

Promoter-mediated Transcriptional Dynamics

Jiajun Zhang and Tianshou Zhou*

Guangdong Province Key Laboratory of Computational Science, School of Mathematics and Computational Science, Sun Yat-Sen University, Guangzhou 510275, People's Republic of China

ABSTRACT Genes in eukaryotic cells are typically regulated by complex promoters containing multiple binding sites for a variety of transcription factors, but how promoter dynamics affect transcriptional dynamics has remained poorly understood. In this study, we analyze gene models at the transcriptional regulation level, which incorporate the complexity of promoter structure (PS) defined as transcriptional exits (i.e., ON states of the promoter) and the transition pattern (described by a matrix consisting of transition rates among promoter activity states). We show that multiple exits of transcription are the essential origin of generating multimodal distributions of mRNA, but promoters with the same transition pattern can lead to multimodality of different modes, depending on the regulation of transcriptional factors. In turn, for similar mRNA distributions in the models, the mean ON or OFF time distributions may exhibit different characteristics, thus providing the supplemental information on PS. In addition, we demonstrate that the transcriptional noise can be characterized by a nonlinear function of mean ON and OFF times. These results not only reveal essential characteristics of promoter-mediated transcriptional dynamics but also provide signatures useful for inferring PS based on characteristics of transcriptional outputs.

INTRODUCTION

Gene expression involves transcription, translation, chromatin remodeling, histone modifications, alternative splicing, and recruitment of transcription factors (TFs), and polymerases. These biochemical processes inevitably lead to stochastic fluctuations (or the noise) in expression levels (1–4). This noise is essential for many cellular functions (5,6) and has been identified as a key factor underlying the observed phenotypic variability of genetically identical cells in homogeneous environments (7). Although recent advances in experimental methods allow direct observations of real-time fluctuations in gene expression levels in individual live cells (8–12), there is considerable interest in theoretically understanding how different molecular mechanisms of gene expression affect variations in mRNA and protein levels across a population of cells. Quantifying the contributions of different sources of noise using stochastic models of gene expression is an important step toward understanding fundamental cellular processes and variations in cell populations (13–30).

Transcription is a key step during gene expression, where the transcription machinery is responsible for transcribing DNA to RNA and initiating mRNA transcripts (31). Biochemical processes associated with transcription often involve a variety of TFs, which bind to multiple sites on regulatory DNA in response to intracellular or extracellular signals. When bound to these sites, the TFs either inhibit or enhance transcription through interactions with RNA polymerase and other TFs. Most regulatory sequences called as “promoters” contain several operator sequences, each of which is often recognized with different affinities by

more than one type of TF. For bacterial cells, the promoters that are viewed as simple can exist in a surprisingly large number of regulatory states. For example, the PRM promoter of phage lambda in *E. coli* is regulated by two different TFs binding to two sets of three operators that can be brought together by looping out the intervening DNA. As a result, the number of regulatory states of the PRM promoter is 128 (32). In contrast, eukaryotic promoters are more complex, involving nucleosomes competing with or being removed by TFs (33). In addition to the conventional regulation by TFs, the eukaryotic promoters can be also epigenetically regulated via histone modifications (34–36). Such regulation may lead to very complex promoter structure (PS) (37). To help readers understand how a PS is formed, we simply introduce three molecular mechanisms (38): 1) *nucleosome occupancy* that promoter-DNA condensation into chromatin may lead to long-lived, silenced or OFF, promoter states, which are followed by rapid, short-lived initiation events; 2) *TATA box* that activates the promoter by helping assemble the pre-initiation complex; 3) *TF binding sites* for which the molecular mechanism has not been well understood.

Transcription takes place often in a bursting manner. Single-cell experimental measurements have provided evidence for transcriptional bursting both in bacteria (8) and in eukaryotic cells (9,39). Although the sources of the transcriptional burst remain poorly understood (40), several lines of evidence (3,4,10,11,41–45) point to transitions among the ON and OFF states of the promoter as an important source of noise in gene expression, which is responsible for generating cell-to-cell heterogeneity in the response of genetically identical cells to the same stimulus. Complex promoters with more than two activity states are not the exception but the rule as combinatorial control of gene

Submitted October 1, 2013, and accepted for publication December 9, 2013.

*Correspondence: mcszts@mail.sysu.edu.cn

Editor: Peter Hunter

© 2014 by the Biophysical Society
0006-3495/14/01/0479/10 \$2.00



regulation by multiple TFs is widespread (46). In two relevant studies, an experiment in yeast cells demonstrated that high levels of cell-to-cell variability, originated by promoter state fluctuations, may confer cell colonies with an enhanced probability of cell survival when subjected to external stress (47); another experiment showed that a stable transcription scaffold that regulates the rate of transitions between ON and OFF states of the promoter can result in “bursts” of gene expression beneficial to increasing cell-to-cell variability (5). In particular, all three of the molecular mechanisms (described above) for the formation of PS can lead to transcriptional bursting. In fact, rapid, short-lived initiation events taking place in nucleosome occupancy can lead to bursting synthesis of mRNA (4); in the TATA box case, it was experimentally demonstrated that mutations that weaken the strength of the TATA box of the PHO5 gene in yeast cells result in a reduction in gene expression noise (4); in the case of TF binding sites, experiments have shown that the number of binding sites for TFs can significantly affect the gene expression noise (10,48).

Given the complexity of most PSs, quantitative models play an important role in testing molecular mechanisms of transcriptional regulation, helping to connect these biochemical models of transcription with experimental measurements of gene expression *in vivo* (26). Thus far, many theoretical models have been developed. A class of gene models developed in response to bulk experiments focused on computing the steady-state occupancies of different operators by TFs (49,50) and can be used to well predict the equilibrium probability of each promoter state and therefore the average transcriptional output. These models, although very useful for computing average gene expression levels at steady state, have nothing to say about the dynamics of gene regulation, that is, which promoter states are kinetically connected, and how often the promoter makes transitions from one state to another. To address these questions, another class of gene models have been also developed during the past decade (24,25,41,51–55), which are specifically tailored to tackle transcription from arbitrarily complex promoters at the single-cell level. In particular, for analytically solvable gene models such as the common ON-OFF model (14,17,19,22,56–58), and multi-OFF models (also called as gene models of promoter progression (59,60), which are often used to model DNA looping), the mechanisms of transcriptional dynamics have been basically revealed from the analytical distributions available in these models (26). However, for multi-ON gene models for which we can find their prototypes in natural and synthetic systems (48), how promoter dynamics affect transcriptional dynamics remain poorly understood, although an experiment combined with model analysis showed that a TF can result in bursty expression, enabling rapid individual cell responses in the transient and increased cell-cell variability at steady state (47).

A related yet interesting question is how multimodality is generated in gene models. As is well known, bimodal or multimodal gene expression (i.e., the mRNA or protein distribution exhibits two or multiple peaks) is a cause of phenotypic diversity in genetically identical cell populations, and it is critical for population survival in a fluctuating environment (61,62). In some instances, the effect of noise can be amplified by the presence of multistability in a genetic network, thus leading to multiple phenotypes coexisting in a cell population. Individual cells can make transitions between those phenotypes driven by fluctuations in the expression of certain key genes in the network, so as to better adapt them to changes in environments. Given this importance, studying the mechanism of generating multimodality including bimodality is of biological significance.

In this study, we investigate a general gene model, which incorporates the complexity of PS, e.g., multiple ON states. By analysis and simulation, we find that unlike the multi-OFF mechanism (i.e., the promoter has more than one OFF states but only one ON state) that can lead to at most two peaks in the mRNA distribution, the multi-ON mechanism (i.e., the promoter has multiple ON states) can lead to mRNA multimodal distributions with different modes depending on transition and transcription rates, implying that multiple exits of transcription are the essential source of multimodality. Similar mRNA distributions do not necessarily imply that the average ON and OFF time distributions have similar characteristics; the PS can tune the mRNA noise in a nonlinear manner where the nonlinearity depends mainly on the transcriptional rates. These results not only uncover essential characteristics of promoter-mediated transcriptional dynamics but also provide signatures useful for inferring PS based on characteristics of transcriptional outputs.

RESULTS

Multiple exits of transcription are the essential mechanism of generating multimodality

In spite of the complex nature of multistate gene models, one can learn many things from the corresponding master equations. For instance, we used a master equation for the mRNA probability density over function to show whether multimodality (i.e., distributions with three or more modes) can emerge in a gene model of multiple OFF states; we did not find it in parameter regimes under our investigation but found that the region of parameter space (defined by the kinetic rates between promoter states) for which bimodality is observed can be made smaller by having multiple promoter states (59). In particular, slow transitions between promoter states, which can lead to a bimodal mRNA distribution in the common two-state gene model, can result in a unimodal mRNA distribution when the number of promoter states is larger than two. This might explain why bimodal protein distributions appear to be rare in nature. In addition, we also

found slight differences in shape between the mRNA distribution generated by multistate promoters and the one generated by two-state promoters, e.g., the former is flatter than the latter. However, we did not find that multimodality can emerge in a gene model where the promoter has one ON state and multiple OFF states that together constitute a loop. These imply that the multi-OFF mechanism is not the main source of generating multimodality including bimodality.

To clearly show how bimodality or multimodality is generated because of PS, we consider only a simple gene model, where the promoter comprises three activity states that may be either ON or OFF but form a loop. The time-dependent distribution of the mRNA number, denoted by $P(m; t)$, can be computed according to the following:

$$P(m; t) = \sum_{n=m}^{\infty} (-1)^{n-m} \binom{n}{m} b_n(t) \quad (1)$$

where $\binom{n}{m}$ is the combinatorial number of choosing m from n ; b_n represent binomial moments, seeing the Model and Method section for their computation. The numerical simulation has verified that the results obtained by Eq. 1 are in good accord with those obtained by the Gillespie algorithm (63) after the time is sufficiently large (data are not shown). Therefore, we may consider steady-state distributions only. In the Supporting Material, we derive analytical distributions in gene models with specific PS, which include all the distributions in the existing literature as their particular cases. Interestingly, we find that if all the promoter states are ON with the same transcriptional rate, then mRNA follows a Poisson distribution, independently of PS; in other cases, the mRNA distributions can be in general expressed as an algebraic sum of confluent hypergeometric functions. These results themselves are interesting facts, not shown in previous references.

