Exact distributions for stochastic gene expression models with arbitrary promoter architecture and translational bursting

Meiling Chen,[1,2,*] Songhao Luo,[1,2,*] Mengfang Cao,[1,2] Chengjun Guo,[3] Tianshou Zhou,[1,2] and Jiajun Zhang[1,2,†]

[1]*Guangdong Province Key Laboratory of Computational Science, Guangzhou 510275, People's Republic of China*

[2]*School of Mathematics, Sun Yat-Sen University, Guangzhou 510275, People's Republic of China*

[3]*School of Mathematics and Statistics, Guangdong University of Technology, Guangzhou 510275, People's Republic of China*

(Received 12 October 2021; revised 10 December 2021; accepted 14 December 2021; published 6 January 2022)

Gene expression in individual cells is inherently variable and sporadic, leading to cell-to-cell variability in mRNA and protein levels. Recent single-cell and single-molecule experiments indicate that promoter architecture and translational bursting play significant roles in controlling gene expression noise and generating the phenotypic diversity that life exhibits. To quantitatively understand the impact of these factors, it is essential to construct an accurate mathematical description of stochastic gene expression and find the exact analytical results, which is a formidable task. Here, we develop a stochastic model of bursty gene expression, which considers the complex promoter architecture governing the variability in mRNA expression and a general distribution characterizing translational burst. We derive the analytical expression for the corresponding protein steady-state distribution and all moment statistics of protein counts. We show that the total protein noise can be decomposed into three parts: the low-copy noise of protein due to probabilistic individual birth and death events, the noise due to stochastic switching between promoter states, and the noise resulting from translational busting. The theoretical results derived provide quantitative insights into the biochemical mechanisms of stochastic gene expression.

DOI: 10.1103/PhysRevE.105.014405

I. INTRODUCTION

Gene expression is a fundamentally programmed and stochastic process. Recent single-cell and single-molecule experiments have demonstrated that gene expression often occurs in a "burst" manner: genes appear to be transcribed or translated during short periods interspersed by silent intervals [1,2]. At the DNA level, transcriptional bursts have been documented among organisms ranging from bacteria to vertebrates [3–12]. Simultaneously, single-molecule studies have demonstrated that individual mRNA's translation output is either sporadic or bursty, called "translational bursts" [13–18]. For example, mRNAs in primary neurons can rapidly switch between a translating state in proximal dendrites and a nontranslating state in distal dendrites, displaying "bursting" translation [17]. These bursting variabilities can be propagated to protein and the downstream target gene, which provides critical functions in cell fate decisions [19–21]. Given its importance, understanding how mRNA and protein bursting influence the variabilities of protein abundance is a critical challenge.

In the past decade, much effort has been invested in theoretical modeling and analysis for characterizing the burst size, burst frequency, and the variabilities of mRNA and protein abundance [22–24]. In a Markovian framework, the gene expression processes such as promoter states switching, transcription, translation, degradation of mRNA or protein are modeled as explicit states-switching with the chemical master equation [25–32]. Alternatively, in a non-Markovian framework, the complex gene expression processes are mapped implicitly into queueing theory models [33–39] or continuous-time random walk models [40] by waiting time distributions and burst size distributions. And the noise in protein abundance can be decomposed into different sources, e.g., molecular memory and bursting [33,34,40].

Despite some progress, many questions remain as to how promoter architecture and translation-bursting kinetics dictate cell-to-cell variability in gene expression. On the one hand, promoter architecture, defined by the number, strength, and regulatory role of the operators governing a gene, maybe contain many states that complicate the model [41]. For example, the number of regulatory states of the Promoter for Repressor Maintenance (PRM) promoter of phage lambda in E. coli may be up to 128 [42]. On the other hand, translation is a complex biochemical process involving multiple factors, such as *cis*-elements encoded in mRNAs, post-transcriptional modifications of mRNAs, kinases, and other signaling molecules [14]. For example, the complex multiprotein regulation leads to sharp protein bursts in contrast to the geometric distributions [18].

To that end, we develop a theoretical model of bursty gene expression which considers the promoter architecture of arbitrary combinatorial complexity governing the variability in mRNA expression and a general distribution characterizing translation burst. With the help of the binomial moments method that we previously developed [43,44], we derive an-

*These authors contributed equally to this work.

†Corresponding author: zhjiajun@mail.sysu.edu.cn

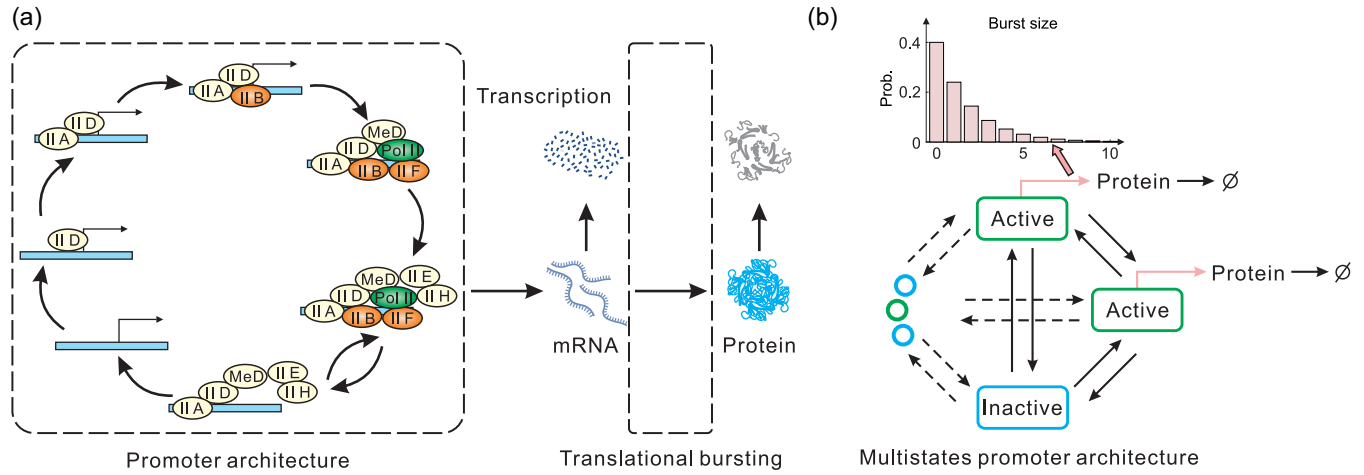

FIG. 1. Gene expression model with arbitrary promoter architecture and general translational bursting. (a) Generalized gene expression process: the promoter switches among different states, transcription occurs during active promoter states, translation from mRNA in bursting manner, mRNA decays faster than protein. (b) Simplified sketch of protein production with a multistate model. The upper figure shows an example of the distribution of protein bursts.

alytical results of all moment statistics of protein counts. The total protein noise can be decomposed into three parts: the low-copy noise of protein due to probabilistic individual birth and death events, the noise due to stochastic switching between promoter states, and the noise resulting from translational bursting. Furthermore, we study the effect of the type of translational burst distribution on the bimodality of protein distribution. Finally, a well-predicted outcome for the protein moment statistics and distribution is verified from real data [45]. The proposed approaches and results derived will help interpret the experiments and probe gene expression bursting and its phenotypic consequences.

II. MODEL FORMULATION

We model the stochastic gene expression processes based on a master equation approach [46]. The transcriptional regulation is regarded as a stochastic process where the promoter transitions between different states [Fig. 1(b)]. The left box in Fig. 1(a) shows an example where the promoter comprises several inactive (OFF) states and one active (ON) state that form a loop. In addition, the translational regulation is modeled as a stochastic bursting process in which each mRNA generates a burst of proteins, whose number is arbitrarily distributed, such as geometric distribution [top inset in Fig. 1(b)]. In our model, the transcription and translation processes can be integrated into one single-step process with translational bursting [Fig. 1(b)]. In fact, mRNA degradation is much faster than that of protein, which ensures that protein production can be approximately done in one step [26,47,48].

