exponential waiting times is that it can focus on a small number of key measurable events. Here, we derive the most common waiting time distributions of biochemical reactions: exponential distribution, Gamma distribution and phase-type distribution, focusing on their biophysical explanations.

Exponential distribution

Exponential distributions can characterize the time between events in a Poisson point process. Exponential waiting times between reactions are a pillar of the classical chemical master equation. Although exponential waiting time distributions can be derived in different ways, an approximate analysis of molecular collision times provides a clear physical explanation.

Assume that all molecules are spheres with diameter D . If the center of one molecule gets the center of another molecule within a distance D , a collision occurs. Specifically, in the next small time interval dt , a molecule will sweep out a volume equal to $\pi D^2 v dt$, in the sense that if another molecule center falls within this volume, a collision will occur in the time interval $(t, t+dt)$. N with

$N \gg 1$ molecules sweep out this volume for N times. For a given molecule center, the probability that it falls within the total volume swept out by the other molecule is then approximately $N\pi D^2 v dt/V$, where V is the total volume and v is the average molecule speed. The probability of no collision for a given molecule is [29],

$$P(\text{no collision in time } dt) = 1 - N\pi D^2 v dt/V \quad (\text{A1})$$

The probability of no collision after N time intervals dt is

$$P(\text{no collision in time } kdt) = (1 - N\pi D^2 v dt/V)^N \sim \exp(-N\pi D^2 v t/V) \quad (\text{A2})$$

where $t = Ndt$. The probability of a collision is $1 - P(\text{no collision in time } Ndt)$, and the probability density of t can be written

$$\psi(t) = \mu e^{-\mu t} \quad \text{with} \quad \mu = N\pi D^2 v/V. \quad (\text{A3})$$

It is the exponential distribution which can be used to modeling the waiting times between single elementary reaction steps. Note that this derivation is based on Markovian assumption and the exponential distribution can also be derived from a master equation [3,4].

Gamma distribution

In contrast to the times until the next molecular event in single-step processes (elementary reactions), the waiting times between reactions are, in general, not exponentially distributed. If a reaction process goes through a series of (identical) exponential steps, the waiting times will be Gamma distributed [51]. Mathematically, this process can be regarded as a sum of n single-step processes, and therefore the over-all waiting time is the n -fold convolution

$$\psi(t) = \underbrace{[\mu e^{-\mu t} * \mu e^{-\mu t} * \dots]}_{n \text{ times}} = \frac{t^{n-1} e^{-\mu t} \mu^n}{(n-1)!} \quad (\text{A4})$$

where, $*$ denotes convolution and $\psi(t)$ is known as the gamma distribution with shape parameter n and rate parameter μ . If n is a positive integer, then the distribution is an Erlang distribution.

Phase-type distribution

A phase-type distribution is a probability distribution constructed by a convolution or mixture of exponential distributions. It results from a system of one or more inter-related Poisson processes occurring in sequence or phases, such as reversible chemical reactions. For example, models

describing enzyme dynamics can be mapped onto phase type models. Phase-type distributions can provide a tractable theoretical framework to understand complex waiting time distributions observed for a variety of different biochemical processes. Formally, the PDF of a phase-type distribution is [51]

$$\psi(t) = \mathbf{u} \exp(\mathbf{S}t) \mathbf{S}_0 \quad (\text{A5})$$

where $\mathbf{S}$ is an $m \times m$ matrix, $\mathbf{S}_0$ is a $m \times 1$ vector which consisting of the negative sums of the columns of matrix $\mathbf{S}$, and $\mathbf{u} = (1, 1, \dots, 1)$ is a $1 \times m$ vector.

There are a lot of probability distributions involving waiting times, such as Weibull distribution, Pareto distribution, and Mittag-Leffler distribution. We do not tend to enumerate all here.

Acknowledgments

This work was supported by grants 91530320, 11775314, 11475273, and 11631005 from Natural Science Foundation of P. R. China; 2014CB964703 from Science and Technology Department, P. R. China; 201707010117 from the Science and Technology Program of Guangzhou, P. R. China.

