Quantitative Biology 2015, 3(4): 168–181
DOI 10.1007/s40484-015-0056-8

Review

Quantitative analysis of gene expression systems

Tianshou Zhou* **and Tuoqi Liu**

School of Mathematics and Computational Science, Sun Yat-Sen University, Guangzhou 510275, China
* Correspondence: mcszhtsh@mail.sysu.edu.cn

Received October 23, 2015; Revised November 19, 2015; Accepted November 24, 2015

Gene expression is a complex biochemical process, involving many specific processes such as transcription, translation, switching between promoter states, and regulation. All these biochemical processes inevitably lead to fluctuations in mRNA and protein abundances. This noise has been identified as an important factor underlying the observed phenotypic variability of genetically identical cells in homogeneous environments. Quantifying the contributions of different sources of noise using stochastic models of gene expression is an important step towards understanding fundamental cellular processes and cell-to-cell variability in expression levels. In this paper, we review progresses in quantitative study of simple gene expression systems, including some results that we have not published. We analytically show how specific processes associated with gene expression affect expression levels. In particular, we derive the analytical decomposition of expression noise, which is important for understanding the roles of the factorial noise in controlling phenotypic variability. We also introduce a new index (called attribute factor) to quantify expression noise, which has more advantages than the commonly-used noise indices such as noise intensity and Fano factor.

Keywords: gene expression; chemical master equation; statistical quantities; binomial moments; expression noise

INTRODUCTION

The latest advances in biological experiments allow us to map not only the protein-coding genes in the genomes of prokaryotic or eukaryotic organisms but also the regulatory sequences present in these genomes. In particular, single-molecule and single-cell measurements allow direct observations of real-time fluctuations in gene expression levels in individual live cells [1–5]. A main challenge in the post-gene epoch is to understand how the regulatory sequences across a genome control the expression spectrum of every gene within a cell and how they collectively determine stochastic behavior of the entire gene regulatory network and further the cell's function.

Traditionally, the regulation of expression spectrum was studied in experiments that usually measured the average expression level in populations consisting of many genetically identical cells (the number of cells used in an experiment would be up to a few millions). These studies related the average expression level of a gene to its regulatory DNA sequence, but averaging over populations often conceals differences in gene expression, which not only may occur between individual cells [6] but also may in turn have consequences for the whole multi-cell system or organism [7]. Therefore, this apparent drawback of the average method makes it necessary to develop other more effective methods to understand gene expression in single cells as well as cell-to-cell variability in expression spectrums.

Within a single cell, gene expression is indigenously stochastic mainly because of the following three reasons: (i) Protein-coding genes are typically present in only one or two copies in a cell; (ii) Transcription initiation is a multi-step biochemical process; (iii) Whether a gene is transcribed at any given moment depends on the arrival of multiple transcription factors to their designated binding sites. From the viewpoint of biochemical reactions, gene expression involves transcription of DNA to mRNA, translation of mRNA to protein, transition or switching between promoter activity states, feedback regulation, alternative splicing, and RNA nuclear retention [8–14].

All these specific biochemical processes are stochastic due to the low copies of the involved reactive species. This stochasticity inevitably results in fluctuations in mRNA and protein abundances, and further cell-to-cell variability in expression levels. This variability is referred to as gene expression noise, which is a main analysis object of this paper.

As a key step of gene expression, transcription takes place often in a bursting fashion. Single-cell measurements have provided evidence for transcriptional bursting in prokaryotic cells [1] and in eukaryotic cells [2,15]. Although the sources of transcriptional bursting remain poorly understood [16], several lines of evidence [3,10,11,17–21] point to stochastic transitions among the active (ON) and inactive (OFF) states of gene promoter as an important source of expression noise, which is responsible for cell-to-cell heterogeneity in homogeneous environments. It has been shown that in contrast to Poissonian transcription, where mRNA is synthesized in random, uncorrelated events with a probability being uniform over time, bursty transcription, where mRNA is produced in episodes of high transcriptional activity (bursts) followed by long periods of inactivity, typically leads to higher expression noise [22].

Given the inherent stochasticity and complexity of gene expression, several questions are raised: (i) Are the stochastic dynamics of gene expression — and therefore the resulting cell-to-cell variability in mRNA and protein levels — encoded by the promoter regulatory sequence, just as the mean expression level of a gene appears to be? (ii) How are reasonable gene models developed and analyzed with results that can interpret experimental phenomena? (iii) What are roles of specific processes associated with gene expression in controlling the expression spectrum of a gene? There are other questions, which either have been partially studied or need to be further studied. Refer to Figure 1, where we list aspects of common interest in quantitative analysis of stochastic gene expression models.

We will address some of the questions listed in Figure 1 (see the deep grey parts) by reviewing progresses in the study of gene expression systems combined with some of our works that were recently finished but have not been published elsewhere. For the understanding convenience of the reader, we will first consider two representative gene models: the one is an extended version of the common ON-OFF model, which considers not only two kinds of regulations (i.e., positive and negative feedbacks) but also promoter leakage (i.e., there is a small transcription rate at the promoter's inactive state in contrast to that at the active state); the other is an extended version of the common gene model of stochastic transcription, where the ON and OFF waiting-time distributions are not exponential but are general. We will derive analytical results that not only can explain some important experimental phenomena but also can make biologically reasonable predictions as well as would provide guidelines for designing functional gene circuits. Then, we will consider a general gene model at the transcription level, where the promoter is assumed to contain multiple ON states and multiple OFF states and that there are stochastic transitions among these states. We will also derive analytical results, including analytical mRNA distribution and mean ON and OFF times. The overall review focuses on generality and accuracy of theoretical results, applicability of analysis methods and elucidation of important mechanisms underlying cell-to-cell variability in expression levels.

QUANTITATIVE ANALYSIS OF TWO REPRESENTATIVE GENE MODELS

First, we introduce a representative model of gene expression, referring to Figure 2A. It considers three

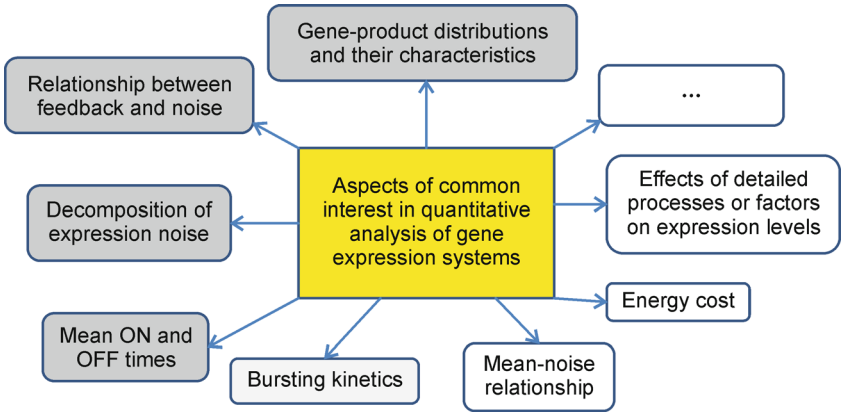

Figure 1. Aspects of common interest in quantitative analysis of gene expression systems in single cells. This paper focuses on research progresses in some of them indicated in deep grey, including our own works unpublished elsewhere.

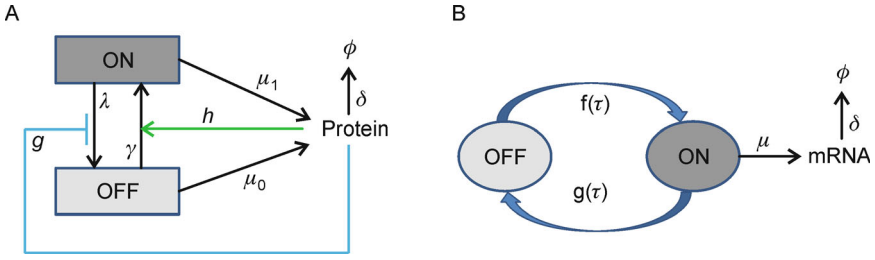

Figure 2. Schematic for two representative ON-OFF models of gene expression. (A) The model considers not only regulations but also promoter leakage; (B) The model considers ON and OFF mechanisms with general waiting-time distributions.

kinds of fundamental biochemical processes: (i) Stochastic switching between two promoter activity states (the active state is denoted by A and the inactive state by I); (ii) Two kinds of regulations, i.e., positive and negative feedbacks. To derive analytical results, however, we assume that feedback regulation is linear although for nonlinear regulation, it is possible to derive some analytical results [23]; (iii) Promoter leakage, i.e., there is a small transcription rate at the OFF state, compared to that at the ON state. Thus, our model contains most of the common ON-OFF models studied in the literature [24–34] as its particular case. For this model, we first derive the analytical distribution of gene product, then give characteristics of statistical quantities such as noise intensity and attribute factor first introduced in this paper, and finally analyze the roles of factorial noise (i.e., a part of the total noise) in inducing the bimodal expression of gene product.

Then, we introduce another representative gene model at the transcription level, where waiting time distributions at ON and OFF are not exponential but are general, referring to Figure 2B. For this kind of model, we first derive an integral equation that can be used to calculate moments of the mRNA distribution, and then give an explicit expression for the common noise index and discuss its characteristics.

We point that the analysis methods used to derive our analytical results are very general, and can be applied to stochastic analysis of any reaction networks.