To make this idea precise, we assume that there are N promoter states, each of which can produce proteins with a translation rate $\mu_i (i = 1, \ldots, N)$ in a bursting manner. The probability distribution of a burst with k proteins is $h(k)(k = 0, 1, 2, \ldots)$ which can be an arbitrary discrete probability distribution. The protein degradation rate is denoted by δ. We denote the number of protein molecules as m and let $\mathbf{P}_i(m;t)$ denote the probability distribution that protein has m molecules at time t at state i of the promoter and let $\mathbf{P}(m;t) = [P_1(m;t), \ldots, P_N(m;t)]^{\mathrm{T}}$ represent the column vector. Denote by λ_{ij} the transition rate from the state j to state i ($\lambda_{ij} = 0$ means that no transition occurs), the size of which may be regulated by transcription factors, and the diagonal elements are $\lambda_{jj} = -\sum_{i \neq j} \lambda_{ij} (j = 1, \ldots, N)$, which represent the sum of the rates flowing in from other states in the current state j. Denote by $\mathbf{\Lambda} = (\lambda_{ij})$ the $N \times N$ transition matrix and $\mathbf{\Lambda} = \mathrm{diag}(\mu_1, \ldots, \mu_N)$ describes exits of translation (called translation matrix) with μ_i representing the translation rate in the state i ($\mu_i = 0$ means that no translation takes place). Note that matrix A is an M matrix. Protein decays with a first-order rate constant δ which is set to 1. Then, the biochemical master equation describing protein dynamics takes the matrix form as

$$\frac{d\mathbf{P}(m;t)}{dt} = \mathbf{A}\mathbf{P}(m;t) + (\mathbf{E} - \mathbf{I})[m\mathbf{P}(m;t)] + \mathbf{\Lambda}\left[\sum_{k=0}^{m} h(k)\mathbf{P}(m-k;t) - \mathbf{P}(m;t)\right], \quad (1)$$

where $\mathbf{E}$ and $\mathbf{E}^{-1}$ are shift operators and $\mathbf{I}$ is the identity operator. The first term in this equation is the rate at which the gene states transition through promoter state switching, and the second term is the net rate at which the gene enters and goes out of the state $\mathbf{P}(m)$ through the degradation of one protein. The last term is the net rate at which the genes enter and go out of the state $\mathbf{P}(m)$ through the production of a burst of proteins distributed by $h(k)$. Note that Eq. (1) represents a set of homogeneous linear ordinary differential equations for the evolution with time of the probability distribution. But analytically and numerically solving these linear equations becomes infeasible because the number of equations in Eq. (1) proliferates exponentially with the number of promoter states. In the next section, we will analyze Eq. (1) using the binomial moment approach.

III. BINOMIAL MOMENTS

The binomial moment approach is a popular moment closure method, with which the moments above a cer-

tain order are expressed in terms of lower-order moments. Thereby the closed moment equations can then be solved either analytically or numerically [43,44]. We set $a_n^{(i)}(t) = \sum_{m\geqslant n} \binom{m}{n} P_i(m,t) (n = 0, 1, 2, \ldots)$, therefore each fixed i, $a_n^{(i)}(t)$ is called a binomial moment corresponding to the probability $P_i(m;t)$; let $\mathbf{a}_n(t) = [a_n^{(1)}(t), \ldots, a_n^{(N)}(t)]^{\mathrm{T}}$ represent the column vector. In particular, $b_n^{\mathrm{protein}}(t) = \sum_{i=1}^{N} a_n^{(i)}(t)$ is called the total binomial moment corresponding to the total probability $P(m,t) = \sum_{i=1}^{N} P_i(m,t)$. The binomial moments $\mathbf{a}_n(t) (n = 1, 2, \ldots)$ satisfy the following equation (see Appendix A for details):

$$\frac{d\mathbf{a}_n(t)}{dt} = \mathbf{A}\mathbf{a}_n(t) - n\mathbf{a}_n(t) + \mathbf{\Lambda}\left[\sum_{i=1}^{n} \mathbf{a}_{n-i}(t) \sum_{k=i}^{\infty} \binom{k}{i} h(k)\right]. \tag{2}$$

Similarly, we define $b_i^{\mathrm{burst}} = \sum_{k=i}^{\infty} \binom{k}{i} h(k)$ as binomial moments corresponding to the probability $h(k)$ of burst size B, with $\langle B\rangle = b_1^{\mathrm{burst}}$ representing the average burst size. Furthermore, the convolution formula is introduced, that is $\mathbf{a}_n(t) * b^{\mathrm{burst}}[n] = \sum_{i=1}^{n} \mathbf{a}_{n-i}(t) b_i^{\mathrm{burst}}$. Thus, Eq. (2) can be rewritten as

$$\frac{d\mathbf{a}_n(t)}{dt} = \mathbf{A}\mathbf{a}_n(t) - n\mathbf{a}_n(t) + \mathbf{\Lambda}\mathbf{a}_n(t) * b^{\mathrm{burst}}[n]. \tag{3}$$

Note that Eq. (3) enables the computation of binomial moments up to any desired order without any approximation because the higher-order moments depend on only the lower-order ones. At steady state, we can immediately know that the steady-state binomial moments $\mathbf{a}_n$ satisfy the following iterative equations

$$\mathbf{a}_n = (n\mathbf{I} - \mathbf{A})^{-1}\mathbf{\Lambda}\mathbf{a}_n * b^{\mathrm{burst}}[n]. \tag{4}$$

Because $\mathbf{A}$ is an M matrix, $b_n^{\mathrm{protein}} = \sum_{i=1}^{N} a_n^{(i)}$ can be rewritten as

$$b_n^{\mathrm{protein}} = \frac{1}{n}\mathbf{u}_N\mathbf{\Lambda}\mathbf{a}_n * b^{\mathrm{burst}}[n], \tag{5}$$

where $\mathbf{u}_N = (1, 1, \ldots, 1), n = 1, 2, \ldots$. Combing Eqs. (4) and (5) yields all binomial moments of protein (b_n^{protein}) iteratively. We can obtain each order raw moment and center moment with Eqs. (4) and (5) in hand. In the next section, we will compute several main statistics: mean, noise intensity, skewness, and kurtosis of protein abundance.

IV. MEAN EXPRESSION AND PROTEIN NOISE

First, we compute the $\mathbf{a}_0$. We let $\mathbf{M}_i$ denote an $(N-1) \times (N-1)$ matrix, which is the minor one of the $N \times N$ matrix $\mathbf{A}$ by crossing out the ith row and ith column of its entry a_{ii}. Denote by $0, -\alpha_1, -\alpha_2, \ldots, -\alpha_{N-1}$ the eigenvalues of matrix $\mathbf{A}$ and by $-\beta_1^{(i)}, -\beta_2^{(i)}, \ldots, -\beta_{N-1}^{(i)}$ the eigenvalues of matrix $\mathbf{M}_i$ [28]. The kth component of $\mathbf{a}_0$ is given by

$$a_0^{(i)} = \prod_{j=1}^{N-1} \frac{\beta_j^{(i)}}{\alpha_j}, \quad 1 \leqslant i \leqslant N. \tag{6}$$

Substituting the expression of $\mathbf{a}_0$ into Eqs. (4) and (5), we obtain the following equations for the mean ($\langle$protein$\rangle$) and noise strength [$\eta^2_{\mathrm{protein}} = \mathrm{Var(Protein)}/\langle\mathrm{Protein}\rangle^2$] of the protein probability distribution in steady state (see Appendix B for details),

$$\langle\mathrm{protein}\rangle = \mathbf{u}_N\mathbf{\Lambda}\mathbf{a}_0\langle B\rangle, \tag{7}$$

$$\eta^2_{\mathrm{protein}} = \underbrace{\frac{1}{\langle\mathrm{Protein}\rangle}}_{\text{Low copy noise}} + \underbrace{\frac{\mathbf{u}_N\mathbf{\Lambda}(\mathbf{I}-\mathbf{A})^{-1}\mathbf{\Lambda}\mathbf{a}_0 - (\mathbf{u}_N\mathbf{\Lambda}\mathbf{a}_0)^2}{(\mathbf{u}_N\mathbf{\Lambda}\mathbf{a}_0)^2}}_{\text{Promoter noise}} + \underbrace{\frac{1}{2}\frac{1}{\langle\mathrm{Protein}\rangle}\left(\langle B\rangle\eta^2_{\mathrm{burst}} + \langle B\rangle - 1\right)}_{\text{Translational burst noise}}, \tag{8}$$

where

$$\eta^2_{\mathrm{burst}} = \frac{\mathrm{Var}(B)}{\langle B\rangle^2} = \frac{1}{b_1^{\mathrm{burst}}} + \frac{2b_2^{\mathrm{burst}}}{\left(b_1^{\mathrm{burst}}\right)^2} - 1$$

is the noise strength of burst size.

We emphasize here that Eqs. (7) and (8) are exact for arbitrary promoter architecture and arbitrary translational bursting dynamics. Equation (7) establishes that the mean steady-state protein level only depends on the promoter architecture and average translational burst size. Equation (8) highlights the different contributions to the protein noise. The first noise term on the right side of Eq. (8) represents the low copy noise of protein due to probabilistic individual birth and death events. If protein abundance $\langle$protein$\rangle$ is low, relative protein levels spontaneously fluctuate, then it has a more significant relative effect on the total. The second term describes the promoter noise, which results from stochastic switching between promoter states. The last term captures the translational burst kinetics, the function of mean burst size and burst noise. Interestingly, Eq. (8) is identical to the previous result apart from the second term corresponding to the promoter noise [33,34]. Our results can capture the promoter noise with an explicit promoter architecture instead of an implicit waiting time distribution.

Next, we consider a promoter architecture with only one active state, as an example, which is a widely studied stochastic gene expression [23]. Assume that the promoter transition matrix is arbitrary, but the transcription matrix is diagonal with only one nonzero element (i.e., only one active state). Without loss of generality, we set $\mathbf{\Lambda} = \mathrm{diag}(0, \ldots, 0, \mu)$. After some algebra (see Appendix C for details), we obtain the protein noise,

$$\eta^2_{\mathrm{protein}} = \frac{1}{\langle\mathrm{Protein}\rangle} + \frac{\prod_{i=1}^{N-1}\left(1+\beta_i^{(N)}\right)\alpha_i - \prod_{i=1}^{N-1}(1+\alpha_i)\beta_i^{(N)}}{\prod_{i=1}^{N-1}(1+\alpha_i)\beta_i^{(N)}} + \frac{1}{2}\frac{1}{\langle\mathrm{Protein}\rangle}\left(\langle B\rangle\eta^2_{\mathrm{burst}} + \langle B\rangle - 1\right). \tag{9}$$

Note that the promoter noise in the second term is independent of the transcription rate matrix $\mathbf{\Lambda}$.