By numerical simulation, we find that the presence of bimodality or multimodality is mainly because of multiple exits of transcription. Moreover, we find that modes of multimodal distribution depend on the transition rates and the transition pattern among promoter activity states. Fig. 1 shows a related example where the promoter comprises several activity states that together form a loop. This gene model can demonstrate six modes of the mRNA distribution, including one peak close to zero, one nonzero peak (hereafter by nonzero peak we mean that the mRNA number corresponding to the peak is an integer of more than one), the combination of the former two, two nonzero peaks, both one peak close to zero and two nonzero peaks, and three nonzero peaks. We observe that the mRNA distribution may exhibit one peak of two different shapes, two peaks of two different shapes, and three peaks with one peak close to zero, depending on transition rates. However, the mRNA distribution in this model cannot exhibit three nonzero peaks as observed in the gene model with three exits of transcription (comparing two distributions indicated by red in the second row of Fig. 1). In other words, the mRNA distribution in models of two ON states exhibits at most two nonzero peaks. The similar conclusion also holds for other similar gene models. All these imply that the multi-ON mechanism is the essential cause of generating multimodality including bimodality.

Mean waiting time distributions exhibit different characteristics albeit similar mRNA distributions

Experiments that reveal the dynamics of transcription initiation at promoters can reveal molecular mechanisms of transcription regulation (64). Several such experiments, where the synthesis of new mRNA molecules was visualized in live cells using single-molecule resolution technology, have been carried out in bacteria or eukaryotic cells

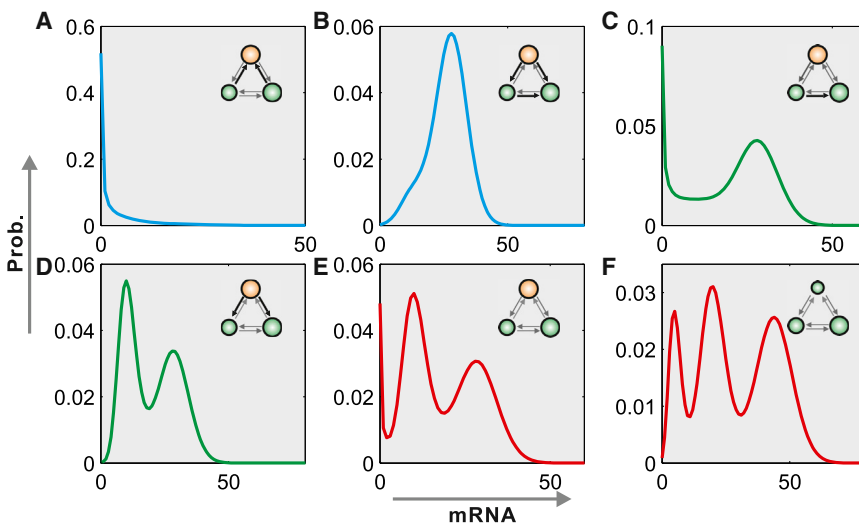


FIGURE 1 Numerical steady-state multimodal distributions of different modes in gene models with the same transition pattern between promoter activity states. The inactive state is denoted by orange whereas active states by green; thick arrows represent large transition rates whereas thin arrows represent small transition rates; large circles correspond to large transcription rates where small circles to small transcription rates. The two distributions indicated by the red imply that multiple exits of transcription are the essential source of generating multimodal distributions. The parameter values used in the computation are found in the Supporting Material. To see this figure in color, go online.

(8,9,39), demonstrating that transcription occurs in a burst fashion. In these experiments, the spatial distribution of polymerases along a gene may bear the fingerprint of the time series of transcription initiation events, which is in turn determined by promoter dynamics. In response to experiments that probe transcription dynamics in single cells, in this study we consider how regulatory architecture modifies the probability distribution of mean waiting times between transcription initiation events. These times are experimentally testable. Here the goal is to provide experimental signatures predicted by theory that are capable of distinguishing between different mechanisms of transcriptional regulation.

First, we give the analytical results. Consider a gene model with the transition matrix among promoter activity states, denoted by $\mathbf{A} = \begin{pmatrix} \mathbf{A}_{00} & \mathbf{A}_{10} \\ \mathbf{A}_{01} & \mathbf{A}_{11} \end{pmatrix}$, where \mathbf{A}_{00} describes the internal transitions between OFF states and the transitions from some ON states to OFF states; \mathbf{A}_{10} and \mathbf{A}_{01} describe transitions from ON to OFF states and from OFF to ON states, respectively; and \mathbf{A}_{11} describes the internal transitions between ON states and the transitions from some OFF states to ON states. Then, the distribution functions for the mean ON and OFF times are computed according to the following formulae:

$$\begin{aligned} f_{off}(\tau) &= \mathbf{u}_L \mathbf{A}_{01} \exp(\mathbf{A}_{00}\tau) \mathbf{A}_{10} \mathbf{u}_L^T, \quad f_{on}(\tau) \\ &= \mathbf{u}_K \mathbf{A}_{10} \exp(\mathbf{A}_{11}\tau) \mathbf{A}_{01} \mathbf{u}_K^T \end{aligned} \quad (2)$$

and, the mean waiting times at OFF and ON states are given by the following:

$$\begin{aligned} \langle \tau_{OFF} \rangle &= \frac{1}{\mathbf{u}_L \mathbf{A}_{10} \mathbf{u}_L^T} \mathbf{u}_L \mathbf{A}_{01} (\mathbf{A}_{00})^{-2} \mathbf{A}_{10} \mathbf{u}_L^T, \quad \langle \tau_{ON} \rangle \\ &= \frac{1}{\mathbf{u}_K \mathbf{A}_{01} \mathbf{u}_K^T} \mathbf{u}_K \mathbf{A}_{10} (\mathbf{A}_{11})^{-2} \mathbf{A}_{01} \mathbf{u}_K^T \end{aligned} \quad (3)$$

where the row vectors $\mathbf{u}_L = (1, 1, \dots, 1)$ and $\mathbf{u}_K = (1, 1, \dots, 1)$ with L and K being the order of \mathbf{A}_{00} and \mathbf{A}_{11} , respectively. See the Model and Method section for their derivation. The above formulae indicate that mean ON and OFF time distributions as well as mean ON and OFF times are easily computed as long as the transition matrix \mathbf{A} is determined. For example, for a gene model where the promoter has two ON states and one OFF state that together constitute a loop, we have the following:

$$\langle \tau_{off} \rangle = \frac{1}{-a_{11}}, \quad \langle \tau_{on} \rangle = \frac{1}{-a_{11}} \frac{c_2 + c_3}{c_1} \quad (4)$$

where c_i is the algebraic complement of the diagonal element a_{ii} of the transition matrix \mathbf{A} ($i = 1, 2, 3$).

Then, we perform numerical simulation. Consider two gene models with different PSs (see insets in Fig. 2 A and B). We observe that two mRNA distributions exhibit the similar bimodal shape with two nontrivial peaks (the bimodality in Fig. 2 A is generated because of two distinct transcription exits whereas the one of two peaks in Fig. 2 B, which is closed to the y axis, is generated because of the cumulating effect of multiple inactive states and the other peak results from the transcription exit), but the mean ON and OFF times display different characteristics. Specifically, for the model in Fig. 2 A, the peak for the mean OFF time is close to zero whereas for the model in Fig. 2 B, the peak for the mean OFF time is away from zero (i.e., nontrivial peak); the mean ON time distributions exhibit different shapes for the time close to zero. Moreover, the time series of the ON states display different dynamical behaviors (compare the insets of Fig. 2 A and B).

The above results indicate that on the one hand, the PS determines the mean ON and OFF times and their distributions; on the other hand, the characteristics of mean ON and

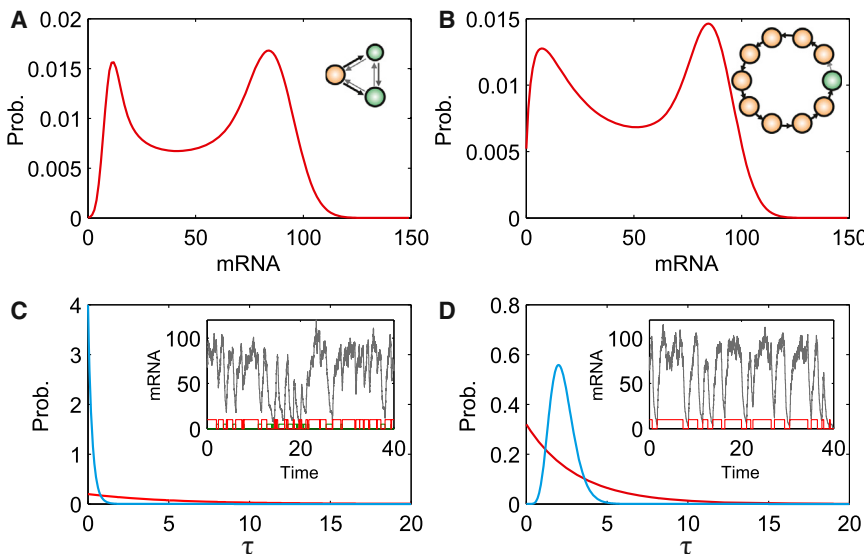


FIGURE 2 Two mRNA number distributions (A, B) look similar, but ON and OFF time distributions (C, D) exhibit different characteristics. In (C) and (D), the red lines correspond to ON whereas the blue lines to OFF. The insets show the time series of mRNAs and describe the time change of ON (red) and OFF states (green). The meaning of symbols in the inset of (A) and (B) is the same as in Fig. 1. The parameter values used in the computation are found in the Supporting Material. To see this figure in color, go online.