Explicit distribution

The distribution of the molecule numbers of the reactive species in a biochemical system of interest is very important for understanding stochastic behavior and properties of this system. Thus, finding this distribution is common interest although it is a challenging task in many cases.

First, we consider model (A) shown in schematic Figure 2. The corresponding chemical master equation (CME) reads

$$\begin{aligned}\frac{\partial P_1(n,\,t)}{\partial t} &= -\lambda P_1(n,\,t)+\gamma P_0(n,\,t)+hnP_0(n,\,t)\\ &\quad -gnP_1(n,\,t)+\mu_1[P_1(n-1,\,t)-P_1(n,\,t)]\\ &\quad +\delta[(n+1)P_1(n+1,\,t)-nP_1(n,\,t)]\\ \frac{\partial P_0(n,\,t)}{\partial t} &= \lambda P_1(n,\,t)-\gamma P_0(n,\,t)-hnP_0(n,\,t)\\ &\quad +gnP_1(n,\,t)+\mu_0[P_0(n-1,\,t)-P_0(n,\,t)]\\ &\quad +\delta[(n+1)P_0(n+1,\,t)-nP_0(n,\,t)]\end{aligned} \tag{1}$$

where $P_i(n,\,t)$ represents the probability that the gene product has n molecules at state-I ($i=1$ stands for ON whereas $i=0$ for OFF) at time t. If the total probability is denoted as $P(n,\,t)$, then $P(n,\,t)=P_0(n,\,t)+P_1(n,\,t)$ according to the sum law of probability. In Equation (1), λ and γ are transition rates from ON to OFF states and vice versa, respectively; μ_1 and μ_0 are transcription rates at ON and OFF states, respectively (We assume $\mu_1 \gg \mu_0$ in this paper since the latter describes promoter leakage); h and g represent strengths of positive and negative feedbacks, respectively; and δ is the degradation rate of gene product. For convenience, we rescale all the parameters by δ in the following, that is, $\tilde{\lambda}=\lambda/\delta$, $\tilde{\gamma}=\gamma/\delta$, $\tilde{h}=h/\delta$, $\tilde{g}=g/\delta$, and $\tilde{\mu}_i=\mu_i/\delta$ with $i=0,\,1$. Our interest is to find the stationary distribution $P(n)$.

There are several efficient methods to find the analytical expression of this distribution [26,31–34]. Here, we adopt the Poisson representation method [35] to solve Equation 1. For this, we introduce two factorial functions $\rho_0(s)$ and $\rho_1(s)$, which are related to two factorial distributions $P_0(n)$ and $P_1(n)$ by

$$P_i(n)=\int_0^{s_{\max}}\rho_i(s)e^{-s}\frac{s^n}{n!}ds, \qquad i=0,\,1 \tag{2}$$

where the total function $\rho(s)=\rho_0(s)+\rho_1(s)$ satisfies the normalization condition $\int\rho(s)ds=1$ due to the probability conservative condition $\sum_{n\geqslant 0}P(n)=1$. By sub-

stituting into Equation (2), we can obtain a group of differential equations regarding $\rho_0(s)$ and $\rho_1(s)$. To guarantee the uniqueness of the corresponding solution, we impose the boundary conditions: $\int_0^{s_{\max}} e^{-s}[(s-\tilde{\mu}_0)/\left((1+\tilde{g}+)\tilde{h}\right)](s^n/n!)ds=0,\ n=0,\ 1,2,\ \ldots$. Solving this equation group with the given boundary conditions, we obtain the following expression of $\rho(s)$

$$\rho(s)=Ce^{\frac{\tilde{g}+\tilde{h}}{\tilde{g}+\tilde{h}+1}s}(s-\tilde{\mu}_0)^{-2}(\alpha_1-s)^{\frac{A(\tilde{g}+\tilde{h})}{\tilde{g}+\tilde{h}+1}+1}(s-\alpha_2)^{\frac{B(\tilde{g}+\tilde{h})}{\tilde{g}+\tilde{h}+1}+1} \tag{3}$$

where C is a normalization constant. In Equation (3), we have denoted

$$A=\frac{(\alpha_1-\beta_1)(\alpha_1-\beta_2)}{\alpha_1-\alpha_2},\qquad B=\frac{(\alpha_2-\beta_1)(\alpha_2-\beta_2)}{\alpha_2-\alpha_1} \tag{3A}$$

$$\alpha_{1,2}=\frac{(1+\tilde{g})\tilde{\mu}_0+\left(1+\tilde{h}\right)\tilde{\mu}_1\pm\sqrt{\left[(1+\tilde{g})\tilde{\mu}_0+\left(1+\tilde{h}\right)\tilde{\mu}_1\right]^2-4\left(1+\tilde{g}+\tilde{h}\right)\tilde{\mu}_0\tilde{\mu}_1}}{2\left(\tilde{g}+\tilde{h}+1\right)} \tag{3B}$$

$$\beta_{1,2}=\frac{\tilde{g}\tilde{\mu}_0+\tilde{h}\tilde{\mu}_1-\tilde{\lambda}-\tilde{\gamma}\pm\sqrt{\left(\tilde{g}\tilde{\mu}_0+\tilde{h}\tilde{\mu}_1-\tilde{\lambda}-\tilde{\gamma}\right)^2+4\left(\tilde{g}+\tilde{h}\right)\left(\tilde{\lambda}\tilde{\mu}_0+\tilde{\gamma}\tilde{\mu}_1\right)}}{2\left(\tilde{g}+\tilde{h}\right)} \tag{3C}$$

As such, the distribution of gene product can be formally expressed as

$$P(n)=\frac{C}{n!}\int_0^{s_{\max}} e^{\frac{-s}{\tilde{g}+\tilde{h}+1}}s^n(s-\tilde{\mu}_0)^{-2}(\alpha_1-s)^{\frac{A(\tilde{g}+\tilde{h})}{\tilde{g}+\tilde{h}+1}+1}(s-\alpha_2)^{\frac{B(\tilde{g}+\tilde{h})}{\tilde{g}+\tilde{h}+1}+1}ds \tag{4}$$

In theory, this expression can reproduce many previously-derived distributions. Here, we give its explicit expressions in two particular cases.

In general, the promoter leakage rate, i.e., the transcription rate at OFF state is much smaller than that at the active state, i.e., $\mu_0 \ll \mu_1$. To derive the analytical expression of the distribution, we assume $\mu_0 \approx 0$ for simplicity. In this case, it is found from Equation (4) that

$$p(n)=\frac{(\xi\widetilde{\mu}_1)^n}{n!}\frac{\Gamma(n+\alpha)\Gamma(\beta)}{\Gamma(\alpha)\Gamma(n+\beta)}{}_1F_1(n+\alpha,\ n+\beta;-\xi\widetilde{\mu}_1) \tag{5}$$

where ${}_1F_1(a,\ b;z)=\sum_{n=0}^{\infty}[(a)_n/(b)_n](z^n/n!)$ with $(c)_n$ being the Pochhammer symbol defined as $(c)_n=\Gamma(c+n)/\Gamma(c)$ is a confluent hypergeometric function, $\alpha=\tilde{\gamma}/\left(\tilde{h}+1\right)$, $\beta=\left(\tilde{\lambda}+\tilde{\gamma}\right)/\left(\tilde{h}+\tilde{g}+1\right)+(\tilde{g}\tilde{\mu}_1)/\left(\tilde{h}+\tilde{g}+1\right)^2$, and $\xi=\left(\tilde{h}+1\right)/\left(\tilde{h}+\tilde{g}+1\right)^2$. Note that in the absence of feedback, i.e., if $\tilde{h}=\tilde{g}=0$, the above expression can reproduce the mRNA distribution in the common ON-OFF model of stochastic transcription [4,26].

Then, we consider model (B) shown in Figure 2. In this case, the mRNA distribution in general cannot be analytically derived since waiting-time distributions are general. In spite of this, we can derive integral equations for both the moment-generating function and moments of the distribution. In fact, denote by $W(z;t)$ the moment-generating function of the distribution at time t, and in particular, denote $W_{init}(z)=W(z;0)$. Suppose that there are N molecules at time $t=0$, where N itself is a random variable. Then, every mRNA at time t, which is also a stochastic variable and is denoted by X_i, $1\leqslant i\leqslant N$, has a survival probability $p(t)=1-D(t)=e^{-\delta t}$, where $D(t)=\int_0^t \delta e^{-\delta s}ds$ represents the cumulative distribution function for the mRNA lifetime. For simplicity, we assume that these variables X_i are independent of one another, each following a Bernoulli distribution with the moment-generating function given by $M(z)=1+p(t)(e^z-1)=1+e^{-\delta t}(e^z-1)$ [36–38]. Note that the total molecule number of mRNA at time t is given by $S=X_1+\cdots+X_N$. Using the law of total expectation, we can know

$$W(z;t)=W_{init}(\log(1+e^{-\delta t}(e^z-1))) \tag{6}$$

where we have assumed that every X_i is independent of N. In particular, at the end of an OFF-state, $W(z;t_{off})$ can be given by integrating $W_{init}(\log(1+e^{-\delta t}(e^z-1)))$ over the interval $(0,\ \infty)$, that is,

$$W(z;t_{off})=\int_{t=0}^{\infty} W_{init}(\log(1+e^{-\delta t}(e^z-1)))f_{off}(t)dt \tag{7}$$

where $f_{off}(s)=\int_0^{\infty}F(\varepsilon,\ s)d\varepsilon$ with $F(u,\ s)$ representing the joint probability density function of OFF and ON times is the distribution of times that the gene dwells at the OFF state (i.e., OFF times).

