An effective method for computing the noise in biochemical networks

Jiajun Zhang,[1] Qing Nie,[2] Miao He,[3,a)] and Tianshou Zhou[1,4,a)]

[1]*School of Mathematics and Computational Science, Sun Yat-Sen University, Guangzhou 510275, People's Republic of China*
[2]*Department of Mathematics, University of California at Irvine, Irvine, California 92697, USA*
[3]*School of Life Science, Sun Yat-Sen University, Guangzhou 510275, People's Republic of China*
[4]*Guangdong Province Key Laboratory of Computational Science and School of Mathematics and Computational Science, Sun Yat-Sen University, Guangzhou 510275, People's Republic of China*

We present a simple yet effective method, which is based on power series expansion, for computing exact binomial moments that can be in turn used to compute steady-state probability distributions as well as the noise in linear or nonlinear biochemical reaction networks. When the method is applied to representative reaction networks such as the ON-OFF models of gene expression, gene models of promoter progression, gene auto-regulatory models, and common signaling motifs, the exact formulae for computing the intensities of noise in the species of interest or steady-state distributions are analytically given. Interestingly, we find that positive (negative) feedback does not enlarge (reduce) noise as claimed in previous works but has a counter-intuitive effect and that the multi-OFF (or ON) mechanism always attenuates the noise in contrast to the common ON-OFF mechanism and can modulate the noise to the lowest level independently of the mRNA mean. Except for its power in deriving analytical expressions for distributions and noise, our method is programmable and has apparent advantages in reducing computational cost. © *2013 American Institute of Physics.* [http://dx.doi.org/10.1063/1.4792444]

I. INTRODUCTION

From the viewpoint of molecular interaction, a biological system can be viewed as a biochemical reaction network. The small number of species molecules participating in the biochemical reactions inevitably gives rise to stochastic fluctuations (or noise) in this species, which would play a significant role in the functioning of this network. It has been shown that the molecular noise might have not only a negative effect on, e.g., the functioning of a synthetic genetic oscillator,[1] but also a beneficial effect on, e.g., stochastic focusing in a signaling system.[2] In particular, the molecular noise is essential for many cellular functions, and has been identified as a key factor underlying the observed phenotypic variability of genetically identical cells in homogeneous environments.[3–5] To quantify contributions of the molecular noise in linear or nonlinear biochemical reaction modules is an important step towards understanding fundamental cellular processes and variations in cell populations.[6–24]

In general, the molecular noise can be characterized or coarsely by the standard difference that is defined as the root mean square deviation of the number of molecules from the mean or precisely by the noise intensity that is defined as the ratio of the standard deviation over the mean, a commonly used index. If the standard deviation is a significant fraction of the mean or if the noise intensity is significantly large, then the noise must be considered and cannot be neglected when analyzing the dynamical behavior of a biochemical system. This is typically associated with biochemical reaction networks where the mean number of species molecules is rather low. Such situations are remarkable characteristic of gene regulatory networks since the concentrations of regulatory proteins are often so low that there are only a few molecules in a given cellular culture. For example, for the *lac* operon in *E. coli*, the *lac* repressor is active at concentrations of only 10–20 molecules per cell.

The linear noise approximation (LNA) has been verified to be an effective method for analyzing stochastic fluctuations (or noise) in species molecules of interest in some biochemical reaction networks. Regarding this approach, there are many studies and applications. The representative results were given by Paulsson, who applied LNA to analyze sources of noise in simple regulatory modules and derived analytical formulae.[15,16] More results were presented in Refs. 6 and 24–29. We point out that LNA is accurate and effective only for linear biochemical networks. However, actual biochemical systems are in general nonlinear, e.g., gene auto-regulatory circuits. For a nonlinear reaction network, LNA fails to capture the stochastic effect in general. Similarly, the spectral analysis method (SAM) proposed by Warren *et al.*[30] is effective for determining the exact noise spectra and related statistic properties for linear biochemical networks, but also fail to capture the stochastic effect in nonlinear networks. This raises a question: how to capture the exact effect of noise in a species of interest in a general biochemical system.

Except for LNA and SAM, the most general description for a biochemical network is based on the master equation that describes the time evolution of the joint probability distribution over copy numbers of all species molecules in this system. Analytically finding such a distribution is a great

a)Author to whom correspondence should be addressed. Electronic addresses: mcszhtsh@mail.sysu.edu.cn and lsshem@mail.sysu.edu.cn.

challenge but some progress has been made by applying approximations to the related master equation.[31–33] For example, a wide class of approximations focuses on limits of large concentrations or small switches.[15, 34, 35] More often, modelers resort to stochastic simulation techniques, e.g., the varying-step Monte Carlo method[36] and the Gillespie algorithm.[37] The former requires a high computational cost (generating many sample trajectories) followed by a difficult statistical task (parameterizing or estimating the relevant probability distribution from which the samples are drawn). The latter is very time-consuming when the number of reactions is large. Recently, there have been significant advances in simulation-based methods to circumvent these shortages.[38] Simulation techniques are especially useful for more detailed studies of experimentally well-characterized systems, including those incorporating DNA looping, nonspecific binding[39] and explicit spatial effects.[40, 41] However, it is often beneficial to gain the first intuition of mechanisms behind complicated biochemical networks from simplified analytical models.

This paper aims to present exact results for computing the noise in biochemical reaction networks governed by chemical master equations. We begin with a general theory, which is based on the power series expansion that is effectively used to solve linear or nonlinear differential equations, and then apply this theory to several common biochemical reaction modules including a two-species reaction network, the common ON-OFF model, a gene model with multiple activity states, gene auto-regulatory modules and a signaling motif. For these modules, we derive exact analytical formulae for computing the noise or steady-state probability distributions, some of which have not been given in the existing literature. In particular, when our method is applied to a stochastic gene model with multiple activity states of the promoter (also called the gene model of promoter progression[42, 43]), we find that the multistep mechanisms always reduce transcriptional noise and can regulate the noise to the lowest level independently from the mean expression of mRNA. In addition, we demonstrate that linear methods such as LNA and SAM over- or under-estimate in general the noise in nonlinear biochemical networks. We emphasize that our method is exact and is more easily programmed in contrast to LNA or SAM since our method can transform questions of quantifying noise into those of solving linear algebraic equations where the complexity of biochemical networks under investigation is reflected in the corresponding iterative matrix. In particular, it is powerful in finding steady-state probability distributions in some nonlinear biochemical reaction networks, which have not been given in the literature yet.

II. GENERAL THEORY

Consider a homogeneous culture of N chemical species $\{X_1, X_2, \ldots, X_N\}$ that undergo M reactions $\{R_1, R_2, \ldots, R_M\}$ in a closed vessel with a fixed volume and a constant temperature. Let the N-dimensional vector $\mathbf{n}$ denote the numbers of species molecules. For each reaction, let $a_k(\mathbf{n})$ represent the propensity function of R_k. More precisely, the probability that the kth reaction will occur in the time interval dt is $a_k(\mathbf{n})dt$

$+ o(\mathbf{n})$, where $o(\mathbf{n})$ satisfies the condition $\lim_{dt \to 0} o(\mathbf{n})/dt = 0$. Let the N-dimensional vector $\mathbf{s}_k$ be the stoichiometry associated with R_k. The joint probability of $\mathbf{n}$ species at time t, denoted by $P(\mathbf{n}; t)$, is given by a chemical master equation of the form

$$\frac{\partial P(\mathbf{n}; t)}{\partial t} = \sum_{k=1}^{M} [a_k(\mathbf{n} - \mathbf{s_k})P(\mathbf{n} - \mathbf{s_k}; t) - a_k(\mathbf{n})P(\mathbf{n}; t)]$$

(1)

subjected to the initial condition $P(\mathbf{n_0}; 0)$. Note that a component of $\mathbf{s_k}$, denoted by s_{kj} with $1 \leq j \leq N$, is actually the changing number of X_j-type molecules participating in the kth reaction.

To solve Eq. (1), the standard method is to introduce the generating function $G(\mathbf{z}; t)$ for $P(\mathbf{n}; t)$, that is

$$G(z; t) = \sum_{\mathbf{n}=(n_1, \ldots, n_N)} P(\mathbf{n}; t)z^{\mathbf{n}} \quad \text{with} \quad z^{\mathbf{n}} \equiv z_1^{n_1} z_2^{n_2} \cdots z_N^{n_N}.$$

(2)

Then, Eq. (1) will become a partial differential equation for $G(\mathbf{z}; t)$

$$\frac{\partial}{\partial t} G(z; t) = L_z[G(z; t)],$$

(3)

where L_z is a linear differential operator determined by the right side of Eq. (1). Solving Eq. (3) is in general difficult, but we can give the formal expression of its solution

$$G(z; t) = \sum_{n_1 \geq 0, \ldots, n_N \geq 0} a_{\boldsymbol{n}}(t)(z - \mathbf{1})^{\boldsymbol{n}}$$

$$= \sum_{n_1 \geq 0, \ldots, n_N \geq 0} \frac{1}{(n_1)! \cdots (n_N)!} \frac{\partial^{n_1 + \cdots + n_N} G}{\partial z_1^{n_1} \cdots \partial z_N^{n_N}} \bigg|_{z=\mathbf{1}}$$

$$\times (z_1 - 1)^{n_1} \cdots (z_N - 1)^{n_N},$$

(4)

if the corresponding series is convergent, or

$$G(z; t) = \sum_{n_1 \geq 0, \ldots, n_N \geq 0} b_{\boldsymbol{n}}(t)(z - z_0)^{\boldsymbol{n}}$$

$$= \sum_{n_1 \geq 0, \ldots, n_N \geq 0} \frac{1}{(n_1)! \cdots (n_N)!} \frac{\partial^{n_1 + \cdots + n_N} G}{\partial z_1^{n_1} \cdots \partial z_N^{n_N}} \bigg|_{z=z_0}$$

$$\times (z_1 - z_{10})^{n_1} \cdots (z_N - z_{N0})^{n_N},$$

(5)

if the corresponding series is convergent. The coefficients of distinct derivatives of $G(\mathbf{z}; t)$ are polynomial in $\mathbf{z}$, so we can obtain a set of linear differential equations with respect to $a_{\boldsymbol{n}}(t)$ or $b_{\boldsymbol{n}}(t)$, which are solvable, by substituting Eq. (4) or (5) into Eq. (3) and comparing coefficients of $(z - \mathbf{1})^{\boldsymbol{n}}$ of the same powers. If $b_{\boldsymbol{n}}(t)$ is analytically given, then $a_{\boldsymbol{n}}(t)$ can be computed according to

$$a_{\boldsymbol{n}}(t) = \sum_{k \geq \boldsymbol{n}} k(k - 1) \cdots (k - \boldsymbol{n} + 1)b_k(t)(1 - z_0)^{k-\boldsymbol{n}}. \quad (6)$$

Often, finding the steady-state generating function $G(z)$ or the steady-state probability distribution $P(\mathbf{n})$ is the common interest. In this case, substituting the expression Eq. (4) or (5) into Eq. (3) will lead to a set of linear algebraic equations of a_n (or b_n) in an iterative manner (see the below examples). Such an iterative format can determine all coefficients

a_n (or b_n) due to the conservative condition of probability $\sum_{\mathbf{n}} P(\mathbf{n}) = G(z = 1) = 1$ that implies $a_0 = 1$. Moreover, by the iterative format we can give the nice analytical expressions of all a_n (or b_n) for some common linear or nonlinear biochemical network modules, which in turn give the analytical expression of $G(z)$ (also see the below examples). We point out: (a) expansion (4) will be used if the corresponding coefficients are more easily computed in contrast to expansion (5), and expansion (5) will be used otherwise. Which expansion will be used and how to choose z_0 depend on the coefficient in the term with the highest-order derivative of function G (refer to examples in Sec. III). (b) If the corresponding deterministic system has more than one steady states, this implies coefficients a_n (or b_n) cannot be uniquely determined in spite of the condition $a_0 = 1$. In this case, we should first choose expansion (5) with different z_0, which determines different sets of b_n, and then compute a_n according to Eq. (6) for each set of b_n. As such, we can obtain more than one steady-state distributions (the details will be discussed elsewhere). In this paper, we only consider the case that a_n (or b_n) has the unique solution.

