Mathematical Biosciences and Engineering

Research article

Modeling stochastic gene expression: From Markov to non-Markov models

Zhenquan Zhang, Junhao Liang, Zihao Wang, Jiajun Zhang and Tianshou Zhou *

Key Laboratory of Computational Mathematics, Guangdong Province, and School of Mathematics, Sun Yat-sen University, Guangzhou 510275, China

* Correspondence: Email: mcszhtsh@mail.sysu.edu.cn.

Abstract: Gene expression is an inherently noisy process due to low copy numbers of mRNA or protein. The involved reaction events may happen in a Markov fashion but also in a non-Markov manner, depending on waiting-time distributions for the occurrence of reaction events. In recent years, many mechanistic models of stochastic gene expression have been developed to forecast fluctuations in mRNA or protein levels. Here we discus commonalities between these models as well as their extensions from Markov to non-Markov models, focusing on the contributions of noisy sources to the protein level. We derive a useful formula for the protein noise quantified by the ratio of the variance over the squared mean. This formula, expressed in terms of the frequencies of the probabilistic events, can be used in the fast evaluation of fluctuations in the protein abundance. Although the detail of the formula may vary from gene to gene, it highlights sources of the protein noise, which can be decomposed into two parts: spontaneous fluctuations resulting from the birth and death of the protein and forced fluctuations originated from switching between the promoter states.

Keywords: gene model; noise; Markov process; non-Markov process

1. Introduction

Mathematical models are a powerful tool for modeling and analysis of complex gene expression processes. An important task of systems biology is to design reasonable mathematical models that faithfully describe the dynamics of gene expression. Any theoretical model of gene expression established based on the first principles needs to make assumptions. This is because the complexity of gene expression would make it difficult to specify what transitions can be condensed into effective reaction steps and what concentrations can be absorbed into reaction rate constants. It seems always difficult to tell if the correct principles are addressed when modeling a complex process of gene

expression. However, it is possible to understand simple mechanistic models of gene expression at the initial stage.

Most published gene models have independently focused on the fundamental genetic information flow described by the central dogma in biology (i.e., the information stored in DNA is first transcribed into mRNAs, which are then translated into functional proteins). These models include (but do not limit to) transcription models [1–5], transcription and translation models [6–9], gene regulatory models [10–14], Markov models of gene expression [15], non-Markov (or queuing) models of gene expression [16,17]. The underlying fundamental principle may be extended, depending on the dissection of molecular details involved in gene expression. A common way of the extension is to implicitly include all other processes in effective reaction rates of gene models [2,4–8,12,13,17]. With the development of single-cell and single-molecule measurement technologies, more molecular details that can significantly impact gene-product levels have been revealed. Here we state key molecular events occurring in gene expression processes, which will be partially incorporated in gene models of this paper.

First, activation and inactivation of a gene can have many different molecular causes, such as alternative splicing, transcription initiation, recruitment of polymerases, dissociation of repressors, association of activators, histone modifications, and chromatin remodeling as well as recruitment of transcription factors (TFs). In particular, the activation of a gene depends on the chromatin template that accumulates over time until the promoter becomes active [2,18–22], implying a multistep process. The details vary from gene to gene and from cell to cell, and can include random transitions between different promoter activity states. As a first approximation, the reaction kinetics of promoters in prokaryotic and eukaryotic cells may be described as a random telegraph process with constant rates or equivalently with exponential waiting-time distributions [23].

Second, transcription or translation is often assumed to follow a Poisson process, i.e., the number of mRNA or protein molecules produced per time unit is often assumed to follow a Poisson distribution. This assumption may or may not be biologically reasonable. Complex regulation may affect transcriptional processes: (1) The binding of TFs to DNA sites may change the promoter structure, thereby impacting the efficiency of transcription; (2) Replication may alter the chromatin structure, giving rise to abrupt changes in expression rates; (3) mRNAs' entry into or exit from the cellular cytoplasm may impact reaction kinetics.

Third, feedback regulation is not exceptional but is ubiquitous in gene expression systems. Positive and negative feedback loops are common regulatory forms found in biological signaling systems. For example, in mammals, more than 3000 signaling proteins and more than 15 second messengers form hundreds and thousands of cell-specific signaling pathways [24]. Many of these signaling pathways may have numerous upstream regulators and downstream targets, which altogether constitute a huge web of connectivity within and between signaling pathways [24]. The presence of multiple feedback loops poses a challenge to understanding how receptor inputs control the cellular behavior.

Fourth, apart from heredity-relevant molecular cues that lead to the so-called non-genetic heterogeneity, other possible contributors include fluctuations in many enzymes and substrates involved, cell cycle effects, and random partitioning of copies at cell division as well as fluctuations in methylation, histone modification, or more generally, epigenetic regulation [25]. These and other complications, which can also affect mRNA or protein abundances, will not be explicitly considered in gene models of this article.
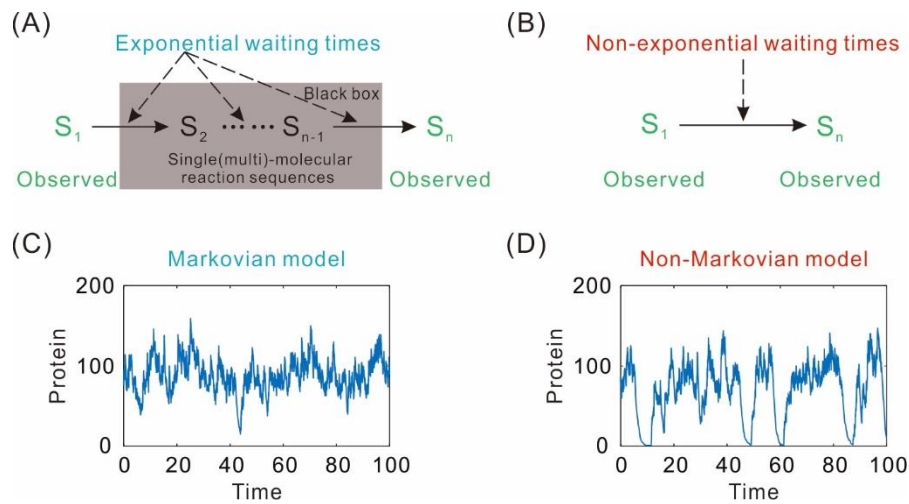
Figure 1. Discrepancy between Markov and non-Markov models: (A) Markov model if all the intermediate processes are specified; (B) non-Markov model where non-exponential waiting-time distribution can model unspecified intermediate processes; (C) unimodal protein expression in the non-Markov on-off model described by $off \underset{b}{\overset{a}{\rightleftharpoons}} on$, $on \xrightarrow{\varepsilon} on + B \cdot X, X \xrightarrow{d} \varnothing$ where parameter values are set as $a = 1, b = 30, d = 1$, $\langle B \rangle = 3, \varepsilon = 0.1$; (D) bimodal protein expression in the non-Markov on-off model described by $off \underset{b}{\overset{f(\tau)}{\rightleftharpoons}} on, on \xrightarrow{\varepsilon} on + B \cdot X, X \xrightarrow{d} \varnothing$, where some parameter values are set as $b = 30, d = 1$, $\langle B \rangle = 3, \varepsilon = 0.1$, and the values of two characteristic parameters in the Gamma waiting-time distribution are set as $k = 1, a = 1$.

Given the above complexity, how gene expression is modeled is challenging. According to the existing literature, we can divide the proposed models of gene expression into two large classes: Markovian and non-Markovian, referring to Figure 1 that shows the discrepancy between a Markov and a non-Markov gene models. For Markov models, it is needed to make the memoryless hypothesis, i.e., the stochastic motion of the reactants is influenced only by the current state but not by previous states. This hypothesis implies that waiting times for reaction events occurring in gene expression obey exponential distributions [26,27], and the reaction kinetics of gene expression are thus Markovian [26–29]. Experimental and theoretical studies have shown that the characteristic parameters in waiting-time distributions can have a significant impact on gene expression levels. However, reaction events occurring in gene expression may happen in a non-Markov fashion. Possible reasons include: (1) the complex control process of transcription initiation, which would involve many kinds of regulators (e.g., repressors, mediators, and TFs) as well as chromatin remodeling and changes in supercoiling, can generate non-exponential time intervals between transcription windows [18,30–35]; (2) The synthesis of an mRNA in general involves multiple intermediate reaction steps that would not have been specified or cannot be specified due to experimental technologies, creating molecular memory between

individual events [18,36]; and (3) medium heterogeneity of gene expression may lead to broad waiting-time distributions between reaction events [37], thus influencing the efficiency of reactants in exploring their environments [38–42]. We will mainly focus on the second case, i.e., molecular memory is created by multistep biochemical processes involved in gene expression.