Except that an ON-period degradation of the mRNA

molecules that have been present at the beginning of the burst continues as described above through the function $1+e^{-\delta u}(e^z-1)$ with $W_{\deg}(z;t_{off}+u)=W_{init}(\log(1+e^{-\delta(u+t_{off})}(e^z-1)))$, mRNAs are also created and degraded according to a birth-death process with an exponential waiting time. Moreover, the distribution for this birth-death process is a Poisson distribution with the average $\rho_u=\frac{\mu}{\delta}(1-e^{-\delta u})$ and the moment-generating function $W_{effect}(z;t_{off}+u)=e^{\rho_u(e^z-1)}$ [38]. Therefore, their combination produces an effective burst size. During an ON-state, the probability distribution of the mRNA number is given by the convolution of the distribution of the number of those molecules that are still present from previous bursts and the effective burst-size distribution. This can be expressed as

$$W(z;t_{off}+u)=\int_{s=0}^{\infty}W_{init}(\log(1+e^{-\delta(s+u)}(e^z-1))) \cdot e^{\rho_u(e^z-1)}f_{off}(s)ds \tag{8}$$

Note that one complete OFF and ON cycle defines a boundary condition:

$$W(z;0)=W(z;t_{on}+t_{off})=\int_{s=0}^{\infty}\int_{t=s}^{\infty}W(z;t)f_{on}(t-s)f_{off}(t)dtds \tag{9}$$

where $f_{on}(u)=\int_0^{\infty}F(u,\ \theta)d\theta$ represents the distribution of ON times. Combining Equation (9) with Equation (6) and Equation (8), we thus arrive at the following integral equation with respect to $W_{init}(z)$

$$W_{init}(z)=W(z;0)=\int_{s=0}^{\infty}\int_{t=0}^{\infty}W_{init}(\log(1+e^{-\delta(s+t)}(e^z-1))) \cdot e^{\rho_u(e^z-1)}f_{on}(t)f_{off}(s)dtds \tag{10}$$

which is a pivot for derivation of analytical results. For example, moments of the total mRNA distribution can be expressed as

$$\langle m^k\rangle=\frac{1}{\tau_{on}+\tau_{off}} \cdot\left[\int_{s=0}^{\infty}\langle m^k\rangle_s(1-f(s))ds+\int_{u=0}^{\infty}\langle m^k\rangle_u(1-g(u))du\right] \tag{11}$$

with $k=1,\ 2,$ where $\tau_{off}=\int_{u=0}^{\infty}\int_{s=0}^{\infty}sF(u,\ s)dsdu$ and $\tau_{on}=\int_{s=0}^{\infty}\int_{u=0}^{\infty}uF(u,\ s)duds$ are the mean OFF and ON times respectively, whereas $f(s)$ and $g(s)$ are the cumulative functions of the duration distributions at OFF- and ON-states respectively. In Equation (11), $\langle m^k\rangle_s$ is the k^{th} -order moment of the mRNA distribution at time $t=s$ during the OFF-state, which can be given by the k^{th} -order derivative of Equation (6) at $z=0$, whereas $\langle m^k\rangle_u$ is the k^{th} moment of the mRNA distribution at time $t=t_{off}+u$ during the ON-state, which can be given by the k^{th} -order derivative of Equation (8) at $z=0$.

In principle, the moments obtained by Equation (11) can be used to approximate the mRNA distribution. In fact, based on these common moments, we can calculate so-called binomial moments [39–41], which, e.g., in the one-dimensional case, are defined as

$$b_k=\sum_{m\geqslant k}\binom{m}{k}P(m)=\frac{1}{k!}\sum_{m\geqslant k}m(m-1)\cdots(m-k+1)P(m), \quad k=0,\ 1,\ 2,\ \cdots \tag{12}$$

where the sum function on the right-hand side can be expressed as the linear combination of the common original moments given above. In turn, these binomial moments can be used to reconstruct the corresponding distribution according to the formula

$$P(m)=\sum_{k\geqslant m}(-1)^{m-k}\binom{k}{m}b_k, \quad m=0,\ 1,\ 2,\ \cdots \tag{13}$$

Note that unlike common moments that are divergent as their orders go to infinity, binomial moments are convergent as their orders go to infinity [39]. In addition, we point out that binomial moments defined above are easily extended to cases of many random variables.

Decomposition and characteristics of expression noise

Noise decomposition is an interesting topic. Some authors studied noise decomposition in gene regulatory systems, e.g., Weinberger, et al., showed that dynamics of protein noise can distinguish between alternate sources of variability in gene expression levels [42]. Other authors studied noise decomposition in general biochemical networks, e.g., Levchenko, et al., gave an empirical decomposition of the total noise (including intrinsic and extrinsic noise) in intracellular biochemical signaling networks using nonequivalent reporters [43], and Bowsher, et al., discussed the issue of noise decomposition in some biological networks and elucidated the biological significance of their noise decomposition [44]. Here, we are interested in accurate decomposition and essential characteristics of the total intrinsic noise in several representative models of gene expression, focusing on the tracing and dissecting of intrinsic noisy sources.

Note that it is in general difficult to see features of the total noise directly from the above analytical distributions. Therefore, we turn to considering statistical indices of a distribution such as noise intensity that is defined as the ratio of the variance over the square of the mean and a new statistical index (called the attribute factor of noise in this paper) that we will introduce. These indices can not only provide intuitive understanding of expression noise in contrast to the distribution bust also simplify stochastic analysis, in particular for complex reaction networks.

First, we consider expression noise in the common sense. Let τ_{on} and τ_{off} represent the mean residence times at the active and inactive states of a gene respectively, and $\langle n\rangle$ represent the mean expression of gene product. Denote by η_n the intensity of common expression noise. Then, in the absence of feedback (i.e., $h=g=0$) and without promoter leakage (i.e., $\mu_0=0$), we find that for each of the above two gene models, the noise intensity can be expressed

$$\eta_{classic}=\frac{1}{\langle n\rangle}+\frac{\tau_{off}^2}{\tau_{off}\tau_{on}+\tau_{off}+\tau_{on}} \tag{14}$$

where the first term on the right-hand side, i.e., $\xi_{birth-death}=1/\langle n\rangle=(\tau_{on}+\tau_{off})/(\mu_1\tau_{on})$, represents the intensity of the factorial noise due to the random birth and death of mRNA whereas the second term, i.e., $\xi_{promoter}=\tau_{off}^2/(\tau_{off}\tau_{on}+\tau_{off}+\tau_{on})$ represents the intensity of the factorial noise from switching between promoter states (i.e., so-called promoter noise). Note that $\eta_{classic}$ calculated by Equation (14) is exact in the case of no feedback but approximate in the presence of feedback since the mean level of gene product is changed in this case (in fact, the mean level is given by $\langle n\rangle=(\tilde{\mu}_1\tau_{on})/\left[(\tilde{h}+\tilde{g}+1)(\tau_{on}+\tau_{off})+\tilde{g}\tilde{h}\tilde{\mu}_1\tau_{on}\tau_{off}\right]$, which depends on feedback strengths). In the latter case, the expression noise (or the total noise) should be modified as [33,45,46]

$$\eta_{exact}=\eta_{classic}+g_{correction} \tag{15}$$

where the first term on the right-hand side represents the approximate noise calculated by Equation (14) whereas the second term represents the feedback-induced additional contribution to the expression noise. By calculation, we find

$$g_{correction}=\frac{1}{\tilde{\gamma}}\left[\frac{(\tilde{h}+\tilde{g}+1)(\tilde{\lambda}+\tilde{h}\tilde{\lambda}-\tilde{g}\tilde{\gamma})+\tilde{\mu}_1\tilde{g}\tilde{h}}{(\tilde{h}+\tilde{g}+1)(\tilde{\lambda}+\tilde{\gamma}+\tilde{h}+\tilde{g}+1)+\tilde{\mu}_1\tilde{g}\tilde{h}}-\frac{\tilde{\lambda}}{\tilde{\lambda}+\tilde{\gamma}+1}\right] \tag{16}$$

which analytically shows how the positive or negative feedback strength impacts this correction. To see this impact more clearly, we plot Figure 3. We observe from this figure that negative feedback reduces the expression noise, referring to the shadowed part in Figure 3A, whereas the positive feedback enlarges this noise, referring to the shadowed part in Figure 3B. In particular, if only positive feedback appears, then the correction becomes

$$g_{correction}=\frac{\tilde{\lambda}}{\tilde{\gamma}}\frac{\tilde{h}(\tilde{\lambda}+\tilde{\gamma})}{(\tilde{\lambda}+\tilde{\gamma}+1)(\tilde{\lambda}+\tilde{\gamma}+\tilde{h}+1)}>0 \tag{17}$$

