

Research Article

Mechanisms of How Random Input Controls Bursting Gene Expression

Sijia Xiao,¹ Yan Wang,² Zhigang Wang,^{1,3} and Haohua Wang ^{1,3,4}

¹School of Sciences, Hainan University, Haikou 570228, Hainan, China

²Department of Neurology, The First Affiliated Hospital, University of South China, Hengyang 421001, Hunan, China

³Hainan University, Key Laboratory of Engineering Modeling and Statistical Computation of Hainan Province, Haikou 570228, Hainan, China

⁴Hainan University, Coll Forestry, Key Laboratory of Genetics & Germplasm Innovation Tropical Special Forest, Ministry of Education, Haikou 570228, Hainan, China

Correspondence should be addressed to Haohua Wang; huazi8112@hainanu.edu.cn

Received 27 April 2022; Revised 8 August 2022; Accepted 8 September 2022; Published 22 September 2022

Academic Editor: Andrea Roli

Copyright © 2022 Sijia Xiao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The process of gene expression is affected by many extracellular stimulus signals, and the stochasticity of these signals reshapes gene expression. To adapt the fluctuation of the extracellular environment, genes have many strategies for augmenting their survival probability, frequency modulation, and amplitude modulation. However, it is unclear how genes utilize the stochasticity of signals to regulate gene expression and which strategy will be chosen to maximize cellular function. Here, we analyze a simple mechanistic model to clarify the effect of extracellular random input on gene expression and burst kinetics at different timescales. We can see that in different contexts, extracellular noise has different effects on downstream gene expression, effects which include the following: (1) extracellular noise will make the ON-OFF-state dwell time drift, which will influence the burst frequency and burst size of downstream gene expression under different modulation paradigms; (2) comparing the burst parameter or gene expression products under different modulation paradigms, we can see that the amplitude signal is more sensitive in the case of extracellular noise input, whereas the signal in noiseless conditions is more sensitive when the random input is a fast process, which indicates that the amplitude signal is a superior and common signal in gene expression; and (3) extracellular random input will change the bimodality for gene expression, but its influence is different for gene expression products under different modulation paradigms. These qualitative results reveal that extracellular random input can prompt the gene to achieve its function quickly under different modulation paradigms.

1. Introduction

Cells live in a very complex environment that sense and respond to extracellular signals to regulate target gene expression (Cai et al. [1]; Acar et al. [2]; Hao et al. [3]). The actual process can be decomposed into two processes: the signal encoding that senses the transcription factors in the concentration and localization, and the decoding process that activates or represses the downstream gene expression to achieve cellular responses (Cai et al. [1]; Acar et al. [2]; Hao et al. [3, 4]). Moreover, the signal encoding falls into two distinct categories: amplitude modulation (AM), in which signal receptors sense the ligand in concentration

continually, and frequency modulation (FM), which senses the frequency information of ligand peaks (Hao and O'Shea [4]; Micali et al. [5]). Essentially, processes occurring in cells sensing ligand signals are stochastic, which leads to stochastic fluctuations in expression product levels, which is also the main source of noise (Ban-Tabou de-Leon and Davidson [6]; Komili and Silver [7]). Tracing transient expression in individual cells with the application of fluorescence microscopy (Locke and Elowitz [8]; Muzzey and van Oudenaarden [9]) is a possible and important step to theoretically quantify the molecular mechanism of how stochastic input signals control intracellular expression and cell-to-cell variability even in genetically identical cell lines

(McAdams and Arkin [10]; Thattai and Van Oudenaarden [11]; Kepler and Elston [12]; Paulsson [13]; Raser and O’Shea [14]; Sanchez and Kondev [15]; Pedraza and Paulsson [16]).