In most situations, biochemical reactions involved in gene expression are essentially single-molecule events, leading to stochastic fluctuations in mRNA and further protein levels. Thus far, there have been two kinetic modes of stochastic transcription that have been experimentally observed in individual cells: one mode is the Poissonian way where mRNAs are synthesized in a probabilistic manner with a probability that is uniform over time [43,44], and the other mode is the bursty way where mRNAs are produced in a bursty fashion with burst size following a distribution [45–48]. Each of the two ways can result in temporal, stochastic fluctuations in mRNA and protein numbers. This cell-to-cell variability is often referred to as gene expression noise.

In order to help the reader understand how stochastic gene expression is modeled and analyzed, we begin with Markov models and then extend them to non-Markov cases. Our analysis focuses on statistical quantities of protein, such as the mean protein level (or expectation), the protein noise defined as the ratio of the variance over the squared mean, and the Fano factor defined as the ratio of the variance over the mean. We will use these statistical quantities to characterize the protein noise even though a random variable is best characterized by its distribution. We highlight the contributions of noisy sources to protein levels by deriving analytical formulae for the protein noise. Additionally, we discus commonalities between these models as well as the possible linking between theoretical predictions and experimental observations.

2. Markov models of stochastic gene expression

As mentioned in the introduction, Markov models assume that waiting times for reaction events are exponentially distributed. Here we review results for Markov models of stochastic gene expression, which are established based on the random telegraph model [18,19,49]. The telegraph model has long been used as a basis for several studies focusing on analysis of probability distributions for gene products and their statistical quantities as well as on inference of gene expression parameters based on experimental observations [50–52]. It is well known that random telegraph model has a good property: in the limit of transcriptional bursting, the arrival of bursts is a Poisson process. We start by the common on-off model of gene expression and then extend it to more realistic cases including multiple off mechanisms and feedback regulation of various forms.

2.1. The common on-off model

In prokaryotic and eukaryotic cells, most genes are expressed in a bursty manner, namely gene product molecules are produced in short periods of high transcriptional activity (on) followed by long periods of inactivity (off). In order to model this expression manner, on-off models have been proposed [48,53–55], which assume that the gene promoter has two activity states: one off state where the gene is not expressed and one on state where the gene is expressed with burst size B following a probability distribution: $\text{prob}\{B = i\}$ with $i = 0, 1, 2, \cdots$. Denote by a and b the mean switching rates from off to off states and vice versa, respectively. As done in previous works, we

assume that proteins are produced instantaneously after mRNAs are synthesized, so that transcription and translation processes can be lumped into a single reaction step. In addition, we assume that proteins are produced with a constant rate denoted by ε (representing the mean synthesis rate) and degrade with another constant rate denoted by d (representing the mean degradation rate). Without loss of generality, we set $d = 1$ throughout this paper. For clarity, we list all the reactions as follows

$$off \underset{b}{\overset{a}{\rightleftarrows}} on, \quad on \xrightarrow{\varepsilon} on + B \cdot X, \quad X \xrightarrow{d} \varnothing, \tag{1}$$

where X represents protein. According to experimental facts [53], we will always assume that B follows a geometric distribution given by $\text{prob}\{B = i\} = \langle B \rangle^i / (1 + \langle B \rangle)^{i+1}$, $i = 0, 1, 2, \cdots$, where $\langle B \rangle$ represents the mean burst size. Then, we can show that the protein expectation, denoted by $\langle X \rangle$, is given by $\langle X \rangle = \langle B \rangle \mu \rho_{on}$, where $\rho_{on} = \tau_{on} / (\tau_{on} + \tau_{off})$ represents the probability that the promoter is at the on state. Here $\tau_{on} = 1/b$ and $\tau_{off} = 1/a$ represent the mean times that the gene dwells at on and off states respectively. The protein noise, denoted by $\eta_X^{(common)}$, is then given by

$$\eta_X^{(common)} = \frac{\langle B \rangle}{\langle X \rangle} \eta_{burst} + \eta_{promoter}^{(common)}, \tag{2}$$

where $\eta_{burst} = (1 + \langle B \rangle) / \langle B \rangle$ represents the burst noise due to the bursty expression manner, and $\eta_{promoter}^{(common)} = (\tau_{off})^2 / (\tau_{on} + \tau_{off} + \tau_{on}\tau_{off})$ represents the promoter noise due to switching between on and off states. In Eq. (2), the first term on the right hand side represents spontaneous fluctuations (i.e., the internal noise) due to the birth and death of the proteins as well as due to bursty expression, whereas the second term represents the forced fluctuations (i.e., the promoter noise), which depends, only, on the two switching rates but is independent of the synthesis and degradation rates of the protein. Note that each of three statistical quantities $\langle B \rangle$, τ_{on} and τ_{off} is experimentally measurable.
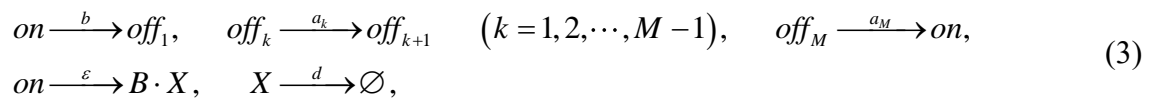
We point out that Eq. (2) is a fundamental formula for the fast evaluation of the protein noise. One will see that if other factors or processes associated with gene expression such as a multistep transition from the off to on state, then term $\eta_{promoter}^{(common)}$ needs to be modified but the first term on the right hand side of Eq. (2) is kept unchanged.

Note that both if $b \to 0$ (or in the limit of small b) and if bursting is not considered, then the common on-off model reduces to the gene model of constitutive expression, which is essentially a birth-death process. In this sense, Eq. (2) includes the result for the gene-product noise in the latter model.

2.2. A gene model of promoter progress

The gene model analyzed in the previous section assumes that the stwiching from inactive to active states is a single-step process. However, biological evidence supports that the chromatin template determining transcription kinetics accumulates over time until the promoter becomes active, implying that the off-to-on transition occurs not in a single-step manner but in a multistep manner. Specifically, the gene activity would proceed sequentially through an on state and several off states (note: in these states, polymerases could be either absent from the promoter or present in a paused or inactive state), and then returns to the on state. As a result, all the promoter's on and off states constitute a chromatin loop. The corresponding model is called the gene model of promoter progress [2,3].

In order to model the gene expression with promoter progress and derive analytical results, we still assume that proteins are produced instantaneously after mRNAs are produced and degrade in a linear manner. The promoter switching from the on state to the off state has been reported to occur essentially with a single rate-limiting step and can thus be modeled by a constant rate [36,56]. Therefore, we introduce the following set of reactions for conveniently writing down the chemical master equation (CME) that describes the time evolution of the probability distribution function for the protein

$$on \xrightarrow{b} off_1, \quad off_k \xrightarrow{a_k} off_{k+1} \quad (k=1,2,\cdots,M-1), \quad off_M \xrightarrow{a_M} on,$$
$$on \xrightarrow{\varepsilon} B \cdot X, \quad X \xrightarrow{d} \varnothing, \tag{3}$$

where b is the mean rate of gene inactivation, a_M is the rate of gene activation, a_k is the transition rate from the k th off state to the $(k+1)$ th off state with $k=1,2,\cdots,M-1$, ε and d are the synthesis and degradation rates of mRNA, respectively. Note that if $M=1$, the corresponding model reduces to the common on-off model analyzed above. In Eq. (3), B represents burst size that is assumed to follow the geometric distribution as specified above.

Denote $\tau_k = 1/a_k$, which represents the time that the gene dwells on the kth off state. Then, $\tau_{off} = \sum_{k=1}^{M} \tau_k$ represents the total time that the gene dwells on the inactive state (called the mean off time). Denote by $\tau_{on} = 1/b$, which represents the mean time that the gene dwells on the on state (called the mean on time). The protein expectation still takes the form of $\langle X \rangle = \langle B \rangle \mu \rho_{on}$ and the protein noise ($\eta_X^{(loop)}$) still takes the form of Eq. (2), but the promoter noise ($\eta_{promoter}^{(loop)}$) should be modified as

$\eta_{promoter}^{(loop)} = \dfrac{(\tau_{on}+\tau_{off})\prod_{k=1}^{M}(1+\tau_k)}{(1+\tau_{on})\prod_{k=1}^{M}(1+\tau_k)-1} - 1$. If the total mean off time τ_{off} is fixed, then we can show $\prod_{k=1}^{M}(1+\tau_k) \geq (\tau_{off}/M+1)^M$, where the equal sign holds if and only if $\tau_k = \tau_{off}/M$ for all k. In this case, the promoter noise reaches the lowest level. For a fixed mean protein level, the protein noise $\eta_X^{(loop)}$ still takes the form Eq. (2) if $\eta_{promoter}^{(common)}$ is replaced with $\eta_{promoter}^{(loop)}$, i.e., the protein noise consists of the term representing spontaneous fluctuations and the term representing the forced fluctuations.