If only negative feedback appears, then it becomes

$$g_{correction}=\frac{\tilde{g}}{\tilde{\gamma}}\frac{-(\tilde{g}+1)(\tilde{\lambda}\tilde{\gamma}+\tilde{\gamma}^2+\tilde{\gamma}+\tilde{\lambda})}{(\tilde{\lambda}+\tilde{\gamma}+1)(\tilde{g}+1)(\tilde{\lambda}+\tilde{\gamma}+\tilde{g}+1)}<0 \tag{18}$$

Then, we introduce an attribute factor to quantify

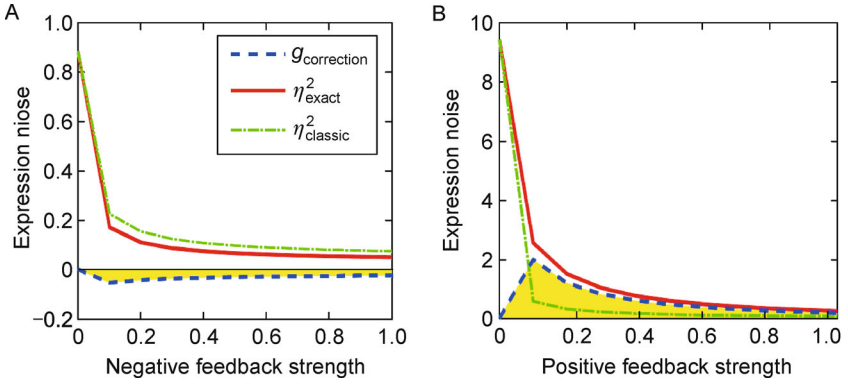

Figure 3. (Color online) Feedback-induced additional contributions to expression noise (A) Negative feedback, where the mean protein is fixed at $\langle n\rangle=20$, and the other parameter values are set as $\gamma=0.1, \mu_1=40, \mu_0=0$ and $h=0$; (B) Positive feedback, where $\langle n\rangle=20, \lambda=16, \mu_1=40, \mu_0=0$ and $g=0$. The solid curve (red) represents the exact noise calculated by Equation (15), the dash-dotted curve (green) represents the approximate noise calculated by Equation (14), and the shadowed area (yellow) with the boundary (the blue dash curve) represents the additional contribution from feedback.

expression noise. Analogous to the definition of common noise intensity, we define the attribute factor as the ratio of the double of the second-order binomial moment over the square of the first-order binomial moment

$$\xi_{attribute}=\frac{2b_2}{b_1^2} \tag{19}$$

One will see that this factor has more advantages than the noise intensity in quantifying characteristics of expression noise.

In order to help understand this factor, let us consider the simplest birth-death process described by $\varnothing \xrightarrow{g} X \xrightarrow{\delta} \varnothing$. For this reaction model, we know that the molecule number of X follows a Poisson distribution, $P(n)=e^{-\lambda}\lambda^n/n!$, where $\lambda=g/\delta$ is a characteristic parameter of this distribution. Recall that the size of the Fano factor defined as the ratio of variance over mean can be used to judge whether a distribution is Poissonian [47]. Specifically, the distribution is sub-Poissonian if the Fano factor is less than 1; it is Poissonian if the Fano factor is equal to 1; and it is sup-Poissonian if the Fano factor is more than 1. Similarly, for the attribute factor introduced above, we have that if $\xi_{attribute}<1$, then the distribution is sub-Poissonian; if $\xi_{attribute}=1$, then the distribution is Poissonian; and if $\xi_{attribute}>1$, then the distribution is sup-Poissonian.

Interestingly, we find that for the common ON-OFF gene model at the transcription level, the attribute factor $\xi_{attribute}$ is given by

$$\begin{aligned}\xi_{attribute}&=\tilde{\xi}_{birth-death}+\xi_{promoter}\\&=1+\frac{\tau_{off}^2}{\tau_{on}+\tau_{off}+\tau_{on}\tau_{off}}>1\end{aligned} \tag{20}$$

which depends only on promoter structure but is irrelative to the transcription rate. Thus, the mRNA distribution is sup-Poissonian for this gene model. In the presence of feedback but without promoter leakage, we can show

$$\begin{aligned}\xi_{attribute}=1+\frac{1}{\tilde{\gamma}}&\frac{\tilde{\lambda}+\tilde{h}\tilde{\lambda}-\tilde{g}\tilde{\gamma}+\dfrac{\tilde{\mu}_1\tilde{g}\left(\tilde{h}+1\right)}{\tilde{h}+\tilde{g}+1}}{\tilde{\lambda}+\tilde{\gamma}+\tilde{h}+\tilde{g}+1+\dfrac{\tilde{\mu}_1\tilde{g}}{\tilde{h}+\tilde{g}+1}}\\&+\frac{\left(\tilde{\lambda}+\tilde{\gamma}\right)\left(\tilde{h}+\tilde{g}+1\right)+\tilde{g}\tilde{\mu}_1}{\tilde{\gamma}\tilde{\mu}_1}>1\end{aligned} \tag{21}$$

This indicates that the mRNA distribution is also sup-Poissonian.

We point out the following three points: (i) the above analysis gives partial reasons why $\xi_{attribute}$ is called the attribute factor; (ii) the attribute factor has many other advantages, e.g., it contains useful information on bursting kinetics; (iii) the above analytical results provide quantitative descriptions of essential intracellular processes. The related results for the first and second points will be published elsewhere.

Roles of factorial noise in controlling phenotypic variability

In this subsection, we investigate the role of factorial noise in controlling cellular phenotype (e.g., bimodality). From the above subsection, we know that the expression noise is composed of two parts: the one is from the birth and death of gene product and the other from stochastic switching between promoter states. Each part is called as factorial noise in this paper. As is well known, the distribution of gene product in an ON-OFF model can exhibit one peak or two distinct peaks, which correspond to different cellular phenotypes. Thus, a natural question is what the role of factorial noise is in inducing unimodality or bimodality.

For simplicity, we consider the common ON-OFF model at the transcription level, implying that we consider neither regulation nor promoter leakage. In this case, we adopt two approximations to elucidate the roles of two noisy sources in inducing bimodality [31]: continuous approximation and adiabatic approximation.

First, we consider continuous approximation. If the characteristic number of gene products (i.e., proteins) is very large as that in the deterministic case (i.e., $M=\mu_1/\delta\gg 1$), then the ratio (or concentration) $x=n/M$ may be considered as a continuous variable. Thus, we can easily derive the corresponding CME. By solving this equation, we obtain

$$P(x)=Ce^{\left(\overline{h}+\overline{g}\right)x}(1-x)^{\tilde{\gamma}+\overline{g}-1}\left(x-\tilde{\mu}_0/\tilde{\mu}_1\right)^{\frac{\overline{h}\tilde{\mu}_0+\tilde{\lambda}\tilde{\mu}_1}{\tilde{\mu}_1}-1} \tag{22}$$

where $\overline{h}=\tilde{h}\tilde{\mu}_1$, $\overline{g}=\tilde{g}\tilde{\mu}_1$, and C is a normalization constant determined by $\int_0^1 P(x)=1$.

Then, we consider adiabatic approximation. If the protein number fluctuations become significant compared to those from switching between promoter activity states (e.g., in prokaryotic cells), the CME will be reduced to another simpler model, where all the gene states are simply integrated by fast equilibrium. In this simplified model, the dominant noise is transcriptional or translational noise, which is generated due to the stochastic birth and death of protein. Moreover, the protein distribution is given by

$$P(n)=\frac{1}{{}_1F_1(\alpha,\ \beta;\widetilde{\mu}_1)}\frac{(\widetilde{\mu}_1)^n}{n!}\frac{(\alpha)_n}{(\beta)_n} \tag{23}$$

where $\alpha=\left(\tilde{\gamma}\tilde{\mu}_1+\tilde{\lambda}\tilde{\mu}_0\right)/\left(\tilde{h}\tilde{\mu}_1\right)$ and $\beta=\tilde{\mu}_1$.

In order to characterize the above two approximations, we introduce a ratio, which is defined as

$$Ratio = \frac{\text{promoter noise}}{\text{synthesis noise}} \tag{24}$$

Apparently, promoter noise is dominant if $Ratio \gg 1$ whereas translational noise is dominant if $Ratio \ll 1$. Thus, the former corresponds to the continuous approximation whereas the latter to the adiabatic approximation. Figure 4 shows how factorial noise can induce bimodality.

It should be pointed out that the above analysis can also obtain the decomposition of expression noise (Actually, we only can the formal expression according to the noise definition). Compared to the derived-above analytical decomposition in which the noisy sources are practically artificial, the former decomosition is essential since the traced sources of noise are releastic. However, the difference between them is not too big under some assumed conditions (detailed discussions are omitted here).

QUANTITATIVE ANALYSIS OF A GENERAL GENE MODEL

Complex promoters with more than two states are not the exception but the rule as combinatorial control of gene regulation by multiple species of transcription factors, and the latter case is widespread in eukaryotic cells [48]. Even those promoters that are regulated by a single transcription factor may have multiple states [2,49]. For bacterial cells, the promoters that are often 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 transcription factors 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 up to 128 [50]. In contrast, eukaryotic promoters are more complex, involving nucleosomes competing with or being removed by transcription factors [51]. In addition to the conventional regulation by transcription factors, the eukaryotic promoters can be also epigenetically regulated via histone modifications [52–54]. Such regulation may lead to very complex promoter kinetics [52].