After having obtained $G(z)$, we can give the analytical expression of the steady-state joint probability according to the formula

$$P(\mathbf{n}) = \frac{1}{(n_1)! \cdots (n_N)!} \left. \frac{\partial^{n_1 + \cdots + n_N} G(z_1, \ldots, z_N)}{\partial z_1^{n_1} \cdots \partial z_N^{n_N}} \right|_{(0, \ldots, 0)}. \quad (7)$$

In particular, after having obtained the first-order derivative $\partial_z G(1, 1, \ldots, 1)$ and the second-order derivative $\frac{1}{2} \partial_z^2 G(1, 1, \ldots, 1)$, we easily compute the mean and variance of n_j, denoted by $\langle n_j \rangle$ and $\sigma_{n_j}^2$, respectively, according to

$$\langle n_j \rangle = \partial_{z_j} G(1, 1, \ldots, 1),$$
$$\sigma_{n_j}^2 = \partial_{z_j z_j}^2 G(1, 1, \ldots, 1) + \partial_{z_j} G(1, 1, \ldots, 1) \quad (8)$$
$$- [\partial_{z_j} G(1, 1, \ldots, 1)]^2.$$

Correspondingly, we can compute the noise intensity for n_j (a common index describing noise), denoted by η_{n_j}, according to the formula

$$\eta_{n_j}^2 \equiv \frac{\sigma_{n_j}^2}{\langle n_j \rangle^2} = \frac{\partial_{z_j z_j}^2 G(1, 1, \ldots, 1)}{[\partial_{z_j} G(1, 1, \ldots, 1)]^2} + \frac{1}{\partial_{z_j} G(1, 1, \ldots, 1)} - 1. \quad (9)$$

We emphasize that the formulae (8) and (9) are exact without any approximation.

For clarity and to show advantages of our method, consider the case of one variable. If we Taylor expand the function $G(z)$ at the point $z = 1$ and assume $G(z) = \sum_{i=0}^{\infty} a_i (z - 1)^i$, then formula (9) of computing the noise intensity will become

$$\eta_n^2 = \frac{2a_2}{a_1^2} + \frac{1}{a_1} - 1. \quad (10)$$

The key is how to determine a_2 and a_1. This can be done by substituting Eq. (4) into the resulting partial differential equation of $G(z)$ and comparing the corresponding coefficients of various same powers $(z - 1)^j$. Note that when solving a_2 and a_1, we need the conservative condition $\sum_{n \geq 0} P(n) = 1$ that gives $a_0 \equiv 1$. We point out that formula (10) is also exact if

both a_2 and a_1 are exactly given. For linear biochemical networks, e.g., for those without feedback, both a_1 and a_2 can be exactly given (see examples in Sec. III), similar to the LNA case. For nonlinear biochemical networks, e.g., for those with feedback loops, formula (6) is in theory used to compute a_1 and a_2, which can be exactly given in some cases (also see examples in Sec. III) but would not in other cases, e.g., a gene regulatory network with multimer repression. In the case that a_1 and a_2 cannot be analytically given, our method is still effective since a_1 and a_2 can be approximately given by truncating initial finite terms in the related series and reach an arbitrarily pre-given accuracy due to the fact that $a_n \to 0$ as $n \to \infty$ (see the next paragraph).

Furthermore, we point out that the above coefficient a_k ($k = 0, 1, 2, \ldots$) have apparent meaning. In fact, they represent binomial moments, thus capable of being directly used to compute the probability distribution. To see this, we let X be a discrete random variable that takes nonnegative integer values, and set $P\{X = n\} = P(n) (n = 0, 1, 2, \ldots)$. Assume that all the binomial moments

$$B_n = E\left\{\binom{X}{n}\right\} = \sum_{k=n}^{\infty} \binom{k}{n} P(k), \quad n = 0, 1, 2, \ldots \quad (11)$$

exist. Note that the series $G(z) = \sum_{n=0}^{\infty} P(n) z^n$ is uniformly convergent and hence regular in the circle $|z| < 1$ (where z may be complex). Using the formula for high-order derivatives, we have

$$G^{(n)}(z) = n! \sum_{k=n}^{\infty} \binom{k}{n} z^{k-n} P(k). \quad (12)$$

Thus, we know

$$P(n) = \frac{1}{n!} G^{(n)}(0), \quad B_n = \frac{1}{n!} G^{(n)}(1), \quad n = 0, 1, 2, \ldots \quad (13)$$

implying that $a_n = B_n$ for all n. On the other hand, if the function $G(z)$ is expanded as the power series at $z = 1$, then we have

$$G^{(n)}(z) = n! \sum_{k=n}^{\infty} \binom{k}{n} a_k (z - 1)^{k-n}. \quad (14)$$

In particular, putting $z = 0$ in Eq. (14) yields

$$P(n) = \sum_{k=n}^{\infty} (-1)^{k-n} \binom{k}{n} a_k, \quad n = 0, 1, 2, \ldots. \quad (15)$$

Since a_n tends to zero as n goes to the infinite, the formula (15) can be used to compute approximate distributions reaching any pre-given accuracy by truncating finite terms in the right side of series (15). We point out that for linear biochemical networks, e.g., for those without feedback, $P(n)$ can be exactly given since a_n can be exactly given, whereas for nonlinear biochemical networks, e.g., for those with feedback loops, $P(n)$ would not but can be approximately given since $a_n \to 0$ as $n \to \infty$. Note that the formula (15) can also be extended to

the multivariate case, that is,

$$P(\mathbf{n}) = \sum_{\mathbf{k}=\mathbf{n}}^{\infty} (-1)^{\mathbf{k}-\mathbf{n}} \binom{\mathbf{k}}{\mathbf{n}} a_{\mathbf{k}}$$

$$\equiv \sum_{k_1=n_1}^{\infty} \cdots \sum_{k_N=n_N}^{\infty} (-1)^{\sum_{i=1}^{N}(k_i-n_i)} \prod_{i=1}^{N} \binom{k_i}{n_i} a_{(k_1,\dots,k_N)}$$

(16)

so the above discussion for the single-variable case is also effective for this case.
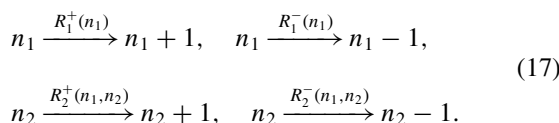
Finally in this section, we emphasize features and applicability of our power series expansion method. Unlike the extensively used moment-closure method,[44–46] which has shortcomings, e.g., it is nonlinear and hence inconvenient to computation; how to approximate high-order moments is unclear; the approximate moments cannot guarantee that a pre-given accuracy is satisfied, our method is linear and hence easier and more convenient to computation. Moreover, it can compute approximate noise and distributions that reach any pre-given accuracies, if $a_{\mathbf{n}}$ or $b_{\mathbf{n}}$ cannot be analytically given. In addition, we point out that for any biochemical reaction network: (1) $\{a_{\mathbf{n}}\}$ can be always iteratively computed, so computing the noise strength of any species of interest as well as the joint probability distribution is programmable; (2) our method can reduce computational cost since solving a set of linear algebraic equations for $a_{\mathbf{n}}$ in general consumes less time than solving a stochastic differential equation described by the master equation.

III. APPLICATIONS

In this section, we present several representative examples to show the power of the above series expansion method. These examples include common gene models (i.e., ON-OFF models), multi-state gene models incorporating chromatin remodeling, gene auto-regulatory models, and some common biochemical reaction network modules in signaling systems, etc. Some of them are linear and the others are nonlinear. By applying our method, we can give the exact analytical expressions for computing noise intensities and even probability distributions. In addition, we adopt the famous Gillespie algorithm[37] simulation to verify our theoretical predictions.

A. A two-species reaction network

Consider a simple but generic toy model, which consist of two chemical species X_1 and X_2, where X_1 provides the randomly fluctuating environment for X_2, e.g., mRNA fluctuations randomize protein synthesis. Let n_1 and n_2 be the number of molecules of species X_1 and X_2. The chemical reactions read

$$n_1 \xrightarrow{R_1^+(n_1)} n_1 + 1, \quad n_1 \xrightarrow{R_1^-(n_1)} n_1 - 1,$$

$$n_2 \xrightarrow{R_2^+(n_1,n_2)} n_2 + 1, \quad n_2 \xrightarrow{R_2^-(n_1,n_2)} n_2 - 1.$$

(17)

Because n_1 affects the rate of X_2 but n_2 does not affect the rate of X_1, this is an example of dynamic disorder.

The joint probability of having n_1 and n_2 molecules per cell of species X_1 and X_2 (for instance messenger RNAs and proteins, but interpretations vary with application) at time t is described by a birth-and-death Markov process with events

$$\frac{\partial P(n_1, n_2; t)}{\partial t}$$

$$= R_1^+(n_1 - 1)P(n_1 - 1, n_2; t) + R_1^-(n_1 + 1)P(n_1 + 1, n_2; t)$$

$$+ R_2^+(n_1, n_2 - 1)P(n_1, n_2 - 1; t) + R_2^-(n_1, n_2 + 1)$$

$$\times P(n_1, n_2 + 1; t) - [R_1^+(n_1) + R_1^-(n_1) + R_2^+(n_1, n_2)$$

$$+ R_2^-(n_1, n_2)]P(n_1, n_2; t).$$

(18)

Now, we introduce the generating function $G(z_1, z_2; t) = \sum_{n_1,n_2=0}^{\infty} z_1^{n_1} z_2^{n_2} P(n_1, n_2; t)$ for $P(n_1, n_2; t)$ and consider the static solution to Eq. (18). Then, Eq. (18) can be converted into a partial differential equation for $G(z_1, z_2)$.

First, consider the case studied in Ref. 15, that is, $R_1^+(n_1) = \lambda_1$, $R_1^-(n_1) = n_1/\tau_1$, $R_2^+(n_1, n_2) = \lambda_2 n_1$, $R_2^-(n_1, n_2) = n_2/\tau_2$. In this case, the static generating function satisfies

$$\lambda_1 s_1 F + \left[-\frac{1}{\tau_1} s_1 + \lambda_2 s_2 + \lambda_2 s_1 s_2 \right] \frac{\partial F}{\partial s_1} - \frac{1}{\tau_2} s_2 \frac{\partial F}{\partial s_2} = 0,$$

(19)

where $s_1 = z_1 - 1$, $s_2 = z_2 - 1$, $F(s_1, s_2) \equiv G(z_1, z_2)$. Now, we expand $F(s_1, s_2)$ as $F(s_1, s_2) = \sum_{i,j=0}^{\infty} a_{ij} s_1^i s_2^j$. Note that $a_{00} = 1$ from the conservative condition for $P(n_1, n_2)$. Substituting such an expansion into Eq. (19) and comparing the coefficients of the same powers of s_1 or s_2 will give

$$a_{10} = \lambda_1 \tau_1 a_{00}, \quad a_{01} = \lambda_2 \tau_2 a_{10}, \quad a_{20} = \frac{\lambda_1 \tau_1}{2} a_{10},$$

$$a_{02} = \frac{\lambda_2 \tau_2}{2} \frac{\lambda_2 + \lambda_1 \lambda_2 (\tau_1 + \tau_2)}{1/\tau_1 + 1/\tau_2} a_{10}.$$

Note that the averages of n_1 and n_2 are $\langle n_1 \rangle = a_{10} = \lambda_1 \tau_1$ and $\langle n_2 \rangle = a_{01} = \lambda_1 \lambda_2 \tau_1 \tau_2$, respectively. According to formula (10), we obtain the analytical expression for the n_2 noise intensities

$$\eta_{n_2}^2 = \frac{1}{\langle n_2 \rangle} + \frac{\tau_1}{\tau_1 + \tau_2} \frac{1}{\langle n_1 \rangle},$$

(20)

which is the same as the previous result obtained by LNA.[15]

Then, still consider the above example but set $R_2^+(n_1, n_2) = \lambda_2 n_1(n_1 - 1)$. In this case, we have the following partial differential equation for $F(s_1, s_2)$:

$$\lambda_1 s_1 F - \frac{1}{\tau_1} s_1 \frac{\partial F}{\partial s_1} - \frac{1}{\tau_2} s_2 \frac{\partial F}{\partial s_2}$$

$$+ \gamma \left(s_2 + 2 s_1 s_2 + s_1^2 s_2 \right) \frac{\partial^2 F}{\partial s_1^2} = 0.$$

(21)

Similar to the above case, we can derive the following exact formula for computing the noise intensity of n_2 (denoted by $\tilde{\eta}_{n_2}^2$):

$$\tilde{\eta}_{n_2}^2 = \eta_{n_2}^2 + \frac{2}{(\lambda_1 \tau_1)^2} \frac{1/\tau_2}{2/\tau_1 + 1/\tau_2},$$

(22)

where $\eta_{n_2}^2$ is given by LNA, that is, $\eta_{n_2}^2 = \frac{1}{(\lambda_1 \tau_1)^2 \lambda_2 \tau_2} + \frac{4}{\lambda_1 \tau_1} \frac{1/\tau_2}{1/\tau_1 + 1/\tau_2}$. The inequality $\eta_{n_2}^2 < \tilde{\eta}_{n_2}^2$ shows that LNA
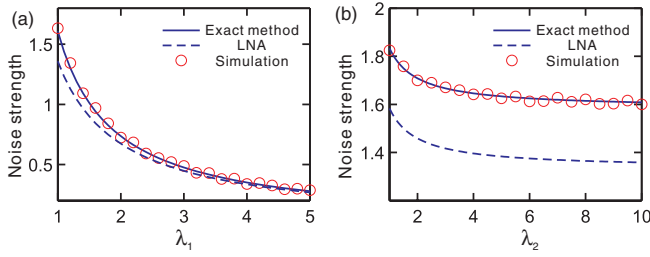
FIG. 1. Shown is that LNA gives approximate noise whereas our method gives exact noise. (a) The noise strength in two-species system (17) with $R_2^+(n_1, n_2) = \lambda_2 n_1(n_1 - 1)$ is taken as a function of the parameter λ_1. The other parameters are $\lambda_2 = 10$, $\tau_1 = 2$, $\tau_2 = 1$. (b) The noise strength in the two-species system (17) with $R_2^+(n_1, n_2) = \lambda_2 n_1(n_1 - 1)$ is taken as a function of the parameter λ_2. The other parameters are $\lambda_1 = 1$, $\tau_1 = 2$, $\tau_2 = 1$.