Moreover, the protein noise in this case equals or is greater than that in the common on-off model. Similarly, we can analyze the cases of multiple off pathways even with crosstalk, and obtain a similar result for the protein noise.

The above analysis indicates that the number (M) of the inactive states can enlarge the protein noise. Moreover, the portein noise in the multi-off model is always higher than that in the common on-off model, implying that the protein noise is underestimated in previous studies. In addition, model (3) is essentially equivalent to model (1) in terms of the mean on and off times.

2.3. A gene model with feedback of any form

Feedback is a ubiquitous mechanism that regulates gene expression in prokaryotic and eukaryotic cells. Here we further introduce feedbacks to the common on-off model analyzed above. Specifically, we assume that the produced proteins regulate, as TFs, two switching rates between on and off states, as well as the protein production rate. Thus, the mean switching rates introduced above, a and b, should be replaced with two functions depending on the number (n) of the protein molecules, denoted by $\Theta_1(n)$ and $\Theta_2(n)$, whereas the mean protein production rate should be replaced with another function that also depends on the number of the protein molecules, denoted by $\Theta_3(n)$. The corresponding gene model includes many gene models (e.g., the common on-off model analyzed above) in the literature as its special case. To that end, the corresponding CME at steady state takes the form

$$-\Theta_1(n)Q_0(n)+\Theta_2(n)Q_1(n)+(E-I)\left[Q_0(n)\right]=0,$$

$$\Theta_1(n)Q_0(n)-\Theta_2(n)Q_1(n)+\sum_{i=0}^{n}g_i\Theta_3(n-i)Q_1(n-i)-\Theta_3(n)Q_1(n) \qquad (4)$$

$$+(E-I)\left[Q_1(n)\right]=0,$$

where E is the step operator whereas I is the unit operator, and we define $g_i=\text{prob}\{B=i\}$ and use the assumption that the protein degradation rate is the unit. In Eq. (4), $Q_0(n)$ and $Q_1(n)$ represent the stationary probability distributions of protein when the gene is in off and on states respectively. Apparently, this CME is an extension of that for the above analyzed common on-off model since each $Q_i(n)$ may be any function of n.

In principle, the total stationary distribution, $Q(n)=Q_0(n)+Q_1(n)$, can be exactly obtained by solving Eq. (4) using an analytical method that we ever developed [57], but the form is very complex. Here we are interested in statistical quantities of the protein, and present approximate results. For example, the stationary mean protein level is approximately given by $\langle X\rangle\approx\dfrac{\langle B\rangle\Theta_3(0)\Theta_1(0)}{\Theta_1(0)+\Theta_2(0)}$ [57], which is apparently an extension of the expression for the mean protein in the Markovian case. Similarly, the steady-state protein noise, denoted by $\eta_X^{(feedback)}$, is approximately given by

$\eta_{promoter}^{(feedback)} = \dfrac{\Theta_2(0)}{\Theta_1(0)} \dfrac{1}{\Theta_1(0) + \Theta_2(0) + 1}$ [57], which is also an extension of the expression for the promoter noise in the Markovian case.

For clarity, let us consider a special case: $\Theta_1(n) = a + \rho n$, where parameter ρ represents feedback strength, $\Theta_2(n) = b$ and $\Theta_3(n) = \varepsilon$. In this case, we can derive the following exact expressions [58,59]

$$\langle X \rangle^{(feedback)} = \frac{\varepsilon \tilde{a} \langle B \rangle}{\tilde{a} + \tilde{b}} \frac{{}_1F_1\left(1 + \tilde{a}, 1 + \tilde{a} + \tilde{b}; \tilde{\varepsilon}\right)}{{}_1F_1\left(\tilde{a}, \tilde{a} + \tilde{b}; \tilde{\varepsilon}\right)}, \tag{5a}$$

for the mean protein level and

$$\eta_X^{(feedback)} = \frac{1 + \langle B \rangle}{\langle X \rangle} + \frac{\left(\tilde{a} + \tilde{b}\right)\left(\tilde{\rho} + \tilde{a}\right)}{\tilde{a}\left(1 + \tilde{a} + \tilde{b}\right)} \frac{{}_1F_1\left(\tilde{a}, \tilde{a} + \tilde{b}; \tilde{\varepsilon}\right){}_1F_1\left(2 + \tilde{a}, 2 + \tilde{a} + \tilde{b}; \tilde{\varepsilon}\right)}{\left[{}_1F_1\left(1 + \tilde{a}, 1 + \tilde{a} + \tilde{b}; \tilde{\varepsilon}\right)\right]^2} - 1, \tag{5b}$$

for the protein noise intensity, where $\tilde{a} = \dfrac{a}{\rho + 1}$, $\tilde{b} = \dfrac{b}{\rho + 1}$, and $\tilde{\varepsilon} = \tilde{\rho}\varepsilon$ with $\tilde{\rho} = \dfrac{\rho}{\rho + 1}$, and

${}_1F_1(\alpha, \beta; z)$ is a confluent hypergeometric function [60]. Note that if $\rho \to 0$, Eqs. (5a) and (5b) can reproduce previous results, e.g., $\langle X \rangle^{(common)} = \dfrac{\varepsilon a}{a + b}$ and $\eta_X^{(common)} = \dfrac{1 + \langle B \rangle}{\langle X \rangle} + \dfrac{b}{a(a + b + 1)}$.

3. Non-Markov models of stochastic gene expression

In the above section, we have reviewed and analyzed Markov models of stochastic gene expression, where waiting-time distributions for reaction events are assumed to be exponential. For many systems of gene expression, however, this assumption is too strong to describe the realistic case. In fact, the waiting times in general follow non-exponential distributions or reaction kinetics is in general non-Markovian, with reasons mentioned as in the introduction. Here, we introduce and analyze several gene models of bursting expression with molecular memory, which can be divided into two classes: one class described in terms of queuing theory [16,36,56] and the other class described in terms of continuous time random walk (CTRW) theory [61–63].

3.1. A queuing model

In queuing theory, queuing models can be categorized into several classes, depending on assumptions on the numbers of customers and servers, and the ways of customer arrival and departure [64]. Here, we map gene expression processes into queuing models in the queuing theory [16,36,56]. In this mapping, proteins are the analogs of customers in a queue, and their production is analogous to arrival of customers whereas their degradation to customers who leave

the queue after receiving service. We assume that waiting-time distribution for protein degradation is exponential. In addition, we assume that each protein degrades independently, implying that there are an infinite number of servers in the queuing model. If we let GI represent that the waiting-time distribution for customer arrival is general, M represent that the distribution for customer service time is exponential, and ∞ represent that there are infinite servers in the queuing system, then our queuing model of gene expression belongs to the $GI/M/\infty$ system in the queuing theory.

Denote by ε (representing the mean production rate) the constant rate for protein synthesis, by d (representing the mean degradation rate) the constant rate for protein degradation, and by b (representing the mean on-to-off switching rate) the constant switching rate from active to inactive states of the promoter. Let $\xi(\tau)$ represent the probability density function for waiting time from inactive to active states. For this queuing model, we can derive exact expressions for the steady-state moments for the number of customers in the system. In this subsection, we do not explicitly consider bursty expression.

For the above queuing model of stochastic gene expression, it seems to us that the CME has not been established so far. In order to help the reader's understanding, here we present the following steady-state CME without providing details of mathematical derivation (note: the CME can be derived using the total probability principle)

$$(\varepsilon + nd + b)Q_1(n) = \varepsilon Q_1(n-1) + (n+1)dQ_1(n+1) + \int_0^\infty Q_0(n,\tau)h(\tau)d\tau,$$

$$\frac{\partial Q_0(n,\tau)}{\partial \tau} = -ndQ_0(n,\tau) + (n+1)dQ_0(n+1,\tau),$$

(6)

where $h(\tau) = \xi(\tau)/\Pi(\tau)$ is a hazard function [65–67], and $\Pi(\tau) = \int_\tau^\infty \xi(t)dt$ represents the cumulant probability. The boundary condition is imposed by $Q_0(n,0) = bQ_1(n)$. Equation (6) can be solved using the characteristic line method combined with the binomial moment method that we ever developed [59].