Based on the above reasons, we introduce a general gene model at the transcription level, where the gene promoter contains many activity states due to different bindings of transcription factors to regulatory sites on DNA or other unspecified mechanisms. These activity states (N in total) are divided as L active states and K ($=N-L$) inactive states. For analysis convenience, we do not consider regulation. We use matrix $\mathbf{A}=(a_{ij})$ to describe promoter activity, diagonal matrix $\mathbf{\Lambda}=\mathrm{diag}(\mu_1, \mu_2, \cdots, \mu_N)$ to describe exits of transcription from DNA to mRNA, and diagonal matrix $\boldsymbol{\delta}=\mathrm{diag}(\delta_1, \delta_2, \cdots, \delta_N)$ to describe degradation of mRNA at promoter activity states. Let $P_k(m)$ represent the distribution that mRNA has m molecules at state-k of the promoter and $\mathbf{P}=(P_1, \cdots, P_N)^{\mathrm{T}}$ represent the column vector consisting of these factorial probabilities. Then, the CME for the corresponding gene model can be expressed as

$$\frac{d\mathbf{P}(m;t)}{dt} = \mathbf{A}\mathbf{P}(m;t) + \delta(\mathbf{E}-\mathbf{I})[m\mathbf{P}(m;t)]$$
$$+\mathbf{\Lambda}(\mathbf{E}^{-1}-\mathbf{I})[\mathbf{P}(m;t)] \tag{25}$$

where $\mathbf{E}$ is a vector of step operators and $\mathbf{I}$ is a vector of unit operators. Interestingly, the time-evolution equations

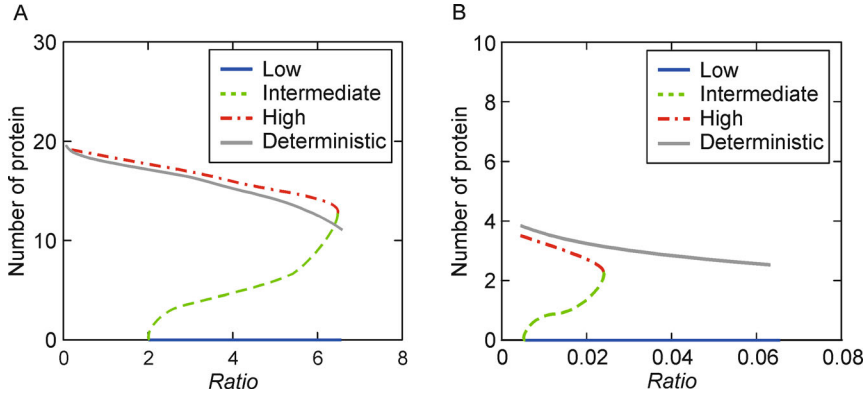

Figure 4. (Color online) Phase diagrams describing how stochastic bimodality is generated (A) Slow switching, where $\lambda=\gamma=0.5$; (B) Fast switching, where $\lambda=20$ and $\gamma=2$. In (A) and (B), the gray curve represents the monstable state in the deterministic case, and blue, dashed green and dot-dashed red curves each representing the noise-induced stable state in the stochastic case correspond respectively to the high state where the protein number is large, the middle state where the protein number is moderate (corresponding to the valley between two peaks of the distribution) and the low state where the protein number is small.

for the binomial moments defined above take the following simpler form in contrast to the CME

$$\frac{d}{dt}\mathbf{b}_k = \mathbf{A}\mathbf{b}_k + \mathbf{\Lambda}\mathbf{b}_{k-1} - k\boldsymbol{\delta}\mathbf{b}_k \tag{26}$$

where $k = 1,\ 2,\ \cdots$. Clearly, the first term on the right-hand side of Equation (26) describes kinetics 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 exits of transcription with the transcription matrix $\mathbf{\Lambda}$, and the third term describes the degradation dynamics of mRNA with the degradation matrix $\boldsymbol{\delta}$ (throughout this paper, we consider only the same degradation rate for simplicity, and denote it as δ). We point out that model (26) includes all previously-studied models of mRNA expression as its particular cases.

mRNA distributions

For simplicity, we consider a particular case, i.e., assume all the degradation rates are equal: $\delta_1 = \delta_2 = \cdots = \delta_N = \delta$ (implying that the degradation matrix takes the form of $\delta = \delta \mathrm{I}_N$ with I_N being the unit matrix). In addition, we rescale all the parameters by δ for convenience. In the following, we are only interested in the steady-state mRNA distribution.

Note that when solving Equation (26), we need to know the expression of $b_0 = (b_0^{(1)},\ b_0^{(2)},\ \cdots,\ b_0^{(N)})^T$, which can be given by noting the fact that the transition matrix $\mathbf{A}$ is an M-matrix (i.e., the sum of every column elements is equal to zero). By solving $\mathbf{u}_N b_0 = 1$ and $\mathbf{A}\mathbf{b}_0 = \mathbf{0}$, where $\mathbf{u}_N = (1,\ 1,\ \cdots,\ 1)$ is an $N-$ dimensional row vector, we find

$$b_0^{(k)} = \prod_{i=1}^{N-1} \frac{\beta_i^{(k)}}{\alpha_i}, \quad 1 \leqslant k \leqslant N \tag{27}$$

where $0,\ -\alpha_1,\ -\alpha_2,\ \cdots,\ -\alpha_{N-1}$ are the characteristic values of the rescaled matrix $\tilde{\mathbf{A}}$ and $-\beta_1^{(k)},\ -\beta_2^{(k)},\ \cdots,\ -\beta_{N-1}^{(k)}$ are the eigenvalues of the matrix $\mathbf{M}_k$ that is the minor one of the $N \times N$ matrix $\tilde{\mathbf{A}}$ by crossing out the k^{th} row and k^{th} column of its entry $\tilde{a}_{kk}$.

Also note that the total distribution is given by $P(m) = \sum_{k=1}^{N} P_k$ and the order-k binomial moment is calculated according to $b_k = \sum_{i=1}^{N} b_k^{(i)} = \mathbf{u}_N \cdot \mathbf{b}_k$. Thus, it follows from Equation (26) that

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

$$n = 1,\ 2,\ \cdots \tag{28}$$

where $\left(k\mathbf{I} - \tilde{\mathbf{A}}\right)^*$ and $\det\left(k\mathbf{I} - \tilde{\mathbf{A}}\right)$ are the adjacency matrix and the determinant of matrix $\left(k\mathbf{I} - \tilde{\mathbf{A}}\right)$, respectively. Once all the binomial moments are given by Equation (28), we can calculate the mRNA distribution according to the above Equation (13).

Now, we consider distributions in several particular cases. If all the rescaled transcription rates are equal, i.e., $\tilde{\mu}_1 = \tilde{\mu}_2 = \cdots = \tilde{\mu}_N = \tilde{\mu}$ (implying that the rescaled transcription matrix takes the form of $\tilde{\mathit{\Lambda}} = \tilde{\mu}\mathrm{I}_N$), then the number of mRNA molecules always follows a Poissonian distribution, independent of promoter structure. This is an interesting fact. If transcription matrix takes the form of $\tilde{\mathbf{\Lambda}} = \tilde{\mu}\begin{pmatrix} \mathbf{0}_{(N-1)} & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix}$, then the mRNA distribution takes the form

$$P(m) = \frac{\widetilde{\mu}^m}{m!} \prod_{i=1}^{N-1} \frac{(\beta_i^{(N)})_m}{(\alpha_i)_m}$$

$$\cdot_{N-1}F_{N-1}\left(\begin{matrix} m+\beta_1^{(N)},\ \cdots,\ m+\beta_{N-1}^{(N)} \\ m+\alpha_1,\ \cdots,\ m+\alpha_{N-1} \end{matrix}\middle|; -\widetilde{\mu}\right),$$

$$m = 0,\ 1,\ 2,\ \cdots \tag{29}$$

where ${}_nF_n\left(\begin{matrix} a_1,\ \cdots,\ a_n \\ b_1,\ \cdots,\ b_n \end{matrix}|; \sigma\right)$ is a confluent hypergeometric function [55]. If the rescaled transcription matrix takes the form: $\tilde{\mathbf{\Lambda}} = \tilde{\mu}\begin{pmatrix} \mathbf{I}_{(N-1)} & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix}$, then the mRNA distribution

$$P(m) = \frac{e^{-\widetilde{\mu}}}{m!} \sum_{k=0}^{m} \binom{m}{k} \widetilde{\mu}^{m-k} (-1)^k$$

$$\cdot \prod_{i=1}^{N-1} \frac{(\beta_1^{(N)})_k}{(\alpha_i)_k} {}_{N-1}F_{N-1}\left(\begin{matrix} k+\beta_1^{(N)},\ \cdots,\ k+\beta_{N-1}^{(N)} \\ k+\alpha_1,\ \cdots,\ k+\alpha_{N-1} \end{matrix}\middle|; \widetilde{\mu}\right) \tag{30}$$

In particular, for the common ON-OFF model of stochastic transcription, we find that the resulting mRNA distribution obtained by Equation (13) combined with Equation (28) can reproduce the one obtained in previous studies [4]. That is,

$$p(m) = \frac{\widetilde{\mu}^m}{m!} \frac{\mathbf{\Gamma}\left(\widetilde{\lambda} + m\right)\mathbf{\Gamma}\left(\widetilde{\lambda} + \widetilde{\gamma}\right)}{\mathbf{\Gamma}\left(\widetilde{\lambda} + \widetilde{\gamma} + m\right)\mathbf{\Gamma}\left(\widetilde{\lambda}\right)}$$

$$\cdot {}_1F_1\left(\widetilde{\lambda} + m,\ \widetilde{\lambda} + \widetilde{\gamma} + m; -\widetilde{\mu}\right) \tag{31}$$

Waiting time distributions and mean waiting times

As is seen from the above, the mean ON and OFF times

are important for calculating the mean mRNA level and studying the mRNA noise. In the case that the gene promoter has multiple ON and OFF states, in order to give mean ON and OFF times, we need to calculate the distributions of ON and OFF times.