Author Contributions

Conceived and designed the experiments: JJ, TS. Performed the experiments: JJ. Analyzed the data: JJ. Contributed reagents/materials/analysis tools: JJ, TS. Wrote the paper: JJ, TS.

Reference

1. Raser JM, O'Shea EK. Noise in gene expression: origins, consequences, and control. *Science*. 2005; 309(5743): 2010-2013. <https://doi.org/10.1126/science.1105891>. PMID: 16179466.
2. Raj A, van Oudenaarden A. Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell*. 2008; 135(2):216-226. <https://doi.org/10.1016/j.cell.2008.09.050>. PMID: 18957198.
3. Gardiner CW. *Stochastic methods-a handbook for the natural and social sciences*. Berlin: Springer; 2009.
4. Van Kampen NG. *Stochastic processes in physics and chemistry*. Amsterdam: North-Holland; 2007.
5. Gillespie DT. A rigorous derivation of the chemical master equation, *Physica A*. 1992.

- 188A:404-425. [https://doi.org/10.1016/0378-4371\(92\)90283-V](https://doi.org/10.1016/0378-4371(92)90283-V).
6. Pardoux E. Markov processes and applications: algorithms, networks, genome and finance. New York: Wiley & Sons; 2008.
7. Haag G. Modelling with the master equation: solution methods and applications in social and natural sciences. Switzerland: Springer; 2017.
8. Weber MF, Frey E. Master equations and the theory of stochastic path integrals, Rep Prog. Phys. 2017; 80(4):046601. <https://doi.org/10.1088/1361-6633/aa5ae2>. PMID: 28306551.
9. Guérin T, Levernier N, Bénichou O, Voituriez R. Mean first-passage times of non-Markovian random walkers in confinement, Nature. 2016; 534(7607):356-359. <https://doi.org/10.1038/nature18272>. PMID: 27306185.
10. Guérin T, Bénichou O, Voituriez R. Non-Markovian polymer reaction kinetics. Nat Chem. 2012; 4(7):568-573. <https://doi.org/10.1038/nchem.1378>. PMID: 22717443.
11. Pedraza JM, Paulsson J. Effects of molecular memory and bursting on fluctuations in gene expression. Science. 2008; 319(5861):339-433. <https://doi.org/10.1126/science.1144331>. PMID: 18202292.
12. Shahrezaei V, Swain PS. Analytical distributions for stochastic gene expression. Proc Natl Acad Sci USA. 2008; 105(45):17256-17261. <https://doi.org/10.1073/pnas.0803850105>. PMID: 18988743.
13. Jia T, Kulkarni RV. Intrinsic noise in stochastic models of gene expression with molecular memory. Phys Rev Lett. 2011; 106(5):058102. <https://doi.org/10.1103/PhysRevLett.106.058102>. PMID: 21405439.
14. Schwabe A, Rybakova KN, Bruggeman FJ. Transcription stochasticity of complex gene regulation models. Biophys J. 2012; 103(6):1152-1161. <https://doi.org/10.1016/j.bpj.2012.07.011>. PMID: 2299548.
15. Kumar N, Singh A, Kulkarni RV. Transcriptional bursting in gene expression: analytical results for general stochastic models. PLoS Comput Biol. 2015; 11(10):e1004292. <https://doi.org/10.1371/journal.pcbi.1004292>. PMID: 26474290.
16. Salathé M, Kazandjieva M, Lee JW, Levis P, Feldman MW, Jones JH. A high resolution human contact network for infectious disease transmission. Proc Natl Acad Sci USA. 2010; 107(51):22020-22025. <https://doi.org/10.1073/pnas.1009094108>. PMID: 21149721.
17. Corral A. Long-term clustering, scaling, and universality in the temporal occurrence of earthquakes. Phys Rev Lett. 2004; 92(10):108501. <https://doi.org/10.1103/PhysRevLett.92.108501>. PMID: 15089251.
18. Stumpf PS, Smith RC, Lenz M, Schuppert A, Müller FJ, Babbie A, et al. Stem cell differentiation as a non-Markov stochastic process. Cell Syst. 2017; 5(3):268-282.e7. <https://doi.org/10.1016/j.cels.2017.08.009>. PMID: 28957659.
19. Suter DM, Molina N, Gatfield D, Schneider K, Schibler U, Naef F. Mammalian genes are