Upstream transcription factors are signal transducers whose actual abundance fluctuates according to the extracellular environment and then affects promoter activity and downstream gene expression (Hao et al. [3]; Hao et al. [4]). Therefore, the statistics for environmental fluctuation and its timescale that the receptors detect precisely are an important basis for adjusting suitable strategies to respond to environmental inputs. To adapt to changes in the environment, stochastic switching is an effective survival strategy in fluctuating environments (Acar et al. [2]; Hung et al. [17]). Although the state-switching model classifies gene expression into models in which the burst size varies, the burst frequency varies, both vary, or the transcription rate varies, it is not a versatile law for gene expression (Larson [18]; Chong et al. [19]). For eukaryotic cells, many genes are expressed in episodic bursts that are interspersed with periods of quiescence (Raj et al. [20]; Munsky and Neuert [21]). Therefore, it is a hot issue to investigate the internal expression mechanism in individual cells induced by the transition between the “ON” and “OFF” states, regulating the burst size and burst frequency to increase or decrease gene expression upon simulations. In particular, Dar et al. demonstrated that transcriptional activators independently regulate the burst size and burst frequency depending on the specific chromosomal location, indicating that the transcriptional burst may be the basic expression pattern across the human genome (Dar et al. [22]). Furthermore, studying the chromatin environment contributed by different genomic locations and integrating the large-scale data to quantify mRNA or protein levels, Dey et al. verified that the mean level and noise of products are uncorrelated at distinct genomic locations, which can be achieved by orthogonal control; that is, the mean expression is controlled by burst size and noise is controlled by burst frequency (Dey et al. [23]; Larson [18]). Hence, it is an important issue to clarify the design principles of the expression network governing time evolution for better development and elucidate how the extracellular environment contributes to the regulation of downstream expression noise and its bursting kinetics, as well as phenotype diversity. Similar to microRNAs that are involved in maintaining expression robustness to gene mutation and environmental perturbations, external stimuli may perform similar functions (Dey et al. [23]; Schmiedel et al. [24]).

Gene expression is a multiscale process. Consequently, there are four timescales involved within even a simple gene network: (i) the degradation scale, (ii) the switching scale between different states, (iii) the transcription scale, and (iv) the dynamic process of extracellular signals (Thomas et al. [25]; Ge et al. [26]; Qian et al. [27], Raser and O’Shea [28]). In general, the abundance of gene products could always reach a certain level if the cell is in an active state, implying that the time scale (i) is usually much slower than (iii). The other scenarios are the focus of many researchers, and the results indicate that if time scale (iii) is much slower than (ii), it is called fast switching, which leads to the classical Poisson distribution by approximating the gene states in the quasi-

steady state (Wang et al. [29]; Zhu et al. [30]). While the time scale (ii) is much less than (i), called slow switching, the steady probability distribution is bimodal even without a positive feedback loop (Thomas et al. [25]; Ge et al. [26]; Ochab-Marcinek and Tabaka [31]). In fact, recent experiments on transcription and translation bursts are ubiquitous, ranging from prokaryotic cells to eukaryotic cells, especially for eukaryotic cells, and the slow switching rate dominates, indicating that slow switching may be a universal phenomenon in gene expression (Raser and O’Shea [28]; Wang et al. [32]). However, not enough attention has been given to the dynamic process of extracellular signals, in which the timescale of signals is measured by its autocorrelation time, also called the memory time, inducing the stochastic process into a non-Markovian process (Lu et al. [33]; Hu et al. [34, 35]). Experiments on the transcription factor Msn2 and its homologue Msn4 under oxidative stress or glucose limitation and the transcription factor Crz1 in response to a calcium stimulus are at least relevant to the underexplored scenario (Cai et al. [1]; Hao et al. [3, 4]; Micali et al. [5]), indicating that stochasticity plays a rather significant role in the kinetics of a single molecule.

Generally, we model gene expression into two common types: the conservative gene model in which the gene is in continuous expression and the so-called two-state gene models in which two adjacent expressions are separated by a long refractory period (Larson et al. [18, 36]; Raj et al. [20]). Here, we focus on the latter model in that it has general results that the larger the burst size is, the higher the expression noise is, whereas the larger the burst frequency is, the lower the expression noise is (Carey et al. [20]; Dar et al. [22]; Hansen and O’Shea [37]). However, extracellular stimuli have multiple roles in gene expression and have been identified as controlling signals in many expression modules, including prokaryotic or eukaryotic genes (Munsky et al. [38]; Munsky et al. [21]). Recently, experiments had illustrated that some transcription factor signals have dual roles in gene expression; that is, they are activators or repressors according to the extracellular environment. For example, the type-specific DNA protein AP-2 regulates the tumour marker gene in dual roles, that is, activator or suppressor, depending on the detected signal characteristics, including the stages of cancer progression or specific tissues (Allouche et al. [39]; Delacroix et al. [40]; Hilger-Eversheim et al. [41]; Matsumoto et al. [42]). Therefore, when the transcription factors fluctuate and the role changes due to extracellular environment input, the conventional two-state gene model must adjust the transition rate of gene switching, which will lead to a nonequilibrium state (Cai et al. [1]; Ge et al. [26]; Hu et al. [34, 35]). In this stable state, the burst kinetics of downstream gene expression are influenced by extracellular stimuli, which further impacts gene expression.