3.1.1. Statistical quantities for the protein noise

Denote by $\eta_X^{(queue)}$ the protein noise. According to the noise definition and noting that the protein degradation rate has been set as unit, we can derive the following analytical expression for the protein noise from Eq. (6)

$$\eta_X^{(queue)} = \frac{1}{\langle X \rangle} + \frac{\tau_{off} - \tilde{\xi}(1)}{\tau_{on} + \tilde{\xi}(1)},$$

(7)

where $\tau_{on} = 1/b$ and $\tau_{off} = \tilde{\xi}(0)$, and $\tilde{\xi}(s)$ is the Laplace transform of function $\xi(\tau)$, i.e., $\tilde{\xi}(s) = \int_0^\infty e^{-s\tau}\xi(\tau)d\tau$. This result is accurate in contrast to previous works that gave approximate expressions for the protein noise [16,36,56]. In Eq. (7), the first term on the right hand side represents spontaneous fluctuations resulting from the production and degradation of the protein whereas the

second term is only related to the promoter kinetics and can then be viewed as forced fluctuations originated from switching between the promoter states. Thus, the decomposition formula in Eq. (7) is an analog of Eq. (2), namely, the protein noise is decomposed into two parts: spontaneous fluctuations resulting from the birth and death of the protein and forced fluctuations originated from switching between on and off states of the promoter.

Next, we consider a special case of distribution $\xi(\tau)$ for the off-to-on switching time, i.e., $\xi(\tau)$ is a Gamma distribution with the form of $\xi(\tau) = \dfrac{a^k}{\Gamma(k)} \tau^{k-1} e^{-a\tau}$, where k is a shape parameter and will be called memory index. Note that $k = 1$ corresponds to the Markov case whereas $k > 1$ to the non-Markov case. If k is a positive integer, then function $\xi(\tau)$ is an Erlang distribution. Also note that the mean of $\xi(\tau)$ is $\dfrac{k}{a}$ and the variance is $\dfrac{k(k+1)}{a^2}$.

For the above setting, the mean protein level $\langle X \rangle$ is exactly given by $\langle X \rangle = \mu \rho_{on}$, where ρ_{on} is the probability that the promoter is in the on state and is exactly given by $\rho_{on} = \dfrac{a/k}{a/k + b}$. This indicates that molecular memory characterized by memory index k can reduce the mean protein level. In addition, the promoter noise, denoted by $\eta_{promoter}^{(queue)}$, is given by $\eta_{promoter}^{(queue)} = \dfrac{\tau_{off} - \tilde{\xi}(1)}{\tau_{on} + \tilde{\xi}(1)}$. If memory index $k = 1$, then we can show that $\eta_{promoter}^{(queue)}$ reduces to $\eta_{promoter}^{(common)} = \dfrac{\tau_{off}\tau_{off}}{\tau_{on} + \tau_{off} + \tau_{on}\tau_{off}}$. If $k > 1$, we can prove that function $\kappa(I) \equiv \dfrac{\tau_{off} - \tilde{\xi}(1)}{\tau_{on} + \tilde{\xi}(1)}$ with $\tau_{off} = \dfrac{k}{a}$ and $\tilde{\xi}(1) = 1 - \left(\dfrac{a}{1+a}\right)^k$ is a monotonically increasing function of index k. Thus, $\kappa(k) \geq \kappa(1) = \dfrac{b}{a(a+b+1)} = \eta_{promoter}^{(common)}$, indicating that the promoter noise in the queuing model is stronger than that in the common on-off model. This monotonicity implies that molecular memory always amplifies the protein noise if the protein mean is fixed.

3.1.2. Bounds for the number of promoter states

The experimental observations [18,33,54] have indicated that the waiting time for promoter switching from inactive to active states in general follows an non-exponential distribution, but the specific form of this distribution is in general unknown since it depends, in general, on the number of promoter activity states, possibly changing from gene to gene. On the other hand, the current single-

cell measurement technologies make it possible to count the number of transcripts in different cells. An interesting yet unsolved question is the estimation on the number of promoter internal states. In order to address this question, we will analyze the Fano factor of protein rather than its noise intensity as done above, focusing on deriving bounds on the number of promoter internal states based on experimentally accessible measurements.

For analysis convenience, we only consider a special case: the number of inactive states is fixed at a positive integer (k). In this case, the distribution for the off-to-on switching time, denoted by $\xi(\tau)$, can be modeled as a phase-type distribution of order k [68]. Note that for a fixed mean switching time, the noise for the phase-type distribution of order k is the minimum if the distribution is set as $\xi(\tau) = \dfrac{a^k}{\Gamma(k)} \tau^{k-1} e^{-a\tau}$ (i.e., an Erlang distribution with shape parameter k [68,69]), where parameter a represents the rate for a single-step reaction whereas parameter k represents the number of reaction steps. This setting will allow us to derive bounds on the Fano factor for protein with details being as follows.

Note that for the $\xi(\tau)$ set above, we can show that the exact expression for the Fano-factor, denoted by $FF(k)$, is given by $FF(k) = 1 + \langle B \rangle (\rho_0)^2 (1 + 1/k)$, where ρ_0 represents the probability that the promoter is in an off state and is a fixed number if the protein expectation is fixed. This expression indicates that the Fano factor of protein is a monotonically decreasing function of k, and reaches the maximum at $k = 1$. Therefore, we have inequality $FF(\infty) \le FF(k) \le FF(1)$. Notice that $2FF(\infty) = 1 + FF(1)$ holds. Thus, both the lower and upper bounds can be specified in terms of $FF(1)$, that is, $(1 + FF(1))/2 \le FF(k) \le FF(1)$. Furthermore, if we make the transform of $\vartheta_{theor} = [FF(1) - 1]/[\langle B \rangle \rho_0^2]$, which can be rewritten as $\vartheta_{theor} = 2 + (1/k)$, then we have $1 \le \vartheta_{theor} \le 2$, which is independent of memory index k.

The significance of the above analysis is as follows. Consider an experimental analog of the FF defined above, which has the experimental Fano factor ϑ_{exper}. In order to set upper and lower bounds on the number of promoter states, it is needed to determine the largest value of integer k such that $\vartheta_{\text{exper}} < \vartheta_{theor}$. Note that the Erlang distribution of order k has the lowest variance among all phase-type distributions of order k. Also note that ϑ_{theor} is a monotonically decreasing function of k. The combination of both implies that the promoter must have at least $k + 1$ off states if $\vartheta_{\text{exper}} < \vartheta_{theor}$ is satisfied. Thus, measurements of the Fano factor for protein can be used in setting bounds on the minimum number of promoter states.

3.2. A CTRW model

As is well known, CTRWs can be further divided into two classes [62]: active CTRWs where waiting times need to be reset; passive CTRWs where waiting times are not reinitialized. The above analyzed queuing model of gene expression belongs to the latter class. From the viewpoint of stochastic simulation, however, it is more convenient to use active CTRWs to describe stochastic processes including gene expression. In particular, physicians frequently use the active CTRW framework to analyze and simulate stochastic processes [61–63]. Here we consider a stochastic gene expression model in the sense of active CTRWs.

Let $\psi_1(t;n)$ and $\psi_2(t;n)$ be intrinsic-event waiting-time distributions from off to on states and vice versa respectively, $\psi_3(t;n)$ and $\psi_4(t;n)$ be intrinsic-event waiting-time distributions for protein synthesis and degradation respectively, where n represents the number of protein molecules. Note that self-regulation exists if $\psi_1(t;n)$ depends on n and there is no feedback otherwise. The similar case holds for $\psi_2(t;n)$. If $\psi_3(t;n)$ depends on n, this implies post-transcriptional or posttranslational regulation. Assume that proteins are generated in bursts with burst size (B) following the geometric distribution described above. To that end, we have finished the setting of our gene model. For convenience, we denote by R_i ($1 \le i \le 4$) four reactions from off to on states and vice versa, synthesis and degradation of protein, respectively.