- transcribed with widely different bursting kinetics. *Science*. 2011; 332(6028):472-474.
<https://doi.org/10.1126/science.1198817>. PMID: 21415320.
20. Harper CV, Finkenstädt B, Woodcock DJ, Friedrichsen S, Semprini S, Ashall L, et al. Dynamic analysis of stochastic transcription cycles. *PLoS Biol*. 2011; 9(4):e1000607.
<https://doi.org/10.1371/journal.pbio.1000607>. PMID: 21532732.
21. English BP, Min W, van Oijen AM, Lee KT, Luo G, Sun H, et al. Ever-fluctuating single enzyme molecules: Michaelis-Menten equation revisited. *Nat Chem Biol*. 2006; 2(2):87-94.
<https://doi.org/10.1038/nchembio759>. PMID: 16415859.
22. Amir A, Kibiler O, Rokney A, Oppenheim AB, Stavans J. Noise in timing and precision of gene activities in a genetic cascade. *Mol Syst Biol*. 2007; 3:71.
<https://doi.org/10.1038/msb4100113>. PMID: 17299413.
23. Larson DR, Zenklusen D, Wu B, Chao JA, Singer RH. Real-time observation of transcription initiation and elongation on an endogenous yeast gene. *Science*. 2011; 332(6028):475-478.
<https://doi.org/10.1126/science.1202142>. PMID: 21512033.
24. Han Q, Bagheri N, Bradshaw EM, Hafler DA, Lauffenburger DA, Love JC. Polyfunctional responses by human T cells result from sequential release of cytokines. *Proc. Natl. Acad. Sci. USA* 2012; 109(5):1607-1612. <https://doi.org/10.1073/pnas.1117194109>. PMID: 22160692.
25. Pilz GA, Bottes S, Betizeau M, Jörg DJ, Carta S, Simons BD, et al. Live imaging of neurogenesis in the adult mouse hippocampus. *Science*. 2018; 359(6376):658-662.
<https://doi.org/10.1126/science.aao5056>. PMID: 29439238.
26. Aquino T, Dentz M. Chemical continuous time random walks. *Phys Rev Lett*. 2017; 119(23):230601. <https://doi.org/10.1103/PhysRevLett.119.230601>. PMID: 29286686.
27. Zhang JJ, Zhou TS. Stationary equations for non-Markovian biochemical systems. *bioRxiv*. 2018; <https://doi.org/10.1101/360123>.
28. Gillespie DT. Monte Carlo simulation of random walks with residence time dependent transition probability rates. *J Comp Phys*. 1978; 28(3):395-407. [https://doi.org/10.1016/0021-9991\(78\)90060-8](https://doi.org/10.1016/0021-9991(78)90060-8).
29. Gillespie DT. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J Comput Phys*. 1976; 22(4):403-434.
[https://doi.org/10.1016/0021-9991\(76\)90041-3](https://doi.org/10.1016/0021-9991(76)90041-3).
30. Bogun áM, Lafuerza LF, Toral R, Serrano MÁ. Simulating non-Markovian stochastic processes. *Phys Rev E*. 2014; 90(4):042108. <https://doi.org/10.1103/PhysRevE.90.042108>. PMID: 25375439.
31. Vestergaard CL, G énois M. Temporal Gillespie algorithm: fast simulation of contagion processes on time-varying networks. *PLoS Comput Biol*. 2015; 11(10):e1004579.
<https://doi.org/10.1371/journal.pcbi.1004579>. PMID: 26517860.
32. Masuda N, Rocha LEC. A Gillespie algorithm for non-Markovian stochastic processes. *SIAM Rev*. 2018; 60(1):95-115. <https://doi.org/10.1137/16M1055876>.