OFF time distributions or those of the time series of ON and OFF states or both can provide the additional information on PS, thus enabling a remedy when the characteristics of mRNA distributions are insufficient to infer the PS.

Mean ON and OFF times together can characterize the transcriptional noise

For a gene model, the mRNA noise has two origins: one from the promoter fluctuations (called as the promoter noise) attributable to stochastic transitions among promoter activity states, and the other from stochastic synthesis and degradation of mRNA. The former is called the mRNA external noise whereas the latter the mRNA internal noise. The noise of the two kinds would together contribute to the generation of cell-to-cell variability.

For a multistate gene model, the mRNA internal noise is easy to describe since it can be characterized by the inverse of the mean mRNA number under our assumption. In contrast, describing the mRNA external noise is more difficult because of the complex PS. Therefore, we will focus on the promoter noise in the following.

Note that for the common ON-OFF model, the formula for computing the mRNA noise intensity, denoted by η_m is the following:

$$\eta_m^2 = \frac{1}{\langle m \rangle} + \frac{\langle \tau_{off} \rangle^2}{\langle \tau_{on} \rangle + \langle \tau_{off} \rangle + \langle \tau_{on} \rangle \langle \tau_{off} \rangle} \quad (5)$$

where the first term on the right side of Eq. 5 represents the internal noise of mRNA from transcription, which is the inverse of the mRNA mean; the second term describes the promoter noise, denoted by $\eta_{promoter}^2$, which is a nonlinear function of mean ON and OFF times, denoted by $F(\langle \tau_{on} \rangle, \langle \tau_{off} \rangle)$. In the Model and Method section, we show that for any gene model, the intensity of the noise in mRNA can be computed according to the following:

$$\eta_m^2 = \frac{1}{b_1} + \frac{2b_2 - b_1^2}{b_1^2} \quad (6)$$

where b_1 and b_2 represent the first- and second-order binomial moments, respectively. Note that b_1 is the mRNA mean $\langle m \rangle$. Therefore, in analog to Eq. 5, it is reasonable to adopt the following formula to compute the intensity of the promoter noise:

$$\eta_{promoter}^2 = \frac{2b_2 - b_1^2}{b_1^2} \quad (7)$$

According to the Model and Method section, this formula indicates that the promoter noise depends not only on transition matrix among promoter activity states but also on the transcription matrix since b_1 and b_2 depends on transition

and transcription rates except in particular cases, e.g., all the transcription rates are precisely the same.

For example, consider the gene model studied in the previous subsection, i.e., the one where the promoter has two ON states and one OFF state. We can show the following:

$$b_1 = \frac{\langle \tau_{off} \rangle}{\langle \tau_{on} \rangle + \langle \tau_{off} \rangle} \left(\mu_1 \frac{c_2}{c_1} + \mu_2 \frac{c_3}{c_1} \right) \quad (8)$$

$$b_2 = \frac{1}{2 \det(\mathbf{I} - \mathbf{A})(c_1 + c_2 + c_3)} \left[c_2 \mu_1^2 A_{22}^{(1)} + \mu_1 \mu_2 (c_2 A_{23}^{(1)} + c_3 A_{32}^{(1)}) + c_3 \mu_2^2 A_{33}^{(1)} \right] \quad (9)$$

where $A_{ij}^{(1)}$ are elements of the transposed matrix $(\mathbf{I} - \mathbf{A})^T$, $i, j = 2, 3$. See the next section for their derivation. Therefore, $\eta_{promoter}^2$ depends in general on transcription rates μ_1 and μ_2 unless $\mu_1 = \mu_2$. Note that Eqs. 4, 8, and 9 are easily extended to the case that the promoter has one OFF and multiple ON states that together form a loop or one ON and multiple OFF states that together constitute a loop as well.

Next, we perform numerical analysis for the promoter noise. Fig. 3 plots the dependence of the promoter noise intensity (η^2) computed by Eq. 7 in combination with Eqs. 8 and 9 on the referred quantity $F(\langle \tau_{on} \rangle, \langle \tau_{off} \rangle)$ with Eq. 4. In this figure, different symbols correspond to different PS whereas the same symbols correspond to different transition patterns with the fixed promoter state number and the fixed transcription rates. More precisely, for a fixed PS and transcription rates, we have different dependences of η^2 on $F(\tau_{ON}, \tau_{OFF})$ as the transition rates are randomly changed.

We observe that the differences among transcription rates have important influences on the deviation of the indicated symbols from the line with the slope equal to 1. Specifically, if all the promoter states are ON and all the transcription rates are precisely the same, then there is no promoter noise. In fact, in this case we can show that the corresponding mRNA number follows a Poisson distribution with the characteristic parameter being the common transcription rate, which is independent of the PS (see the Supporting Material). If a part of promoter states are ON and all the corresponding transcription rates are precisely the same, then the dependence of η^2 on $F(\tau_{ON}, \tau_{OFF})$ is basically orientated on this slope, depending on transition rates regardless of the transition pattern among promoter activity states. This case implies that Eq. 7 can be used to quantify the promoter noise whose level is determined completely by the mean ON and OFF times. In other cases, we find that the larger the differences among transcription rates are because of, e.g., the regulation of TFs, the greater is the deviation of the indicated symbols from the line with the slope equal to 1, implying that the Eq. 7 cannot be used to approximate the promoter noise intensity. Such a nonlinear relationship can provide useful information on PS.

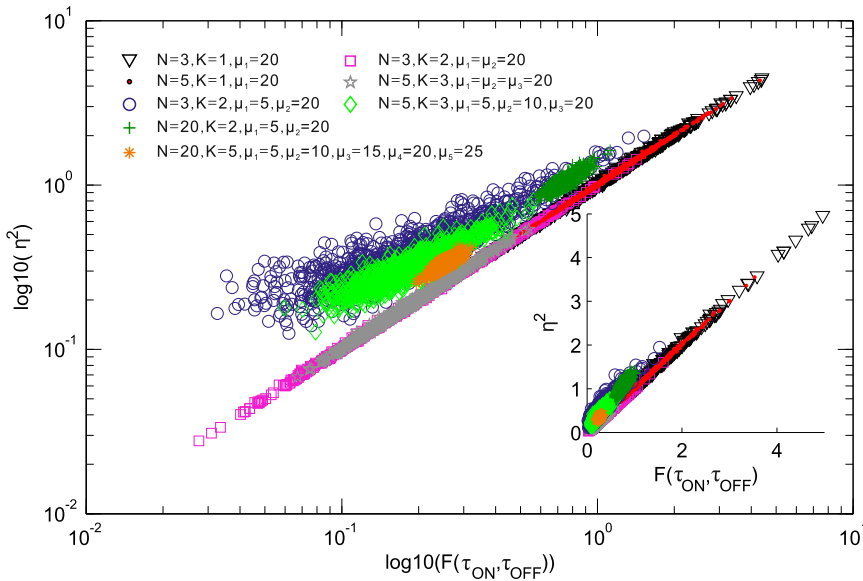


FIGURE 3 As an example of promoter structure, if all the transcription rates are precisely the same, then $\log_{10}(\eta^2)$ vs. $\log_{10} F(\tau_{ON}, \tau_{OFF})$ is located on a line and will deviate from the line otherwise. Here the abscissa represents the logarithm of the quantity $F(\tau_{ON}, \tau_{OFF})$ computed according to Eq. 5, whereas the y axis represents the logarithm of the square of the promoter noise intensities computed by Eq. 7. Different symbols correspond to different promoter structures, whereas the same symbols correspond to different transition patterns with a fixed promoter state number and fixed transcription rates. The inset corresponds to the case without the logarithm. To see this figure in color, go online.

In other words, the promoter noise and also the mRNA noise depend in general on the transition pattern among the promoter activity states as well as the transcription exits; the mean ON and OFF times that are experimentally measurable can be used to characterize the promoter noise and the mRNA noise, thus providing signatures useful for inferring PS.

MODEL AND METHOD

For the convenience of applications, in this section we consider a general gene model where the promoter has several ON and OFF states among which transitions may exist, and we give general formulae for computing steady-state mRNA distributions, mean waiting times, and waiting time distributions. These formulas are very useful and contain previous results obtained in simple gene models (e.g., the common ON-OFF model (14,17,19,22,56–58) and gene models of promoter-progression (25,26)) as their particular cases.

Model description

Assume that the promoter has N states, L states of which are active (denoted by A_i) and the other $K = N - L$ are inactive (denoted by I_j). Let matrices \mathbf{A}_{11} and \mathbf{A}_{00} describe transitions among the active states and among the inactive states, respectively. The matrix \mathbf{A}_{10} describes how the active states transition to the inactive states, and similarly for matrix \mathbf{A}_{01} . These matrices, called as the transition matrices, together describe the PS partially. Denote by m , the number of mRNA molecules, and let $P_k(m)$ represent the distribution that mRNA has m molecules at state- k of the promoter and let $\mathbf{P} = (P_1, \dots, P_N)^T$ represent the column vector. Denote by λ_{ij} , the transition rate from state- j to state- i ($\lambda_{ij} = 0$ means that no transition occurs), the size of which may be regulated by TFs. Denote by $\mathbf{A} = (\lambda_{ij})$, the $N \times N$ transition matrix, which consist of four block matrices \mathbf{A}_{11} , \mathbf{A}_{00} , \mathbf{A}_{10} , and \mathbf{A}_{01} ; and let $\mathbf{\Lambda} = \text{diag}(\mu_1, \dots, \mu_N)$ describe the exits of transcription (called transcription matrix) with μ_i representing the transcription rate of mRNA in state- i ($\mu_i = 0$ means that no transcription takes place). Two matrices \mathbf{A} and $\mathbf{\Lambda}$ together describe the PS completely. Then, the biochemical master equation describing mRNA dynamics takes the following form:

$$\frac{d\mathbf{P}(m; t)}{dt} = \mathbf{A}\mathbf{P}(m; t) + \delta(\mathbf{E} - \mathbf{I})[\mathbf{m}\mathbf{P}(m; t)] + \mathbf{\Lambda}(\mathbf{E}^{-1} - \mathbf{I})[\mathbf{P}(m; t)] \quad (10)$$

where \mathbf{E} and \mathbf{E}^{-1} are shift operators, and \mathbf{I} is the identity operator. Clearly, the first term in Eq. 10 describes dynamics of the promoter with the transition matrix \mathbf{A} that is actually an M-matrix (since the sum of every column elements is equal to zero); the second term describes the degradation dynamics of mRNA with the degradation matrix δ that is a diagonal matrix (throughout this paper, we consider only the same degradation rate for simplicity, and denote it as δ); and the third term describes the exits of transcription with the transcription matrix $\mathbf{\Lambda}$. We point out that the model in Eq. 10 includes all previously studied mRNA expression models as its particular cases.