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. Similarly, we can introduce matrix $\mathbf{A}_{01}$. Denote by $\mathbf{A}=(a_{ij})$ the $N\times N$ transition matrix, which consist of four block matrices $\mathbf{A}_{11}$, $\mathbf{A}_{00}$, $\mathbf{A}_{10}$ and $\mathbf{A}_{01}$. Matrix $\mathbf{\Lambda}=\mathrm{diag}(\mu_1, \cdots, \mu_N)$ describes exits of transcription with μ_i representing the transcription rate of mRNA in state-i ($\mu_i=0$ means that no transcription takes place). Thus, two matrices $\mathbf{A}$ and $\mathbf{\Lambda}$ altogether determine the promoter structure completely

Assume that the promoter states begin to transition from OFF (ON) to ON (OFF) at time $t=0$. Define $Q_i^{(1)}(\tau)$ $(i=1, \cdots, L)$ and $Q_k^{(0)}(\tau)$ $(k=1, \cdots, K)$ as the subsequent "survival" probabilities that the promoter is still at the i^{th} ON and at the k^{th} OFF state at time $t=\tau>0$, respectively. If we denote $\mathbf{Q}^{(0)}(\tau)=(Q_1^{(0)}(\tau), \cdots, Q_K^{(0)}(\tau))^{\mathrm{T}}$ and $\mathbf{Q}^{(1)}(\tau)=(Q_1^{(1)}(\tau), \cdots, Q_L^{(1)}(\tau))^{\mathrm{T}}$, then according to the corresponding CME, we can show

$$\begin{aligned}\mathbf{Q}^{(0)}(\tau)&=\exp(\mathbf{A}_{00}\tau)\mathbf{Q}^{(0)}(0)\\ \mathbf{Q}^{(1)}(\tau)&=\exp(\mathbf{A}_{11}\tau)\mathbf{Q}^{(1)}(0)\end{aligned} \tag{32}$$

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

$$\begin{aligned}\tilde{f}_{off}(\tau)&=\mathbf{u}_L\mathbf{A}_{01}\mathbf{Q}^{(0)}(\tau)=\mathbf{u}_L\mathbf{A}_{01}\exp(\mathbf{A}_{00}\tau)\mathbf{Q}^{(0)}(0)\\ \tilde{f}_{on}(\tau)&=\mathbf{u}_K\mathbf{A}_{10}\mathbf{Q}^{(1)}(\tau)=\mathbf{u}_K\mathbf{A}_{10}\exp(\mathbf{A}_{11}\tau)\mathbf{Q}^{(1)}(0)\end{aligned} \tag{33}$$

From Equation (33), we can see that each of two distribution functions is in general a linear combination of exponential functions of the form $e^{\lambda_j\tau}$, so the result here is an extension of that in [56–58]. Furthermore, the OFF and ON times can be computed by substituting $\tilde{f}_{off}(\tau)$, $\tilde{f}_{on}(\tau)$ into the general expression $\tilde{\tau}=\int_0^\infty \tau f(\tau)d\tau$, that is,

$$\begin{aligned}\tilde{\tau}_{off}&=\int_0^\infty \tau\mathbf{u}_L\mathbf{A}_{01}\exp(\mathbf{A}_{00}\tau)\mathbf{Q}^{(0)}(0)d\tau\\ &=\mathbf{u}_L\mathbf{A}_{01}(\mathbf{A}_{00})^{-2}\mathbf{Q}^{(0)}(0)\end{aligned}$$

$$\begin{aligned}\tilde{\tau}_{on}&=\int_0^\infty \tau\mathbf{u}_K\mathbf{A}_{10}\exp(\mathbf{A}_{11}\tau)\mathbf{Q}^{(1)}(0)d\tau\\ &=\mathbf{u}_K\mathbf{A}_{10}(\mathbf{A}_{11})^{-2}\mathbf{Q}^{(1)}(0)\end{aligned} \tag{34}$$

Note that Equation (34) 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 promoter structure, to obtain the total OFF and ON dwell times, we require 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 $Q_i^{(1j)}(0)=(\sum_{l=1}^{L}a_{il}^{(0\to1)}/\sum_{k=1}^{K}\sum_{l=1}^{L}a_{kl}^{(0\to1)})\delta_{ij}$ $(j, i=1, \cdots, K)$ as the initial conditions, where δ_{ij} is the Kronecker delta, and for clarity, we let $a_{ik}^{(0\to1)}$ represent the transition rate from the k^{th} OFF state to the i^{th} ON state (similarly, $a_{ik}^{(1\to0)}$, $a_{ik}^{(0\to0)}$ and $a_{ik}^{(1\to1)}$). The resulting distribution functions for the mean ON and OFF times are given by

$$\begin{aligned}f_{off}(\tau)&=\mathbf{u}_L\mathbf{A}_{01}\exp(\mathbf{A}_{00}\tau)\mathbf{A}_{10}\mathbf{u}_L^T\\ f_{on}(\tau)&=\mathbf{u}_K\mathbf{A}_{10}\exp(\mathbf{A}_{11}\tau)\mathbf{A}_{01}\mathbf{u}_K^T\end{aligned} \tag{35}$$

Correspondingly, the resulting mean dwell times at OFF and ON states are given by

$$\begin{aligned}\tau_{off}&=\frac{1}{\mathbf{u}_L\mathbf{A}_{10}\mathbf{u}_L^T}\mathbf{u}_L\mathbf{A}_{01}(\mathbf{A}_{00})^{-2}\mathbf{A}_{10}\mathbf{u}_L^T\\ \tau_{on}&=\frac{1}{\mathbf{u}_K\mathbf{A}_{01}\mathbf{u}_K^T}\mathbf{u}_K\mathbf{A}_{10}(\mathbf{A}_{11})^{-2}\mathbf{A}_{01}\mathbf{u}_K^T\end{aligned} \tag{36}$$

One can use the common ON-OFF model to verify the correctness of the above analytical expressions.

Decomposition and characteristics of the mRNA noise

First, note that the common noise (i.e., it is defined as the ratio of variance over the square of mean) in a reactive species of interest in any reaction network can be calculated using the first two binomial moments. That is, we have the following general formula

$$\eta=\frac{1}{b_1}+\frac{2b_2-b_1^2}{b_1^2} \tag{37}$$

Second, for the above gene model with general promoter structure, we find that the mRNA noise is given by

$$\eta_m = \frac{1}{\langle m\rangle} + \frac{\prod_{i=1}^{N-1}\frac{\beta_i^{(N)}}{\alpha_i}\left[\prod_{i=1}^{N-1}\frac{1+\beta_i^{(N)}}{1+\alpha_i} - \prod_{i=1}^{N-1}\frac{\beta_i^{(N)}}{\alpha_i}\right]}{\left(\prod_{i=1}^{N-1}\frac{\beta_i^{(N)}}{\alpha_i}\right)^2} \quad (38)$$

where $\langle m\rangle = (\tilde{\mu}\tau_{on})/(\tau_{off}+\tau_{on})$. Like the case of the common ON-OFF model, the first term on the right-hand side of Equation (38) represents the birth-death noise of mRNA whereas the second term represents the promoter noise. In other words, we have the following decomposition formula for the mRNA noise in any case

$$\eta_m = \xi_{birth-death} + \xi_{promoter} \quad (38)$$

\ Third, we give the decomposition of the mRNA noise using the attribute factor introduced above. It is easy to verify that the attribute factor can be expressed as

$$\xi_{attribute} = 1 + \frac{\prod_{i=1}^{N-1}\frac{\beta_i^{(N)}}{\alpha_i}\left[\prod_{i=1}^{N-1}\frac{1+\beta_i^{(N)}}{1+\alpha_i} - \prod_{i=1}^{N-1}\frac{\beta_i^{(N)}}{\alpha_i}\right]}{\left(\prod_{i=1}^{N-1}\frac{\beta_i^{(N)}}{\alpha_i}\right)^2} \quad (39)$$

which is independent of the transcription rate μ. Recall that $(2b_2)/b_1^2$ represents the burst size [59]. Thus, for the above general model of stochastic transcription, the attribute factor can describe not only the level of the expression noise but also the burst size.