33. Rimón N, Schuldiner M. Getting the whole picture: combining throughput with content in microscopy. *J Cell Sci.* 2011; 124(Pt 22):3743-3751. <https://doi.org/10.1242/jcs.087486>. PMID: 22124141.
34. Bodenmiller B, Zunder ER, Finck R, Chen TJ, Savig ES, Bruggner RV, et al. Multiplexed mass cytometry profiling of cellular states perturbed by small-molecule regulators. *Nat Biotechnol.* 2012; 30(9):858-867. <https://doi.org/10.1038/nbt.2317>. PMID: 22902532.
35. Taniguchi K, Kajiya T, Kambara H. Quantitative analysis of gene expression in a single cell by qPCR. *Nat Methods.* 2009; 6(7):503-506. <https://doi.org/10.1038/nmeth.1338>. PMID: 19525960.
36. Shapiro E, Biezuner T, Linnarsson S. Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat Rev Genet.* 2013; 14(9):618-630. <https://doi.org/10.1038/nrg3542>. PMID: 23897237.
37. Hallen M, Li B, Tanouchi Y, Tan C, West M, You L. Computation of steady-state probability distributions in stochastic models of cellular networks. *PLoS Comput Biol.* 2011; 7(10):e1002209. <https://doi.org/10.1371/journal.pcbi.1002209>. PMID: 2202252.
38. S Karim S, Buzzard GT, Umulis DM. Efficient calculation of steady state probability distribution for stochastic biochemical reaction network. *BMC Genomics.* 2012; 13 Suppl 6:S10. <https://doi.org/10.1186/1471-2164-13-S6-S10>. PMID: 23134718.
39. Gupta A, Mikelson J, Khammash M. A finite state projection algorithm for the stationary solution of the chemical master equation. *J Chem Phys.* 2017; 147(15):154101. <https://doi.org/10.1063/1.5006484>. PMID: 29055349.
40. Munsky B, Li G, Fox ZR, Shepherd DP, Neuert G. Distribution shapes govern the discovery of predictive models for gene regulation. *Proc Natl Acad Sci USA.* 2018; 115(29):7533-7538. <https://doi.org/10.1073/pnas.1804060115>. PMID: 29959206.
41. Wang J, Huang M, Torre E, Dueck H, Shaffer S, Murray J, et al. Gene expression distribution deconvolution in single-cell RNA sequencing. *Proc Natl Acad Sci USA.* 2018; 115(28):E6437-E6446. <https://doi.org/10.1073/pnas.1721085115>. PMID: 29946020.
42. Montroll E, Weiss G. Random walks on lattices. II. *J Math Phys.* 1965; 6:167. <https://doi.org/10.1063/1.1704269>.
43. Kenkre VM, Montroll EW, Shlesinger MF. Generalized master equations for continuous-time random walks. *J Stat Phys.* 1973; 9(1):45-50. <https://doi.org/10.1007/BF01016796>
44. Scher H, Lax M. Stochastic transport in a disordered solid. I. theory. *Phys Rev B.* 1973; 7(10):4491-4502. <https://doi.org/10.1103/PhysRevB.7.4491>.
45. Munsky B, Khammash M. The finite state projection algorithm for the solution of the chemical master equation. *Chem Phys.* 2006; 124(4):044104. <https://doi.org/10.1063/1.2145882>. PMID: 16460146.
46. Henzinger T, Mateescu M, Wolf V. Sliding window abstraction for infinite markov chains. *Proc CAV.* 2009; 5643:337-352. https://doi.org/10.1007/978-3-642-02658-4_27.