In summary, extracellular stimuli or random signal input is an inevitable event in gene expression that has been confirmed by large numbers of biological experiments; however, the mechanisms by which random input controls bursting are not clear. Here, we employ a mechanistic model to elucidate how the extracellular stimuli signal contributes to the burst kinetics of downstream gene expression even in

a population of genetically identical cells when the transcription factor is an activator and repressor on different timescales, that is, under two modulation cases: AM and FM, respectively. Without losing generality, our model is simple but still representative, including the process of random synthesis, random degradation, and stochastic switching between the so-called active state (“ON” state) and inactive state (“OFF” state), as well as extracellular stimuli and multiple timescales. By theoretical analysis, we focus mainly on the roles of random noise input in regulating burst kinetics, including the following: (1) the random noise input will induce the mean of ON-OFF dwell time focus, which will further affect the burst frequency and burst size of the downstream expression process; (2) if the modulation of the downstream gene expression is FM, the activators need the extracellular random input to be a fast process to maximize the cellular function, that is, maximal mean gene expression product. In contrast to activators, a repressed signal requires the extracellular stimulus to be an amplitude signal to minimize the downstream gene expression product. Meanwhile, if the modulation of the downstream gene expression is AM, the amplitude signal will activate and make the downstream gene expression optimum. However, the extracellular signal to repress downstream gene expression requires that the extracellular stimulus is a fast process, and its strength is greater than a certain threshold to accomplish cellular functions most quickly. The results reveal that the amplitude modulation signal is superior in the gene expression process. (3) The upstream signal will affect the distribution of downstream gene expression. It will induce the bimodality to be more obvious when the transcription factor is an activator, while it will suppress the bimodality if the transcription factor is a repressor. The sensitivity of the peak of high expression in the gene product distribution further reveals the relationship between the dynamic decoding process and gene function. These results are qualitative because they are independent of the choice of system parameters, revealing that the extracellular random input would control the bursting gene expression.

2. Gene Model and Analytical Distribution

2.1. Model Formula. In the following, we investigate the effect of extracellular stimuli on gene expression as well as burst kinetics. Without the loss of generality, we focus on the transcriptional level process in gene expression and ignore the other processes, for example, the translation process, to emphasize the regulation of the extracellular random input. Here, we also assume that there are two activity states in the gene promoter, called the active state (“ON”) and the inactive state (“OFF”), and the ON state is much more efficient than the OFF state (i.e., there is a small leakage in the OFF state) (Baler et al. [43]; Reed et al. [43]; Huang et al. [44]). Additionally, the gene switches between these two promoter states, leading to the burst transcription of mRNA (Pedraza and Paulsson [16]; Chong et al. [19]), which is also the main source of cell heterogeneity (Paulsson [13]; Raser and O’Shea [14]; Sanchez and Kondev [15]). Moreover, we also assume that the degradation process is linear; that is, it is a first-order

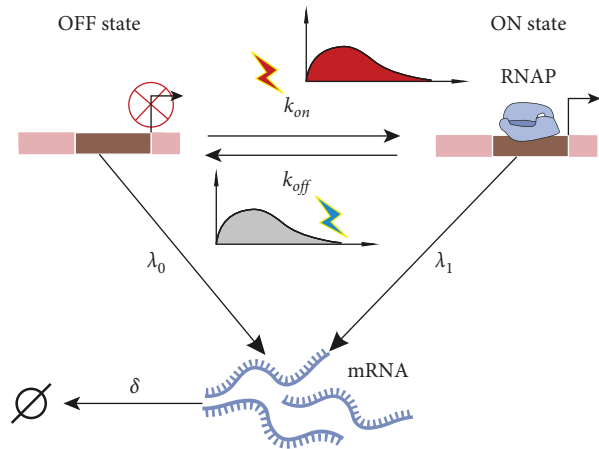


FIGURE 1: Schematic illustration of the two-state gene model stimulated by the extracellular signals (OU process whose stationary distribution obeys a gamma distribution, red or grey distribution) regulating the activity of the promoter through transfer function under different modulation paradigms, and the transcription rate is much larger than the leakage of transcription rate ($\lambda_1 \gg \lambda_0$).

process, although the results regarding multistep processes in gene expression experiments have been verified in many cases (Pedraza and Paulsson [16]; Chong et al. [19]).