As the simplest promoter circle model, the telegraph model is widely studied and can effectively simulate the experimental observation. This model considers two promoter states: an OFF state an ON state. Protein is produced at the rate μ when the gene is activated. Protein decays with a first-order rate constant δ which is set to 1. The model consists of the following four reactions:

$$\begin{aligned} &\text{OFF} \xrightarrow{k_{\text{on}}} \text{ON}, \\ &\text{ON} \xrightarrow{k_{\text{off}}} \text{OFF}, \\ &\text{ON} \xrightarrow{\mu} \text{ON} + B \cdot \text{Protein}, \\ &\text{Protein} \xrightarrow{\delta} \emptyset, \end{aligned} \tag{10}$$

where B is the burst size which can be an arbitrary discrete random variable, e.g., geometric random variable. After some algebra, we obtain the analytical expressions for the mean and noise of proteins,

$$\langle \text{protein} \rangle = \frac{\mu k_{\text{on}}}{k_{\text{on}} + k_{\text{off}}} \langle B \rangle, \tag{11}$$

$$\eta^2_{\text{protein}} = \frac{1}{\langle \text{Protein} \rangle} + \frac{k_{\text{off}}}{k_{\text{on}}(1 + k_{\text{on}} + k_{\text{off}})} + \frac{1}{2} \frac{1}{\langle \text{Protein} \rangle} \left(\langle B \rangle \eta^2_{\text{burst}} + \langle B \rangle - 1 \right). \tag{12}$$

In addition, the higher-order moments associated with the probability distribution of protein numbers, such as skewness and kurtosis, can also be computed using Eqs. (4) and (5). To verify the analytic results and examine the effect of the bursting stochasticity on the protein dynamics, we consider three different probability distributions for burst size B: (1) geometric distribution with probability distribution $h(k) = \langle B \rangle^k / (1 + \langle B \rangle)^k$, $(k = 0, 1, 2, \ldots)$, (2) Poisson distribution with probability distribution $h(k) = \langle B \rangle^k \exp(-\langle B \rangle)/k!$, $(k = 0, 1, 2, \ldots)$, and (3) negative binomial distribution with probability distribution $h(k) = C^k_{k+r-1}(1-p)^k p^r$, $(k = 0, 1, 2, \ldots)$ and average burst size $\langle B \rangle = (1-p)r/p$. Many previous experimental observations and mathematical models support these three burst size distributions. We usually consider the burst size distribution as a geometric distribution under the assumption that mRNA is degraded in a single step [4,16,49,50]. However, in some studies, a nongeometric distribution of burst size is predicted in eukaryotic cells [35,51,52] and the degradation of mRNA is considered as a multistep process (i.e., nonexponential mRNA lifetime) [5,53]. Therefore, assuming that the mRNA lifetime corresponds to an Erlang distribution, this results in a negative binomial distribution for the burst size [52,54]. In the particular case where mRNA lifetime is a deterministic constant, the burst size obeys a Poisson distribution [52]. We confirm the analytical results using Gillespie simulations [55], as shown in Fig. 2. Figure 2(a) shows that the mean steady-state protein level depends linearly on the average burst size. Figures 2(b)–2(d) illustrate that noise strength, skewness, and kurtosis for protein number decrease as mean expression increases, respectively. These results support the canonical mean-noise inverse correlation [56–59] and extend the statistical relationship to skewness or kurtosis-noise inverse correlations. Interestingly, we find that Poissonian bursting can produce less protein noise than geometric and negative binomial bursting for the same average burst size.

It implies that the dynamics of burst kinetics can be propagated into protein and encode the stochastic protein dynamics. In the next section, we will show how the bursting distribution affects the probability distribution of protein abundance.

V. PROBABILITY DISTRIBUTION

For a gene model with arbitrary promoter architecture and arbitrary bursting dynamics, the steady-state probability distribution of protein counts can be reconstructed according to the following binomial moment method [43,44]:

$$P(m) = \sum_{n \geqslant m} (-1)^{n-m} \binom{n}{m} b_n^{\text{protein}}, \quad m = 0, 1, 2, \ldots, \tag{13}$$

where the binomial moments $b_n^{\text{protein}} (n = 0, 1, 2, \ldots)$ are computed with Eqs. (4) and (5).

While Eq. (13) is valid for general gene expression models, it is interesting to consider specific examples. We consider the case where the promoter has only one ON and multiple OFFs and the protein burst size is geometrically distributed. Specifically, the transition matrix is $\mathbf{\Lambda} = \text{diag}(0, \ldots, 0, \mu)$, and the probability distribution of burst size B with k proteins is $h(k) = \langle B \rangle^k / (1 + \langle B \rangle)^k$, $(k = 0, 1, 2, \ldots)$, with $\langle B \rangle$ representing the average burst size. In this case, the binomial moments corresponding to burst size are $b_i^{\text{burst}} = \langle B \rangle^i$. And setting $\mathbf{Q}_n = (n\mathbf{I} - \mathbf{A})^{-1}\mathbf{\Lambda}$, $n = 1, 2, \ldots$, we can obtain an expression for $\mathbf{a}_n$ without iteration from Eq. (4) (see Appendix C for details),

$$\mathbf{a}_n = \mathbf{Q}_n \prod_{i=1}^{n-1} (\mathbf{Q}_i + \mathbf{I}) \mathbf{a}_0 \langle B \rangle^n. \tag{14}$$

Substituting Eq. (14) into (5) yields

$$b_n^{\text{protein}} = \frac{1}{n} \mathbf{u}_N \mathbf{\Lambda} \prod_{i=1}^{n-1} (\mathbf{Q}_i + \mathbf{I}) \mathbf{a}_0 \langle B \rangle^n. \tag{15}$$

To compute the explicit expression b_n^{protein}, we introduce two characteristic polynomials $f_{\mathbf{\Lambda}}(n) = \det(n\mathbf{I}_N - \mathbf{A}) = n \prod_{i=1}^{N-1} (n + \alpha_i)$ and $f_{\mathbf{M}_k}(n) = \det(n\mathbf{I}_{N-1} - \mathbf{M}_k) = \prod_{i=1}^{N-1} (n + \beta_i^{(k)})$. Assume that the sum of two polynomials can be factorized as $f_{\mathbf{\Lambda}}(n) + \mu f_{\mathbf{M}_N}(n) = \prod_{i=1}^{N} (n + \gamma_i)$. Then we can obtain (see Appendix C for details)

$$b_n^{\text{protein}} = \frac{\langle B \rangle^n}{n!} \frac{\prod_{i=1}^{N} (\gamma_i)_n}{\prod_{i=1}^{N-1} (\alpha_i)_n}, \tag{16}$$

where $(\square)_n$ is the (rising) Pochhammer symbol. Substituting Eq. (16) into (13), we get the probability distribution of protein abundance,

$$P(m) = \frac{\langle B \rangle^m}{m!} \frac{\prod_{i=1}^{N} (\gamma_i)_m}{\prod_{i=1}^{N-1} (\alpha_i)_m} {}_N F_{N-1} \times \left(\begin{matrix} m + \gamma_1, \ldots, m + \gamma_N \\ m + \alpha_1, \ldots, m + \alpha_{N-1} \end{matrix} \middle| ; -\langle B \rangle \right). \tag{17}$$