Computation of mRNA distributions

To solve Eq. 10, we introduce probability-generating functions of the vector form $\mathbf{G} = (G_1, \dots, G_N)^T$ with every component $G_k(z; t) = \sum_{m \geq 0} z^m P_k(m; t)$ for the distributions of the vector form $\mathbf{P}(m; t)$. Then, from Eq. 10, we can derive the following linear system of partial differential equations:

$$\frac{\partial}{\partial t} \mathbf{G} = \mathbf{A}\mathbf{G} - s \frac{\partial}{\partial s} \mathbf{G} + s\mathbf{\Lambda}\mathbf{G} \quad (11)$$

where $s = z - 1$ is taken as a new variable, and all the system parameters are rescaled by δ . Note that Eq. 11 is an equivalent version of the biochemical master equation in Eq. 10 because of the relationship between the probability distribution and the generating function. This equivalence can help us find solutions to Eq. 11. Now, we expand every generating function as $G_k(s) = \sum_{n=0}^{\infty} a_n^{(k)} s^n$. Then, we can show the following:

$$a_k^{(i)}(t) = \sum_{m \geq k} \binom{m}{k} P_i(m; t) \quad (12)$$

where $\binom{m}{k}$ represents the common binomial coefficient. Therefore, the fixed $i, a_k^{(i)}(t)$ are called as binomial moments (65,66) corresponding to the probability $P_i(m; t)$. In particular, $b_k = \sum_{i=1}^N a_k^{(i)}$ are the total binomial

moments corresponding to the total probability $P = \sum_{i=1}^N P_i$. Note that $b_0 = 1$ because of the conservative condition $\sum_{m=0}^{\infty} P(m) = 1$. Denote $\mathbf{a}_n = (a_n^{(1)}, a_n^{(2)}, \dots, a_n^{(N)})^T$. Then, from Eq. 11 we attain the following:

$$\frac{d}{dt} \mathbf{a}_n = -(\mathbf{nI} - \mathbf{A})\mathbf{a}_n + \mathbf{\Lambda}\mathbf{a}_{n-1} \quad (13)$$

which is a linear ordinary differential equation that is easily solved. Once all $a_k^{(i)}(t)$ are given, then the distribution $P_i(m; t)$ is computed according to the following:

$$P_i(m; t) = \sum_{k \geq m} (-1)^{k-m} \binom{k}{m} a_k^{(i)}(t) \quad (14)$$

In particular, we can give analytical results at steady state. In fact, if we denote the total steady-state generating function $G(s) = \sum_{k=1}^N G_k(s)$ with $G(0) = 1$ and Taylor expand $G(s) = \sum_{n=0}^{\infty} b_n s^n$, then it is not difficult to show $\mu_1 G_1(s) + \dots + \mu_N G_N(s) = G'(s)$ since \mathbf{A} is an M-matrix. We find that b_n takes the following form:

$$b_n = \frac{1}{n} \sum_{i=1}^N \mu_i a_{n-1}^{(i)} = \frac{1}{n} \mathbf{u}_N \mathbf{\Lambda} \mathbf{a}_{n-1} \quad (15)$$

which is useful for deriving analytical distributions, where $\mathbf{u}_N = (1, 1, \dots, 1)$ is a row vector and $\mathbf{a}_n = (a_n^{(1)}, a_n^{(2)}, \dots, a_n^{(N)})^T$ is a column vector. Substituting the expansions of $G_k(s)$ into Eq. 11 at steady state, we see that the vector \mathbf{a}_n satisfies the following algebraic equations:

$$(\mathbf{nI} - \mathbf{A})\mathbf{a}_n = \mathbf{\Lambda}\mathbf{a}_{n-1}, \quad n = 1, 2, \dots \quad (16)$$

where \mathbf{a}_0 can be analytically given (see the Supporting Material). When Eq. 16 is combined with Eq. 15, then b_n can be formally expressed as the following:

$$b_n = \frac{1}{n! \prod_{k=1}^n \det(k\mathbf{I} - \mathbf{A})} \mathbf{u}_N \prod_{k=n}^1 [(k\mathbf{I} - \mathbf{A})^* \mathbf{\Lambda}] \mathbf{a}_0, \quad (17)$$

$$n = 1, 2, \dots$$

where $(k\mathbf{I} - \mathbf{A})^*$ and $\det(k\mathbf{I} - \mathbf{A})$ are the adjacency matrix and the determinant of matrix $(k\mathbf{I} - \mathbf{A})$, respectively; \mathbf{a}_0 is given in the Supporting Material. Such a formal expression of b_n does not impose any condition on the transition matrix \mathbf{A} and the transcription matrix $\mathbf{\Lambda}$.

Equation 17 indicates that Eq. 11 at steady state is solvable. Furthermore, Eq. 10 at steady state is also solvable. In some cases, the steady-state distributions can be expressed by confluent hypergeometric functions (67–70). Refer to the Supporting Material. In any case, $P(m)$ can be approximately computed up to a desired accuracy because $b_n \rightarrow 0$ as $n \rightarrow \infty$ (65). We point out that such a binomial moment method can be generalized to the stochastic analysis of any reaction networks. The details will be published elsewhere.

Computation of waiting time distributions and mean waiting times

Given a transition matrix \mathbf{A} that is expressed as a block matrix of the form $\begin{pmatrix} \mathbf{A}_{00} & \mathbf{A}_{10} \\ \mathbf{A}_{01} & \mathbf{A}_{11} \end{pmatrix}$, where \mathbf{A}_{00} , \mathbf{A}_{11} , \mathbf{A}_{10} and \mathbf{A}_{01} are the $K \times K$, $L \times L$, $K \times L$ and $L \times K$ matrix, respectively, we derive the waiting-time distribution functions for ON and OFF states. Assume that the promoter states begin to transition from OFF (ON) to ON (OFF) at time $t = 0$, and we define $Q_i^{(1)}(\tau)$ ($i = 1, \dots, L$) and $Q_k^{(0)}(\tau)$ ($k = 1, \dots, K$) as the subsequent “survival” probability that the promoter is still at the i^{th} ON and at the k^{th}

OFF state at the time $t = \tau > 0$, respectively. Then, the master equations for $Q_i^{(0)}(\tau)$ and $Q_i^{(1)}(\tau)$ take the following form:

$$\frac{d\mathbf{Q}^{(0)}(\tau)}{d\tau} = \mathbf{A}_{00}\mathbf{Q}^{(0)}(\tau) \quad (18)$$

$$\frac{d\mathbf{Q}^{(1)}(\tau)}{d\tau} = \mathbf{A}_{11}\mathbf{Q}^{(1)}(\tau)$$

respectively, where $\mathbf{Q}^{(0)}(\tau) = (Q_1^{(0)}(\tau), \dots, Q_K^{(0)}(\tau))^T$, $\mathbf{Q}^{(1)}(\tau) = (Q_1^{(1)}(\tau), \dots, Q_L^{(1)}(\tau))^T$. The solution to Eq. 18 can be expressed as the following:

$$\mathbf{Q}^{(0)}(\tau) = \exp(\mathbf{A}_{00}\tau)\mathbf{Q}^{(0)}(0) \quad (19)$$

$$\mathbf{Q}^{(1)}(\tau) = \exp(\mathbf{A}_{11}\tau)\mathbf{Q}^{(1)}(0)$$

Thus, for two given sets of initial survival probabilities $\{Q_1^{(0)}(0), \dots, Q_K^{(0)}(0)\}$ and $\{Q_1^{(1)}(0), \dots, Q_L^{(1)}(0)\}$, the distribution functions for the dwell times τ at the OFF and ON states are given by the following:

$$\tilde{f}_{off}(\tau) = \mathbf{u}_L \mathbf{A}_{01} \mathbf{Q}^{(0)}(\tau) = \mathbf{u}_L \mathbf{A}_{01} \exp(\mathbf{A}_{00}\tau) \mathbf{Q}^{(0)}(0) \quad (20)$$

$$\tilde{f}_{on}(\tau) = \mathbf{u}_K \mathbf{A}_{10} \mathbf{Q}^{(1)}(\tau) = \mathbf{u}_K \mathbf{A}_{10} \exp(\mathbf{A}_{11}\tau) \mathbf{Q}^{(1)}(0)$$