FIG. 2. Schematic of gene expression models considering promoter activity (active or inactive): (a) two-stage gene model and (b) three-stage gene model.

underestimates the noise. Refer to Fig. 1. This example indicates that LNA is an approximate method.

Since all a_{ij} can be uniquely determined, deriving analytical expression for the steady-state joint probability distribution $P(n_1, n_2)$ is possible according to formula (16) for this system. The detail is here omitted due to complexity.

B. Common gene models

The above two-component generic model may be used to describe gene expression with the promoter being always in the active state, but we know that the promoter of a gene may have both one active and one inactive state or several active states or inactive states, depending on chromatin template that accumulates over time until the promoter becomes active. In this section, to show the effectiveness of our method, we shall consider two common gene models and derive analytical expressions for the mRNA noise and even the analytical distribution for the ON-OFF model. In Sec. III C, we shall consider more general gene models, i.e., gene models of promoter progression, and give analytical results.

Case 1: Two-stage gene model

Consider a stochastic model consisting of a gene that transitions randomly between an active state A where transcription of DNA into mRNA molecules (denoted M) is very efficient, and an inactive state I where transcription is not possible. Refer to Fig. 2(a). Assume that the time during which the gene is in the active state A is the "burst" of mRNA synthesis.[11, 12, 18, 47, 48] The reactions describing this process read

$$I \xrightarrow{\lambda} A, \ A \xrightarrow{\gamma} I,$$
$$A \xrightarrow{\mu} A + M, \ M \xrightarrow{\delta} \varnothing, \tag{23}$$
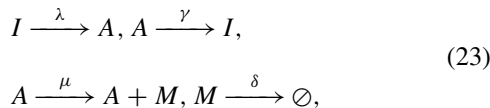
where λ is the rate of gene activation, γ the rate of gene inactivation, μ the rate of transcription when the gene is in the active state, and δ the rate of mRNA decay. We will henceforth use m to denote the number of mRNA molecules.

Let $P_0(m)$ and $P_1(m)$ represent the probability of having m mRNA molecules at time t when the gene is at the ON (active) and OFF (inactive) states, respectively. Then, the cor-

responding discrete master equation reads

$$\frac{dP_0(m; t)}{dt} = -\gamma P_0(m; t) + \lambda P_1(m; t) + \delta(\boldsymbol{E} - \boldsymbol{I})$$
$$\times [m P_0(m; t)] + \mu(\boldsymbol{E}^{-1} - \boldsymbol{I})[P_0(m; t)],$$
$$\tag{24}$$
$$\frac{dP_1(m; t)}{dt} = -\lambda P_1(m; t) + \gamma P_0(m; t) + \delta(\boldsymbol{E} - \boldsymbol{I})[m P_1(m; t)],$$

where $\boldsymbol{I}$ is the identity operator, $\boldsymbol{E}$ and $\boldsymbol{E}^{-1}$ are shift operators, i.e., $\boldsymbol{E}[f(n)] = f(n + 1)$ and $\boldsymbol{E}^{-1}[f(n)] = f(n - 1)$ for any function f and any integer n. By introducing generating functions $G_i(z; t) = \sum_{m=0}^{\infty} z^m P_i(m; t)$ with $i = 0, 1$ and considering the steady-state solution, we can transform Eq. (24) into the following partial differential equations:

$$-\gamma G_0(z) + \lambda G_1(z) + (1 - z)\frac{\partial G_0(z)}{\partial z} + \mu(z - 1)G_0(z) = 0,$$
$$\tag{25}$$
$$-\lambda G_1(z) + \gamma G_0(z) + (1 - z)\frac{\partial G_1(z)}{\partial z} = 0,$$

where all the parameters are normalized by the parameter δ, that is, $\lambda/\delta \to \lambda$, $\gamma/\delta \to \gamma$, and $\mu/\delta \to \mu$.

To solve Eq. (25), we Taylor expand two functions $G_0(z)$ and $G_1(z)$ at the point $z = 1$, that is, $G_i(z) \equiv F_i(s) = \sum_{n=0}^{\infty} a_n^{(i)} s^n$ ($i = 0, 1$) with $s = z - 1$. Let $F = F_0 + F_1$. Then, we have $F(s) = \sum_{n=0}^{\infty} a_n s^n$ with $a_n \equiv a_n^{(0)} + a_n^{(1)}$. Note that the conservative condition for probability gives $a_0(0) = 1$. Substituting these expansions into Eq. (25) will yield the following set of algebraic equations:

$$(n + \gamma)a_n^{(0)} = \lambda a_n^{(1)} + \mu a_{n-1}^{(0)},$$
$$(n + \lambda)a_n^{(1)} = \gamma a_n^{(0)}, \tag{26}$$

where $n = 1, 2, \dots$. It follows from Eq. (26) that

$$a_n = a_n^{(0)} + a_n^{(1)} = \frac{\mu^n}{n!} a_0^{(0)} \prod_{k=1}^{n} \frac{k + \lambda}{k + \lambda + \gamma} = \frac{\mu^n}{n!} \frac{(\lambda)_n}{(\lambda + \gamma)_n}, \tag{27}$$

where $(c)_n$ is the Pochhammer symbol, which is defined as $(c)_n = \Gamma(c + n)/\Gamma(c)$. In particular, we have

$$a_1 = \frac{\mu\lambda}{\lambda + \gamma}, \quad a_2 = \frac{\mu^2\lambda(\lambda + 1)}{2(\lambda + \gamma)(\lambda + \gamma + 1)}. \tag{28}$$

Therefore, according to Eq. (10) and by recovering the original parameters, we obtain the explicit expression for the static noise intensity

$$\eta^2 = \frac{\gamma\delta}{\lambda(\delta + \lambda + \gamma)} + \frac{\delta(\lambda + \gamma)}{\mu\lambda}, \tag{29}$$

which is the same as a previous result.[18] Furthermore, according to Eq. (15) and using Eq. (27) with recovered parameters,

we obtain the analytical expression for the probability distribution of mRNA

$$P(m) = \frac{\mu^m}{m!} \frac{\left(\frac{\lambda}{\delta}\right)_m}{\left(\frac{\lambda}{\delta} + \frac{\gamma}{\delta}\right)_m} {}_1F_1\left(m + \frac{\lambda}{\delta}; m + \frac{\lambda}{\delta} + \frac{\gamma}{\delta}; -\mu\right),$$
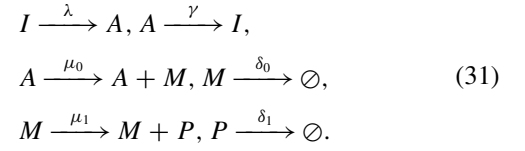$$m = 0, 1, 2, \ldots, \tag{30}$$

where ${}_1F_1(\alpha; \beta; x)$ is a confluent hypergeometric function.[49] Such an expression was before derived under some assumptions of reaction rates[47] but is here derived without any assumption.

Case 2: Three-stage gene model

Consider the full model of a gene, which includes three stages: transitions between active and inactive states, transcription from DNA to mRNA, and translation from mRNA to protein. Refer to Fig. 2(b). The corresponding biochemical reactions read[11, 48, 50–54]

$$I \xrightarrow{\lambda} A, \quad A \xrightarrow{\gamma} I,$$
$$A \xrightarrow{\mu_0} A + M, \quad M \xrightarrow{\delta_0} \oslash, \tag{31}$$
$$M \xrightarrow{\mu_1} M + P, \quad P \xrightarrow{\delta_1} \oslash.$$

Let $P_0(m, n; t)$ be the probability of having m mRNAs and n proteins when the gene is inactive at time t, and $P_1(m, n; t)$ be the probability of having m mRNAs and n proteins when the gene is active at time t. Then, we have two coupled equations

$$\frac{\partial P_0(m, n; \tau)}{\partial \tau} = \kappa_1 P_1(m, n; \tau) - \kappa_0 P_0(m, n; \tau) + (n + 1)P_0(m, n + 1; \tau) - n P_0(m, n; \tau)$$
$$+ \omega \{(m + 1)P_0(m + 1, n; \tau) - m P_0(m, n; \tau) + bm[P_0(m, n - 1; \tau) - P_0(m, n; \tau)]\},$$

$$\frac{\partial P_1(m, n; \tau)}{\partial \tau} = -\kappa_1 P_1(m, n; \tau) + \kappa_0 P_0(m, n; \tau) + (n + 1)P_1(m, n + 1; \tau) - n P_1(m, n; \tau)$$
$$+ \omega\{(m + 1)P_1(m + 1, n; \tau) - m P_1(m, n; \tau) + bm[P_1(m, n - 1; \tau) - P_1(m, n; \tau)]\}$$
$$+ a[P_1(m - 1, n; \tau) - P_1(m, n; \tau)], \tag{32}$$

where $\kappa_0 = \gamma/\delta_1$, $\kappa_1 = \lambda/\delta_1$, $a = \mu_0/\delta_1$, $b = \mu_1/\delta_0$, $\omega = \delta_0/\delta_1$, $\tau = t/\delta_1$. If we consider the steady-state case and introduce two generating functions $G_i(z_1, z_2) = \sum_{m,n} z_1^m z_2^n P_i(m, n)$ $(i = 1, 2)$, then Eq. (32) will become two coupled partial differential equations

$$\kappa_1 F_1 - \kappa_0 F_0 - s_2 \frac{\partial F_0}{\partial s_2} + \omega[bs_2(1 + s_1) - s_1]\frac{\partial F_0}{\partial s_1} = 0,$$
$$-\kappa_1 F_1 + \kappa_0 F_0 - s_2 \frac{\partial F_1}{\partial s_2} + as_1 F_1 + \omega[bs_2(1 + s_1) - s_1]\frac{\partial F_1}{\partial s_1} = 0, \tag{33}$$

where $s_1 = z_1 - 1$, $s_2 = z_2 - 1$ and $F_k(s_1, s_2) \equiv G_k(z_1, z_2)$. Denote $F \equiv F_0 + F_1$. Now, Taylor expand $F_k(s_1, s_2)$ and $F(s_1, s_2)$ at the origin and write $F_k(s_1, s_2) = \sum_{i,j=0}^{\infty} a_{ij}^{(k)} s_1^i s_2^j$ and $F(s_1, s_2) = \sum_{i,j=0}^{\infty} a_{ij} s_1^i s_2^j$. Substituting the expansions of $F_k(s_1, s_2)$ into Eq. (33), and comparing coefficients of variables of the same powers, we obtain

$$\kappa_1 a_{ij}^{(1)} - \kappa_0 a_{ij}^{(0)} - (\omega i + j)a_{ij}^{(0)} + b\omega i\left[a_{(i+1)(j-1)}^{(0)} + a_{i(j-1)}^{(0)}\right] = 0,$$
$$-\kappa_1 a_{ij}^{(1)} + \kappa_0 a_{ij}^{(0)} - (\omega i + j)a_{ij}^{(1)} + b\omega i\left[a_{(i+1)(j-1)}^{(1)} + a_{i(j-1)}^{(1)}\right] + aa_{(i-1)j}^{(0)} = 0, \tag{34}$$

where $i, j = 0, 1, 2, \ldots$. The iteration (34) combined with the conservative condition $a_{00} = 1$ can uniquely determine all $a_{ij}^{(0)}$ and $a_{ij}^{(1)}$. In particular, we have

$$a_{10} = \frac{a\kappa_0}{\omega(\kappa_0 + \kappa_1)}a_{00}, \quad a_{01} = b\omega a_{10}, \quad a_{20} = \frac{a(\omega + \kappa_0)}{2\omega(\kappa_0 + \kappa_1 + 1)}a_{10}$$

and

$$a_{02} = \frac{b^2\omega^2\left[\left(\frac{a}{\omega} + a + 1\right)\kappa_0 + \kappa_1 + 2\omega + a - 1 + \frac{a(\kappa_0 + \omega)(2 - \omega)}{\omega(\kappa_0 + \kappa_1 + 1)}\right]}{2[(\omega + 1)(\kappa_0 + \kappa_1) + 1 - 2\omega]}a_{10}.$$

To that end, according to Eq. (10), we obtain the analytical expressions for the mRNA and protein noise intensities

$$\eta_m^2 = \frac{(\omega - 1)\kappa_0 + \omega\kappa_1}{\kappa_0(\kappa_0 + \kappa_1 + 1)} + \frac{\omega(\kappa_0 + \kappa_1)}{a\kappa_0} \tag{35}$$

and

$$\eta_n^2 = \frac{[a + \omega(a+1)]\kappa_0 + \omega(\kappa_1 + 2\omega + a - 1) + \frac{a(\kappa_0+\omega)(2-\omega)}{\kappa_0+\kappa_1+1}}{a\kappa_0\left(\omega + 1 + \frac{1-2\omega}{\kappa_0+\kappa_1}\right)} + \frac{(1-ab)\kappa_0 + \kappa_1}{ab\kappa_0}, \tag{36}$$

respectively. These formulae for computing noise intensities are not approximate but exact.