3.2.1. Effective transition rates

Since $\psi_i(t;n)$ ($i = 1, 2, 3, 4$) may be general distributions, which may lead to molecular memory, each of 4 reactions R_i ($1 \le i \le 4$) has a memory function denoted by $M_i(t;n)$ [62,73,74]. Interestingly, we can prove that for arbitrary waiting-time distributions $\psi_k(t;n)$ for reaction R_k ($1 \le k \le 4$), the limit $\lim_{s \to 0} \tilde{M}_i(s;n)$ ($1 \le i \le 4$) always exists [70]. If this limit is denoted by $\Theta_i(n)$, then we have [70]

$$\Theta_i(n) = \frac{\int_0^{+\infty} \bar{\psi}_i(t;n) \prod_{j \ne i} \left[1 - \int_0^t \bar{\psi}_j(t';n) dt' \right] dt}{\int_0^{+\infty} \prod_{j=1}^{4} \left[1 - \int_0^t \bar{\psi}_j(t';n) dt' \right] dt} , \tag{8}$$

where $\bar{\psi}_k(t;n)$ is a modification of $\psi_k(t;n)$ due to the consideration of on and off states. Specifically, $\bar{\psi}_k(t;n) = \psi_k(t;n)$ ($k = 1, 4$) and $\bar{\psi}_k(t;n) = 0$ ($k = 2, 3$) when we calculate $\Theta_1(n)$

according to $\Theta_1(n) = \lim\limits_{s\to 0} \tilde{M}_1(s;n)$; $\bar{\psi}_1(t;n) = 0$, $\bar{\psi}_k(t;n) = \psi_k(t;n)$ ($k = 2,3,4$) when we calculate $\Theta_2(n)$ or $\Theta_3(n)$ according to $\Theta_k(n) = \lim\limits_{s\to 0} \tilde{M}_k(s;n)$ ($i = 2,3$); and $\bar{\psi}_i(t;n) = \psi_i(t;n)$ ($1 \le i \le 4$) when we calculate $\Theta_2(n)$ according to $\Theta_4(n) = \lim\limits_{s\to 0} \tilde{M}_4(s;n)$.

Function $\Theta_i(n)$ is called the effective transition rate for the ith reaction [70,71], where $1 \le i \le 4$. Note that if $\psi_i(t;n)$ is an exponential distribution of the form $\psi_i(t;n) = \lambda_i(n) e^{-\lambda_i(n)t}$, where $\lambda_i(n)$ should be understood as the reaction propensity function for R_i, we can show $\Theta_i(n) = \lambda_i(n)$. This indicates that effective transition rates are extensions of reaction propensity function s. If all waiting-time distributions are exponential, which corresponds to the Markov reaction case, the model reduces to the common CME for the on-off model with feedback and bursting in the sense of Laplace transform. We point out that the introduction of effective transition rates will be a key for analyzing non-Markov behavior. In addition, we emphasize that our gene model includes almost the on-off models of gene expression in the existing literature as its special cases.

The promoter switching from the on state to the off state has been reported to occur essentially with a single rate-limiting step and can thus be modeled by a constant rate [39]. Therefore, $\psi_2(t;n)$ can be assumed to be an exponential distribution, i.e., $\psi_2(t;n) = be^{-bt}$, where b represents the mean switching rate from on to off states. Assume that proteins degrade in a linear manner with a constant rate denoted by d (implying that the waiting-time distribution for degradation takes the form $\psi_4(t;n) = ne^{-nt}$ due to the assumption of $d = 1$).

In order to show the explicit effect of molecular memory, we consider the following two special cases.

Case 1: $\psi_1(t;n) = \left[a^{k_1} / \Gamma(k_1) \right] t^{k_1 - 1} e^{-at}$, $\psi_2(t;n) = be^{-bt}$ and $\psi_3(t;n) = \varepsilon e^{-\varepsilon t}$, where positive constants a, b and ε represent an average switching rates from off to on and vice versa, and an average transcriptional or translational rate, respectively. Note that $k_1 = 1$ corresponds to the Markov reaction case whereas $k_1 > 1$ to the non-Markov reaction case. Therefore, k_1 is also called memory index. According to Eq. (8), we can show $\Theta_1(n) = na^{k_1} / \left[(a+n)^{k_1} - a^{k_1} \right]$ with $\Theta_1(0) = \lambda_1(0)/k_1$, $\Theta_2(n) = b$, $\Theta_3(n) = \varepsilon$, and $\Theta_4(n) = n$. Therefore, the effect of molecular memory is equivalent to the introduction of a negative feedback if $I_1 > 1$.

Case 2: $\psi_1(t;n) = ae^{-at}$, $\psi_2(t;n) = b\,e^{-b\,t}$, and $\psi_3(t;n) = \left[\varepsilon^{k_3}/\Gamma(k_3)\right]t^{k_3-1}e^{-\varepsilon t}$, where k_3 is

also called a memory index. Note that $k_3 > 1$ corresponds to the non-Markov case and $k_3 = 1$

corresponds to the Markov case. In this case, we can show $\Theta_3(n) = \varepsilon^{k_3}\dfrac{b+n}{(b+\varepsilon+n)^{k_3} - \varepsilon^{k_3}}$, $\Theta_1(n) = a$,

$\Theta_2(n) = b$, and $\Theta_4(n) = n$. Moreover, the effect of molecular memory is equivalent to the

introduction of a negative posttranscriptional or posttranslational regulation if $k_3 > 1$.

Similarly, we can analyze the combination of the above two cases. In a word, effective transition rates not only explicitly decode the effect of molecular memory but also can give us useful information on the underlying system. In particular, the use of effective transition rates can transform a non-Markov issue into a Markov one [70]. Therefore, we can use the results in section 2.3 to directly give results for the non-Markov gene model.

3.2.2. Statistical quantities for the protein noise

Based on Eq. (4) above, we can first show that the stationary mean protein level $\langle X \rangle$ is

approximately given by $\langle X \rangle \approx \langle B \rangle \dfrac{\Theta_1(0)\Theta_3(0)}{\Theta_1(0) + \Theta_2(0)}$, which is apparently an extension of the expression

for the mean protein in the Markov case. Then, we can show that the noise intensity for the protein at

steady state, denoted by $\eta_X^{(CTRW)}$, still takes the form of Eq. (2) if the expression of $\eta_{promoter}^{(on-off)}$ is replaced

with $\eta_{promoter}^{(CTRW)} = \dfrac{\Theta_2(0)}{\Theta_1(0)}\dfrac{1}{\Theta_1(0) + \Theta_2(0) + 1}$, which is also an extension of the expression for the

promoter noise in the Markov case.

In order to analytically show the effects of global delay and molecular memory, we consider the above two special cases.

For **Case 1**, the mean level and the noise intensity for the protein are approximated as

$\langle X \rangle \approx \langle B \rangle \dfrac{\varepsilon a}{a + bk_1}$ and $\eta_X^{(CTRW)} \approx \dfrac{1 + \langle B \rangle}{\langle X \rangle} + \dfrac{b}{a}\dfrac{k_1^2}{a + (b+1)k_1}$, respectively. We observe that a larger value

of memory index k_1 leads to a lower mean protein level but to the stronger protein noise (including

the stronger promoter noise) if the mean protein level is fixed.

For **Case 2**, the mean level and the noise intensity for the protein can be approximated as

$$\langle X \rangle \approx \langle B \rangle \frac{a}{a+b} \frac{a\varepsilon^{k_3}}{\left(a+\varepsilon\right)^{k_3} - \varepsilon^{k_3}} \quad \text{and} \quad \eta_X^{(CTRW)} \approx \frac{1+\langle B \rangle}{\langle X \rangle} + \frac{b}{a}\frac{1}{a+b+1} ,$$ respectively. Apparently, $\langle X \rangle$ is

a monotonically decreasing function of memory index k_3.

The above analysis indicates that molecular memory is an unneglectable factor affecting the gene expression level and noise.

4. Conclusion and discussion

Although gene expression is complex biochemical process, the noise in mRNA or protein can be decomposed into two parts: spontaneous fluctuations resulting from the birth and death of the protein and the forced fluctuations originated from switching between the promoter states, as shown in Eq. (2). This formula is universal in form, suitable to various gene-expression models established based on the central dogma in Biology, which can be categorized into two classes: Markov and non-Markov models. But specific processes may fine-tune the level of the protein noise.

4.1. On upstream and downstream dynamics of gene expression

The models analyzed in this paper were simplified, namely they only considered downstream dynamics and neglected upstream dynamics. However, chromatin regulators play a major role in establishing and maintaining gene expression states. For these regulators, silencing and reactivation occur in all-or-none events, enabling the regulators to modulate the fraction of cells silenced rather than the amount of gene expression. The regulators operate over different time scales and generate distinct types of epigenetic memory, through their individual transition rates. Relevant dynamics can be described by a three-state model involving stochastic transitions between active, reversibly silent, and irreversibly silent states [25].

All-or-none stochastic switching of gene expression states can occur at two levels: chromatin-mediated switching and transcriptional bursting. Active and silent chromatin states may transition to each other. However, even within an active chromatin state, promoters in general switch stochastically between transcriptionally active and inactive states (this phenomenon is called transcriptional bursting [72]). In ref. [25], the authors discussed the connection between these two modes of gene regulation. They also analyzed the dynamic regimes that lead to either graded or fractional all-or-none responses at the protein level. Interestingly, they found that chromatin-mediated switching generally produces bimodal protein distributions similar to those observed experimentally, but can also produce graded protein level distributions in some parameter regimes.