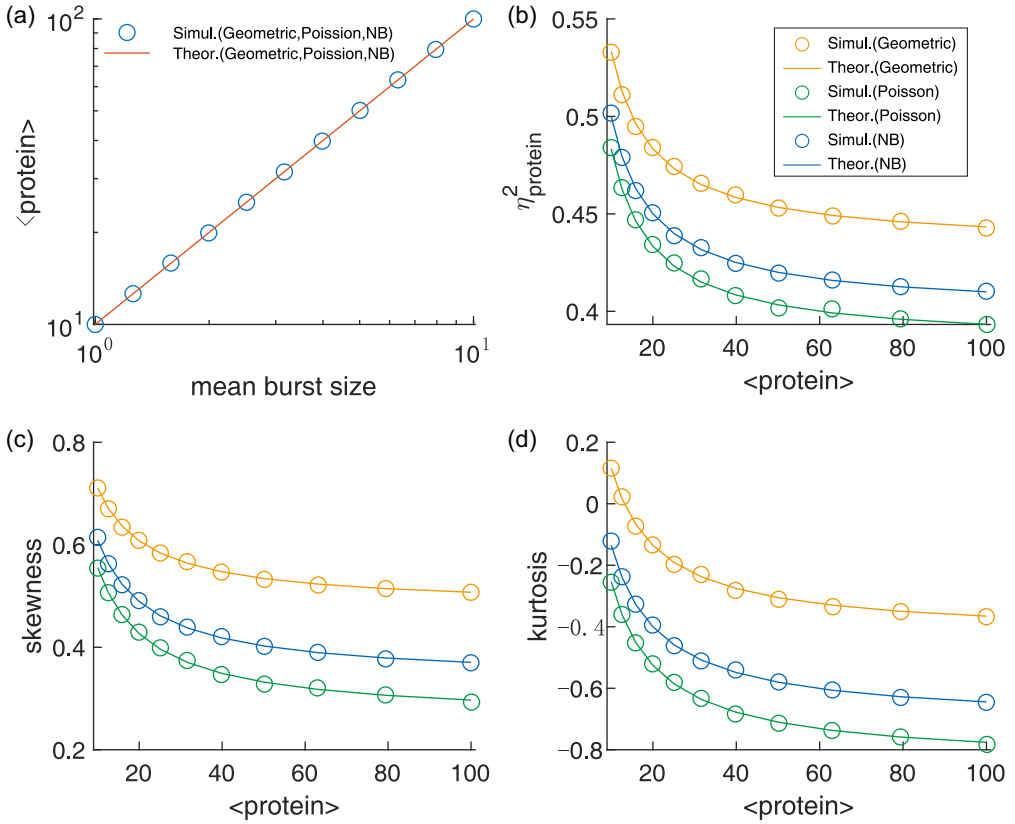

FIG. 2. Moments statistics of protein abundance for different bursting kinetics. (a) Mean protein abundance as a function of average burst size. Noise (b), skewness (c), and kurtosis (d) as a function of mean protein abundance with different bursting distributions. Parameters used for calculations are $k_{\rm on} = 1,\ k_{\rm off} = 1,\ \mu = 20,\ \delta = 1$, and $r = 3$ for negative binomial distribution. Negative binomial distribution is abbreviated to NB.

where ${}_nF_n(\begin{smallmatrix}a_1,\ldots,a_n\\ b_1,\ldots,b_n\end{smallmatrix}|;\sigma)$ is a generalized hypergeometric function. And the corresponding generating function is $G(z) = {}_NF_{N-1}(\begin{smallmatrix}\gamma_1,\ldots,\gamma_N\\ \alpha_1,\ldots,\alpha_{N-1}\end{smallmatrix}|;\langle B\rangle(z-1))$. As a concrete example, the probability distribution of protein number of telegraph model (10) is

$$P(m) = \frac{\langle B\rangle^m}{m!}\frac{(\gamma_1)_m(\gamma_2)_m}{(\alpha)_m}\,{}_2F_1\left(\begin{matrix}m+\gamma_1, m+\gamma_2\\ m+\alpha\end{matrix}\middle|;-\langle B\rangle\right), \tag{18}$$

where $\gamma_{1,2} = \frac{1}{2}(\mu + k_{\rm off} + k_{\rm on} \pm \sqrt{(\mu + k_{\rm off} + k_{\rm on})^2 - 4\mu k_{\rm on}})$, $\alpha = k_{\rm on} + k_{\rm off}$. Note that Eq. (18) is the same as a result obtained in Ref. [24] in the absence of feedback.

We verified Eqs. (13) and (18) with stochastic simulations by using the Gillespie algorithm [55]. Specifically, we obtain the probability distribution using three different methods: (1) computing the analytic probability distribution using Eqs. (17) and (18), (2) approximating the probability distribution using the binomial moment method [Eq. (13)], and (3) estimating the probability distribution with stochastic sim-

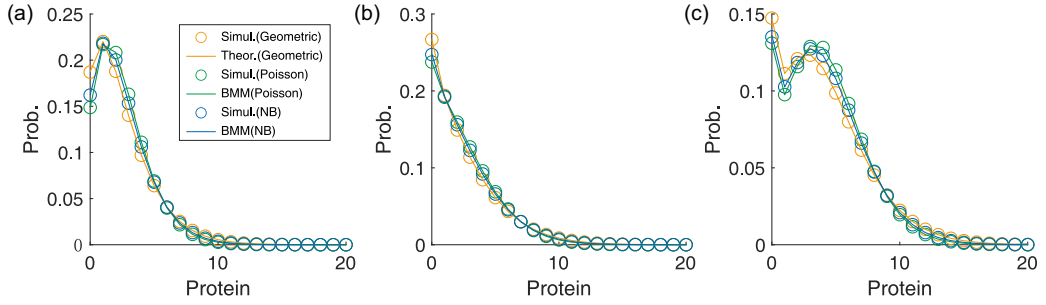

FIG. 3. Stationary distributions of protein molecules in various state-switching rates. (a) Fast state-switching case. Parameters used for calculations are: $k_{\rm on} = 10,\ k_{\rm off} = 10,\ \mu = 5,\ \delta = 1,\ \langle B\rangle = 1$. (b) Middle state-switching case. $k_{\rm on} = 1,\ k_{\rm off} = 1,\ \mu = 5,\ \delta = 1,\ \langle B\rangle = 1.$; (c) Slow state-switching case. Parameters used for calculations are $k_{\rm on} = 0.4,\ k_{\rm off} = 0.1,\ \mu = 5,\ \delta = 1,\ \langle B\rangle = 1,\ r = 3$. Binomial moments method is abbreviated to BMM.

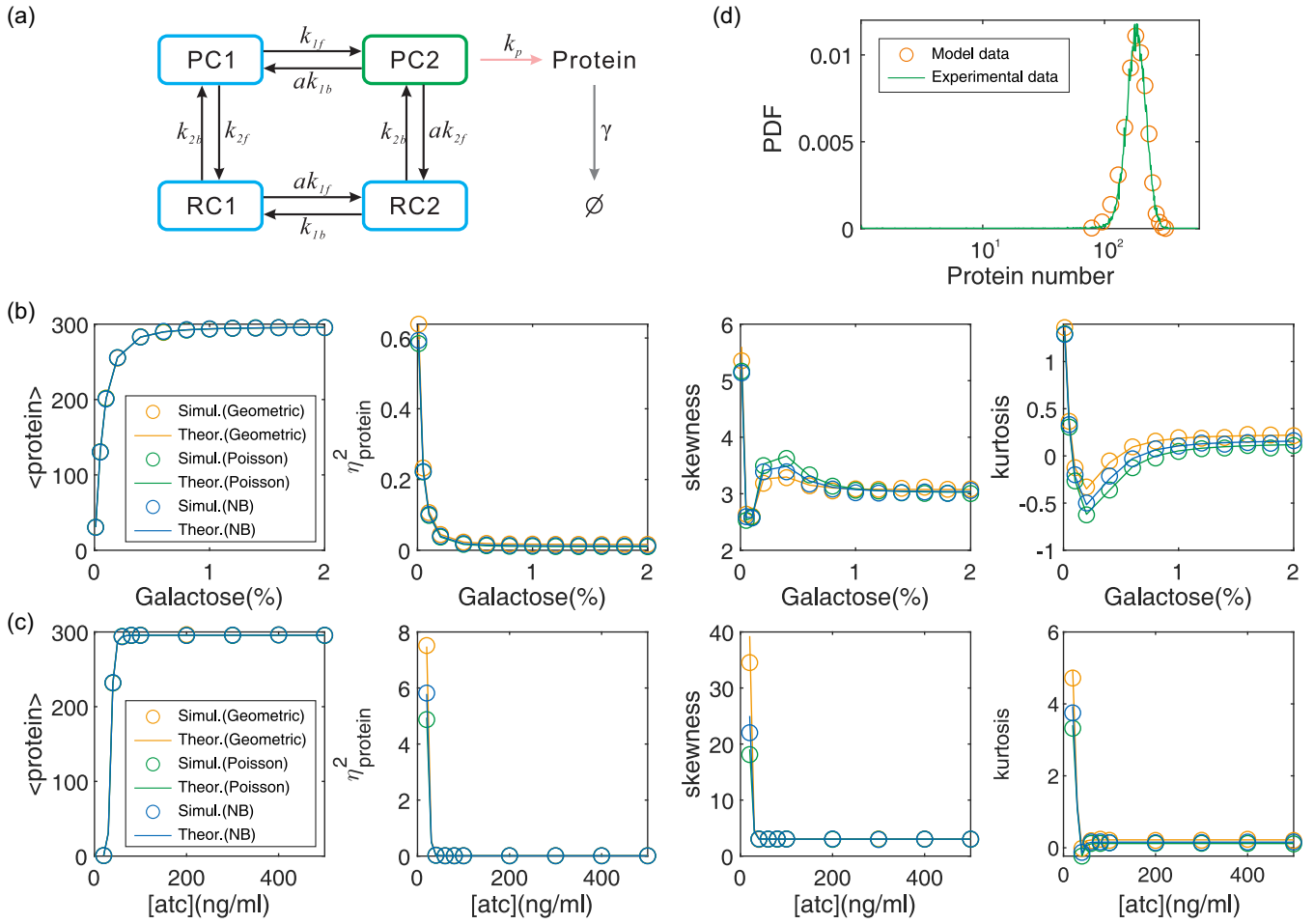

FIG. 4. Analysis results from GAL1* promoter in yeast. (a) The illustration of four-state promoter architecture kinetic model with translational burst for GAL1* promoter. PC1 is an inactive or OFF state during the transcription cycle and PC2 is active state which can transcribe mRNA and translate it into protein with bursting. RC1 and RC2 are two states with different degrees of repression when tetR binds to the promoter. We have used the kinetic parameters for calculations provided by [25,45] $k_{1f} = 0.02 + 0.2[\text{gal}]$, $k_{2f} = 200[\text{tet}]^2/(1 + c_l^4[\text{atc}]^4)^2$, $k_{1b} = 0.01 + 0.1[\text{gal}] + 0.077/[\text{gal}]$, $k_{2b} = 10$, $a = 0.025$, $c_l = 0.1$, $[\text{tet}] = 100$, $k_p = 1$, $\gamma = 0.0125$. (b) Moment statistics comparison between data simulated by Gillespie algorithm (circle) and theoretical value (solid line) under three different distributions of burst size: geometric (orange), Poisson (green), and NB (blue) with mean burst size $\langle B\rangle = 3.75$. Mean expression, noise, skewness, and kurtosis of protein abundance is computed as a function of [gal] at the condition of [atc] = 500 ng/ml. (c) Similar to (b), mean expression, noise, skewness and kurtosis of protein abundance is computed as a function of [atc] at the condition of [gal] = 2%. (d) A comparison between the probability density function of the data generated by the kinetic model (orange circle) with geometric burst size distribution ($\langle B\rangle = 3$) and the experimentally observed distribution of fluorescence (green line) at steady state, induced with 2% galactose and 40 ng/ml of Atc for 440 min [45].