We also consider the extracellular random stimulus in our model. Generally, cell sensing is achieved depending on whether the transmembrane receptors can bind and unbind the ligand signal molecules in a fluctuating environment. From the viewpoint of statistics, the signal dynamics mainly tend to infer the statistical characteristics of the information contents or its dynamical features, such as the amplitude or frequency of the signal, which will lead to different modulation paradigms on signals, that is, AM or FM (Behar and Hoffmann [45]; Aquino et al. [46]). Therefore, encoding the stimulus signals into the information that the cell can detect in different paradigms will form the so-called receptor-associated signalling network mechanisms, such as linear pathways and the Hill function. Here, we assume that the transfer function is the Hill function (Behar and Hoffmann [45]; Ochab-Marcinek and Tabaka [31]). Without the loss of generality, we denote the input signal as the Ornstein–Uhlenbeck (OU) process by $X(t)$, representing the concentration of ligand molecules at time, which can directly govern the switching rates between the ON and OFF states by the transfer method. In fact, the OU process is a representative stochastic process in gene expression due to its analytical simplicity and mean-reverting property (Hu et al. [35]; Gardiner [47]; Van Kampen [48]), as shown in Figure 1.

For both mathematical convenience and biophysical constraints, the OU process $X(t)$ encapsulates a square-root diffusion process for modeling chemical concentrations without negative values (Hu et al. [35]; Gardiner [47]; Van Kampen [48]):

$$dX(t) = \lambda[u - X(t)]dt + \sigma\sqrt{X(t)}dW_t, \quad (1)$$

TABLE 1: Transfer function of gene expression.

Transcription factors' role	Transfer function	Modulation paradigm
Activator	$TF_1(X) = X^n / (k_d^n + X^n)$	k_{on} modulation
	$TF_2(X) = k_d^n / (k_d^n + X^n)$	k_{off} modulation
Repressor	$TF_3(X) = X^n / (k_d^n + X^n)$	k_{off} modulation
	$TF_4(X) = k_d^n / (k_d^n + X^n)$	k_{on} modulation

where λ represents the rate of the process reverting to its mean value u , σ is the noise intensity, and W_t denotes the standard Brownian motion.

Transforming equation (1) into the equivalent Fokker-Planck equation for solving, we can obtain the stationary probability distribution of the input signal (for more details, see Appendix 1):

$$P_{ss}(X = x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{(\alpha-1)} e^{-x/\beta}, \quad (2)$$

which is a gamma distribution with parametric values $\alpha = 2\lambda u / \sigma^2$, $1/\beta = 2\lambda / \sigma^2$, with mean value $\langle x \rangle = u = \alpha\beta$ and variance $\sigma_x^2 = \alpha\beta^2$.

Of note, the molecular signal illustrated by the OU process can often be fitted by a gamma distribution in many biological experiments, which is another fascinating aspect of the OU process (Friedman et al. [49]), with the parametric value α representing the burst frequency (the mean number of bursts per cell cycle) and β denoting the burst size (the mean number of molecules produced per burst). Defining the noise strength as the ratio of the variance over the square of the mean, the noise strength of the input signal is equal to $1/\alpha$, indicating that the extracellular noise is only affected by the burst frequency but independent of other factors. Therefore, the OU process is also a gene product that can regulate downstream gene expression, as shown in Figure 1. When the concentration of the ligand molecule changes, the upstream gene product can be an activator or repressor, which will promote or suppress downstream gene expression. This phenomenon can be observed in AP-2 tumorigenesis modulation proteins (Allouche et al. [39]; Delacroix et al. [40]; Hilger-Eversheim et al. [41]; Matsumoto et al. [42]). It increases the transition rate k_{on} or decreases the transition rate k_{off} to promote downstream gene expression, while it decreases the transition rate k_{on} or increases the transition rate k_{off} to suppress downstream gene expression, as shown in Figure 1. Additionally, we define the

corresponding transfer function listed in Table 1, in which n is the Hill coefficient.

For $\alpha \geq 1$, it is guaranteed that the zero point is inaccessible for the OU process $X(t)$. Using Ito calculus (Gardiner [47]; Oksendal [50]) and letting the time tend to infinity, we have $\lim_{t \rightarrow \infty} \text{cov}[X(t), x(t+s)] = \sigma_x^2 e^{-\lambda|s|}$ for the steady-state covariance; that is, the OU process $X(t)$ is a stationary stochastic process with correlation time λ^{-1} .