The two levels of gene regulation can be combined in a single model (Figure 2) [72]. In this model, a gene can switch between active and silent chromatin states. In the silent chromatin state, the promoter is always in an off state, so the mRNA level is zero. By contrast, the active chromatin state is permissive, allowing the promoter to switch between periods of active transcription, which produce bursts of multiple mRNA molecules, and periods of inactivity [4,73,74].

(A)
Upstream(slow)
Chromatin level

(B)
Downstream(fast)
Transcription level
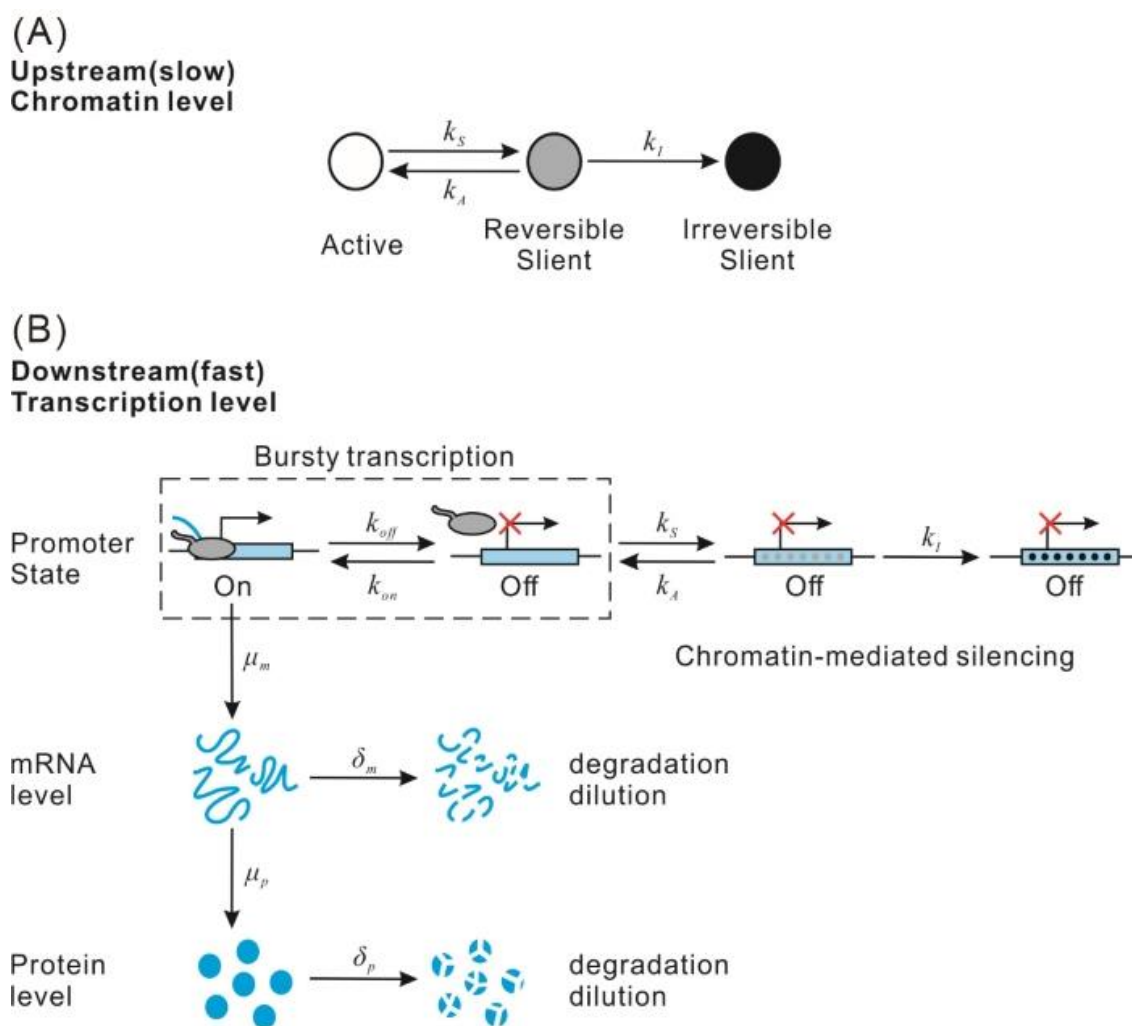
Figure 2. (A) Schematic diagram for a chromatin model; (B) Schematic diagram for a gene model considering chromatin regulation.

In general, transitions in chromatin state and transcriptional bursting are associated with different timescales. At the transcription level, switching between on and off promoter states has been attributed to short-lived interactions of TFs and core machinery with the promoter, occurring usually on the timescale of seconds to minutes [72]. By contrast, at the chromatin level, switching between active and silent chromatin states occur usually on the timescale of hours to days, an order of magnitude much slower than the timescales involved in transcriptional bursting.

The timescales of switching can determine the noise in protein levels at the population level. When switching timescales are fast in contrast to mRNA and protein half-lives, which is typically the case for transcription factor regulation, the level of protein expressed from the gene reaches a stationary unimodal distribution [75]. The mean value of this distribution depends, in a graded manner, on the occupancy of the TF at the promoter [76,77]. When switching timescales are slow relative to mRNA and protein half-lives, as is generally the case for chromatin-mediated regulation, one expects bimodal protein distributions. The relative ratio of the two peaks in this bimodal distribution depends not only on the occupancy of the chromatin remodeling at the promoter, but also on the time it has been recruited there.

4.2. On division of noisy sources

Previous studies have used different classifications to clarify different sources of gene expression noise. Here, we focus on the distinction between intrinsic and extrinsic noise by considering four aspects: The statistical nature of fluctuations, correlations between different proteins in a single cell, how central a process is to gene expression, and experimental strategies for measuring noise:

(I) Terms 'intrinsic' and 'extrinsic' have no specific meaning except that they mean 'inside' and 'outside' respectively. In fact, the classification of intrinsic and extrinsic noise depends on the definition of system versus environment and is therefore relative. For example, if proteins are taken as system, the first term in Eq. (2) is intrinsic and the promoter noise is extrinsic [53,78]. Some studies [16,31] define the two terms on the right hand side of Eq. (2) as the intrinsic noise to distinguish them from the extrinsic noise in the overall state of the cell. However, spontaneous fluctuations in the protein level are largely different from the noise that originates in enslavement by the promoter, but the latter noisy sources are not fundamentally different, e.g., ribosome-mediated noise also enslaves proteins. If we consider the statistical nature of fluctuations only, there is thus no reason to label the two componental noise terms in Eq. (2) as intrinsic.

(II) Protein noise can be classified based on correlations between different types of proteins. Some noisy sources are shared by multiple genes in a single cell whereas others are exclusive to a particular gene or a small set of genes. The most specific componental noise comes from having low protein numbers, originating in the random births and deaths of individual molecules. additionally, there are many other specific sources of gene expression noise, which include: (a) spontaneous mRNA fluctuations are quite specific although some transcripts encode several different proteins; (b) in contrast, operator fluctuations are less specific and typically affect all genes in an operon; and (c) many DNA and RNA binding proteins act as repressors and activators, some of which are specific to a particular gene and others would regulate large classes of genes. This can form complicated correlation structures. For example, two genes may be regulated by the same repressor, but their mRNAs may be degraded by different RNases. A few central factors, e.g., ribosomes, core polymerases, tRNAs, and amino acids, are shared universally. However, different proteins are still affected differently by such fluctuations because they have different sequences and lifetimes.

From this perspective, the two terms in Eq. (2) or its other forms could all be intrinsic. However, it should be emphasized that such a classification can be made in many different ways. The mRNA term would be more intrinsic than the promoter term since the former is directly related to translation, but the latter is not necessarily more intrinsic than fluctuations in regulator concentrations.

(III) From a biological viewpoint, sources of gene expression noise could be categorized according to how central the corresponding component is to gene expression. But classifying fluctuations in promoter activity as intrinsic and fluctuations in ribosomes as extrinsic would not separate the central parts of gene expression from more peripheral cell processes. The opposite classification seems more appropriate, e.g., ribosomes are inherent to gene expression whereas spontaneous changes in promoter activity can indirectly reflect regulation.