From Eq. 20, we can see that each of two distribution functions is in general a linear combination of exponential functions of the form $e^{\lambda\tau}$, so the result is an extension of that found in previous studies (71–73). Furthermore, the mean OFF and ON times can be computed by substituting $\tilde{f}_{off}(\tau)$, $\tilde{f}_{on}(\tau)$ into the general expression $\langle \tau \rangle = \int_0^{\infty} \tau f(\tau) d\tau$, which attains the following:

$$\langle \tilde{\tau}_{OFF} \rangle = \int_0^{\infty} \tau \mathbf{u}_L \mathbf{A}_{01} \exp(\mathbf{A}_{00}\tau) \mathbf{Q}^{(0)}(0) d\tau = \mathbf{u}_L \mathbf{A}_{01} (\mathbf{A}_{00})^{-2} \mathbf{Q}^{(0)}(0) \quad (21)$$

$$\langle \tilde{\tau}_{ON} \rangle = \int_0^{\infty} \tau \mathbf{u}_K \mathbf{A}_{10} \exp(\mathbf{A}_{11}\tau) \mathbf{Q}^{(1)}(0) d\tau = \mathbf{u}_K \mathbf{A}_{10} (\mathbf{A}_{11})^{-2} \mathbf{Q}^{(1)}(0)$$

However, Eq. 21 is not the resulting mean waiting times at OFF and ON states since the initial survival probabilities $\mathbf{Q}^{(0)}(\tau)$ and $\mathbf{Q}^{(1)}(\tau)$ depend on the transition pattern among ON and OFF states. For a given PS, to obtain the total OFF and ON dwell times, we have to average $\langle \tilde{\tau}_{OFF} \rangle$ or $\langle \tilde{\tau}_{ON} \rangle$ over all such ON states that transition to OFF states or over all such OFF states that transition to ON states. For example, to compute the resulting $f_{on}(\tau)$, one should choose $Q_i^{(1)}(0) = (\sum_{l=1}^L a_{il}^{(0 \rightarrow 1)} / \sum_{k=1}^K \sum_{l=1}^L a_{kl}^{(0 \rightarrow 1)}) \delta_{ij}$ ($j, i = 1, \dots, K$) as the initial conditions, where δ_{ij} is the Kronecker delta; and for clarity and convenience, we let $a_{ik}^{(0 \rightarrow 1)}$ represent the transition rate from the k^{th} OFF state to the i^{th} ON state (similarly, $a_{ik}^{(1 \rightarrow 0)}$, $a_{ik}^{(0 \rightarrow 0)}$, and $a_{ik}^{(1 \rightarrow 1)}$). The resulting distribution functions for the mean ON and OFF times are computed according to the following:

$$f_{off}(\tau) = \mathbf{u}_L \mathbf{A}_{01} \exp(\mathbf{A}_{00}\tau) \mathbf{A}_{10} \mathbf{u}_L^T \quad (22)$$

$$f_{on}(\tau) = \mathbf{u}_K \mathbf{A}_{10} \exp(\mathbf{A}_{11}\tau) \mathbf{A}_{01} \mathbf{u}_K^T$$

whereas the resulting mean dwell times at OFF and ON states are given by the following:

$$\langle \tau_{OFF} \rangle = \frac{1}{\mathbf{u}_L \mathbf{A}_{10} \mathbf{u}_L^T} \mathbf{u}_L \mathbf{A}_{01} (\mathbf{A}_{00})^{-2} \mathbf{A}_{10} \mathbf{u}_L^T \quad (23)$$

$$\langle \tau_{ON} \rangle = \frac{1}{\mathbf{u}_K \mathbf{A}_{01} \mathbf{u}_K^T} \mathbf{u}_K \mathbf{A}_{10} (\mathbf{A}_{11})^{-2} \mathbf{A}_{01} \mathbf{u}_K^T$$

CONCLUSIONS AND DISCUSSIONS

Transcription is a complex biochemical process, involving recruitment of TFs and DNA polymerases, chromatin

remodeling, and a sequence of transitions between activity states of the promoter. Previous studies have shown that transcription occur either as pulsatile bursts or as Poisson-like accumulations, but how promoter dynamics quantitatively and qualitatively affect transcriptional dynamics remains to be fully explored. In this study, we have analyzed gene models that incorporate the complexity of PS, focusing on the effects of the multi-ON mechanism on transcriptional dynamics. We have shown that multiple exits of transcription are the essential source of generating multimodal mRNA distributions (Fig. 1). In addition, we have demonstrated that in the $(\log F, \eta^2)$ plane, the larger the differences among transcription rates are, the higher is the nonlinearity describing the dependence of the transcriptional noise on PS (Fig. 3). These qualitative characteristics that still hold in more complex gene models reveal the essential mechanism of how promoter dynamics affect transcriptional dynamics.

In a previous study (59), we demonstrated that slow transitions between promoter states that can lead to a bimodal distribution of the mRNA copy number as observed in the two-state promoter model can also result in a unimodal distribution when the number of promoter states is larger than two. In this study, we have shown that multiple exits of transcription can lead to multimodal mRNA distributions, but there are exceptions, e.g., when all the promoter states are ON with the same transcription rate, which result in a Poisson distribution independent of PS. A more careful investigation of this issue would clarify whether distributions from multistate promoters can be clearly distinguished from those produced by two-state ON/OFF promoters, and how many events would be needed to see the difference in the experimental data. If it were found that two-state and multistate promoters cannot be clearly distinguished by virtue of the distribution, it would call into the question of how seriously we should take the fitting parameters extracted from fitting experimental distributions of mRNA numbers. We would like to point out that our model cannot exhibit bistability in the deterministic case but can exhibit stochastic multimodality including bistability in the stochastic case according to the definition in a previous study (74).

For gene models with more complex PS, which would correspond to complicated mathematical forms expressed in the master equation, the first- and second-order moments of mRNA can be derived using the same simple method used to compute the moments in the two-state promoter model. In fact, the formulas for the moments apply to any other promoter model as well, regardless of the number of promoter states. In addition, the focus of the analysis in this study is on mRNA noise, but the mathematical approaches to compute mRNA statistics can be easily extended to compute protein distributions and their moments as well. For instance, for a promoter switching between two different states, in the limit when mRNA lifetime is much shorter than protein lifetime, the ratio of

variance over mean for protein takes the form very similar to that for mRNA. At least in this limit, all of the conclusions about how complex promoter dynamics affect mRNA noise would be qualitatively true for protein noise. In other words, all the related extension and computation are easily carried out in the case that TFs do not temporally regulate transition rates among the promoter activity states (47). However, the regulation of TFs is often dynamic and noisy. In this case, what is the dependence relationship between transcriptional output and TFs (as inputs) or how the latter affects the former deserves further study since analytical results (e.g., analytical distributions as derived in this paper) are in general unavailable. If the gene promoter has only one ON and one OFF state (i.e., the common gene model), one can use the input-associated Signed Activation Time (iSAT) index introduced in another study (75), which concisely captures an intrinsic temporal property at either the ON or OFF state, to characterize the input-output relation including the transcriptional noise. In the case of complex PS (i.e., the promoter has multiple ON or OFF states or both), the iSAT index would be still effective in quantifying this relation but its definition seems to need modification. By analyzing this modified index, it is possible to reveal which of the multiple ON and OFF mechanisms is dominant in buffering the transcriptional noise. In particular, it is possible that there is a tradeoff between achieving good noise buffering in the ON versus the OFF states as shown in a previous study (75). The further study is under way.

Finally, as with any quantitative model, especially one attempting to describe processes within a living cell, it is important to understand the limitations of the chemical master equation description of transcription presented in this paper. Particular care has to be taken when using mathematical models in conjunction with experimental data to test specific hypotheses about biological mechanisms. Even for most informative models, there would be a discrepancy between the model predictions and experimental data. This offers the opportunity to discard incorrect assumptions about the mechanisms of interest. But to reach such strong conclusions, we must have some degree of certainty that the discrepancy between the model predictions and experimental outcomes is because of the deficiencies in our understanding of the underlying biological mechanism, and not because of spurious experimental effects that have nothing to do with the biological process of interest. This is a particularly challenging problem when using cell-to-cell or temporal variability of cellular outputs as the experimental signature of a biological process because there are typically multiple sources of this variability, and we are interested in the only one of them. For example, if the variability of the transcriptional output of a single cell is used as the experimental signature of promoter dynamics, then we must make sure that this is indeed the dominant source of the observed fluctuations.

SUPPORTING MATERIAL

Supplementary analyses and computations are available at [http://www.biophysj.org/biophysj/supplemental/S0006-3495\(13\)05800-1](http://www.biophysj.org/biophysj/supplemental/S0006-3495(13)05800-1).

This work was supported by grants 91230204, 2014CB964703, 11005162, 20120171110047, 2010CB945400, 2012J2200017, 20100171120039, and 201003383 from the National Scientific Foundation and Department of Science and Technology, P.R. China.