We point out that we have not found the nice expression for the joint distribution in the three-stage model as in the two-stage gene model, but our series method can give an approximate distribution up to any pre-given accuracy by truncating initial finite terms of the expansion $F(s_1, s_2) = \sum_{i,j=0}^{\infty} a_{ij} s_1^i s_2^j$. In particular, our method is convenient to numerically computing the distribution.
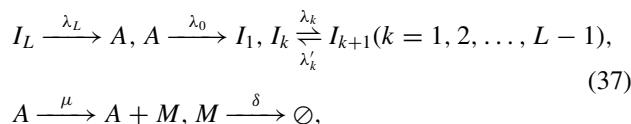
C. Gene models of promoter progression

Except for transcription and translation, gene expression also involves the recruitment of transcription factors and polymerases, transitioning between active and inactive states of promoter, and chromatin remodeling (CR). In particular, it has been verified that CR is an important factor that influences gene expression. In fact, several recent experimental studies have implicated fluctuations in chromatin state between transcriptionally active and inactive conformations as a major source of cell-to-cell variability in gene expression.[55–59] Here, we introduce multi-state stochastic models of gene expression, which extend the previous gene models by incorporating slow dynamics of inactivation or activation. These models assume that the gene activity proceeds sequentially through "ON" states (only here can mature mRNAs be produced) and/or OFF states (here no transcription occurs). Refer to Figs. 3(a) and 3(b), where Fig. 3(a) considers the case that the promoter has one active state and several inactive states (hence the corresponding model is called the multi-OFF model or the multi-OFF mechanism), whereas Fig. 3(b) considers the case that the promoter has one inactive state and several active states (hence the corresponding model is called the multi-ON model or multi-ON mechanism). In these models, the activity states of the promoter form a loop and the transitions between two states may be reversible or irreversible.

Case 1: The multi-OFF mechanism

Assume that the gene activity proceeds sequentially through the ON state, several reversible and irreversible OFF states, and returns to the ON state, forming a loop. Our model considers two main processes: dynamic transitions between active (ON) and inactive (OFF) states of the promoter, denoted by A and I, respectively, and transcription of the DNA sequence into an mRNA molecule (denoted by M). Here, transcription is very efficient in the active state, and is not possible in the inactive state. Moreover, the time spent in the active state is the "burst" of mRNA synthesis.

The corresponding biochemical reactions read

$$I_L \xrightarrow{\lambda_L} A, \; A \xrightarrow{\lambda_0} I_1, \; I_k \underset{\lambda'_k}{\overset{\lambda_k}{\rightleftharpoons}} I_{k+1}(k = 1, 2, \ldots, L-1),$$
$$\tag{37}$$
$$A \xrightarrow{\mu} A + M, \; M \xrightarrow{\delta} \varnothing,$$

where $\lambda_0 \equiv \gamma$ is the rate of gene inactivation, λ_L is the rate of gene activation, λ_k is the transition rate from the kth OFF state to the $(k+1)$th OFF state with $k = 1, 2, \ldots, L-1$, λ'_k is the transition rate from the $(k+1)$th OFF state to the kth OFF state with $k = 1, 2, \ldots, L-1$, μ is the synthesis rate of mRNA, and δ is the degradation rate of mRNA. Note that if $L = 1$, the corresponding model will become our familiar two-state gene model or the telegraph model. In addition, for convenience, we call the multi-OFF process irreversible if all $\lambda'_k = 0$; reversible if all $\lambda'_k \neq 0$; and partially reversible in other cases. Now, we introduce the chemical master equation for mRNA, and henceforth use m to denote the number of mRNAs. The mRNA master equation reads

$$\frac{dP_0(m,t)}{dt} = -\lambda_0 P_0(m,t) + \lambda_L P_L(m,t) + \mu(\boldsymbol{E}^{-1} - \boldsymbol{I})$$
$$\times P_0(m,t) + \delta(\boldsymbol{E} - \boldsymbol{I})[mP_k(m,t)],$$

$$\frac{dP_k(m,t)}{dt} = -(\lambda_k + \lambda'_{k-1})P_k(m,t) + \lambda_{k-1}P_{k-1}(m,t)$$
$$+ \lambda'_k P_{k+1}(m,t) + \delta(\boldsymbol{E} - \boldsymbol{I})[mP_k(m,t)]$$
$$(1 \leq k \leq L-1), \tag{38}$$

$$\frac{dP_L(m,t)}{dt} = -(\lambda_L + \lambda'_{L-1})P_L(m,t) + \lambda_{L-1}P_{L-1}(m,t)$$
$$+ \delta(\boldsymbol{E} - \boldsymbol{I})[mP_L(m,t)],$$

where we define $\lambda'_0 \equiv 0$. In Eq. (36), P_0 represents the probability of having m mRNA in the ON state at time t, and P_k represent the probability of having m mRNAs in the kth OFF state at time t ($1 \leq k \leq L$). The first equation describes how the mRNA is generated in the active state, the subsequent $L - 1$ equations describe the transitions between inactive state of the promoter, and the final equation describes the transitions between the Lth and $(L-1)$th inactive states as well as between the Lth inactive and active states.

For analysis convenience, we consider static solutions only. Introduce generating functions $G_k(z)$ for $P_k(m)$ and $G(z) = \sum_{k=0}^{L} G_k(z)$ for $P(m) = \sum_{k=0}^{L} P_k(m)$. This will lead to the following differential equations:

$$-\lambda_0 F_0(s) + \lambda_L F_L(s) + s F_0(s) - s F'_0(s) = 0,$$
$$-(\lambda_k + \lambda'_{k-1})F_k(s) + \lambda_k F_{k-1}(s) + \lambda'_{k+1}F_{k+1}(s) - s F'_k(s)$$
$$= 0(k = 1, 2, \ldots, L-1), \tag{39}$$
$$-(\lambda_L + \lambda'_{L-1})F_L(s) + \lambda_{L-1}F_{L-1}(s) - s F'_L(s) = 0,$$

where $F_j(s) \equiv G_j(z)$ with $s = \mu(z-1)$ and $0 \leq j \leq L$. Now, we Taylor expand every function $F_j(s)$ by $F_j(s) = \sum_{n=0}^{\infty} b_n^{(j)} s^n$. Note that if we Taylor expand $F(s) = \sum_{n=0}^{\infty} c_n s^n$, then $c_n = \sum_{j=0}^{L} b_n^{(j)}$ with $n \geq 0$. Substituting the expansion of $F_j(s)$

FIG. 3. Multi-state gene expression models and chromatin template-controlled noise in mRNA: (a) Schematic of a multi-inactive-state model of gene expression; (b) schematic of a multi-active-state model of gene expression; (c) the mean noise strength ratio $\eta^2_{multi-on}/\eta^2_{multi-off}$ as a function of two experimentally-measureable indices τ_{on} and τ_{off}, where both $\eta^2_{multi-on}$ and $\eta^2_{multi-off}$ are obtained from one of 1000 sets of randomly sampling $\tau_1, \ldots, \tau_K$ with the fixed τ_{on} and τ_{off} for the fixed $K \equiv L = R = 10$ (where L and R represent the number of OFF states and ON states, respectively). The other parameters are $\mu = 20$, $\delta = 1$, $K = 10$; (d), (e), and (f) shown is the histogram of $\eta^2_{multi-on}$ (red) and $\eta^2_{multi-off}$ (green), where (d), (e) and (f) correspond to points D, E and F labeled in (c), respectively.

into Eq. (39) yields

$$(n\mathbf{I} - \mathbf{A})\mathbf{b}_n = \left(b_{n-1}^{(0)}, 0, \ldots, 0\right)^T, \qquad (40)$$

where the matrix $\mathbf{A} = \begin{pmatrix} -\lambda_0 & & & & \lambda_L \\ \lambda_0 & -\lambda_1 & \lambda_1' & & \\ & \lambda_1 & -(\lambda_2 + \lambda_1') & \ddots & \\ & & \ddots & \ddots & \lambda_{L-1}' \\ & & & \lambda_{L-1} & -(\lambda_L + \lambda_{L-1}') \end{pmatrix}$

describes the pattern of transitions between activity states of the promoter, and $\mathbf{b}_n = (b_n^{(0)}, b_n^{(1)}, \ldots, b_n^{(L)})^T$, $n = 0, 1, 2, \ldots$. Equation (40) is a set of algebraic equations for $b_n^{(k)}$, which can be iteratively solved. In fact, note that A is an M-matrix, so we can express the determinant $\det(n\mathbf{I} - \mathbf{A})$ and the adjacency matrix $(n\mathbf{I} - \mathbf{A})^*$ as $\det(n\mathbf{I} - \mathbf{A}) = n(n + \alpha_1) \cdots (n + \alpha_L)$ and $(n\mathbf{I} - \mathbf{A})^* = \begin{pmatrix} M_1(n) & * \\ * & * \end{pmatrix}$, respectively, where $M_1(n)$ is the polynomial of order L with $M_1(n) = (n + \beta_1) \ldots (n + \beta_L)$, and all $\alpha_k \geq 0$, $\beta_k \geq 0$ are constants depending on the reaction rates. Thus, it follows from Eq. (40) that

$$b_n^{(0)} = \frac{1}{n} \frac{(n + \beta_1) \cdots (n + \beta_L)}{(n + \alpha_1) \cdots (n + \alpha_L)} b_{n-1}^{(0)}$$

$$= \frac{1}{n!} \left[\prod_{k=1}^{n} \frac{(k + \beta_1) \cdots (k + \beta_L)}{(k + \alpha_1) \cdots (k + \alpha_L)} \right] b_0^{(0)}. \qquad (41)$$

In addition, it follows from Eq. (39) that $F(s) = F'_0(s)$, implying that $c_n = \frac{1}{n} b_{n-1}^{(0)}$ for any n. In other word, we can give the expressions of all c_n. Also, we can show $b_0^{(0)} = \frac{\alpha_1 \cdots \alpha_L}{\beta_1 \cdots \beta_L}$ according to the conservative condition $F(0) = 1$ that implies $c_0 = 1$. To that end, according to Eq. (15), we obtain the analytical expression of the steady-state probability distribution

$$P(m) = \frac{\mu^m}{m!} \frac{(\beta_1)_m \cdots (\beta_L)_m}{(\alpha_1)_m \cdots (\alpha_L)_m} {}_L F_L(m + \beta_1, \ldots, m + \beta_L; m$$

$$+ \alpha_1, \ldots, m + \alpha_L; -\mu). \qquad (42)$$

In particular, we can give the explicit expression of c_1 and c_2. For example, in the irreversible case, i.e., all $\lambda'_k = 0$, we have

$$c_1 = \sum_{j=0}^{L} b_1^{(j)} = \frac{\prod_{k=1}^{L} \lambda_k}{\prod_{k=1}^{L} \lambda_k + \gamma \sum_{k=1}^{L-1} \left[\prod_{i=1}^{L-k-1} \lambda_i \cdot \prod_{j=1}^{k} \lambda_{L-j+1} \right] + \gamma \left(\prod_{k=1}^{L-1} \lambda_k \right)},$$

(43)

$$c_2 = \sum_{j=0}^{L} b_2^{(j)} = \frac{1}{2} \frac{\sum_{k=1}^{L-1} \left[\prod_{i=1}^{L-k-1} \lambda_i \cdot \prod_{j=1}^{k} (1 + \lambda_{L-j-1}) \right] + \prod_{k=1}^{L-1} \lambda_k + \prod_{k=1}^{L} \lambda_k}{\prod_{k=1}^{L} (1 + \lambda_k) + \gamma \sum_{k=1}^{L-1} \left[\prod_{i=1}^{L-k-1} \lambda_i \cdot \prod_{j=1}^{k} (1 + \lambda_{L-j+1}) \right] + \gamma \left(\prod_{k=1}^{L-1} \lambda_k \right)},$$

where if $k = L - 1$, then we define $\sum_{i=1}^{L-k-1} \lambda_i = 1$. According to $G'(z)|_{z=1} = \mu F'(s)|_{s=0} = \mu c_1$, $G''(z)|_{z=1} = \mu^2 F''(s)|_{s=0} = 2\mu^2 c_2$ and formula (10), we obtain the analytical expression for the square of the noise intensity in the irreversible case

$$\eta_m^2 = \frac{1}{\langle m \rangle} + \frac{1 + (\tau_{off} - 1) \prod_{k=1}^{L} (1 + \tau_k)}{(1 + \tau_{on}) \prod_{k=1}^{L} (1 + \tau_k) - 1}, \quad (44)$$

where $\langle m \rangle = \frac{b}{\tau_{on} + \tau_{off}} = \frac{\mu \tau_{on}}{\tau_{on} + \tau_{off}}$, $\tau_{on} = 1/\gamma$, $\tau_{off} = \sum_{k=1}^{L} (1/\lambda_k)$, $b = \mu/\gamma$, and $\tau_k = 1/\lambda_k$. Recall that τ_{on} and τ_{off} represent the total ON and OFF time, respectively, and b represents the mean burst size that characterizes the burst statistics. Note that $L = 1$ corresponds to the common ON-OFF model of gene expression. Also note that for fixed τ_{on} and τ_{off}, the noise intensity η_m monotonically decreases with regard to the number of OFF states (L), implying that an important fact that the multi-OFF mechanism plays a role of attenuating the noise. Such a qualitative conclusion still holds in a general case including reversible and mixing cases.