(IV) The classification of noisy sources can be tailor-made to the experimental methods available. A good strategy for separating noisy sources would be based on correlations between the expressions of two physically separate but identically regulated fluorescent reporter genes [74,79]. If the two proteins are kinetically independent, e.g., if the underlying mechanisms are linear, then the normalized covariance between them is equal to the sum of all the common noise terms [14]. Fluctuations in

ribosomes, polymerases, RNases, etc., together end up in the 'extrinsic' covariance category while the two componential noise terms in Eq. (2) or its other forms are 'intrinsic' because each green fluorescent protein has its own operators and transcripts. The appeal of this approach is not only that some noisy sources are separated from others but also that the separation to some extent relates to specificity (see subsection 4.2.2 above). The only risk for this separation is that the terms are over-explained, comprehending biological or physical meaning where none is intended: The 'intrinsic noise' only partially relates to specificity, as demonstrated by experiments where most extrinsic noise comes from a repressor that is specific to that particular gene [74]. Realizing that the distinction is largely a side effect of the experimental set-up can in turn open the doors for other applications. If the two reporters were placed under the same operator, then operator fluctuations would become shared between the corresponding two genes, thus moving from the intrinsic to the extrinsic category. But if the reporters were encoded on the same transcript, the mRNA term would become extrinsic. In addition, if the reporters were regulated by different repressors, repressor fluctuations would move from the extrinsic to the intrinsic category. The dual reporter strategy does not separate fluctuations in target proteins based on a priori physical or biological principles. It is much more general and useful than that.

4.3. On experimental measures of gene expression noise

There are many ways to quantify gene expression noise, e.g., autocorrelation analysis is an analytically tractable way since autocorrelations conveniently summarize both the magnitude and the frequency of fluctuations. However, most models of stochastic gene expression have focused on stationary averages and variances so far.

The above analytical results for the protein noise are formulated in terms of the ratio of the variance over squared average, which allows for the fast evaluation and a clear separation of different noise sources as long as the corresponding gene models are weakly nonlinear. Another common measure is the Fano factor, which is defined as the ratio of the variance over the expectation and has a good property: it equals one for Poisson distributions. The Fano factor can work well only for univariate discrete random processes where the variance is proportional to the average with a proportional constant that reflects the overall nature of the process. For a multivariate random process, however, the Poisson distribution holds no special position and using the Fano factor would be misleading. To illustrate this with a more extreme example than the bursts in Eq. (4), we assume that fluctuations in the protein abundance come, partly, from fluctuations in ribonuclease or protease concentrations. The ratio of the variance over the squared average would then contain an extra term that is more or less independent of transcription or translation rates. Multiplying by the average to obtain the Fano factor would thus force the measure to depend on anything that impacts the average.

The results given by Eq. (5a,5b) or obtained by Eq. (8) suggest that no measures work well for all types of fluctuations: the componential noise due to non-exponential waiting-time distribution cannot be separated from the promoter noise; spontaneous fluctuations depend on the number of molecules whereas forced fluctuations do not. In spite of this, the ratio of the variance over the squared mean is still a suitable basis for various possible experimental interpretations. This is because (a) the average number of proteins per cell is too high to contribute substantial spontaneous fluctuations in most experimental studies so far; (b) by plotting the ratio of the variance over the squared average as a function of the inverse average, any univariate scaling behavior can easily be identified without introducing scaling problems for any extrinsic source of noise; and (c) the relevance of fluctuations

depends, typically, on the size of the underlying system. The variance (i.e., a second-order moment) must be normalized by the squared average.

Acknowledgments

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. A. Sanchez, S. Choubey, J. Kondev, Stochastic models of transcription: From single molecules to single cells, *Methods*, **62** (2013), 13–25.
2. J. Zhang, T. Zhou, Promoter-mediated transcriptional dynamics, *Biophys. J.*, **106** (2014), 479–488.
3. T. Zhou, J. Zhang, Analytical results for a multistate gene model, *SIAM J. Appl. Math.*, **72** (2012), 789–818.
4. L. Cai, N. Friedman, X. S. Xie, Stochastic protein expression in individual cells at the single molecule level, *Nature*, **440** (2006), 358–362.
5. T. Liu, J. Zhang, T. Zhou, Effect of interaction between chromatin loops on cell-to-cell variability in gene expression, *PLoS Comput. Biol.*, **12** (2016), e1004917.
6. D. R. Rigney, W. C. Schieve, Stochastic model of linear, continuous protein-synthesis in bacterial populations, *J. Theor. Biol.*, **69** (1977), 761–766.
7. O. G. Berg, A model for the statistical fluctuations of protein numbers in a microbial population, *J. Theor. Biol.*, **71** (1978), 587–603.
8. P. K. Tapaswi, R. K. Roychoudhury, T. Prasad, A stochastic model of gene activation and RNA synthesis during embryogenesis, *Sankhyā Indian J. Statist. Ser. B*, **49** (1987), 51–67.
9. J. Peccoud, B. Ycart, Markovian modeling of gene-product synthesis, *Theor. Popul. Biol.*, **48** (1995), 222–234.
10. D. R. Rigney, Stochastic model of constitutive protein levels in growing and dividing bacterial cells, *J. Theor. Biol.*, **76** (4), 453–480.
11. D. R. Rigney, Stochastic models of cellular variability. In *Kinetic Logic A Boolean Approach to the Analysis of Complex Regulatory Systems*, Springer, Berlin, Heidelberg, 1979.
12. T. B. Kepler, T. C. Elston, Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations, *Biophys. J.*, **81** (2001), 3116–3136.
13. M. Thattai, A. Van Oudenaarden, Intrinsic noise in gene regulatory networks, *Proc. Natl. Acad. Sci. U.S.A.*, **98** (2001), 8614–8619.
14. P. S. Swain, M. B. Elowitz, E. D. Siggia, Intrinsic and extrinsic contributions to stochasticity in gene expression, *Proc. Natl. Acad. Sci. U.S.A.*, **99** (2002), 12795–12800.
15. M. Sasai, P. G. Wolynes, Stochastic gene expression as a many-body problem, *Proc. Natl. Acad. Sci. U.S.A.*, **100** (2003), 2374–2379.

16. T. Jia, R. V. Kulkarni, Intrinsic noise in stochastic models of gene expression with molecular memory and bursting, *Phys. Rev. Lett.*, **106** (2011), 058102.

17. Z. Cao, R. Grima, Linear mapping approximation of gene regulatory networks with stochastic dynamics, *Nat. Commun.*, **9** (2018), 1–15.

18. C. V. Harper, B. Finkenstädt, D. J. Woodcock, S. Friedrichsen, S. Semprini, L. Ashall, et al., Dynamic analysis of stochastic transcription cycles, *PLoS Biol.*, **9** (2011), e1000607.

19. M. R. Green, Eukaryotic transcription activation: Right on target, *Mol. Cell*, **18** (2005), 399–402.

20. J. Paulsson, Models of stochastic gene expression, *Phys. Life Rev.*, **2** (2005), 157–175.

21. G. Hornung, R. Bar-Ziv, D. Rosin, N. Tokuriki, D. S. Tawfik, M. Oren, et al., Noise–mean relationship in mutated promoters, *Genom. Res.*, **22** (2012), 2409–2417.

22. Q. Li, G. Barkess, H. Qian, Chromatin looping and the probability of transcription, *Trend. Genet.*, **22** (2006), 197–202.

23. C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, Springer, Berlin, Heidelberg, 2004.

24. J. D. Jordan, E. M. Landau, R. Iyengar, Signaling networks: The origins of cellular multitasking, *Cell*, **103** (2000), 193–200.

25. L. Bintu, J. Yong, Y. E. Antebi, K. McCue, Y. Kazuki, N. Uno, et al., Dynamics of epigenetic regulation at the single-cell level, *Science*, **351** (2016), 720–724.

26. C. W. Gardiner, *Stochastic Methods: a handbook for the natural and social sciences*, Springer, New York, 2009.

27. N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry*, North-Holland, Amsterdam, 2007.

28. E. Pardoux, *Markov Processes and Applications: Algorithms, Networks, Genome and Finance*, vol 796, John Wiley & Sons, New York, 2008.

29. H. Andersson, T. Britton, *Stochastic epidemic models and their statistical analysis*, vol. 151, Springer Science & Business Media, 2012.

30. M. Salathé, M. Kazandjieva, J. W. Lee, P. Levis, M. W. Feldman, J. H. Jones, A high-resolution human contact network for infectious disease transmission, *Proc. Natl. Acad. Sci. U.S.A.*, **107** (2010), 22020–22025.

31. A. Corral, Long-term clustering, scaling, and universality in the temporal occurrence of earthquakes, *Phys. Rev. Lett.*, **92** (2004), 108501.