REFERENCES

- Ozbudak, E. M., M. Thattai, ..., A. van Oudenaarden. 2002. Regulation of noise in the expression of a single gene. *Nat. Genet.* 31:69–73.
- Elowitz, M. B., A. J. Levine, ..., P. S. Swain. 2002. Stochastic gene expression in a single cell. *Science*. 297:1183–1186.
- Blake, W. J., M. KAern, ..., J. J. Collins. 2003. Noise in eukaryotic gene expression. *Nature*. 422:633–637.
- Raser, J. M., and E. K. O’Shea. 2004. Control of stochasticity in eukaryotic gene expression. *Science*. 304:1811–1814.
- Boettiger, A. N., and M. Levine. 2009. Synchronous and stochastic patterns of gene activation in the *Drosophila* embryo. *Science*. 325:471–473.
- Raj, A., S. A. Rifkin, ..., A. van Oudenaarden. 2010. Variability in gene expression underlies incomplete penetrance. *Nature*. 463:913–918.
- Eldar, A., and M. B. Elowitz. 2010. Functional roles for noise in genetic circuits. *Nature*. 467:167–173.
- Golding, I., J. Paulsson, ..., E. C. Cox. 2005. Real-time kinetics of gene activity in individual bacteria. *Cell*. 123:1025–1036.
- Raj, A., C. S. Peskin, ..., S. Tyagi. 2006. Stochastic mRNA synthesis in mammalian cells. *PLoS Biol.* 4:e309.
- Suter, D. M., N. Molina, ..., F. Naef. 2011. Mammalian genes are transcribed with widely different bursting kinetics. *Science*. 332:472–474.
- Harper, C. V., B. Finkenstädt, ..., M. R. White. 2011. Dynamic analysis of stochastic transcription cycles. *PLoS Biol.* 9:e1000607.
- Spiller, D. G., C. D. Wood, ..., M. R. White. 2010. Measurement of single-cell dynamics. *Nature*. 465:736–745.
- Berg, O. G. 1978. A model for the statistical fluctuations of protein numbers in a microbial population. *J. Theor. Biol.* 71:587–603.
- Peccoud, J., and B. Ycart. 1995. Markovian modeling of gene product synthesis. *Theor. Popul. Biol.* 48:222–234.
- McAdams, H. H., and A. Arkin. 1997. Stochastic mechanisms in gene expression. *Proc. Natl. Acad. Sci. USA*. 94:814–819.
- Thattai, M., and A. van Oudenaarden. 2001. Intrinsic noise in gene regulatory networks. *Proc. Natl. Acad. Sci. USA*. 98:8614–8619.
- Kepler, T. B., and T. C. Elston. 2001. Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations. *Biophys. J.* 81:3116–3136.
- Paulsson, J. 2004. Summing up the noise in gene networks. *Nature*. 427:415–418.
- Paulsson, J. 2005. Models of stochastic gene expression. *Phys. Life Rev.* 2:157–175.
- Raser, J. M., and E. K. O’Shea. 2005. Noise in gene expression: origins, consequences, and control. *Science*. 309:2010–2013.
- Friedman, N., L. Cai, and X. S. Xie. 2006. Linking stochastic dynamics to population distribution: an analytical framework of gene expression. *Phys. Rev. Lett.* 97:168302.
- Shahrezaei, V., and P. S. Swain. 2008. Analytical distributions for stochastic gene expression. *Proc. Natl. Acad. Sci. USA*. 105:17256–17261.
- Pedraza, J. M., and J. Paulsson. 2008. Effects of molecular memory and bursting on fluctuations in gene expression. *Science*. 319:339–343.
- Sánchez, A., and J. Kondev. 2008. Transcriptional control of noise in gene expression. *Proc. Natl. Acad. Sci. USA*. 105:5081–5086.
- Sánchez, A., H. G. Garcia, ..., J. Kondev. 2011. Effect of promoter architecture on the cell-to-cell variability in gene expression. *PLoS Comput. Biol.* 7:e1001100.
- Sánchez, A., S. Choubey, and J. Kondev. 2013. Stochastic models of transcription: From single molecules to single cells. *Methods*. 62:13–25.
- Jia, T., and R. V. Kulkarni. 2011. Intrinsic noise in stochastic models of gene expression with molecular memory and bursting. *Phys. Rev. Lett.* 106:058102.
- Dobrzynski, M., and F. J. Bruggeman. 2009. Elongation dynamics shape bursty transcription and translation. *Proc. Natl. Acad. Sci. USA*. 106:2583–2588.
- Rinott, R., A. Jaimovich, and N. Friedman. 2011. Exploring transcription regulation through cell-to-cell variability. *Proc. Natl. Acad. Sci. USA*. 108:6329–6334.
- Huh, D., and J. Paulsson. 2011. Non-genetic heterogeneity from stochastic partitioning at cell division. *Nat. Genet.* 43:95–100.
- Hager, G. L., J. G. McNally, and T. Misteli. 2009. Transcription dynamics. *Mol. Cell*. 35:741–753.
- Vilar, J. M. G., and L. Saiz. 2010. CplexA: a Mathematica package to study macromolecular-assembly control of gene expression. *Bioinformatics*. 26:2060–2061.
- Hornung, G., R. Bar-Ziv, ..., N. Barkai. 2012. Noise-mean relationship in mutated promoters. *Genome Res.* 22:2409–2417.
- Halme, A., S. Bumgarner, ..., G. R. Fink. 2004. Genetic and epigenetic regulation of the FLO gene family generates cell-surface variation in yeast. *Cell*. 116:405–415.
- Octavio, L. M., K. Gedeon, and N. Maheshri. 2009. Epigenetic and conventional regulation is distributed among activators of FLO11 allowing tuning of population-level heterogeneity in its expression. *PLoS Genet.* 5:e1000673.
- Weinberger, L., Y. Voicheck, ..., N. Barkai. 2012. Expression noise and acetylation profiles distinguish HDAC functions. *Mol. Cell*. 47:193–202.
- Stavreva, D. A., L. Varticovski, and G. L. Hager. 2012. Complex dynamics of transcription regulation. *Biochim. Biophys. Acta*. 1819:657–666.
- Sánchez, A., S. Choubey, and J. Kondev. 2013. Regulation of noise in gene expression. *Annu. Rev. Biophys.* 42:469–491.
- Chubb, J. R., T. Trcek, ..., R. H. Singer. 2006. Transcriptional pulsing of a developmental gene. *Curr. Biol.* 16:1018–1025.
- Chubb, J. R., and T. B. Liverpool. 2010. Bursts and pulses: insights from single cell studies into transcriptional mechanisms. *Curr. Opin. Genet. Dev.* 20:478–484.
- Boeger, H., J. Griesenbeck, and R. D. Kornberg. 2008. Nucleosome retention and the stochastic nature of promoter chromatin remodeling for transcription. *Cell*. 133:716–726.
- Larson, D. R. 2011. What do expression dynamics tell us about the mechanism of transcription? *Curr. Opin. Genet. Dev.* 21:591–599.
- Mao, C., C. R. Brown, ..., H. Boeger. 2010. Quantitative analysis of the transcription control mechanism. *Mol. Syst. Biol.* 6:431.
- Mariani, L., E. G. Schulz, ..., T. Höfer. 2010. Short-term memory in gene induction reveals the regulatory principle behind stochastic IL-4 expression. *Mol. Syst. Biol.* 6:359.
- Miller-Jensen, K., S. S. Dey, ..., A. P. Arkin. 2011. Varying virulence: epigenetic control of expression noise and disease processes. *Trends Biotechnol.* 29:517–525.
- Gama-Castro, S., H. Salgado, ..., J. Collado-Vides. 2011. RegulonDB version 7.0: transcriptional regulation of *Escherichia coli* K-12 integrated within genetic sensory response units (Sensor Units). *Nucleic Acids Res. Suppl.* D98–D105.

47. Blake, W. J., G. Balázsi, ..., J. J. Collins. 2006. Phenotypic consequences of promoter-mediated transcriptional noise. *Mol. Cell.* 24:853–865.
48. To, T.-L., and N. Maheshri. 2010. Noise can induce bimodality in positive transcriptional feedback loops without bistability. *Science.* 327:1142–1145.
49. Bintu, L., N. E. Buchler, ..., R. Phillips. 2005. Transcriptional regulation by the numbers: models. *Curr. Opin. Genet. Dev.* 15:116–124.
50. Bintu, L., N. E. Buchler, ..., R. Phillips. 2005. Transcriptional regulation by the numbers: applications. *Curr. Opin. Genet. Dev.* 15:125–135.
51. Coulon, A., O. Gandrillon, and G. Beslon. 2010. On the spontaneous stochastic dynamics of a single gene: complexity of the molecular interplay at the promoter. *BMC Syst. Biol.* 4:2.
52. Simpson, M. L., C. D. Cox, and G. S. Saylor. 2004. Frequency domain chemical Langevin analysis of stochasticity in gene transcriptional regulation. *J. Theor. Biol.* 229:383–394.
53. Höfer, T., and M. J. Rasch. 2005. On the kinetic design of transcription. *Genome Inform.* 16:73–82.
54. Carey, L. B., D. van Dijk, ..., E. Segal. 2013. Promoter sequence determines the relationship between expression level and noise. *PLoS Biol.* 11:e1001528.
55. Brown, C. R., C. Mao, ..., H. Boeger. 2013. Linking stochastic fluctuations in chromatin structure and gene expression. *PLoS Biol.* 11:e1001621.
56. Karmakar, R., and I. Bose. 2004. Graded and binary responses in stochastic gene expression. *Phys. Biol.* 1:197–204.
57. Iyer-Biswas, S., F. Hayot, and C. Jayaprakash. 2009. Stochasticity of gene products from transcriptional pulsing. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 79:031911.
58. Mugler, A., A. M. Walczak, and C. H. Wiggins. 2009. Spectral solutions to stochastic models of gene expression with bursts and regulation. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 80:041921.
59. Zhang, J., L. Chen, and T. Zhou. 2012. Analytical distribution and tunability of noise in a model of promoter progress. *Biophys. J.* 102:1247–1257.
60. Zhou, T., and J. Zhang. 2012. Analytical results for a multistate gene model. *SIAM J. Appl. Math.* 72:789–818.
61. Fraser, D., and M. Kaern. 2009. A chance at survival: gene expression noise and phenotypic diversification strategies. *Mol. Microbiol.* 71:1333–1340.
62. Acar, M., J. T. Mettetal, and A. van Oudenaarden. 2008. Stochastic switching as a survival strategy in fluctuating environments. *Nat. Genet.* 40:471–475.
63. Gillespie, D. T. 1977. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* 81:2340–2361.
64. Friedman, L. J., and J. Gelles. 2012. Mechanism of transcription initiation at an activator-dependent promoter defined by single-molecule observation. *Cell.* 148:679–689.
65. Takacs, L. 1965. A moment problem. *J. Austral. Math. Soc.* 5:487–490.
66. Zhang, J., Q. Nie, ..., T. Zhou. 2013. An effective method for computing the noise in biochemical networks. *J. Chem. Phys.* 138:084106.
67. Slater, L. J. 1960. *Confluent Hypergeometric Functions*. Cambridge University Press, Cambridge, UK.
68. Okubo, K. 1971. Connection problems for systems of linear differential equations. *Lect. Notes in Math.* 243:238–248.
69. Schäfke, R. 1998. Confluence of several regular singular points into an irregular singular point. *J. Dyn. Contr. Syst.* 4:401–424.
70. Balsler, W., and C. Röscheisen. 2009. Solving the hypergeometric system of Okubo type in terms of a certain generalized hypergeometric function. *J. Differ. Equ.* 247:2485–2494.
71. Tu, Y. 2008. The nonequilibrium mechanism for ultrasensitivity in a biological switch: sensing by Maxwell’s demons. *Proc. Natl. Acad. Sci. USA.* 105:11737–11741.
72. Li, G., and H. Qian. 2002. Kinetic timing: a novel mechanism that improves the accuracy of GTPase timers in endosome fusion and other biological processes. *Traffic.* 3:249–255.
73. Qian, H. 2007. Phosphorylation energy hypothesis: open chemical systems and their biological functions. *Annu. Rev. Phys. Chem.* 58:113–142.
74. Bishop, L. M., and H. Qian. 2010. Stochastic bistability and bifurcation in a mesoscopic signaling system with autocatalytic kinase. *Biophys. J.* 98:1–11.
75. Chen, M., L. Wang, ..., Q. Nie. 2013. Noise attenuation in the ON and OFF states of biological switches. *ACS Synth Biol.* 2:587–593.