Except that η_m is monotonic with regard to L, we numerically find that if τ_{on} is dominant, then the noise level in the case of the multi-OFF mechanism is always lower than that in the case of the multi-ON mechanism, and in particular, if τ_{on} is equal to τ_{off}, then the former is equal to the latter. Refer to Fig. 3(c). For clarity, we also demonstrate probabilities of the mRNA number corresponding to three representative points labeled in Fig. 3(c), referring to Figs. 3(d)–3(f). These numerical results indicate that the ratio of τ_{off} over τ_{on} has important influences on the mRNA noise in the case of the multi-OFF mechanism.

Case 2: The multi-ON mechanism

Assume that the gene activity proceeds sequentially through the OFF state, several irreversible ON states (below, only the irreversible case is considered for simplicity but other cases such as partially irreversible or reversible may be similarly treated), and returns to the OFF state, forming a loop. In this case, the corresponding biochemical reactions read

$$A_k \xrightarrow{\gamma_k} A_{k+1} \ (k = 1, 2, \ldots, R-1), \ A_R \xrightarrow{\gamma_R} I, \ I \xrightarrow{\lambda} A_1,$$

(45)

$$A_k \xrightarrow{\mu} A_k + M \ (k = 1, 2, \ldots, R), \ M \xrightarrow{\delta} \emptyset,$$
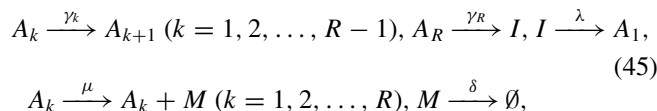
where the meaning of all the symbols is the same as in Case 1. The chemical master equation corresponding to this set of reactions can be expressed as

$$\frac{d P_0(m, t)}{dt} = -\gamma_0 P_0(m, t) + \gamma_R P_R(m, t) + \delta(\mathbf{E} - \mathbf{I})[m P_0(m, t)],$$

$$\frac{d P_k(m, t)}{dt} = -\gamma_k P_k(m, t) + \gamma_{k-1} P_{k-1}(m, t) + \delta(\mathbf{E} - \mathbf{I})[m P_k(m, t)] + \mu(\mathbf{E}^{-1} - \mathbf{I})[P_k(m, t)]$$

$$(k = 1, 2, \ldots, R)$$

(46)

with $\gamma_0 \equiv \lambda$. Introduce generating functions $G_i(z, t) = \sum_{m=0}^{\infty} z^m P_i(m, t)$ $(i = 0, 1, \ldots, R)$ and new variable $s \equiv \mu(z - 1)$, and consider the static solution. Then,

$$-\gamma_0 G_0(s) + \gamma_R G_R(s) - s G_0'(s) = 0,$$

$$-\gamma_1 G_1(s) + \gamma_0 G_0(s) - s G_1'(s) + s G_1(s) = 0, \quad (47)$$

$$-\gamma_k G_k(s) + \gamma_{k-1} G_{k-1}(s) - s G_k'(s) + s G_k(s)$$

$$= 0 (k = 2, \ldots, R),$$

where all the parameters including μ are rescaled by δ. Furthermore, to solve Eq. (47), we introduce the transformation $F_i(s) = e^{-s} G_i(s)$. Then, Eq. (47) becomes

$$-\gamma_0 F_0(s) + \gamma_R F_R(s) - s \frac{d}{ds} F_0(s) - s F_0(s) = 0,$$

$$-\gamma_1 F_1(s) + \gamma_0 F_0(s) - s \frac{d}{ds} F_1(s) = 0, \quad (48)$$

$$-\gamma_k F_k(s) + \gamma_{k-1} F_{k-1}(s) - s \frac{d}{ds} F_k(s) = 0 (k = 2, \ldots, R).$$

Note that Eq. (48) has the form of Eq. (39). Therefore, we can give the expression of the sum function $F \equiv F_0 + F_1 + \ldots + F_R$. Furthermore, we can obtain the analytical expression for the steady-state probability distribution of mRNA. That is,

$$P(m) = \frac{\mu^m e^{-\mu}}{m!} \sum_{k=0}^{m} \binom{m}{k} (-1)^k \frac{(\beta_1)_k \cdots (\beta_R)_k}{(\alpha_1)_k \cdots (\alpha_R)_k} {}_R$$

$$\times F_R(k + \beta_1, \ldots, k + \beta_R; k + \alpha_1, \ldots, k + \alpha_R; \mu),$$

(49)

where all α_i and β_i are constants depending only on the reaction rates.

In particular, we can obtain the analytical mean and variance of mRNA in the irreversible

$$\langle \tilde{m} \rangle = \frac{\mu \tilde{\tau}_{on}}{\tilde{\tau}_{on} + \tilde{\tau}_{off}}, \quad \sigma_{\tilde{m}}^2 = \langle \tilde{m} \rangle + \langle \tilde{m} \rangle^2 \left(\frac{\tilde{\tau}_{off}}{\tilde{\tau}_{on}} \right)^2$$

$$\times \frac{1 + (\tilde{\tau}_{on} - 1) \prod_{k=1}^{R} (1 + \tilde{\tau}_k)}{(\tilde{\tau}_{off} + 1) \prod_{k=1}^{R} (1 + \tilde{\tau}_k) - 1}, \tag{50}$$

where $\tilde{\tau}_{off} = 1/\lambda$, $\tilde{\tau}_{on} = \sum_{k=1}^{L} (1/\gamma_k)$, and $\tilde{\tau}_k = 1/\gamma_k$. Furthermore, according to formula (10), we obtain the analytical expression for the square of the noise intensity in the irreversible case

$$\eta_{\tilde{m}}^2 = \frac{1}{\langle \tilde{m} \rangle} + \left(\frac{\tilde{\tau}_{off}}{\tilde{\tau}_{on}} \right)^2 \frac{\tilde{\tau}_{on} - 1 + \prod_{k=1}^{R} \frac{1}{1 + \tilde{\tau}_k}}{\tilde{\tau}_{off} + 1 - \prod_{k=1}^{R} \frac{1}{1 + \tilde{\tau}_k}}, \tag{51}$$

which is a monotonically decreasing function of R if $\tilde{\tau}_{on}$, $\tilde{\tau}_{off}$, and $\langle \tilde{m} \rangle$ are fixed. This implies that the multi-ON mechanism also always reduces noise like the multi-OFF mechanism. In particular, when $R = 1$ that corresponds to the common ON-OFF model (denote by η_{on-off} the noise intensity), $\eta_{\tilde{m}}^2$ reaches the maximum. Moreover, $\eta_{on-off}^2 = 1/\langle m \rangle + \tau_{off}^2 / [(\tau_{off} + 1)(\tau_{on} + 1) - 1]$. In addition, when $\tilde{\tau}_1 = \tilde{\tau}_2 = \cdots = \tilde{\tau}_R = \tilde{\tau}_{on}/R$, $\eta_{\tilde{m}}^2$ reaches its minimum. Moreover,

we have

$$\eta_{\tilde{m}}^2 \geq \frac{1}{\langle \tilde{m} \rangle} + \left(\frac{\tilde{\tau}_{off}}{\tilde{\tau}_{on}} \right)^2 \frac{\tilde{\tau}_{on} - 1 + \left(1 + \frac{\tilde{\tau}_{on}}{R}\right)^{-R}}{\tilde{\tau}_{off} + 1 - \left(1 + \frac{\tilde{\tau}_{on}}{R}\right)^{-R}} > \frac{1}{\langle \tilde{m} \rangle}$$

$$+ \left(\frac{\tilde{\tau}_{off}}{\tilde{\tau}_{on}} \right)^2 \frac{\tilde{\tau}_{on} - 1 + e^{-\tilde{\tau}_{on}}}{\tilde{\tau}_{off} + 1 - e^{-\tilde{\tau}_{on}}}. \tag{52}$$

Some numerical results are shown in Figs. 3(c)–3(f). We observe that if τ_{on} is dominant over τ_{off}, then the noise level in the case of the multi-ON mechanism is always higher than that in the case of the multi-OFF mechanism, and in particular, if $\tau_{on} = \tau_{off}$, then the mean noise level is the same in two cases, independently from the detailed transitions between activity states of the promoter. Refer to Fig. 3(c). For clarity, we also demonstrate the histograms of the mRNA number corresponding to three representative points labeled in Fig. 3(c), referring to Figs. 3(d)–3(f). These numerical results indicate that the ratio of τ_{on} over τ_{off} has also important influences on the mRNA noise in the case of the multi-ON mechanism as the case of the multi-OFF mechanism.

In addition, combining analysis of the above two cases, we obtain low boundaries of the mRNA noise intensity in the common ON-OFF model as

$$\eta_{on-off}^2 \geq \max \left\{ \frac{1}{\langle m \rangle} + \frac{\tau_{off} - 1 + \prod_{k=1}^{L} \frac{1}{1 + \tau_k}}{\tau_{on} + 1 - \prod_{k=1}^{L} \frac{1}{1 + \tau_k}}, \frac{1}{\langle \tilde{m} \rangle} + \left(\frac{\tilde{\tau}_{off}}{\tilde{\tau}_{on}} \right)^2 \frac{\tilde{\tau}_{on} - 1 + \prod_{k=1}^{R} \frac{1}{1 + \tilde{\tau}_k}}{\tilde{\tau}_{off} + 1 - \prod_{k=1}^{R} \frac{1}{1 + \tilde{\tau}_k}} \right\}$$

$$> \min \left\{ \frac{1}{\langle m \rangle} + \frac{\tau_{off} - 1 + e^{-\tau_{off}}}{\tau_{on} + 1 - e^{-\tau_{off}}}, \frac{1}{\langle \tilde{m} \rangle} + \left(\frac{\tilde{\tau}_{off}}{\tilde{\tau}_{on}} \right)^2 \frac{\tilde{\tau}_{on} - 1 + e^{-\tilde{\tau}_{on}}}{\tilde{\tau}_{off} + 1 - e^{-\tilde{\tau}_{on}}} \right\}. \tag{53}$$

Case 3: Other mechanisms

In general, the pattern of transitions between active and inactive states of the promoter is complex, in particular in eukaryotic cells. Here, we analyze an example only, which consider one inactive state and three active states of the gene promoter but there exist transitions between inactive and active states or between active states, referring to Fig. 4. The corresponding master equations are described as follows:

$$\frac{\partial P_0}{\partial t} = \lambda_1 P_1 + \lambda_2 P_2 + \lambda_3 P_3 - (\gamma_1 + \gamma_2 + \gamma_3)P_0 + \delta(\boldsymbol{E} - \boldsymbol{I})[m P_0(m, t)],$$

$$\frac{\partial P_1}{\partial t} = \gamma_1 P_0 + \lambda_4' P_2 + \lambda_5' P_3 - (\lambda_1 + \lambda_4 + \lambda_5)P_1 + \delta(\boldsymbol{E} - \boldsymbol{I})[m P_1(m, t)] + \mu(\boldsymbol{E}^{-1} - \boldsymbol{I})[P_1(m, t)],$$

$$\frac{\partial P_2}{\partial t} = \gamma_2 P_0 + \lambda_4 P_1 + \lambda_6' P_3 - (\lambda_2 + \lambda_4' + \lambda_6)P_2 + \delta(\boldsymbol{E} - \boldsymbol{I})[m P_2(m, t)] + \mu(\boldsymbol{E}^{-1} - \boldsymbol{I})[P_2(m, t)],$$

$$\frac{\partial P_3}{\partial t} = \gamma_3 P_0 + \lambda_5 P_1 + \lambda_6 P_2 - (\lambda_3 + \lambda_5' + \lambda_6')P_2 + \delta(\boldsymbol{E} - \boldsymbol{I})[m P_3(m, t)] + \mu(\boldsymbol{E}^{-1} - \boldsymbol{I})[P_3(m, t)].$$
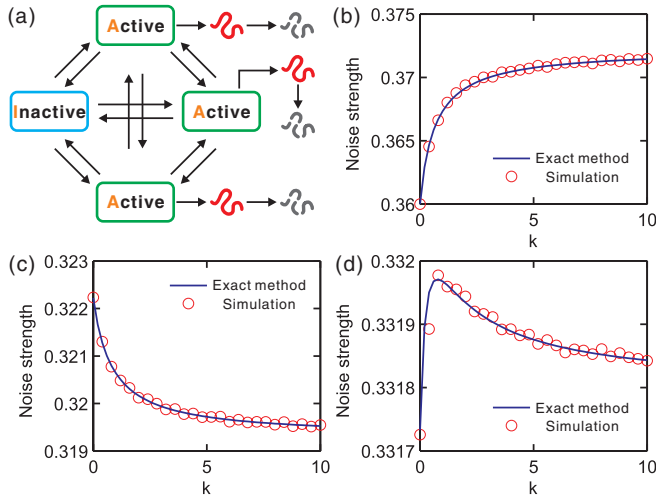
$$\tag{54}$$