ulations. Figure 3 shows that all the results from different methods agree well. And the bursting kinetics with the same average burst size can shape the probability distribution of protein abundance. Interestingly, geometric bursting can produce a heavier-tailed distribution than Poissonian and negative binomial bursting, implying more protein noise, as shown in Fig. 2.

VI. ANALYSIS RESULTS FOR THE GAL1* PROMOTER IN YEAST

In the actual biological process, the abundance of expressed proteins is usually influenced by regulatory factors, and consequently, kinetic models are often used to characterize the gene regulation process. Depending on the regulators' different regulatory effects (promotion or repression), we assume promoters at different concentrations of regulators as multiple states, which can switch between each other, and the translation process is carried out as a bursting.

To confirm that our model and conclusions can predict and analyze the regulatory impact of those with complex promoter architectures, we apply the analytical results of the previously mentioned mathematical models of arbitrary promoter architectures and translational bursting to a realistic example: the GAL1* promoter in yeast. As the concentration of galactose increases, GAL1* is activated by the transcription factor Gal4; however, it carries the TetO operator upstream from the promoter, which inhibits the initiation of transcription

of Gal1* when Tet repressor (TetR) binds to it. Conversely, anhydrotetracycline (Atc) attenuates the binding of TetR and thus indirectly upregulates the transcription of GAL1. This complex activation process constitutes the multistate promoter architecture of GAL1* [45].

Because the mRNA degradation rate is much greater than the protein degradation rate ($\gamma_R = 1, \gamma = 0.0125$) [25,45], we can approximately eliminate mRNA production and use a model with a four-state promoter with translational bursting to portray the regulation of GAL1*. The model is illustrated in Fig. 4(a), where PC1, RC1, and RC2 are inactive states. Translation occurs only in the activated state (PC2), and the burst size obeys a geometric, Poisson, or negative binomial distribution. We predict the mean expression, noise, skewness, and kurtosis of proteins from the theoretical results using the parameters reported in Refs. [25,45] and compared them with stochastic simulations as shown in Figs. 4(b) and 4(c). The theoretical and simulated values fit perfectly for a wide range of regulator concentration variations ([gal] and [Atc]). Further, we observe the mean protein expression as a function of the concentration of regulatory factors with the mean burst, $\langle B\rangle = 3.75$, in excellent agreement with that reported by Ref. [45]. Finally, our model also produces a steady-state distribution consistent with real experimental observations [45], as shown in Fig. 4(d).

VII. DISCUSSION

Delineating gene regulatory mechanisms *in vivo* are continuing challenges for systems biology. The rapid development of single-cell and single-molecule technologies is greatly accelerating such research, given its power to provide comprehensive descriptions of gene expression dynamics and the kinetics of the underlying molecular processes. By keeping pace with these experimental advances, theoretical models enable us to reveal the dynamics of stochastic gene expression.

In this paper, we have studied an exactly solvable stochastic model that integrates two key features of gene regulation, specifically: arbitrary promoter architecture and general translational bursting kinetic, in a single model. We derive the analytical expression for the corresponding protein steady-state distribution and all moment statistics of protein counts. We show that the total protein noise can be decomposed into three parts: low-copy noise, promoter noise, and translational busting noise. We also show that the bursting kinetics with the same average burst size can shape the probability distribution of protein abundance. We also applied the results of our analysis to a real observed example, GAL1*promoter, and obtained excellent prediction results. Our work has the following special features: (1) the promoter states transition matrix is an arbitrary network structure which can model the wide variety of promoter architectures; (2) the translational bursting distribution assumed by the model are of a very general form which can fit experimentally measured distributions; (3) the binomial moments method used allows the derivation of protein distribution and all moment statistics of protein counts rather than the mean and variance.

We emphasize that some special cases of our model can be found in the literature: (1) for an arbitrarily complex *cis*-regulatory motif, the production of a burst of proteins is geometrically distributed, and expressions for the mean and variance have been obtained in Refs. [25,60]; (2) for a promoter architecture with only one state modeled by queuing theory, distribution of protein burst is arbitrary, expressions for the mean and variance have been derived in Refs. [33,34,36]. Note that promoter architecture models can capture the regulation details, but they are complicated with many parameters. On the contrary, queuing models concentrate the promoter architecture into a probability distribution but lose promoter switching information and have difficulty modeling the feedback regulations. Furthermore, we omitted the process of mRNA production based on the assumption that mRNA degradation is much faster than protein degradation. In fact, mRNA production has the same burst process as proteins. And it is interesting to discuss the combined effect on the mean and noise of protein counts when transcriptional burst and protein burst are present simultaneously. In addition, some important biological processes such as cell division, replication, and mRNA maturation are ignored in the model [61–64]. How to overcome these limitations and establish effective models is the subject of ongoing research.

In conclusion, as outlined in this work, the theoretical analysis and results are essential ingredients for linking stochastic gene expression mechanisms with single-cell measurement data and gaining mechanistic insights into gene regulation.

ACKNOWLEDGMENTS

This work was supported by Grants No. 12171494, No. 11931019, No. 11631005, and No. 11775314 from the Natural Science Foundation of People's Republic of China; Grant No. 201707010117 from the Science and Technology Program of Guangzhou, Grant No. 2019B0233002 from the Guangdong Key Research and Development Project, and Grant No. 2018A030313871 from the Natural Science Foundation of Guangdong Province.

APPENDIX A: DERIVATION OF EQ. (2) IN THE MAIN TEXT

We can obtain the following equation representing the time evolution of the binomial moment equations by multiplying both sides of Eq. (1) by $\binom{m}{n}$ and taking summation over all m,

$$\frac{d\sum_{m\geqslant n}\binom{m}{n}\mathbf{P}(m;t)}{dt} = \sum_{m\geqslant n}\binom{m}{n}\left\{\mathbf{A}\mathbf{P}(m;t) + (\mathbf{E}-\mathbf{I})[m\mathbf{P}(m;t)] + \mathbf{\Lambda}\left[\sum_{k=0}^{m} h(k)\mathbf{P}(m-k;t) - \mathbf{P}(m;t)\right]\right\}. \tag{A1}$$

The left side of Eq. (A1) can be written as

$$\frac{d\sum_{m\geqslant n}\binom{m}{n}\mathbf{P}(m;t)}{dt}=\frac{d\mathbf{a}_n(t)}{dt}. \tag{A2}$$

The first term on the right side of Eq. (A1) is given by

$$\sum_{m\geqslant n}\binom{m}{n}\mathbf{A}\mathbf{P}(m;t)=\mathbf{A}\mathbf{a}_n(t). \tag{A3}$$

The second term on the right side of Eq. (A1) can be written as

$$\begin{aligned}\sum_{m\geqslant n}\binom{m}{n}(\mathbf{E}-\mathbf{I})[m\mathbf{P}(m;t)]&=\sum_{m\geqslant n}\binom{m}{n}m[\mathbf{P}(m+1;t)-\mathbf{P}(m;t)]\\&=\sum_{m-1\geqslant n}\binom{m-1}{n}m\mathbf{P}(m;t)-\sum_{m\geqslant n}\binom{m}{n}m\mathbf{P}(m;t)=-n\sum_{m\geqslant n}\binom{m}{n}\mathbf{P}(m;t)=-n\mathbf{a}_n(t).\end{aligned} \tag{A4}$$

To obtain the last term on the right side of Eq. (A1), we first change the order of terms in the summation,

$$\sum_{m\geqslant n}\binom{m}{n}\sum_{k=0}^{m}h(k)\mathbf{P}(m-k;t)=\sum_{k\geqslant 0}h(k)\sum_{m\geqslant n}\binom{m}{n}\mathbf{P}(m-k;t). \tag{A5}$$