Supporting Online Material for

Promoter-mediated transcriptional dynamics

Jiajun Zhang and Tianshou Zhou*

School of Mathematics and Computational Science, Sun Yat-Sen University,

Guangzhou 510275, P. R. China

In this supplementary material, we first derive the analytical expression of \mathbf{a}_0 used in the main text, then give some detailed processes for the derivation of analytical distributions in particular cases, and finally list some parameter values used in figures shown in the main text.

A. Analytical expression of \mathbf{a}_0

Based on the iterative Eq. (16) with $n=0$ in the main article and using the conservative condition for probability, we have the following algebraic equation

$$\begin{aligned} \mathbf{u}_N \mathbf{a}_0 &= 1 \\ \mathbf{A} \mathbf{a}_0 &= \mathbf{0} \end{aligned} \tag{A1}$$

To give the analytical expressions of components of the vector \mathbf{a}_0 , we let \mathbf{M}_k represent an $(N-1) \times (N-1)$ matrix, which is the minor one of the $N \times N$ matrix \mathbf{A} by crossing out the k th row and k th column of its entry a_{kk} . Denote by $0, -\alpha_1, -\alpha_2, \dots, -\alpha_{N-1}$ the eigenvalues of matrix \mathbf{A} and by $-\beta_1^{(k)}, -\beta_2^{(k)}, \dots, -\beta_{N-1}^{(k)}$ the eigenvalues of matrix \mathbf{M}_k . Note that $-\mathbf{A}$ is an M-matrix, so the real part of each α_i is positive. Therefore, with loss of generality, we assume $\alpha_i > 0$. With the similar reason, we also assume $\beta_i^{(k)} > 0$. In addition, we introduce two symbols

$$f_{\mathbf{A}}(n) = \det(n\mathbf{I}_N - \mathbf{A}) = n \prod_{i=1}^{N-1} (n + \alpha_i), \quad f_{\mathbf{M}_k}(n) = \det(n\mathbf{I}_{N-1} - \mathbf{M}_k) = \prod_{i=1}^{N-1} (n + \beta_i^{(k)}) \tag{A2}$$

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

$$a_0^{(k)} = \prod_{i=1}^{N-1} \frac{\beta_i^{(k)}}{\alpha_i}, \quad 1 \leq k \leq N \quad (\text{A3})$$

In fact, $\mathbf{A}\mathbf{A}^* = \det(\mathbf{A})\mathbf{I} = \mathbf{0}$ due to $\det(\mathbf{A}) = 0$. In addition, as a consequence

of Laplace's formula for the determinant of matrix \mathbf{A} , we have $\mathbf{A} \begin{pmatrix} \det(\mathbf{M}_1) \\ \vdots \\ \det(\mathbf{M}_N) \end{pmatrix} = \mathbf{0}$,

where $\det(\mathbf{M}_k) = (-1)^{N-1} \prod_{i=1}^{N-1} \beta_i^{(k)}$. Note that the null space of \mathbf{A} is of one dimension

due to $\text{rank}(\mathbf{A}) = N-1$. Thus, combining $\mathbf{A}\mathbf{a}_0 = \mathbf{0}$, we know that $\mathbf{a}_0 = c \begin{pmatrix} \det(\mathbf{M}_1) \\ \vdots \\ \det(\mathbf{M}_N) \end{pmatrix}$,

where c is a constant. The condition $\mathbf{u}_N \mathbf{a}_0 = 1$ gives $c = 1 / \sum_{k=1}^N \det(\mathbf{M}_k)$. A direct

computation yields $\left. \frac{df_{\mathbf{A}}(\mu)}{d\mu} \right|_{\mu=0} = \prod_{i=1}^{N-1} \alpha_i$. On the other hand, Jacobi's formula gives

$$\left. \frac{df_{\mathbf{A}}(\mu)}{d\mu} \right|_{\mu=0} = \left. \frac{d \det(\mu \mathbf{I}_N - \mathbf{A})}{d\mu} \right|_{\mu=0} = \text{tr} \left((\mu \mathbf{I}_N - \mathbf{A})^* \frac{d(\mu \mathbf{I}_N - \mathbf{A})}{d\mu} \right) \Big|_{\mu=0} = (-1)^{N-1} \sum_{k=1}^N \det(\mathbf{M}_k) = \sum_{k=1}^N \prod_{i=1}^{N-1} \beta_i^{(k)}$$

The combination of both yields the equality $\sum_{k=1}^N \prod_{i=1}^{N-1} \beta_i^{(k)} = \prod_{i=1}^{N-1} \alpha_i$. Using this fact

combined with the expressions of c and $\det(\mathbf{M}_k)$, we immediately know that Eqn.

(A3) holds.

Finally, we point out that if matrix \mathbf{A} is symmetric, then the adjacency matrix

$(n\mathbf{I}_N - \mathbf{A})^*$ is also symmetric for any n .

B. Derivation of analytical distributions in particular cases

Case 1 $\Lambda = \mu \mathbf{I}_N$

Such a case corresponds to a model of constitutive gene expression with the

same transcription rate. Using the above symbols, we have $b_n = \mathbf{u}_N \mathbf{a}_n$ due to $b_n = \sum_{i=1}^N a_n^{(i)}$, and $b_n = \frac{1}{n} \mathbf{u}_N \Lambda \mathbf{a}_{n-1} = \frac{\mu}{n} \mathbf{u}_N \mathbf{a}_{n-1}$. Thus, we obtain $b_n = \frac{\mu}{n} b_{n-1} = \frac{\mu^n}{n!}$. According to Eq. (1) in the main text, we compute and obtain

$$P(m) = \sum_{k=m}^{\infty} (-1)^{k-m} \binom{k}{m} b_k = \sum_{k=m}^{\infty} (-1)^{k-m} \binom{k}{m} \frac{\mu^k}{k!} = e^{-\mu} \frac{\mu^m}{m!}, \quad m = 0, 1, 2, \dots \quad (\text{B1})$$

which is a Poisson distribution with characteristic parameter μ . Eqn. (B1) indicates that the mRNA distribution is independent of the pattern of transitions between activity states of the promoter. In other words, whatever transitions among activity states are, does the static mRNA obey the same Poisson distribution determined completely by the transcription rate.

$$\text{Case 2 } \Lambda = \mu \begin{pmatrix} \mathbf{0}^{(N-1)} & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix}$$

Such a case corresponds to the gene model where the promoter has one ON and multiple OFFs. It follows from Eq. (16) in the main text that

$$\begin{aligned} \begin{pmatrix} * \\ a_n^{(N)} \end{pmatrix} &= \mathbf{a}_n = \frac{1}{\det(n\mathbf{I} - \mathbf{A})} (n\mathbf{I} - \mathbf{A})^* \Lambda \mathbf{a}_{n-1} = \frac{\mu}{\det(n\mathbf{I} - \mathbf{A})} \begin{pmatrix} * & * \\ * & \det(n\mathbf{I}_{N-1} - \mathbf{M}_N) \end{pmatrix} \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix} \begin{pmatrix} * \\ a_{n-1}^{(N)} \end{pmatrix} \\ &= \frac{\mu}{f_A(n)} \begin{pmatrix} * \\ f_{M_N}(n) a_{n-1}^{(N)} \end{pmatrix} = \begin{pmatrix} * \\ \frac{\mu f_{M_N}(n)}{f_A(n)} a_{n-1}^{(N)} \end{pmatrix} \end{aligned}$$

Using the expressions of $f_A(n)$ and $f_{M_N}(n)$, we can obtain

$$a_n^{(N)} = \frac{\mu^n}{n!} \frac{\prod_{i=1}^{N-1} (\beta_i^{(N)})_{n+1}}{\prod_{i=1}^{N-1} (\alpha_i)_{n+1}} \frac{\prod_{i=1}^{N-1} \alpha_i}{\prod_{i=1}^{N-1} \beta_i^{(N)}} a_0^{(N)} \quad (\text{B2})$$

Furthermore, according to $b_n = \frac{1}{n} \mathbf{u}_N \Lambda \mathbf{a}_{n-1}$, we compute and obtain

$$b_n = \frac{\mu^n}{n!} \frac{\prod_{i=1}^{N-1} (\beta_i^{(N)})_n}{\prod_{i=1}^{N-1} (\alpha_i)_n} \frac{\prod_{i=1}^{N-1} \alpha_i}{\prod_{i=1}^{N-1} \beta_i^{(N)}} a_0^{(N)} \quad (\text{B3})$$