FIG. 4. The effect of multi-state mechanism on the noise in gene expression. (a) Schematic of multi-inactive-active-state model of gene expression. (b)–(d) The noise strength as a function of the parameter $k \equiv \lambda_4 = \lambda'_4 = \lambda_5 = \lambda'_5 = \lambda_6 = \lambda'_6$. The other parameter values used in computation are $\lambda_1 = 2$, $\lambda_2 = 1$, $\lambda_3 = 1$, $\gamma_2 = 1$, $\gamma_3 = 1$, $\mu = 5$ and (b) $\gamma_1 = 1$, (c) $\gamma_1 = 2$, (d) $\gamma_1 = 1.7$.

To solve Eq. (54), we introduce generating functions $G_i(z, t) = \sum_{m=0}^{\infty} z^m P_i(m, t)$ ($i = 0, 1, 2, 3$) and new variable $s \equiv \mu(z - 1)$, and consider the steady-state case. This will lead to a following set of coupled ordinary differential equations:

$$\lambda_1 G_1 + \lambda_2 G_2 + \lambda_3 G_3 - (\gamma_1 + \gamma_2 + \gamma_3)G_0 - s\frac{dG_0}{ds} = 0,$$

$$\gamma_1 G_0 + \lambda'_4 G_2 + \lambda'_5 G_3 - (\lambda_1 + \lambda_4 + \lambda_5)G_1$$
$$- s\frac{dG_1}{ds} + sG_1 = 0,$$

$$\gamma_2 G_0 + \lambda_4 G_1 + \lambda'_6 G_3 - (\lambda_2 + \lambda'_4 + \lambda_6)G_2 - s\frac{dG_2}{ds}$$
$$+ sG_2(s) = 0,$$

(55)

$$\gamma_3 G_0 + \lambda_5 G_1 + \lambda_6 G_2 - (\lambda_3 + \lambda'_5 + \lambda'_6)G_3 - s\frac{dG_3}{ds}$$
$$+ sG_3(s) = 0,$$

where all the parameters including μ are rescaled by δ. Furthermore, for Eq. (55), we introduce the transformation $F_i(s) = e^{-s}G_i(s)$, which transform Eq. (55) into the following:

$$\lambda_1 F_1 + \lambda_2 F_2 + \lambda_3 F_3 - (\gamma_1 + \gamma_2 + \gamma_3)F_0 - s\frac{dF_0}{ds}$$
$$- sF_0 = 0,$$

$$\gamma_1 F_0 + \lambda'_4 F_2 + \lambda'_5 F_3 - (\lambda_1 + \lambda_4 + \lambda_5)F_1 - s\frac{dF_1}{ds} = 0,$$

(56)

$$\gamma_2 F_0 + \lambda_4 F_1 + \lambda'_6 F_3 - (\lambda_2 + \lambda'_4 + \lambda_6)F_2 - s\frac{dF_2}{ds} = 0,$$

$$\gamma_3 F_0 + \lambda_5 F_1 + \lambda_6 F_2 - (\lambda_3 + \lambda'_5 + \lambda'_6)F_3 - s\frac{dF_3}{ds} = 0.$$

Summing the above four equations gives $F_0(s) = -F'(s)$, where $F \equiv F_0 + F_1 + F_2 + F_3$. Now, we Taylor expand every function $F_j(s)$ by $F_j(s) = \sum_{n=0}^{\infty} b_n^{(j)} s^n$. Note that if we Taylor expand $F(s) = \sum_{n=0}^{\infty} c_n s^n$, then $c_n = \sum_{j=0}^{3} b_n^{(j)}$ with $n \geq 0$. Substituting such expansions into Eq. (56) yields

$$\lambda_1 b_n^{(1)} + \lambda_2 b_n^{(2)} + \lambda_3 b_n^{(3)} - (n + \gamma_1 + \gamma_2 + \gamma_3)b_n^{(0)} = b_{n-1}^{(0)},$$

$$\gamma_1 b_n^{(0)} + \lambda'_4 b_n^{(2)} + \lambda'_5 b_n^{(3)} - (n + \lambda_1 + \lambda_4 + \lambda_5)b_n^{(1)} = 0,$$

(57)

$$\gamma_2 b_n^{(0)} + \lambda_4 b_n^{(1)} + \lambda'_6 b_n^{(3)} - (n + \lambda_2 + \lambda'_4 + \lambda_6)b_n^{(2)} = 0,$$

$$\gamma_3 b_n^{(0)} + \lambda_5 b_n^{(1)} + \lambda_6 b_n^{(2)} - (n + \lambda_3 + \lambda'_5 + \lambda'_6)b_n^{(3)} = 0,$$

where $b_{-1}^{(0)} = 0$, $n = 0, 1, 2$, and $c_0 = b_0^{(0)} + b_0^{(1)} + b_0^{(2)} + b_0^{(3)} = 1$ due to $F(0) = 1$. For convenience, introduce the following matrixes:

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ \gamma_1 & -(\lambda_1 + \lambda_4 + \lambda_5) & \lambda'_4 & \lambda'_5 \\ \gamma_2 & \lambda_4 & -(\lambda_2 + \lambda'_4 + \lambda_6) & \lambda'_6 \\ \gamma_3 & \lambda_5 & \lambda_6 & -(\lambda_3 + \lambda'_5 + \lambda'_6) \end{pmatrix},$$

$$\mathbf{B} = \begin{pmatrix} -(\gamma_1 + \gamma_2 + \gamma_3 + 1) & \lambda_1 & \lambda_2 & \lambda_3 \\ \gamma_1 & -(\lambda_1 + \lambda_4 + \lambda_5 + 1) & \lambda'_4 & \lambda'_5 \\ \gamma_2 & \lambda_4 & -(\lambda_2 + \lambda'_4 + \lambda_6 + 1) & \lambda'_6 \\ \gamma_3 & \lambda_5 & \lambda_6 & -(\lambda_3 + \lambda'_5 + \lambda'_6 + 1) \end{pmatrix}.$$

Then, according to $F_0(s) = -F'(s)$, we know

$$c_1 = -b_0^{(0)} = -\frac{\det(\mathbf{A}[2, 3, 4; 2, 3, 4])}{\det(\mathbf{A})}.$$

(58)

Using Eq. (57) in the case of $n = 1$ and again according to $F_0(s) = -F'(s)$, we obtain

$$c_2 = -\frac{1}{2}b_1^{(0)} = -\frac{b_0^{(0)}}{2}\frac{\det(\mathbf{B}[2, 3, 4; 2, 3, 4])}{\det(\mathbf{B})}.$$

(59)

Note that

$$G'(z)\big|_{z=1} = \mu G'(s)\big|_{s=0} = \mu(c_0 + c_1) = \mu(1 + c_1),$$

$$G''(z)\big|_{z=1} = \mu^2 G''(s)\big|_{s=0} = \mu^2(c_0 + 2c_1 + 2c_2)$$
$$= \mu^2(1 + 2c_1 + 2c_2). \tag{60}$$

Thus, we obtain the analytical expression of noise intensity

$$\eta_m^2 = \frac{\sigma_m^2}{\langle m \rangle^2} = \frac{G''(z)\big|_{z=1} + G'(z)\big|_{z=1} - \left[G'(z)\big|_{z=1}\right]^2}{\left[G'(z)\big|_{z=1}\right]^2}$$

$$= \frac{1}{\langle m \rangle} + \frac{2c_2 - c_1^2}{(1 + c_1)^2}, \tag{61}$$

where c_1 and c_2 are given by Eqs. (58) and (59), respectively. We point out that Eq. (61) can reproduce the result obtained in the above multi-on mechanism with $R = 3$ if some rate constants related to transitions between active and inactive states are set as zero. Especially, we find that if $\lambda_1 = \lambda_2 = \lambda_3 \equiv \lambda$, then $\langle m \rangle = \frac{\mu(\gamma_1 + \gamma_2 + \gamma_3)}{(\gamma_1 + \gamma_2 + \gamma_3 + \lambda)}$ add $\eta_m^2 = \frac{1}{\langle m \rangle}$ $+ \frac{\lambda}{(\gamma_1 + \gamma_2 + \gamma_3)(\gamma_1 + \gamma_2 + \gamma_3 + \lambda)}$. This indicates that the mean and the noise level are independent of the reversible transitions between active states, i.e., they do not rely on parameters λ_4', λ_5', λ_6'. In addition, the analytical distribution can be also given using Eq. (15) and computing Eq. (57), and can be expressed by a confluent hypergeometric function. The detail is omitted here.

Now, we perform numerical simulations. For clarity, we consider a particular case: $k \triangleq \lambda_4 = \lambda_4' = \lambda_5 = \lambda_5' = \lambda_6 = \lambda_6'$. In this case, Figs. 4(c)–4(e) show three representative patterns of the noise intensity square (η_m^2) vs k, where Fig. 4(c) shows that η_m^2 monotonically increases with regard to k with fixed $\gamma_1 = 1$, Fig. 4(d) shows that η_m^2 monotonically decreases with regard to k with fixed $\gamma_1 = 2$, and Fig. 4(e) shows that η_m^2 has one maximum with regard to k with fixed $\gamma_1 = 1.7$.

This example indicates that transition patterns between activity states of the promoter are a factor that can influence the noise level.

D. Gene auto-regulatory models

In this subsection, we consider two common gene auto-regulatory models: the auto-activation model and the auto-repression model, which are nonlinear biochemical networks, and derive analytical expressions for gene expression noise. Regarding self-regulatory systems, there have been some analytical results, e.g., seeing Refs. 60 and 61.

Case 1: Auto-activation

This model is described by the following biochemical reactions with the protein switching the gene from the OFF-state to the ON-state or vice versa

$$I \xrightarrow{\lambda} A, A \xrightarrow{\gamma} I, I + P \xrightarrow{a} A + P,$$
$$A \xrightarrow{\mu} A + P, P \xrightarrow{\delta} \varnothing, \tag{62}$$
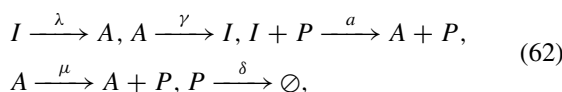
Denote the protein probability distributions in the I and A states as $P_0(m; t)$ and $P_1(m; t)$, respectively. Then, the cor-

responding master equations read[62]

$$\frac{dP_0(m;t)}{dt} = -\lambda P_0(m;t) + \gamma P_1(m;t) - am P_0(m;t)$$
$$+ \delta\left[(m+1)P_0(m+1;t) - mP_0(m;t)\right], \tag{63}$$

$$\frac{dP_1(m;t)}{dt} = \lambda P_0(m;t) - \gamma P_1(m;t) + am P_0(m;t)$$
$$+ \delta\left[(m+1)P_1(m+1;t) - mP_1(m;t)\right]$$
$$+ \mu\left[P_1(m-1;t) - P_1(m;t)\right].$$

Define generating functions $G_k = \sum_{m=0}^{\infty} z^m P_k$ and consider static solutions. Then, it follows from Eq. (63) that

$$-\lambda G_0 + \gamma G_1 - az\frac{dG_0}{dz} - \delta(z-1)\frac{dG_0}{dz} = 0,$$
$$\tag{64}$$

$$\lambda G_0 - \gamma G_1 + az\frac{dG_0}{dz} - \delta(z-1)\frac{dG_1}{dz} + \mu(z-1)G_1 = 0.$$

Let $G = G_0 + G_1$. Then, we have

$$\frac{\delta(a+\delta)}{\mu}\left[z - \frac{\delta}{a+\delta}\right]G''$$

$$- \left[(a+\delta)\left(z - \frac{\delta}{a+\delta}\right) - \frac{(\lambda+\gamma)\delta}{\mu}\right]G' - \lambda G = 0,$$
$$\tag{65}$$

where G satisfies the normalization condition $G(1) = 1$ due to the conservative condition $\sum_{m=0}^{\infty} P(m) = 1$.