Using the combinatorial equality $\binom{m+k}{n}=\binom{k}{0}\binom{m}{n}+\binom{k}{1}\binom{m}{n-1}+\cdots+\binom{k}{k}\binom{m}{n-k}$, we obtain

$$\begin{aligned}h(k)\sum_{m\geqslant n}\binom{m}{n}\mathbf{P}(m-k;t)&=h(k)\sum_{m\geqslant n-k}\binom{m+k}{n}\mathbf{P}(m;t)\\&=h(k)\sum_{m\geqslant n-k}\left[\binom{k}{0}\binom{m}{n}+\binom{k}{1}\binom{m}{n-1}+\cdots+\binom{k}{k}\binom{m}{n-k}\right]\mathbf{P}(m;t)=h(k)\sum_{i=0}^{k}\binom{k}{i}\mathbf{a}_{n-i}(t).\end{aligned} \tag{A6}$$

Substituting Eq. (A6) into Eq. (A5) and changing the order of terms in the summation yields

$$\sum_{k\geqslant 0}h(k)\sum_{m\geqslant n}\binom{m}{n}\mathbf{P}(m-k;t)=\sum_{k\geqslant 0}h(k)\sum_{i=0}^{k}\binom{k}{i}\mathbf{a}_{n-i}(t)=\sum_{i=0}^{n}\mathbf{a}_{n-i}(t)\sum_{k\geqslant i}\binom{k}{i}h(k), \tag{A7}$$

and substituting Eq. (A6) into the last term on the right side of Eq. (A1), we obtain

$$\sum_{m\geqslant n}\binom{m}{n}\mathbf{\Lambda}\left[\sum_{k=0}^{m}h(k)\mathbf{P}(m-k;t)-\mathbf{P}(m;t)\right]=\mathbf{\Lambda}\left[\sum_{i=0}^{n}\mathbf{a}_{n-i}(t)\sum_{k\geqslant i}\binom{k}{i}h(k)-\mathbf{a}_n(t)\right]=\mathbf{\Lambda}\sum_{i=1}^{n}\mathbf{a}_{n-i}(t)\sum_{k\geqslant i}\binom{k}{i}h(k). \tag{A8}$$

Combining Eqs. (A2)–(A4) with (A8), we arrive at the resulting binomial moment equation in the main text,

$$\frac{d\mathbf{a}_n(t)}{dt}=\mathbf{A}\mathbf{a}_n(t)-n\mathbf{a}_n(t)+\mathbf{\Lambda}\left[\sum_{i=1}^{n}\mathbf{a}_{n-i}(t)\sum_{k=i}^{\infty}\binom{k}{i}h(k)\right]. \tag{A9}$$

APPENDIX B: DERIVATION OF EQS. (8) IN THE MAIN TEXT

Using the equation $b_n^{\text{protein}}=\frac{1}{n}\mathbf{u}_N\mathbf{\Lambda}\mathbf{a}_n * b^{\text{burst}}[n]$, we can get $\langle\text{protein}\rangle=b_1^{\text{protein}}=\mathbf{u}_N\mathbf{\Lambda}\mathbf{a}_0\langle B\rangle$ [i.e., Eq. (7) in the main text] immediately, and

$$b_2^{\text{protein}}=\frac{1}{2}\sum_{i=1}^{N}\mu_i\left[a_1^{(i)}b_1^{\text{burst}}+a_0^{(i)}b_2^{\text{burst}}\right]=\frac{1}{2}\langle B\rangle\sum_{i=1}^{N}\mu_i a_1^{(i)}+\langle\text{Protein}\rangle\frac{b_2^{\text{burst}}}{\langle B\rangle}=\frac{1}{2}\langle B\rangle^2\mathbf{u}_N\mathbf{\Lambda}(\mathbf{I}-\mathbf{A})^{-1}\mathbf{\Lambda}\mathbf{a}_0+\langle\text{Protein}\rangle\frac{b_2^{\text{burst}}}{\langle B\rangle}. \tag{B1}$$

The protein noise can be computed according to the first two binomial moments,

$$\eta^2_{\text{protein}}=\frac{1}{b_1^{\text{protein}}}+\frac{2b_2^{\text{protein}}}{\left(b_1^{\text{protein}}\right)^2}-1. \tag{B2}$$

Substituting the expressions of b_1^{protein} and b_2^{protein} into Eq. (B2) yields

$$\begin{aligned}\eta^2_{\text{protein}} &= \frac{1}{\langle\text{Protein}\rangle} + \frac{\langle B\rangle^2 \mathbf{u}_N \mathbf{\Lambda}(\mathbf{I}-\mathbf{A})^{-1}\mathbf{\Lambda}\mathbf{a}_0 + \langle\text{Protein}\rangle \frac{b_2^{\text{burst}}}{\langle B\rangle}}{\langle\text{Protein}\rangle^2} - 1 \\ &= \frac{1}{\langle\text{Protein}\rangle} + \frac{\langle B\rangle^2 \mathbf{u}_N \mathbf{\Lambda}(\mathbf{I}-\mathbf{A})^{-1}\mathbf{\Lambda}\mathbf{a}_0}{(\mathbf{u}_N\mathbf{\Lambda}\mathbf{a}_0\langle B\rangle)^2} + \frac{b_2^{\text{burst}}/\langle B\rangle}{\langle\text{Protein}\rangle} - 1 \\ &= \frac{1}{\langle\text{Protein}\rangle} + \frac{\mathbf{u}_N \mathbf{\Lambda}(\mathbf{I}-\mathbf{A})^{-1}\mathbf{\Lambda}\mathbf{a}_0 - (\mathbf{u}_N\mathbf{\Lambda}\mathbf{a}_0)^2}{(\mathbf{u}_N\mathbf{\Lambda}\mathbf{a}_0)^2} + \frac{1}{2}\frac{1}{\langle\text{Protein}\rangle}\big(\langle B\rangle\eta^2_{\text{burst}} + \langle B\rangle - 1\big). \end{aligned} \tag{B3}$$

APPENDIX C: DERIVATION OF EQ. (9) IN THE MAIN TEXT

We introduce two functions,

$$f_{\mathbf{A}}(n) = \det(n\mathbf{I}_N - \mathbf{A}) = n\prod_{i=1}^{N-1}(n+\alpha_i), \quad f_{\mathbf{M}_k}(n) = \det(n\mathbf{I}_{N-1} - \mathbf{M}_k) = \prod_{i=1}^{N-1}\big(n+\beta_i^{(k)}\big), \tag{C1}$$

which are practically the characteristic polynomial of $\mathbf{A}$ and $\mathbf{M}_k$, respectively.

Because the transition matrix is $\mathbf{\Lambda} = \text{diag}(0,\ldots,0,\mu)$, we have

$$\begin{aligned}\mathbf{a}_n &= \begin{pmatrix} * \\ a_n^{(N)} \end{pmatrix} = \frac{1}{\det(n\mathbf{I}-\mathbf{A})}(n\mathbf{I}-\mathbf{A})^*\mathbf{\Lambda}\mathbf{a} * b^{\text{burst}}[n] \\ &= \frac{\mu}{\det(n\mathbf{I}-\mathbf{A})}\begin{pmatrix} * & * \\ * & \det(n\mathbf{I}_{N-1}-\mathbf{M}_N)\end{pmatrix}\begin{pmatrix}\mathbf{O} & \mathbf{O} \\ \mathbf{O} & 1\end{pmatrix}\begin{pmatrix} * \\ \sum_{i=1}^n a_{n-i}^{(N)} b_i^{\text{burst}}\end{pmatrix} \\ &= \frac{\mu}{f_{\mathbf{A}}(n)}\begin{pmatrix} * \\ f_{M_N}(n)\sum_{i=1}^n a_{n-i}^{(N)} b_i^{\text{burst}}\end{pmatrix} = \begin{pmatrix} * \\ \frac{\mu f_{M_N}(n)}{f_{\mathbf{A}}(n)}\sum_{i=1}^n a_{n-i}^{(N)} b_i^{\text{burst}}\end{pmatrix}. \end{aligned} \tag{C2}$$

Then we obtain

$$a_1^{(N)} = \frac{\mu f_{M_N}(1)}{f_{\mathbf{A}}(1)} a_0^{(N)} b_1^{\text{burst}} = \mu\prod_{i=1}^{N-1}\frac{\big(1+\beta_i^{(k)}\big)}{(1+\alpha_i)}\prod_{i=1}^{N-1}\frac{\beta_i^{(N)}}{\alpha_i} b_1^{\text{burst}} = \mu\langle B\rangle\prod_{i=1}^{N-1}\frac{\beta_i^{(N)}\big(1+\beta_i^{(N)}\big)}{\alpha_i(1+\alpha_i)}. \tag{C3}$$

Combining Eqs. (5) and (6) in the main text and (C3), we get the second order binomial moment,

$$b_2^{\text{protein}} = \frac{1}{2}\mu\sum_{k=1}^{2} a_{n-k}^{(N)} b_k^{\text{burst}} = \frac{1}{2}\mu\big[a_1^{(N)} b_1^{\text{burst}} + a_0^{(N)} b_2^{\text{burst}}\big] = \frac{1}{2}\mu\left[\mu\langle B\rangle\prod_{i=1}^{N-1}\frac{\beta_i^{(N)}\big(1+\beta_i^{(N)}\big)}{\alpha_i(1+\alpha_i)} b_1^{\text{burst}} + \prod_{i=1}^{N-1}\frac{\beta_i^{(N)}}{\alpha_i} b_2^{\text{burst}}\right]. \tag{C4}$$