In addition, if the gene promoter has one active state and L inactive states, which altogether form a loop with unidirectional transcription between every two neighboring states, then the mRNA noise intensity can be expressed as

$$\eta_m = \frac{1}{\langle m\rangle} + \frac{(\tau_{on}+\tau_{off})\prod_{k=1}^{L}(1+\tau_k)}{(1+\tau_{on})\prod_{k=1}^{L}(1+\tau_k)-1} - 1 \quad (40)$$

Finally, we show how the number of the inactive states impacts the noise intensity in the common sense. If the total OFF time is fixed, i.e., $\sum_{k=1}^{N-1}\tau_k = \mathrm{cosntant}$, then we have

$$\eta_m \geqslant \eta_{\min} = \frac{1}{\langle m\rangle} + \frac{(\tau_{on}+\tau_{off})(\tau_{off}/L+1)^L}{(1+\tau_{on})(\tau_{off}/L+1)^L-1} - 1 \quad (41)$$

Note that the function $f(L) = (\tau_{off}/L+1)^L/[(1+\tau_{on})(\tau_{off}/L+1)^L-1]$ is monotonically decreasing with the increase of L, so the noise intensity (η_m) achieves the maximum at $L=1$ that corresponds to the common two-state gene model (we denote by η_{on-off} the corresponding noise intensity). Therefore,

$$\eta_m < \eta_{on-off} \quad (42)$$

unless $L=1$. Since the actual OFF mechanism corresponds to $L>1$ [10], we obtain an important biological conclusion, i.e., the multi-OFF mechanism always reduces the noise in contrast to the common ON-OFF mechanism. This implies that the common ON-OFF model overestimates the noise in gene expression in the real case.

SUMMARY AND DISCUSSION

Gene expression is one of the important research contents of systems biology since it is the core of intracellular processes. While recent advances in experimental methods allow direct observations of real-time fluctuations in gene expression levels in individual live cells [1–5], there is considerable interest in theoretically understanding how different molecular mechanisms of gene expression impact variations in mRNA and protein levels across a population of cells. By analyzing two representative gene models (Figure 2) and a general gene model at the transcription level, we have shown that the molecule number of gene product in general follows a distribution expressed by a confluent hypergeometric function. We have also shown that in the absence of feedback, expression noise can be decomposed into the simple sum of the birth-death noise and transcription noise. In the presence of feedback, however, we have found that the feedback can induce additional contribution to expression noise. In particular, the multi-OFF mechanism always plays a role of reducing expression noise in contrast to the common ON-OFF mechanism. These results are independent of choice of system parameters and are therefore qualitative.

As is pointed out in the introduction, gene expression involves other biochemical processes such as alternative splicing [50,61] and RNA nuclear retention [62], apart from transcription, translation and feedback regulation. In fact, gene expression processes are becoming clearer due to the occurrence of new experimental technologies. One can imagine that these detailed processes would impact expression levels in their own ways. For quantitative analysis of this impact, one may take some methods and indices used in this paper, such as binomial moment method and attribute factor. In addition, we point out that this paper has focused on analysis of intrinsic noise, but extrinsic noise can also exist in gene expression systems. Analyzing contributions of extrinsic noise to cell-to-cell variability and dissecting decomposition principles of the total noise as done in this paper are challenging tasks since source of extrinsic noise may be complex.

In this paper, we have reviewed some progresses in the study of several gene expression systems, focusing on modeling and analysis as well as elucidation of the related

mechanisms. Even for these systems, however, there are some other questions that are also interesting but unsolved. Here, we list only partial and unsolved questions, and give frameworks for their quantitative analysis.

Differences between transient and stationary dynamics of gene expression

Many dynamical systems may exhibit very different steady-state and transient dynamics. In particular, there are big differences between stationary and transient behaviors of gene expression systems, mainly because of stochastic switching between promoter activity states. For example, for the common ON-OFF model of gene expression, the time-evolutional distribution may exhibit bimodalities of different modes although the corresponding steady-state distribution is unimodal [46]. For steady-state dynamics, this paper has given nice, analytical descriptions. For transient dynamics, however, it seems impossible to give analytical descriptions. In spite of this, numerical simulation based on the Gillespie stochastic algorithm [63] or on the binomial moment method mentioned in this paper may give quantitative descriptions for differences between two distinct behaviors.

Inferring promoter structure based on expression spectrums

While expression spectrums observed in experiments are comprehensive consequences of gene expression, the number of promoter activity states in eukaryotic organisms may be up to 128 [50]. A question naturally arises: how is promoter structure inferred from experimental data? This question is interesting but challenging. A possible way of solving the question is to analyze and compare all the possible modes of steady-state and transient distributions and find differences between them. For example, for the common ON-OFF model at the transcription level, all the possible modes of the mRNA distribution are only those: two modes of unimodality, where the peak is close to the origin and the peak is away from the origin; one mode of bimodality, where one peak is close to the origin whereas the other peak is away from the origin [46]. If the promoter has activity states of more than one, then the modes of the mRNA distribution may be complex but different from those in the case of only two activity states.

The mean-noise relationship

Mathematically, the mean-noise relationship is formulated as

$$\log(\sigma^2/\mu^2) = \beta \log\mu + \log\alpha \qquad (43)$$

where σ^2 and μ represent the variance and the mean of mRNA or protein, respectively. In Equation (43), both α and β are two constants depending on the parameters of a stochastic system of interest. For this formulation, the key is to determine the sign and size of β since β represents the slope of the line in the $(\log\mu, \log(\sigma^2/\mu^2))$ plane. For systems of gene expression, there are many works to study the relationship between mean and noise [64–66], some of which showed that β is negative [64] whereas others showed it may be positive or negative [65]. An unsolved question is what mechanisms govern the mean-noise relationship, in particular the sign and size of β. Owing to potential applications of this relationship in, e.g., disease systems [66], this question deserves study. Note that σ^2/μ^2 represents the noise intensity. Therefore, one may use the results given in this paper to analyze the mean-noise relationship in some cases.

ACKNOWLEDGEMENTS

This work was partially supported by Grant Nos. 91230204 (T. Z.), 91530320 (T. Z.), 2014CB964703 and 20120171110047 (T. Z.).

COMPLIANCE WITH ETHICS GUIDELINES

The authors Tianshou Zhou and Tuoqi Liu declare that they have no conflict of interests.

This article does not contain any studies with human or animal subjects performed by any of the authors.

REFERENCES

1. Golding, I., Paulsson, J., Zawilski, S. M. and Cox, E. C. (2005) Real-time kinetics of gene activity in individual bacteria. Cell, 123, 1025–1036
2. Raj, A., Peskin, C. S., Tranchina, D., Vargas, D. Y. and Tyagi, S. (2006) Stochastic mRNA synthesis in mammalian cells. PLoS Biol., 4, e309
3. Suter, D. M., Molina, N., Gatfield, D., Schneider, K., Schibler, U. and Naef, F. (2011) Mammalian genes are transcribed with widely different bursting kinetics. Science, 332, 472–474
4. Harper, C. V., Finkenstädt, B., Woodcock, D. J., Friedrichsen, S., Semprini, S., Ashall, L., Spiller, D. G., Mullins, J. J., Rand, D. A., Davis, J. R., et al. (2011) Dynamic analysis of stochastic transcription cycles. PLoS Biol., 9, e1000607
5. Spiller, D. G., Wood, C. D., Rand, D. A. and White, M. R. H. (2010) Measurement of single-cell dynamics. Nature, 465, 736–745
6. Raj, A. and van Oudenaarden, A. (2008) Nature, nurture, or chance: stochastic gene expression and its consequences. Cell, 135, 216–226
7. Blake, W. J., Balázsi, G., Kohanski, M. A., Isaacs, F. J., Murphy, K. F., Kuang, Y., Cantor, C. R., Walt, D. R. and Collins, J. J. (2006) Phenotypic consequences of promoter-mediated transcriptional noise. Mol. Cell, 24, 853–865
8. Ozbudak, E. M., Thattai, M., Kurtser, I., Grossman, A. D. and van Oudenaarden, A. (2002) Regulation of noise in the expression of a single gene. Nat. Genet., 31, 69–73
9. Elowitz, M. B., Levine, A. J., Siggia, E. D. and Swain, P. S. (2002)