32. P. S. Stumpf, R. C. Smith, M. Lenz, A. Schuppert, F. J. Müller, A. Babtie, et al., Stem cell differentiation as a non-Markov stochastic process, *Cell Syst.*, **5** (2017), 268–282.

33. D. M. Suter, N. Molina, D. Gatfield, K. Schneider, U. Schibler, F. Naef, Mammalian genes are transcribed with widely different bursting kinetics, *Science*, **332** (2011), 472–474.

34. T. Guérin, O. Bénichou, R. Voituriez, Non-Markovian polymer reaction kinetics, *Nat. Chem.*, **4** (2012), 568–573.

35. A. L. Barabasi, The origin of bursts and heavy tails in human dynamics, *Nature*, **435** (2005), 207–211.

36. J. M. Pedraza, J. Paulsson, Effects of molecular memory and bursting on fluctuations in gene expression, *Science*, **319** (2008), 339–343.

37. G. Srinivasan, D. M. Tartakovsky, B. A. Robinson, A. B. Aceves, Quantification of uncertainty in geochemical reactions, *Water Resour. Res.*, **43** (2007), W12415.

38. S. Condamin, O. Bénichou, V. Tejedor, R. Voituriez, J. Klafter, First-passage times in complex scale-invariant media, *Nature*, **450** (2007), 77–80.

39. G. Guigas, M. Weiss, Sampling the cell with anomalous diffusion—the discovery of slowness, *Biophys. J.*, **94** (2008), 90–94.

40. Y. Meroz, I. M. Sokolov, J. Klafter, Distribution of first-passage times to specific targets on compactly explored fractal structures, *Phys. Rev. E*, **83** (2011), 020104.

41. M. Dentz, A. Russian, P. Gouze, Self-averaging and ergodicity of subdiffusion in quenched random media, *Phys. Rev. E*, **93** (2016), 010101.

42. A. A. Ovchinnikov, Y. B. Zeldovich, Role of density fluctuations in bimolecular reaction kinetics, *Chem. Phys.*, **28** (1978), 215–218.

43. M. Dobrzyński, F. J. Bruggeman, Elongation dynamics shape bursty transcription and translation, *Proc. Natl. Acad. Sci. U.S.A.*, **106** (2009), 2583–2588.

44. D. R. Larson, D. Zenklusen, B. Wu, J. A. Chao, R. H. Singer, Real-time observation of transcription initiation and elongation on an endogenous yeast gene, *Science*, **332** (2011), 475–478.

45. S. Yunger, L. Rosenfeld, Y. Garini, Y. Shav-Tal, Single-allele analysis of transcription kinetics in living mammalian cells, *Nat. Methods*, **7** (2010), 631–633.

46. I. Golding, J. Paulsson, S. M. Zawilski, E. C. Cox, Real-time kinetics of gene activity in individual bacteria, *Cell*, **123** (2005), 1025–1036.

47. T. Muramoto, D. Cannon, M. Gierliński, A. Corrigan, G. J. Barton, J. R. Chubb, Live imaging of nascent RNA dynamics reveals distinct types of transcriptional pulse regulation, *Proc. Natl. Acad. Sci. U.S.A.*, **109** (2012), 7350–7355.

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

49. D. G. Spiller, C. D. Wood, D. A. Rand, M. R. White, Measurement of single-cell dynamics, *Nature*, **465** (2010), 736–745.

50. A. Eldar, M. B. Elowitz, Functional roles for noise in genetic circuits, *Nature*, **467** (2010), 167–173.

51. B. Zoller, D. Nicolas, N. Molina, F. Naef, Structure of silent transcription intervals and noise characteristics of mammalian genes, *Mol. Syst. Biol.*, **11** (2015), 823.

52. T. R. Sokolowski, T. Erdmann, P. R. Ten Wolde, Mutual repression enhances the steepness and precision of gene expression boundaries, *PLoS Comput. Biol.*, **8** (2012), e1002654.

53. J. Paulsson, Summing up the noise in gene networks, *Nature*, **427** (2001), 415–418.

54. D. R. Larson, What do expression dynamics tell us about the mechanism of transcription?, *Curr. Opin. Gen. Dev.*, **21** (2011), 591–599.

55. V. Shahrezaei, P. S. Swain, Analytical distributions for stochastic gene expression, *Proc. Natl. Acad. Sci. U.S.A.*, **105** (2008), 17256–17261.

56. N. Kumar, A. Singh, R. V. Kulkarni, Transcriptional bursting in gene expression: analytical results for general stochastic models, *PLoS Comput. Biol.*, **11** (2015), e1004292.

57. Z. Wang, Z. Zhang, T. Zhou, Exact distributions for stochastic models of gene expression with arbitrary regulation, *Sci. China Math.*, **63** (2020), 485–500.

58. P. Liu, Z. Yuan, L. Huang, T. Zhou, Roles of factorial noise in inducing bimodal gene expression, *Phys. Rev. E*, **91** (2015), 062706.

59. J. Zhang, Q. Nie, T. Zhou, A moment-convergence method for stochastic analysis of biochemical reaction networks, *J. Chem. Phys.*, **144** (2016), 194109.

60. A. B. O. Daalhuis, *Confluent hypergeometric functions*, *NIST Handb. Math. Funct.*, 2010.

61. T. Aquino, M. Dentz, Chemical continuous time random walks, *Phys. Rev. Lett.*, **119** (2017), 230601.

62. N. Masuda, M. A. Porter, R. Lambiotte, Random walks and diffusion on networks, *Phys. Rep.*, **716** (2017), 1–58.

63. R. Kutner, J. Masoliver, The continuous time random walk, still trendy: fifty-year history, state of art and outlook, *Eur. Phys. J. B*, **90** (2017), 50.

64. L. Liu, B. R. K. Kashyap, J. G. C. Templeton, On the GIX/G/∞ system, *J. Appl. Prob.*, **27** (1990), 671–683.

65. A. R. Stinchcombe, C. S. Peskin, D. Tranchina, Population density approach for discrete mRNA distributions in generalized switching models for stochastic gene expression, *Phys. Rev. E*, **85** (2012), 061919.

66. N. Masuda, L. E. Rocha, A Gillespie algorithm for non-Markovian stochastic processes, *SIAM Rev.*, **60** (2018), 95–115.

67. C. Deneke, R. Lipowsky, A. Valleriani, Complex degradation processes lead to non-exponential decay patterns and age-dependent decay rates of messenger RNA, *PloS One*, **8** (2013), e55442.

68. B. C. Arnold, Majorization: Here, there and everywhere, *Statist. Sci.*, **22** (2007), 407–413.

69. A. David, S. Larry, The least variable phase type distribution is Erlang, *Stochastic Models*, **3** (1987), 467–473.

70. J. Zhang, T. Zhou, Markovian approaches to modeling intracellular reaction processes with molecular memory, *Proc. Natl. Acad. Sci. U.S.A.*, **116** (2019), 23542–23550.

71. H. Qiu, B. Zhang, T. Zhou, Analytical results for a generalized model of bursty gene expression with molecular memory, *Phys. Rev. E*, *100* (2019), 012128.

72. A. Coulon, C. C. Chow, R. H. Singer, D. R. Larson, Eukaryotic transcriptional dynamics: From single molecules to cell populations, *Nat. Rev. Genet.*, **14** (2013), 572–584.

73. W. J. Blake, M. Kærn, C. R. Cantor, J. J. Collins, Noise in eukaryotic gene expression, *Nature*, **422** (2003), 633–637.

74. J. M. Raser, E. K. O'Shea, Control of stochasticity in eukaryotic gene expression, *Science*, **304** (2004), 1811–1814.

75. N. Friedman, L. Cai, X. S. Xie, Linking stochastic dynamics to population distribution: an analytical framework of gene expression, *Phys. Rev. Lett.*, **97** (2006), 168302.

76. A. M. Kringstein, F. M. Rossi, A. Hofmann, H. M. Blau, Graded transcriptional response to different concentrations of a single transactivator, *Proc. Natl. Acad. Sci. U.S.A.*, **95** (1998), 13670–13675.

77. J. Stewart-Ornstein, C. Nelson, J. DeRisi, J. S. Weissman, H. El-Samad, Msn2 coordinates a stoichiometric gene expression program, *Curr. Biol.*, **23** (2013), 2336–2345.

78. J. Paulsson, M. Ehrenberg, Noise in a minimal regulatory network: plasmid copy number control, *Quart. Rev. Biophys.*, **34** (2001), 1–59.

79. M. B. Elowitz, A. J. Levine, E. D. Siggia, P. S. Swain, Stochastic gene expression in a single cell, *Science*, **297** (2002), 1183–1186.