Then we obtain the protein noise [i.e., Eq. (9) in the main text], given by

$$\begin{aligned}\eta^2_{\text{protein}} &= \frac{1}{b_1^{\text{protein}}} + \frac{2b_2^{\text{protein}}}{\big(b_1^{\text{protein}}\big)^2} - 1 \\ &= \frac{1}{\langle\text{Protein}\rangle} + \frac{\mu\big[\mu\langle B\rangle\prod_{i=1}^{N-1}\frac{\beta_i^{(N)}(1+\beta_i^{(N)})}{\alpha_i(1+\alpha_i)} b_1^{\text{burst}} + \prod_{i=1}^{N-1}\frac{\beta_i^{(N)}}{\alpha_i} b_2^{\text{burst}}\big]}{\langle\text{Protein}\rangle^2} - 1 \\ &= \frac{1}{\langle\text{Protein}\rangle} + \prod_{i=1}^{N-1}\frac{\alpha_i\big(1+\beta_i^{(N)}\big)}{\beta_i^{(N)}(1+\alpha_i)} + \frac{b_2^{\text{burst}}/\langle B\rangle}{\langle\text{Protein}\rangle} - 1 \\ &= \frac{1}{\langle\text{Protein}\rangle} + \prod_{i=1}^{N-1}\frac{\alpha_i\big(1+\beta_i^{(N)}\big)}{\beta_i^{(N)}(1+\alpha_i)} + \frac{1}{2}\frac{1}{\langle\text{Protein}\rangle}\big(\langle B\rangle\eta^2_{\text{burst}} + \langle B\rangle - 1\big) - 1 \\ &= \frac{1}{\langle\text{Protein}\rangle} + \frac{\prod_{i=1}^{N-1}\big(1+\beta_i^{(N)}\big)\alpha_i - \prod_{i=1}^{N-1}(1+\alpha_i)\beta_i^{(N)}}{\prod_{i=1}^{N-1}(1+\alpha_i)\beta_i^{(N)}} + \frac{1}{2}\frac{1}{\langle\text{Protein}\rangle}\big(\langle B\rangle\eta^2_{\text{burst}} + \langle B\rangle - 1\big). \end{aligned} \tag{C5}$$

APPENDIX D: DERIVATION OF EQ. (17) IN THE MAIN TEXT

Using the specific characteristics of the transition matrix $\mathbf{\Lambda} = \mathrm{diag}(0, \ldots, 0, \mu)$, matrix $\mathbf{Q}_n$ can be expressed as the following:

$$\mathbf{Q}_n = \frac{\mu}{\det(n\mathbf{I} - \mathbf{A})}\begin{pmatrix} * & * \\ * & \det(n\mathbf{I}_{N-1} - \mathbf{M}_N) \end{pmatrix}\begin{pmatrix} \mathbf{O} & \mathbf{O} \\ \mathbf{O} & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{O} & * \\ \mathbf{O} & \frac{\mu f_{M_N}(n)}{f_{\mathbf{A}}(n)} \end{pmatrix}. \tag{D1}$$

Then, we have

$$\mathbf{Q}_n + \mathbf{I} = \begin{pmatrix} \mathbf{I} & * \\ \mathbf{O} & \frac{f_{\mathbf{A}}(n) + \mu f_{M_N}(n)}{f_{\mathbf{A}}(n)} \end{pmatrix}. \tag{D2}$$

Using the equation $f_{\mathbf{A}}(n) + \mu f_{\mathbf{M}_N}(n) = \prod_{i=1}^{N}(n + \gamma_i)$, we obtain the binomial moments for protein, given by

$$b_n^{\text{protein}} = \frac{1}{n}\mathbf{u}_N \mathbf{\Lambda} \prod_{k=1}^{n-1}(\mathbf{Q}_i + \mathbf{I})\mathbf{a}_0 \langle B\rangle^n = \frac{\mu}{n}\frac{\prod_{i=1}^{N-1}\gamma_i}{\prod_{i=1}^{N-1}\alpha_i}\langle B\rangle^n \prod_{k=1}^{n-1}\frac{\prod_{i=1}^{N}(k+\gamma_i)}{k\prod_{i=1}^{N-1}(k+\alpha_i)} = \frac{\langle B\rangle^n}{n!}\frac{\prod_{i=1}^{N}(\gamma_i)_n}{\prod_{i=1}^{N-1}(\alpha_i)_n}. \tag{D3}$$

Substituting Eq. (D3) into Eq. (B2) with a bit of algebra yields

$$\begin{aligned} P(m) &= \sum_{n\geqslant m}(-1)^{n-m}\binom{n}{m} b_n^{\text{protein}} = \sum_{n\geqslant m}(-1)^{n-m}\binom{n}{m}\frac{\langle B\rangle^n}{n!}\frac{\prod_{i=1}^{N}(\gamma_i)_n}{\prod_{i=1}^{N-1}(\alpha_i)_n} \\ &= \frac{\langle B\rangle^n}{m!}\frac{\prod_{i=1}^{N}(\gamma_i)_n}{\prod_{i=1}^{N-1}(\alpha_i)_n}{}_N F_{N-1}\left(\begin{matrix} m+\gamma_1, \ldots, m+\gamma_N \\ m+\alpha_1, \ldots, m+\alpha_{N-1} \end{matrix}\middle| ; -\langle B\rangle\right), \end{aligned} \tag{D4}$$

where ${}_nF_n\left(\begin{smallmatrix} a_1, \ldots, a_n \\ b_1, \ldots, b_n \end{smallmatrix}\middle| ; \sigma\right)$ is a generalized hypergeometric function. Using the relationship between generating function and binomial moments, we obtain the corresponding generating function

$$G(z) = {}_N F_{N-1}\left(\begin{matrix} \gamma_1, \ldots, \gamma_N \\ \alpha_1, \ldots, \alpha_{N-1} \end{matrix}\middle| ; \langle B\rangle(z-1)\right). \tag{D5}$$

[1] J. Rodriguez and D. R. Larson, Transcription in living cells: Molecular mechanisms of bursting, Annu. Rev. Biochem. **89**, 189 (2020).

[2] E. Tunnacliffe and J. R. Chubb, What is a transcriptional burst? Trends Genet. **36**, 288 (2020).

[3] J. P. Bothma, H. G. Garcia, E. Esposito, G. Schlissel, T. Gregor, and M. Levine, Dynamic regulation of eve stripe 2 expression reveals transcriptional bursts in living Drosophila embryos, Proc. Natl. Acad. Sci. USA **111**, 10598 (2014).

[4] I. Golding, J. Paulsson, S. M. Zawilski, and E. C. Cox, Real-time kinetics of gene activity in individual bacteria, Cell **123**, 1025 (2005).

[5] D. M. Suter, N. Molina, D. Gatfield, K. Schneider, U. Schibler, and F. Naef, Mammalian genes are transcribed with widely different bursting kinetics, Science **332**, 472 (2011).

[6] A. Coulon, C. C. Chow, R. H. Singer, and D. R. Larson, Eukaryotic transcriptional dynamics: From single molecules to cell populations, Nat. Rev. Genet. **14**, 572 (2013).

[7] D. R. Larson, D. Zenklusen, B. Wu, J. A. Chao, and R. H. Singer, Real-time observation of transcription initiation and elongation on an endogenous yeast gene, Science **332**, 475 (2011).

[8] A. Raj, C. S. Peskin, D. Tranchina, D. Y. Vargas, and S. Tyagi, Stochastic mRNA synthesis in mammalian cells, PLoS Biol. **4**, e309 (2006).

[9] S. Chong, C. Chen, H. Ge, and X. S. Xie, Mechanism of transcriptional bursting in bacteria, Cell **158**, 314 (2014).

[10] B. Zoller, S. C. Little, and T. Gregor, Diverse spatial expression patterns emerge from unified kinetics of transcriptional bursting, Cell **175**, 835 (2018).

[11] J. R. Chubb, T. Trcek, S. M. Shenoy, and R. H. Singer, Transcriptional pulsing of a developmental gene, Curr. Biol. **16**, 1018 (2006).

[12] C. V. Harper, B. Finkenstädt, D. J. Woodcock, S. Friedrichsen, S. Semprini, L. Ashall, D. G. Spiller, J. J. Mullins, D. A. Rand, and J. R. Davis, Dynamic analysis of stochastic transcription cycles, PLoS Biol. **9**, e1000607 (2011).

[13] X. Yan, T. A. Hoek, R. D. Vale, and M. E. Tanenbaum, Dynamics of translation of single mRNA molecules in vivo, Cell **165**, 976 (2016).

[14] C. Wang, B. Han, R. Zhou, and X. Zhuang, Real-time imaging of translation on single mRNA transcripts in live cells, Cell **165**, 990 (2016).

[15] V. Tatavarty, M. F. Ifrim, M. Levin, G. Korza, E. Barbarese, J. Yu, and J. H. Carson, Single-molecule imaging of translational output from individual RNA granules in neurons, Mol. Biol. Cell **23**, 918 (2012).

[16] J. Yu, J. Xiao, X. Ren, K. Lao, and X. S. Xie, Probing gene expression in live cells, one protein molecule at a time, Science **311**, 1600 (2006).