47. Cao Y, Liang J. Optimal enumeration of state space of finitely buffered stochastic molecular networks and exact computation of steady state landscape probability. *BMC Syst Biol.* 2008; 2:30. <https://doi.org/10.1186/1752-0509-2-30>. PMID: 18373871.
48. Cao Y, Lu HM, Liang J. Probability landscape of heritable and robust epigenetic state of lysogeny in phage lambda. *Proc Natl Acad Sci USA.* 2010; 107(43):18445-18450. <https://doi.org/10.1073/pnas.1001455107>. PMID: 20937911.
49. Kazeev V, Khammash M, Nip M, Schwab C. Direct solution of the Chemical Master Equation using quantized tensor trains. *PLoS Comput Biol.* 2014; 10(3):e1003359. <https://doi.org/10.1371/journal.pcbi.1003359>. PMID: 24626049.
50. Novozhilov AS, Karev GP, Koonin EV. Biological applications of the theory of birth-and-death processes. *Brief Bioinform.* 2006; 7(1):70-85. <https://doi.org/10.1093/bib/bbk006>. PMID: 16761366.
51. Feller W. An introduction to probability theory and its applications vol. 2. New York: John Wiley & Sons; 2008.
52. Smits WK, Kuipers OP, Veening JW. Phenotypic variation in bacteria: the role of feedback regulation. *Nat Rev Microbiol.* 2006; 4(4):259-271. <https://doi.org/10.1038/nrmicro1381>. PMID: 16541134.
53. Assaf M, Roberts E, Luthey-Schulten Z, Goldenfeld N. Extrinsic noise driven phenotype switching in a self-regulating gene. *Phys Rev Lett.* 2013; 111(5):058102. <https://doi.org/10.1103/PhysRevLett.111.058102>. PMID: 23952448.
54. Gardner TS, Cantor CR, Collins JJ. Construction of a genetic toggle switch in *Escherichia coli*. *Nature.* 2000; 403(6767):339-342. <https://doi.org/10.1038/35002131>. PMID: 10659857.
55. A Lipshtat A, Loinger A, Balaban NQ, Biham O. Genetic toggle switch without cooperative binding. *Phys Rev Lett.* 2006; 96(18):188101. <https://doi.org/10.1103/PhysRevLett.96.188101>. MID: 16712399.
56. Biancalani T, Assaf M. Genetic toggle switch in the absence of cooperative binding: exact results. *Phys Rev Lett.* 2015; 115(20):208101. <https://doi.org/10.1103/PhysRevLett.115.208101>, PMID: 26613474.
57. Peccoud J, Ycart B. Markovian modeling of gene-product synthesis. *Theor Popul Biol.* 1995; 48(2):222-234. <https://doi.org/10.1006/tpbi.1995.1027>.
58. Zhang JJ, Zhou TS. Promoter-mediated transcriptional dynamics. *Biophys J.* 2014; 106(2):479-488. <https://doi.org/10.1016/j.bpj.2013.12.011>. PMID: 24461023.
59. Schnoerr D, Sanguinetti G, Grima R. Approximation and inference methods for stochastic biochemical kinetics-a tutorial review. *J Phys A Math Theor.* 2017; 50(9):093001. <https://doi.org/10.1088/1751-8121/aab38b>.
60. Starnini M, Gleeson JP, Boguñá M. Equivalence between non-Markovian and Markovian dynamics in epidemic spreading processes. *Phys Rev Lett.* 2017; 118(12):128301. <https://doi.org/10.1103/PhysRevLett.118.128301>. PMID: 28388191.

61. Gleeson JP, O'Sullivan KP, Baños RA, Moreno Y. Effects of network structure, competition and memory time on social spreading phenomena. *Phys Rev X*. 2016; 6:021019. <https://doi.org/10.1103/PhysRevX.6.021019>.
62. Masuda N, Porter MA, Lambiotte R. Random walks and diffusion on networks *Phys Rep*. 2017; 716-717:1-58. <https://doi.org/10.1016/j.physrep.2017.07.007>.
63. D Delvenne JC, Lambiotte R, Rocha LE. Diffusion on networked systems is a question of time or structure. *Nat Commun*. 2015; 6:7366. <https://doi.org/10.1038/ncomms8366>. PMID: 26054307.