If we Taylor expand $G(z)$ at THE POINT $z = 1$, then the coefficients in the corresponding series cannot be analytically given. However, if we Taylor expand $G(z)$ at THE POINT $z = \delta/(a+\delta)$ due to the characteristic of the coefficient in the second order derivative of $G(z)$ in Eq. (65), then the corresponding coefficients can be explicitly given. In fact, substituting the expansion

$$G(z) = \sum_{n=0}^{\infty} b_n \left(z - \frac{\delta}{a+\delta}\right)^n$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} G^{(n)}(z)\big|_{z=\delta/(a+\delta)} \left(z - \frac{\delta}{a+\delta}\right)^n \tag{66}$$

into Eq. (65) and comparing the coefficients of various same powers of $z - \delta/(a+\delta)$ will yield

$$b_n = \frac{a_0}{n!}\left(\frac{\mu}{\delta}\right)^n \frac{(\lambda/(a+\delta))_n}{((\lambda+\gamma)/(a+\delta))_n}, \quad n = 1, 2, 3, \ldots. \tag{67}$$

Thus, we obtain the analytical expression for generating function $G(z)$

$$G(z) = \sum_{n=0}^{\infty} \frac{b_0}{n!}\left(\frac{\mu}{\delta}\right)^n \frac{(\lambda/(a+\delta))_n}{((\lambda+\gamma)/(a+\delta))_n}\left(z - \frac{\delta}{a+\delta}\right)^n$$

$$= b_0 {}_1F_1\left(\frac{\lambda}{a+\delta}; \frac{\lambda+\gamma}{a+\delta}; \frac{\mu}{\delta}z - \frac{\mu}{a+\delta}\right), \tag{68}$$

where the condition $G(1) = 1$ gives $b_0 = 1/{}_1F_1\left(\frac{\lambda}{a+\delta}; \frac{\lambda+\gamma}{a+\delta}; \frac{a\mu}{\delta(a+\delta)}\right)$. Furthermore, we obtain the
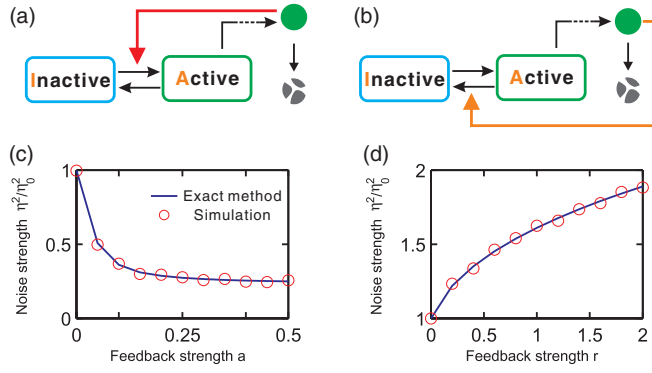
FIG. 5. The effect of feedback mechanisms on the noise in gene expression. (a) Schematic of auto-activation model of gene expression. (b) Schematic of auto-repression model of gene expression. (c) The noise strength as a function of the positive feedback strength (a) for system (62), where the parameters are $\lambda = \gamma = 1$, $\mu = 5$, $\delta = 0.1$. (d) The noise strength as a function of the negative feedback strength (r) for system (71), where the parameters are $\lambda = \gamma = 1$, $\mu = 5$, $\delta = 0.1$.

FIG. 6. (a) Schematic of the enzymatic futile cycle reaction mechanism. (b) The noise strength as a function of the total number N_T, showing that there is an extreme, where the parameters are $k_1 = 1$, $k_{-1} = 1$, $k_2 = 10$, $k_{-2} = 0.01$.

analytical expression for the steady-state distribution

$$P(m) = \frac{1}{m!}\left(\frac{\mu}{\delta}\right)^m \frac{\left(\frac{\lambda}{a+\delta}\right)_m}{\left(\frac{\lambda+\gamma}{a+\delta}\right)_m} \frac{{}_1F_1\left(m+\frac{\lambda}{a+\delta};m+\frac{\lambda+\gamma}{a+\delta};-\frac{\mu}{a+\delta}\right)}{{}_1F_1\left(\frac{\lambda}{a+\delta};\frac{\lambda+\gamma}{a+\delta};\frac{a\mu}{\delta(a+\delta)}\right)}.$$
(69)

In addition, we can give the analytical formula for mean and variance according to $\langle m \rangle = G'(1)$ and $\sigma_m^2 = G''(1) + G'(1) - [G'(1)]^2$. Moreover, the formula can show that when $a \to 0$, the analytical expression for the noise intensity without feedback is

$$\lim_{a\to 0} \eta^2(a) = \frac{\gamma\delta}{\lambda(\delta+\lambda+\gamma)} + \frac{\delta(\lambda+\gamma)}{\mu\lambda},$$
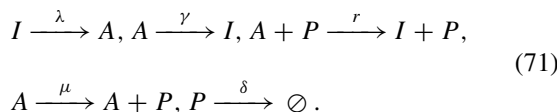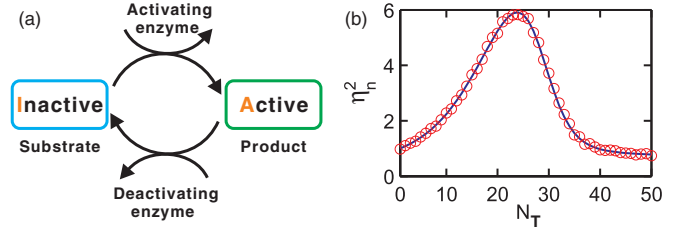(70)

which is the same as the known results.[18] In addition, by computation we know

$$\lim_{a\to\infty} \eta^2(a) = \frac{\gamma+\lambda e^{\mu/\delta}}{(\lambda+\gamma)e^{\mu/\delta}}\frac{(\mu+\delta)(\lambda+\gamma)}{\mu\lambda}$$
$$-1 < \frac{\left[\frac{\mu(\lambda+\delta)}{(\lambda+\gamma+\delta)}+\delta\right](\lambda+\gamma)}{\mu\lambda} - 1 = \lim_{a\to 0}\eta^2(a).$$

Using the property on derivatives of confluent hypergeometric functions, we can show that the noise intensity computed by $\eta_m = \sigma_m/\langle m \rangle$ is not a monotonically increasing but monotonically decreasing function in a (referring to Fig. 5(c)). This indicates that the traditional conclusion that positive feedback enlarges noise[15] is not correct for the circuit described by biochemical reactions (60).

Case 2: Auto-repression

Consider the following biochemical reactions, which describe the auto-repression model of gene expression:

$$I \xrightarrow{\lambda} A, A \xrightarrow{\gamma} I, A + P \xrightarrow{r} I + P,$$
(71)
$$A \xrightarrow{\mu} A + P, P \xrightarrow{\delta} \oslash.$$

Let $P_0(m; t)$ and $P_1(m; t)$ be the probability density function of having m proteins at the t time in the I and A states, respec-

tively. Then, the corresponding master equations read

$$\frac{dP_0(m;t)}{dt} = -\lambda P_0(m;t) + \gamma P_1(m;t) + rm P_1(m;t)$$
$$+ \delta\left[(m+1)P_0(m+1;t) - mP_0(m;t)\right],$$
(72)

$$\frac{dP_1(m;t)}{dt} = \lambda P_0(m;t) - \gamma P_1(m;t) - rm P_1(m;t)$$
$$+ \delta\left[(m+1)P_1(m+1;t) - mP_1(m;t)\right]$$
$$+ \mu\left[P_1(m-1;t) - P_1(m;t)\right],$$

Introduce two generating functions $G_k = \sum_{m=0}^{\infty} z^m P_k$ and the total generating function $G = G_0 + G_1$, and consider the static solutions to Eq. (72). Similar to the case of auto activation, we can obtain the analytical expression for $G(z)$, that is

$$G(z) = \frac{{}_1F_1\left(\frac{\lambda}{\delta};\frac{\lambda+\gamma}{r+\delta}+\frac{r\mu}{(r+\delta)^2};\frac{\mu}{r+\delta}z-\frac{\mu\delta}{(r+\delta)^2}\right)}{{}_1F_1\left(\frac{\lambda}{\delta};\frac{\lambda+\gamma}{r+\delta}+\frac{r\mu}{(r+\delta)^2};\frac{\mu r}{(r+\delta)^2}\right)}.$$
(73)

Furthermore, we can obtain the analytical expression for the steady-state probability distribution

$$P(m) = \frac{1}{m!}\left(\frac{\mu}{r+\delta}\right)^m \frac{\left(\frac{\lambda}{\delta}\right)_m}{\left(\frac{\lambda+\gamma}{r+\delta}+\frac{r\mu}{(r+\delta)^2}\right)_m}$$
$$\times \frac{{}_1F_1\left(m+\frac{\lambda}{\delta};m+\frac{\lambda+\gamma}{r+\delta}+\frac{r\mu}{(r+\delta)^2};-\frac{\mu\delta}{(r+\delta)^2}\right)}{{}_1F_1\left(\frac{\lambda}{\delta};\frac{\lambda+\gamma}{r+\delta}+\frac{r\mu}{(r+\delta)^2};\frac{\mu r}{(r+\delta)^2}\right)}.$$
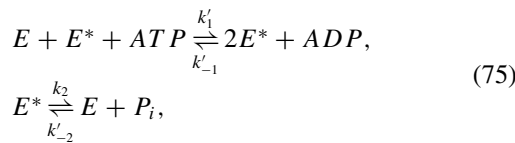(74)

In addition, according to the formulae $\langle m \rangle = G'(1)$ and $\sigma_m^2 = G''(1) + G'(1) - [G'(1)]^2$, we can compute the mean and variance and further the noise intensity by $\eta_m = \sigma_m/\langle m \rangle$. Note that if $r \to 0$, then we obtain the same analytical expression for the noise intensity as Eq. (70) in the case of no feedback.

Again using the property on derivatives of hypergeometric functions, we can show that the noise intensity computed by $\eta_m = \sigma_m/\langle m \rangle$ is not a monotonically decreasing but monotonically increasing function in r (referring to Fig. 5(d)). This indicates that the traditional conclusion that negative feedback reduces noise[15] is not correct for the circuit described by biochemical reactions (71).

Our numerical results show that feedback has a counterintuitive effect (referring to Figs. 5(c) and 5(d)). The similar conclusion was also obtained in Ref. 6, wherein more numerical results were demonstrated.

E. A signaling system with autocatalytic kinase

Here we consider a simple enzymatic futile cycle, which is schematically shown in Fig. 6(a). This mechanism is ubiquitous throughout biological systems and is encountered as a recurrent control motif in such diverse regulatory processes as metabolism, GTPase cycles, mitogen-activated protein kinase (MAPK) cascades, glucose mobilization, cell division/apoptosis, checkpoint control, actin tread-milling, and membrane transport as well as two-component systems and phosphorelays in microbial stress-response signaling pathways (refer to Refs. 63 and 64 and references therein). All the above enzymatic futile cycle examples can be summarized into the following simple biochemical kinetic system

$$E + E^* + ATP \underset{k'_{-1}}{\overset{k'_1}{\rightleftharpoons}} 2E^* + ADP,$$
$$E^* \underset{k'_{-2}}{\overset{k_2}{\rightleftharpoons}} E + P_i, \tag{75}$$

where the reaction scheme for PdPC with autocatalytic kinase depicts the molecular elements E, E^*, and ATP combining in the forward reaction at a rate of k'_1 to produce $2E^*$ and ADP. The backward reaction occurs at a rate of k'_{-1} and is autocatalytic because E^* serves as a catalyst for itself. The second reaction can be interpreted in a similar manner.