Furthermore, combined with $a_0^{(N)} = \frac{\prod_{i=1}^{N-1} \beta_i^{(N)}}{\prod_{i=1}^{N-1} \alpha_i}$, Eq. (B3) implies that the total static

generating function is

$$G(s) = {}_{N-1}F_{N-1} \left(\beta_1^{(N)}, \dots, \beta_{N-1}^{(N)} \middle| \mu s \right) \quad (\text{B4})$$

Again according to Eq. (1) in the main text, we compute and obtain

$$\begin{aligned} P(m) &= a_0^{(N)} \frac{\prod_{i=1}^{N-1} \alpha_i}{\prod_{i=1}^{N-1} \beta_i^{(N)}} \sum_{k=m}^{\infty} (-1)^{k-m} \binom{k}{m} \frac{\mu^k}{k!} \frac{\prod_{i=1}^{N-1} (\beta_i^{(N)})_k}{\prod_{i=1}^{N-1} (\alpha_i)_k} \\ &= a_0^{(N)} \frac{\prod_{i=1}^{N-1} \alpha_i}{\prod_{i=1}^{N-1} \beta_i^{(N)}} \sum_{k=0}^{\infty} (-1)^k \binom{k+m}{m} \frac{\mu^{k+m}}{(k+m)!} \frac{\prod_{i=1}^{N-1} (\beta_i^{(N)})_{k+m}}{\prod_{i=1}^{N-1} (\alpha_i)_{k+m}} \\ &= a_0^{(N)} \frac{\mu^m}{m!} \frac{\prod_{i=1}^{N-1} \alpha_i}{\prod_{i=1}^{N-1} \beta_i^{(N)}} \frac{\prod_{i=1}^{N-1} (\beta_i)_m}{\prod_{i=1}^{N-1} (\alpha_i)_m} \sum_{k=0}^{\infty} \frac{\prod_{i=1}^{N-1} (m + \beta_i^{(N)})_k}{\prod_{i=1}^{N-1} (m + \alpha_i)_k} \frac{(-\mu)^k}{k!} \end{aligned}$$

Using Eq. (A3), we obtain the analytical expression of mRNA distribution as

$$P(m) = \frac{\mu^m}{m!} \prod_{i=1}^{N-1} \frac{(\beta_i^{(N)})_m}{(\alpha_i)_m} {}_{N-1}F_{N-1} \left(m + \beta_1^{(N)}, \dots, m + \beta_{N-1}^{(N)} \middle| -\mu \right), \quad m = 0, 1, 2, \dots \quad (\text{B5})$$

where ${}_nF_n \left(a_1, \dots, a_n \middle| b_1, \dots, b_n \right); \sigma$ is a confluent hypergeometric function (1).

It is worth being pointed out that if $N=2$ that corresponds to the common ON-OFF model, then Eqn. (B5) can reproduce the result obtained previously (2,3).

That is,

$$P(m) = \frac{\mu^m}{m!} \frac{\Gamma(m + \beta_1^{(2)})}{\Gamma(m + \alpha_1)} \frac{\Gamma(\alpha_1)}{\Gamma(\beta_1)} {}_1F_1 \left(m + \beta_1^{(2)} \middle| -\mu \right), \quad m = 0, 1, 2, \dots \quad (\text{B6})$$

where $\alpha_1 = \lambda_{12} + \lambda_{21}$, $\beta_1^{(1)} = \lambda_{12}$, $\beta_1^{(2)} = \lambda_{21}$

$$\text{Case 3 } \Lambda = \mu \begin{pmatrix} \mathbf{I}_{(N-1)} & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix}$$

Such a case corresponds to the gene model where the promoter has $(N-1)$ active states with the same transcription rate. Correspondingly, Eq. (11) at steady

state in the main article becomes

$$\begin{aligned} \sum_{k=1, \neq i}^N \lambda_{ki} G_k - \sum_{k=1, \neq i}^N \lambda_{ik} G_i - s \frac{\partial G_i}{\partial s} + \mu s G_i &= 0, \quad i = 1, 2, \dots, N-1 \\ \sum_{k=1}^{N-1} \lambda_{kN} G_k - \sum_{k=1}^{N-1} \lambda_{Nk} G_N - s \frac{\partial G_N}{\partial s} &= 0 \end{aligned} \quad (\text{B7})$$

One will see that solving Eq. (B7) can become **Case 2**. In fact, the transformations

$G_i = e^{\mu s} F_i$ ($1 \leq i \leq N$) will transform Eq. (B7) into

$$\begin{aligned} \sum_{k=1, \neq i}^N \lambda_{ki} F_k - \sum_{k=1, \neq i}^N \lambda_{ik} F_i - \tau \frac{\partial F_i}{\partial \tau} &= 0, \quad i = 1, 2, \dots, N-1 \\ \sum_{k=1}^{N-1} \lambda_{kN} F_k - \sum_{k=1}^{N-1} \lambda_{Nk} F_N - \tau \frac{\partial F_N}{\partial \tau} + \mu \tau F_N &= 0 \end{aligned} \quad (\text{B8})$$

with $\tau = -s$, which is nothing but the gene model with one ON with the transcription

rate μ and $(N-1)$ OFFs. According to discussion in **Case 2**, we have

$$F(\tau) = {}_{N-1}F_{N-1} \left(\begin{matrix} \beta_1^{(N)}, \dots, \beta_{N-1}^{(N)} \\ \alpha_1, \dots, \alpha_{N-1} \end{matrix} \middle| ; \mu \tau \right) \quad (\text{B9})$$

Note that $G(z) = e^{\mu s} F(-s) = e^{\mu(z-1)} {}_{N-1}F_{N-1} \left(\begin{matrix} \beta_1^{(N)}, \dots, \beta_{N-1}^{(N)} \\ \alpha_1, \dots, \alpha_{N-1} \end{matrix} \middle| ; \mu(1-z) \right)$. Therefore, using the

relationship between probability distribution and generating function, we have

$$\begin{aligned} P(m) &= \frac{1}{m!} \frac{d^m}{dz^m} G(z) \Big|_{z=0} \\ &= \frac{e^{-\mu}}{m!} \sum_{k=0}^m \binom{m}{k} \mu^{m-k} (-1)^k \prod_{i=1}^{N-1} \frac{(\beta_i^{(N)})_k}{(\alpha_i)_k} {}_{N-1}F_{N-1} \left(\begin{matrix} k + \beta_1^{(N)}, \dots, k + \beta_{N-1}^{(N)} \\ k + \alpha_1, \dots, k + \alpha_{N-1} \end{matrix} \middle| ; \mu \right) \end{aligned} \quad (\text{B10})$$

C. The parameter values used in computation in Fig. 1

$$\text{(A) } \mathbf{A} = \left(\begin{array}{c|cc} -0.2 & 1 & 1 \\ \hline 0.1 & -1.1 & 0.1 \\ \hline 0.1 & 0.1 & -1.1 \end{array} \right), \mu = (0, 10, 30), \delta = 1;$$

$$\text{(B) } \mathbf{A} = \left(\begin{array}{c|cc} -2 & 0.1 & 0.1 \\ \hline 1 & -1.1 & 0.1 \\ \hline 1 & 1 & -0.2 \end{array} \right), \mu = (0, 10, 30), \delta = 1;$$

$$(C) \mathbf{A} = \left(\begin{array}{c|cc} -0.3 & 0.1 & 0.1 \\ \hline 0.2 & -1.1 & 0.1 \\ 0.1 & 1 & -0.2 \end{array} \right), \mu = (0,10,30), \delta = 1;$$

$$(D) \mathbf{A} = \left(\begin{array}{c|cc} -2 & 0.02 & 0.02 \\ \hline 1 & -0.12 & 0.1 \\ 1 & 0.1 & -0.12 \end{array} \right), \mu = (0,10,30), \delta = 1;$$

$$(E) \mathbf{A} = \left(\begin{array}{c|cc} -0.2 & 0.02 & 0.02 \\ \hline 0.1 & -0.12 & 0.1 \\ 0.1 & 0.1 & -0.12 \end{array} \right), \mu = (0,10,30), \delta = 1;$$

$$(F) \mathbf{A} = \begin{pmatrix} -0.09 & 0.02 & 0.02 \\ 0.04 & -0.07 & 0.04 \\ 0.05 & 0.05 & -0.06 \end{pmatrix}, \mu = (5,20,45), \delta = 1.$$

D. The parameter values used in computation in Fig. 2

$$(A)(C) \mathbf{A} = \left(\begin{array}{c|cc} -4 & 0.2 & 0.2 \\ \hline 2 & -0.6 & 0.2 \\ 2 & 0.4 & -0.4 \end{array} \right), \mu = (0,10,90), \delta = 1;$$

$$(B)(D) \mathbf{A} = \left(\begin{array}{cccccccc|c} -4 & 4 & & & & & & & \\ & -4 & 4 & & & & & & \\ & & -4 & 4 & & & & & \\ & & & -4 & 4 & & & & \\ & & & & -4 & 4 & & & \\ & & & & & -4 & 4 & & \\ & & & & & & -4 & 4 & \\ & & & & & & & -4 & 4 \\ \hline & & & & & & & & -4 & 0.08 \\ 4 & & & & & & & & & -0.08 \end{array} \right), \mu = (0,0,0,0,0,0,0,0,0,90),$$

$\delta = 1.$

REFERENCES

1. Slater L. J. 1960. Confluent Hypergeometric Functions. Cambridge University Press, Cambridge.
2. Berg, O. G. 1978. A model for the statistical fluctuations of protein numbers in a

microbial population. *J. Theor. Biol.* 71:587–603.

3. Raj, A., C. S. Peskin, ..., S. Tyagi. 2006. Stochastic mRNA synthesis in mammalian cells. *PLoS Biol.* 4:e309.