Stochastic gene expression in a single cell. Science, 297, 1183–1186
10. Blake, W. J., KAErn, M., Cantor, C. R. and Collins, J. J. (2003) Noise in eukaryotic gene expression. Nature, 422, 633–637
11. Raser, J. M. and O'Shea, E. K. (2004) Control of stochasticity in eukaryotic gene expression. Science, 304, 1811–1814
12. McAdams, H. H. and Arkin, A. (1997) Stochastic mechanisms in gene expression. Proc. Natl. Acad. Sci. USA, 94, 814–819
13. Thattai, M. and van Oudenaarden, A. (2001) Intrinsic noise in gene regulatory networks. Proc. Natl. Acad. Sci. USA, 98, 8614–8619
14. Kepler, T. B. and Elston, T. C. (2001) Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations. Biophys. J., 81, 3116–3136
15. Chubb, J. R., Trcek, T., Shenoy, S. M. and Singer, R. H. (2006) Transcriptional pulsing of a developmental gene. Curr. Biol., 16, 1018–1025
16. Chubb, J. R. and Liverpool, T. B. (2010) Bursts and pulses: insights from single cell studies into transcriptional mechanisms. Curr. Opin. Genet. Dev., 20, 478–484
17. Boeger, H., Griesenbeck, J. and Kornberg, R. D. (2008) Nucleosome retention and the stochastic nature of promoter chromatin remodeling for transcription. Cell, 133, 716–726
18. Larson, D. R. (2011) What do expression dynamics tell us about the mechanism of transcription? Curr. Opin. Genet. Dev., 21, 591–599
19. Mao, C., Brown, C. R., Falkovskaia, E., Dong, S., Hrabeta-Robinson, E., Wenger, L. and Boeger, H. (2010) Quantitative analysis of the transcription control mechanism. Mol. Syst. Biol., 6, 431
20. Mariani, L., Schulz, E. G., Lexberg, M. H., Helmstetter, C., Radbruch, A., Löhning, M., Höfer, T. and Höfer, T. (2010) Short-term memory in gene induction reveals the regulatory principle behind stochastic IL-4 expression. Mol. Syst. Biol., 6, 359
21. Miller-Jensen, K., Dey, S. S., Schaffer, D. V. and Arkin, A. P. (2011) Varying virulence: epigenetic control of expression noise and disease processes. Trends Biotechnol., 29, 517–525
22. Sanchez, A. and Golding, I. (2013) Genetic determinants and cellular constraints in noisy gene expression. Science, 342, 1188–1193
23. Assaf, M., Roberts, E. and Luthey-Schulten, Z. (2011) Determining the stability of genetic switches: explicitly accounting for mRNA noise. Phys. Rev. Lett., 106, 248102
24. Peccoud, J. and Ycart, B. (1995) Markovian modelling of gene product synthesis. Theor. Popul. Biol., 48, 222–234.
25. Paulsson, J. (2005) Models of stochastic gene expression. Phys. Life Rev., 2, 157–175.
26. Shahrezaei, V. and Swain, P. S. (2008) Analytical distributions for stochastic gene expression. Proc. Natl. Acad. Sci. USA, 105, 17256–17261
27. Karmakar, R. and Bose, I. (2004) Graded and binary responses in stochastic gene expression. Phys. Biol., 1, 197–204
28. Iyer-Biswas, S., Hayot, F. and Jayaprakash, C. (2009) Stochasticity of gene products from transcriptional pulsing. Phys. Rev. E: Stat. Nonlin. Soft Matter Phys., 79, 031911
29. Mugler, A., Walczak, A. M. and Wiggins, C. H. (2009) Spectral solutions to stochastic models of gene expression with bursts and regulation. Phys. Rev. E: Stat. Nonlin. Soft Matter Phys., 80, 041921
30. Friedman, N., Cai, L. and Xie, X. S. (2006) Linking stochastic dynamics to population distribution: an analytical framework of gene expression. Phys. Rev. Lett., 97, 168302
31. Liu, P., Yuan, Z., Huang, L. and Zhou, T. (2015) Roles of factorial noise in inducing bimodal gene expression. Phys. Rev. E: Stat. Nonlin. Soft Matter Phys., 91, 062706
32. Huang, L., Yuan, Z., Liu, P. and Zhou, T. (2015) Effects of promoter leakage on dynamics of gene expression. BMC Syst. Biol., 9, 16
33. Huang, L., Yuan, Z., Liu, P. and Zhou, T. (2014) Feedback-induced counterintuitive correlations of gene expression noise with bursting kinetics. Phys. Rev. E: Stat. Nonlin. Soft Matter Phys., 90, 052702
34. Liu, P. J., Yuan, Z. J., Huang, L. F. and Zhou, T. S. (2015) Feedback-induced variations of distribution in a representative gene model. Int. J. Bifurcat. Chaos, 25, 1540008.
35. Iyer-Biswas, S. and Jayaprakash, C. (2014) Mixed Poisson distributions in exact solutions of stochastic autoregulation models. Phys. Rev. E: Stat. Nonlin. Soft Matter Phys., 90, 052712
36. Jia, T. and Kulkarni, R. V. (2011) Intrinsic noise in stochastic models of gene expression with molecular memory and bursting. Phys. Rev. Lett., 106, 058102
37. Liu, L., Kashyap, B. R. K. and Templeton, J. G. C. (1990) On the $GI^X/G/\infty$. J. Appl. Probab., 27, 671–683.
38. Schwabe, A., Rybakova, K. N. and Bruggeman, F. J. (2012) Transcription stochasticity of complex gene regulation models. Biophys. J., 103, 1152–1161
39. Zhang, J. J., Huang, L. F. and Zhou, T. S. (2014) Comment on 'Binomial moment equations for chemical reaction networks'. Phys. Rev. Lett., 112, 088901.
40. Barzel, B. and Biham, O. (2011) Binomial moment equations for stochastic reaction systems. Phys. Rev. Lett., 106, 150602
41. Barzel, B. and Biham, O. (2012) Stochastic analysis of complex reaction networks using binomial moment equations. Phys. Rev. E: Stat. Nonlin. Soft Matter Phys., 86, 031126
42. Singh, A., Razooky, B. S., Dar, R. D. and Weinberger, L. S. (2012) Dynamics of protein noise can distinguish between alternate sources of gene-expression variability. Mol. Syst. Biol., 8, 607
43. Rhee, A., Cheong, R. and Levchenko, A. (2014) Noise decomposition of intracellular biochemical signaling networks using nonequivalent reporters. Proc. Natl. Acad. Sci. USA, 111, 17330–17335
44. Voliotis, M., Perrett, R. M., McWilliams, C., McArdle, C. A. and Bowsher, C. G. (2014) Information transfer by leaky, heterogeneous, protein kinase signaling systems. Proc. Natl. Acad. Sci. USA, 111, E326–E333
45. Zhang, J., Chen, L. and Zhou, T. (2012) Analytical distribution and tunability of noise in a model of promoter progress. Biophys. J., 102, 1247–1257
46. Zhang, J. and Zhou, T. (2014) Promoter-mediated transcriptional dynamics. Biophys. J., 106, 479–488
47. He, Y. and Barkai, E. (2005) Super- and sub-Poissonian photon statistics for single molecule spectroscopy. J. Chem. Phys., 122, 184703
48. Bintu, L., Buchler, N. E., Garcia, H. G., Gerland, U., Hwa, T., Kondev, J., Phillips, R. and Phillips, R. (2005) Transcriptional regulation by the numbers: models. Curr. Opin. Genet. Dev., 15, 116–124
49. Huh, D. and Paulsson, J. (2011) Non-genetic heterogeneity from stochastic partitioning at cell division. Nat. Genet., 43, 95–100
50. Vilar, J. M. G. and Saiz, L. (2010) CplexA: a Mathematica package to study macromolecular-assembly control of gene expression. Bioinformatics, 26, 2060–2061
51. Hornung, G., Bar-Ziv, R., Rosin, D., Tokuriki, N., Tawfik, D. S., Oren, M. and Barkai, N. (2012) Noise-mean relationship in mutated promoters. Genome Res., 22, 2409–2417
52. Halme, A., Bumgarner, S., Styles, C. and Fink, G. R. (2004) Genetic and epigenetic regulation of the *FLO* gene family generates cell-surface

variation in yeast. Cell, 116, 405–415
53. Octavio, L. M., Gedeon, K. and Maheshri, N. (2009) Epigenetic and conventional regulation is distributed among activators of FLO11 allowing tuning of population-level heterogeneity in its expression. PLoS Genet., 5, e1000673
54. Weinberger, L., Voichek, Y., Tirosh, I., Hornung, G., Amit, I. and Barkai, N. (2012) Expression noise and acetylation profiles distinguish HDAC functions. Mol. Cell, 47, 193–202
55. Slater, L. J. (1960) Confluent Hypergeometric Functions. Cambridge: Cambridge University Press
56. 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
57. Li, G. and Qian, H. (2002) Kinetic timing: a novel mechanism that improves the accuracy of GTPase timers in endosome fusion and other biological processes. Traffic, 3, 249–255
58. Qian, H. (2007) Phosphorylation energy hypothesis: open chemical systems and their biological functions. Annu. Rev. Phys. Chem., 58, 113–142
59. Singh, A. and Hespanha, J. P. (2010) Stochastic hybrid systems for studying biochemical processes. Philos. Trans. A Math. Phys. Eng. Sci., 368, 4995–5011
60. Black, D. L. and Douglas, L. (2003) Mechanisms of alternative pre-messenger RNA splicing. Annu. Rev. Biochem., 72, 291–336
61. Wang, Q. and Zhou, T. (2014) Alternative-splicing-mediated gene expression. Phys. Rev. E: Stat. Nonlin. Soft Matter Phys., 89, 012713
62. Prasanth, K. V., Prasanth, S. G., Xuan, Z., Hearn, S., Freier, S. M., Bennett, C. F., Zhang, M. Q. and Spector, D. L. (2005) Regulating gene expression through RNA nuclear retention. Cell, 123, 249–263
63. Gillespie, D. (1977) Exact stochastic simulation of coupled chemical reactions. J. Phys. Chem., 81, 2340–2361.
64. Vallania, F. L. M., Sherman, M., Goodwin, Z., Mogno, I., Cohen, B. A. and Mitra, R. D. (2014) Origin and consequences of the relationship between protein mean and variance. PLoS One, 9, e102202
65. Carey, L. B., van Dijk, D., Sloot, P. M. A., Kaandorp, J. A. and Segal, E. (2013) Promoter sequence determines the relationship between expression level and noise. PLoS Biol., 11, e1001528
66. Dar, R.D., Hosmane, N.N., Arkin, M.R., Siliciano, R.F. and Weinberger, L.S. (2014) Screening for noise in gene expression identifies drug synergies. Science, 344, 1932–1936