To avoid cluttering, we let ATP, ADP, and Pi be absorbed into the rate constants. Thus, the above reactions can be rewritten as

$$E + E^* \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} 2E^*,$$
$$E^* \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} E, \tag{76}$$
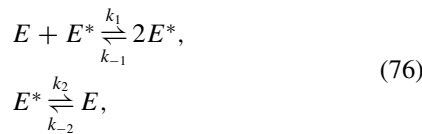
which satisfy the conservative condition $[E] + [E^*] = E_T$ (the total concentration, which is assumed as a constant). Let n represent the number of E^* molecules, N_T represent the total number of kinase molecules (a constant), and $P(n; t)$ represent the probability of having n molecules at time t. Then, the corresponding master equation reads

$$\frac{\partial P(n,t)}{\partial t} = -[k_1 n(N_T - n) + k_{-1} n(n-1) + k_2 n$$
$$+ k_{-2}(N_T - n)]P(n,t) + (N_T - n + 1)[k_1(n-1)$$
$$+ k_{-2}]P(n-1,t) + (n+1)[k_{-1}n + k_2]$$
$$\times P(n+1,t). \tag{77}$$

By introducing the generating function $G(z;t) = \sum_{n=0}^{N_T} z^n P(n;t)$ for $P(n; t)$, we have a partial differential equation for $G(z; t)$

$$\frac{\partial G}{\partial t} = N_T(-k_1 + k_{-2}z)G + \{k_2 + [k_1(N_T - 1) + k_2 - k_{-2}]z$$
$$+ [k_1(N_T - 1) - k_{-2}]z^2\}\frac{\partial G}{\partial z}$$
$$+ [k_{-1}z + (k_{-1} - k_1)z^2 + k_1 z^3]\frac{\partial^2 G}{\partial z^2}. \tag{78}$$

Consider the static solution to Eq. (78). Now, we Taylor expand $G(z)$ as $G(z) = \sum_{n=0}^{N_T} a_n(z - 1)^n$. Substituting this expansion into the above equation and comparing the coefficients of $(z - 1)$ of the zero-order and the first-order powers lead to

$$a_1 = C \sum_{n=1}^{N_T} \frac{1}{(n-1)!} \prod_{j=0}^{n-1} \frac{(k_1 j + k_{-2})(N_T - j)}{(k_{-1}j + k_2)},$$

$$a_2 = \frac{C}{2} \sum_{n=2}^{N_T} \frac{1}{(n-2)!} \prod_{j=0}^{n-1} \frac{(k_1 j + k_{-2})(N_T - j)}{(k_{-1}j + k_2)},$$

where C is a normalized constant, satisfying $C \sum_{n=1}^{N_T} \prod_{j=0}^{n-1} \frac{(k_1 j + k_{-2})(N_T - j)}{(k_{-1}j + k_2)(j+1)} = 1$. According to the formula (10), we obtain the analytical expression for the noise intensity

$$\eta_n^2 = \frac{2a_2 + a_1}{a_1^2} - 1$$

$$= \frac{C}{a_1^2} \sum_{n=1}^{N_T} n(2n-1) \prod_{j=0}^{n-1} \frac{(k_1 j + k_{-2})(N_T - j)}{(k_{-1}j + k_2)(j+1)} - 1. \tag{79}$$

In addition, we can obtain the following analytical steady-state distribution

$$p_{ss}(n) = C \prod_{j=0}^{n-1} \frac{(k_1 j + k_{-2})(N_T - j)}{(k_{-1}j + k_2)(j+1)}. \tag{80}$$

The similar expression was also derived in Ref. 64.

Interestingly, we numerically find that there exists an optimal total number of kinase molecules, N_T, such that the noise reaches the highest level, referring to Fig. 6(b). In addition, we find that the total number of kinase molecules has important influences on the generation of bistabiliy. More results for this system will be published elsewhere.

IV. CONCLUSION AND DISCUSSION

Molecular noise in biochemical systems or networks is inevitable due to the low copy numbers of species molecules participating in the reactions. Here, we have proposed a general yet exact method for computing the noise and probability function, which can be applied to arbitrarily linear or nonlinear biochemical reaction networks. By examples, we have shown the strong power of such a power series expansion method, which is particularly effective for linear biochemical systems since a_n in the related series expansion of the generating function $G(z)$ is iteratively computed and hence can be analytically given. As such, the noise strength for any species of interest and the joint probability can be explicitly given. For nonlinear biochemical networks, a_n is also iteratively computed in all cases, and can be analytically given in some cases and cannot in other cases, but can be approximately given and the approximate results can reach any pre-given accuracy. In addition, we have given exact analytical formulae for the noise intensities in several common biochemical networks by

applying this method. We have shown that LNA is effective only in linear network cases but would overestimate or underestimate the noise in nonlinear network cases. Moreover, positive (negative) feedback does not always enlarge (reduce) noise but has a counter effect, implying that the previous conclusion on the effect of feedback on noise needs to be modified. In addition, interesting analytical distributions for two common nonlinear biochemical modules have also been given for the first time.

We point out that our method is linear, and hence has many advantages in contrast to the following several existing methods: (1) The moment-closure method[44–46, 62] is in general nonlinear since low-order moments depend often on higher-order moments, implying that nonlinear approximations are required. Moreover, how to set an approximate scheme does not have a general rule but depends on specific biochemical networks, and the common setting of approximate schemes do not guarantee that a pre-given accuracy is satisfied since high-order moments do not tend to zero as its order goes to the infinity. Our method can overcome these shortcomings mainly because our method is linear and the related series is convergent but how to choose points at which power series are expanded is worth being further studied. (2) LNA[6, 24–29] can only be used to obtain low-order approximation to noise, while our method can be used to numerically compute arbitrarily high-order approximations for noise and probability distributions and the resulting approximations can reach arbitrary accuracies. (3) Although linear, the variational method proposed in Ref. 32 would meet difficulties in handling high-dimensional systems and only give approximate results. In contrast, our method, which is also linear, can be applied to an arbitrarily dimensional linear or nonlinear biochemical network for which we need only to solve a set of linear algebraic equations.

Our method can also be used to study noise propagation in signaling systems where analytical solutions to the corresponding master equations may not exist, and can give exact rather than approximate results for the propagating noise in the signaling systems. Moreover, the corresponding method is programmable since our analytical formulae for computing noise strengths depend only on the first and second order derivatives of generating functions at a certain point on the argument. Our method can even be used to study noise propagation in signaling networks in terms of probability distributions.

Owing to the exact formulae available to computing noise, our next task is to elucidate roles and functions of noise in controlling stochastic fluctuations of biochemical network modules, e.g., gene auto-regulatory motifs, feedback forward loops, *cis*-regulatory modules, etc. As such, we can reckon that some previous results on noise obtained by LNA would require to be modified. In addition, by applying the power series expansion method proposed here, it is possible to derive analytical probability distributions described by biochemical master equations in more general biochemical reaction networks. This is significant since a distribution in general contains more accurate random information than noise intensity, a commonly used index.

[1] M. B. Elowitz and S. Leibler, Nature (London) **403**, 335 (2000).

[2] J. Paulsson, O. G. Berg, and M. Ehrenberg, Proc. Natl. Acad. Sci. U.S.A. **97**, 7148 (2000).

[3] A. N. Boettiger and M. Levine, Science **325**, 471 (2009).

[4] A. Raj *et al.*, Nature (London) **463**, 913 (2010).

[5] A. Eldar and M. B. Elowitz, Nature (London) **467**, 167 (2010).

[6] J. M. Pedraza and J. Paulsson, Science **319**, 339 (2008).

[7] A. Sanchez and J. Kondev, Proc. Natl. Acad. Sci. U.S.A. **105**, 5081 (2008).

[8] A. Sanchez *et al.*, PLOS Comput. Biol. **7**, e1001100 (2011).

[9] O. G. Berg, J. Theor. Biol. **71**, 587 (1978).

[10] J. Peccoud and B. Ycart, Theor. Popul. Biol. **48**, 222 (1995).

[11] N. Friedman, L. Cai, and X. S. Xie, Phys. Rev. Lett. **97**, 168302 (2006).

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

[13] T. Jia and R. V. Kulkarni, Phys. Rev. Lett. **106**, 058102 (2011).

[14] H. H. McAdams and A. Arkin, Proc. Natl. Acad. Sci. U.S.A. **94**, 814 (1997).

[15] J. Paulsson, Nature (London) **427**, 415 (2004).

[16] J. Paulsson, Phy. Life Rev. **2**, 157 (2005).

[17] M. Thattai and A. van Oudenaarden, Proc. Natl. Acad. Sci. U.S.A. **98**, 8614 (2001).

[18] T. B. Kepler and T. C. Elston, Biophys. J. **81**, 3116 (2001).

[19] E. M. Ozbudak *et al.*, Nat. Genet. **31**, 69 (2002).

[20] J. M. Raser and E. K. O'Shea, Science **309**, 2010 (2005).

[21] M. Dobrzynski and F. J. Bruggeman, Proc. Natl. Acad. Sci. U.S.A. **106**, 2583 (2009).

[22] R. Rinott, A. Jaimovich, and N. Friedman, Proc. Natl. Acad. Sci. U.S.A. **108**, 6329 (2011).

[23] D. Huh and J. Paulsson, Nat. Genet. **43**, 95 (2011).

[24] G. Hornung, and N. Barkai, PLOS Comput. Biol. **4**, e8 (2008).

[25] S. Hooshangi, S. Thiberge, and R. Weiss, Proc. Natl. Acad. Sci. U.S.A. **102**, 3581 (2005).

[26] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, Amsterdam, 2007).

[27] M. Scott, T. Hwa, and B. Ingalls, Proc. Natl. Acad. Sci. U.S.A. **104**, 7402 (2007).

[28] J. Elf and M. Ehrenberg, Genome Res. **13**, 2475 (2003).

[29] J. J. Zhang, Z. J. Yuan, and T. Zhou, Phys. Biol. **6**, 046009 (2009).

[30] P. B. Warren, S. Tanase-Nicola, and P. R. ten Wolde, J. Chem. Phys. **125**, 144904 (2006).

[31] Y. H. Lan and G. A. Papoian, J. Chem. Phys. **125**, 154901 (2006).

[32] Y. H. Lan, P. G. Wolynes, and G. A. Papoian, J. Chem. Phys. **125**, 124106 (2006).

[33] Y. H. Lan and G. A. Papoian, Phys. Rev. Lett. **98**, 228301 (2007).

[34] S. Tanase-Nicola, P. B. Warren, and P. R. ten Wolde, Phys. Rev. Lett. **97**, 068102 (2006).

[35] A. M. Walczak, M. Sasai, and P. G. Wolynes, Biophys. J. **88**, 828 (2005).

[36] A. B. Bortz, M. H. Kalos, and J. L. Lebowitz, J. Comput. Phys. **17**, 10 (1975).

[37] D. T. Gillespie, J. Phys. Chem. **81**, 2340 (1977).

[38] D. T. Gillespie, Annu. Rev. Phys. Chem. **58**, 35 (2007).

[39] A. M. Walczak, A. Mugler, and C. H. Wiggins, Proc. Natl. Acad. Sci. U.S.A. **106**, 6529 (2009).

[40] J. S. van Zon *et al.*, Biophys. J. **91**, 4350 (2006).

[41] J. S. van Zon and P. R. ten Wolde, J. Chem. Phys. **123**, 234910 (2005).

[42] J. J. Zhang, L. N. Chen, and T. S. Zhou, Biophys. J. **102**, 1247 (2012).

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

[44] R. Grima, J. Chem. Phys. **136**, 154105 (2012).

[45] C. Zechner *et al.*, Proc. Natl. Acad. Sci. U.S.A. **109**, 8340 (2012).

[46] C. S. Gillespie, Syst. Biol. **3**, 52–58 (2009).

[47] A. Raj *et al.*, PLoS Biol. **4**, e309 (2006).

[48] J. Paulsson and M. Ehrenberg, Phys. Rev. Lett. **84**, 5447 (2000).

[49] L. J. Slater, *Confluent Hypergeometric Functions* (Cambridge University Press, Cambridge, 1960).

[50]R. Karmakar and I. Bose, Phys. Biol. **1**, 197 (2004).

[51]S. Iyer-Biswas, F. Hayot, and C. Jayaprakash, Phys. Rev. E **79**, 031911 (2009).

[52]A. Mugler, A. M. Walczak, and C. H. Wiggins, Phys. Rev. E **80**, 041921 (2009).

[53]M. C. Mackey, M. Tyran-Kaminska, and R. Yvinec, J. Theor. Biol. **274**, 84 (2011).

[54]M. Assaf, E. Roberts, and Z. Luthey-Schulten, Phys. Rev. Lett. **106**, 248102 (2011).

[55]D. M. Suter *et al.*, Science **332**, 472 (2011).

[56]C. V. Harper *et al.*, PLoS Biol. **9**, e1000607 (2011).

[57]C. Mao *et al.*, Mol. Syst. Biol. **6**, 431 (2010).

[58]L. Mariani *et al.*, Mol. Syst. Biol. **6**, 359 (2010).

[59]K. Miller-Jensen *et al.*, Trends Biotechnol. **29**, 517 (2011).

[60]J. E. M. Hornos *et al.*, Phys. Rev. E. **72**, 051907 (2005).

[61]H. D. Feng, B. Han, and J. Wang, J. Phys. Chem. B **115**, 1254 (2011).

[62]A. Singh and J. P. Hespanha, Biophys. J. **96**, 4013 (2009).

[63]M. Samoilov, S. Plyasunov, and A. P. Arkin, Proc. Natl. Acad. Sci. U.S.A. **102**, 2310 (2005).

[64]L. M. Bishop and H. Qian, Biophys. J. **98**, 1 (2010).