[17] B. Wu, C. Eliscovich, Y. J. Yoon, and R. H. Singer, Translation dynamics of single mRNAs in live cells and neurons, Science **352**, 1430 (2016).

[18] L. Cai, C. K. Dalal, and M. B. Elowitz, Frequency-modulated nuclear localization bursts coordinate gene regulation, Nature (London) **455**, 485 (2008).

[19] H. H. Chang, M. Hemberg, M. Barahona, D. E. Ingber, and S. Huang, Transcriptome-wide noise controls lineage choice in mammalian progenitor cells, Nature (London) **453**, 544 (2008).

[20] A. Raj, S. A. Rifkin, E. Andersen, and A. Van Oudenaarden, Variability in gene expression underlies incomplete penetrance, Nature (London) **463**, 913 (2010).
[21] W. J. Blake, G. Balázsi, M. A. Kohanski, F. J. Isaacs, K. F. Murphy, Y. Kuang, C. R. Cantor, D. R. Walt, and J. J. Collins, Phenotypic consequences of promoter-mediated transcriptional noise, Mol. Cell **24**, 853 (2006).
[22] A. M. Corrigan, E. Tunnacliffe, D. Cannon, and J. R. Chubb, A continuum model of transcriptional bursting, eLife **5**, e13051 (2016).
[23] J. Peccoud and B. Ycart, Markovian modeling of gene-product synthesis, Theor. Popul. Biol. **48**, 222 (1995).
[24] N. Kumar, T. Platini, and R. V. Kulkarni, Exact Distributions for Stochastic Gene Expression Models with Bursting and Feedback, Phys. Rev. Lett. **113**, 268105 (2014).
[25] Á. Sánchez and J. Kondev, Transcriptional control of noise in gene expression, Proc. Natl. Acad. Sci. USA **105**, 5081 (2008).
[26] V. Shahrezaei and P. S. Swain, Analytical distributions for stochastic gene expression, Proc. Natl. Acad. Sci. USA **105**, 17256 (2008).
[27] J. Zhang, L. Chen, and T. Zhou, Analytical distribution and tunability of noise in a model of promoter progress, Biophys. J. **102**, 1247 (2012).
[28] J. Zhang and T. Zhou, Promoter-mediated transcriptional dynamics, Biophys. J. **106**, 479 (2014).
[29] M. C. Mackey, M. Tyran-Kaminska, and R. Yvinec, Dynamic behavior of stochastic gene expression models in the presence of bursting, SIAM J. Appl. Math. **73**, 1830 (2013).
[30] U. Herbach, Stochastic gene expression with a multistate promoter: Breaking down exact distributions, SIAM J. Appl. Math. **79**, 1007 (2019).
[31] C. Jia, Kinetic foundation of the zero-inflated negative binomial model for single-cell RNA sequencing data, SIAM J. Appl. Math. **80**, 1336 (2020).
[32] L. Ham, D. Schnoerr, R. D. Brackston, and M. P. Stumpf, Exactly solvable models of stochastic gene expression, J. Chem. Phys. **152**, 144106 (2020).
[33] J. M. Pedraza and J. Paulsson, Effects of molecular memory and bursting on fluctuations in gene expression, Science **319**, 339 (2008).
[34] T. Jia and R. V. Kulkarni, Intrinsic Noise in Stochastic Models of Gene Expression with Molecular Memory and Bursting, Phys. Rev. Lett. **106**, 058102 (2011).
[35] A. Schwabe, K. N. Rybakova, and F. J. Bruggeman, Transcription stochasticity of complex gene regulation models, Biophys. J. **103**, 1152 (2012).
[36] N. Kumar, A. Singh, and R. V. Kulkarni, Transcriptional bursting in gene expression: Analytical results for general stochastic models, PLoS Comput. Biol. **11**, e1004292 (2015).
[37] M. Thattai, Universal Poisson statistics of mRNAs with complex decay pathways, Biophys. J. **110**, 301 (2016).
[38] J. Zhang and T. Zhou, Stationary moments, distribution conjugation and phenotypic regions in stochastic gene transcription, Math. Biosci. Eng. **16**, 6134 (2019).
[39] S. Choubey, Nascent RNA kinetics: Transient and steady state behavior of models of transcription, Phys. Rev. E **97**, 022402 (2018).
[40] J. Zhang and T. Zhou, Markovian approaches to modeling intracellular reaction processes with molecular memory, Proc. Natl. Acad. Sci. USA **116**, 23542 (2019).
[41] D. L. Jones, R. C. Brewster, and R. Phillips, Promoter architecture dictates cell-to-cell variability in gene expression, Science **346**, 1533 (2014).
[42] J. M. Vilar and L. Saiz, CplexA: A Mathematica package to study macromolecular-assembly control of gene expression, Bioinformatics **26**, 2060 (2010).
[43] J. Zhang, Q. Nie, M. He, and T. Zhou, An effective method for computing the noise in biochemical networks, J. Chem. Phys. **138**, 084106 (2013).
[44] J. Zhang, Q. Nie, and T. Zhou, A moment-convergence method for stochastic analysis of biochemical reaction networks, J. Chem. Phys. **144**, 194109 (2016).
[45] W. J. Blake, M. Kærn, C. R. Cantor, and J. J. Collins, Noise in eukaryotic gene expression, Nature (London) **422**, 633 (2003).
[46] C. W. Gardiner, *Stochastic Methods: A Handbook for the Natural and Social Sciences* (Springer, Berlin, Heidelberg 2009).
[47] R. Milo and R. Phillips, *Cell Biology by the Numbers* (Garland Science, New York, 2015) .
[48] J. Holehouse, Z. Cao, and R. Grima, Stochastic modeling of autoregulatory genetic feedback loops: A review and comparative study, Biophys. J. **118**, 1517 (2020).
[49] L. Cai, N. Friedman, and X. S. Xie, Stochastic protein expression in individual cells at the single molecule level, Nature (London) **440**, 358 (2006).
[50] Y. Taniguchi, P. J. Choi, G.-W. Li, H. Chen, M. Babu, J. Hearn, A. Emili, and X. S. Xie, Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells, Science **329**, 533 (2010).
[51] V. Elgart, T. Jia, A. T. Fenley, and R. Kulkarni, Connecting protein and mRNA burst distributions for stochastic models of gene expression, Phys. Biol. **8**, 046001 (2011).
[52] K. R. Ghusinga and A. Singh, Effect of gene-expression bursts on stochastic timing of cellular events, in *2017 American Control Conference (ACC)*, edited by A. Astolfi (IEEE, Seattle, WA, USA, 2017), p. 2118.
[53] T. C. Voss and G. L. Hager, Dynamic regulation of transcriptional states by chromatin and transcription factors, Nat. Rev. Genet. **15**, 69 (2014).
[54] R. R. Delongchamp, C. R. Valentine, and H. V. Malling, Estimation of the average burst size of ΦX174 am3, cs70 for use in mutation assays with transgenic mice, Environ. Mol. Mutag. **37**, 356 (2001).
[55] D. T. Gillespie, Exact stochastic simulation of coupled chemical reactions, J. Phys. Chem. **81**, 2340 (1977).
[56] R. D. Dar, S. M. Shaffer, A. Singh, B. S. Razooky, M. L. Simpson, A. Raj, and L. S. Weinberger, Transcriptional bursting explains the noise–versus–mean relationship in mRNA and protein levels, PLoS One **11**, e0158298 (2016).
[57] S. S. Dey, J. E. Foley, P. Limsirichai, D. V. Schaffer, and A. P. Arkin, Orthogonal control of expression mean and variance by epigenetic features at different genomic loci, Mol. Syst. Biol. **11**, 806 (2015).
[58] G. Hornung, R. Bar-Ziv, D. Rosin, N. Tokuriki, D. S. Tawfik, M. Oren, and N. Barkai, Noise–mean relationship in mutated promoters, Genome Res. **22**, 2409 (2012).
[59] F. L. M. Vallania, M. Sherman, Z. Goodwin, I. Mogno, B. A. Cohen, and R. D. Mitra, Origin and consequences of the relationship between protein mean and variance, PLoS One **9**, e102202 (2014).

[60] A. Sanchez, H. G. Garcia, D. Jones, R. Phillips, and J. Kondev, Effect of promoter architecture on the cell-to-cell variability in gene expression, PLoS Comput. Biol. **7**, e1001100 (2011).

[61] Z. Cao and R. Grima, Analytical distributions for detailed models of stochastic gene expression in eukaryotic cells, Proc. Natl. Acad. Sci. USA **117**, 4682 (2020).

[62] C. H. Beentjes, R. Perez-Carrasco, and R. Grima, Exact solution of stochastic gene expression models with bursting, cell cycle and replication dynamics, Phys. Rev. E **101**, 032403 (2020).

[63] J. R. Peterson, J. A. Cole, J. Fei, T. Ha, and Z. A. Luthey-Schulten, Effects of DNA replication on mRNA noise, Proc. Natl. Acad. Sci. USA **112**, 15886 (2015).

[64] T. Filatova, N. Popovic, and R. Grima, Statistics of nascent and mature rna fluctuations in a stochastic model of transcriptional initiation, elongation, pausing, and termination, Bull. Math. Biol. **83**, 1 (2021).