2.2. Analysis Method and Some Analytic Distributions. Furthermore, we denote the two factorial probabilities $P_0(m; t)$ and $P_1(m; t)$, representing the copy number of gene products (mRNA) at time t , and the gene is in inactive and active states, respectively; that is, $P = P_0 + P_1$ denotes the total probability of gene products. The transcription rate is denoted as λ_1 and λ_0 ($\lambda_1 \gg \lambda_0$); that is, the gene has a leakage of transcription rate λ_0 . The degradation rate of products is denoted as δ . To obtain further insight into how random input signals regulate downstream gene expression, we focus on the two cases when the transcription factor is an activator under the k_{on} modulation paradigm, that is, two timescales. One case is called the ‘‘slow switch’’ limit, in which the timescale of the input signal is far less than that of the downstream expression system, that is, $\lambda^{-1} \ll T_x$. Here, $T_x = (k_{\text{on}} TF_1(u) + k_{\text{off}})^{-1}$ is also the correlation time for the noiseless input model (Lu et al. [33]; Hu et al. [34, 35]). When the upstream process fluctuates rapidly, the transcription factor will reach the quasi-steady state first compared with the downstream gene expression; that is, the transcription factor concentration is approximately a constant (Friedman [49]; Ochab-Marcinek [31]; Lu [33]), called a constant stimulus or noiseless input, whose value is equal to the mean value $\langle x \rangle = u = \alpha\beta$. Therefore, the chemical master equation (CME) of the reaction network illustrated in Figure 1 can be rewritten as follows:

$$\begin{aligned} \frac{\partial P_0(m; t)}{\partial t} &= -k_{\text{on}} TF_1(u) P_0(m; t) + k_{\text{off}} P_1(m; t) + \lambda_0 (E^{-1} - 1) [P_0(m; t)] + \delta (E - I) [m P_0(m; t)], \\ \frac{\partial P_1(m; t)}{\partial t} &= k_{\text{on}} TF_1(u) P_0(m; t) - k_{\text{off}} P_1(m; t) + \lambda_1 (E^{-1} - 1) [P_1(m; t)] + \delta (E - I) [m P_1(m; t)], \end{aligned} \quad (3)$$

where I represents the unity operator and E with the inverse E^{-1} is the common step operator, defined as $Ef(m) = f(m+1)$ for any function f . Setting $k_{\text{on}}^i = k_{\text{on}}TF_i(u)$ and $k_{\text{off}}^i = k_{\text{off}}TF_i(u)$, ($i = 1, 2, 3, 4$), we also obtain the CME under other modulation paradigms by revising the transition rate in (3); for more details, see Appendix 2.

Another case is called the “fast switch” limit, where extracellular stimulus signal $X(t)$ fluctuates so slowly that $\lambda^{-1} \gg T_x$, where $T_x = (k_{\text{on}}TF_1(u) + k_{\text{off}})^{-1}$; that is, the input signals will affect the downstream switching rate by a stochastic process (Friedman [49]; Ochab-Marcinek [31]; Lu [33]). In this case, the transcription factor concentration is essentially a random variable or random input, with the noise strength also equal to $1/\alpha$ and the steady-state distribution being a gamma distribution with mean value $\langle x \rangle = u = \alpha\beta$ and variance $\sigma_x^2 = \alpha\beta^2$ (reference to Figure 1, the red or grey signal distribution).

Then, we can also establish our mathematical model in the form of discrete CME based on conditional probability:

$$\begin{aligned} \frac{\partial P_0(m; t|X)}{\partial t} &= -k_{\text{on}}TF_1(X)P_0(m; t|X) + k_{\text{off}}P_1(m; t|X) \\ &\quad + \lambda_0(E^{-1} - 1)[P_0(m; t|X)] \\ &\quad + \delta(E - I)[mP_0(m; t|X)], \\ \frac{\partial P_1(m; t|X)}{\partial t} &= k_{\text{on}}TF_1(X)P_0(m; t|X) - k_{\text{off}}P_1(m; t|X) \\ &\quad + \lambda_1(E^{-1} - 1)[P_1(m; t|X)] \\ &\quad + \delta(E - I)[mP_1(m; t|X)], \end{aligned} \quad (4)$$

where E represents the step operator, I represents the identity operator, and $P_i(m; t|X)$, ($i = 0, 1$) is the conditional probability. That is, $P_i(m; t) = \int_X P_i(m; t|X)P_{\text{ss}}(X)dX$, ($i = 0, 1$) is also factorial probability, representing the probability of the expression system being in active and inactive states and having m molecules at time t . Obviously, when the transcription factor role and modulation paradigm change, the CME for downstream gene expression can also be established according to the two timescales. Setting $k_{\text{on}}^i = k_{\text{on}}TF_i(X)$, $k_{\text{off}}^i = k_{\text{off}}TF_i(X)$, ($i = 1, 2, 3, 4$), we also obtain the CME under other modulation paradigms by revising the transition rate in (4); for more details, see Appendix 3.

Next, we focus on how to solve the CME for its steady-state solution under the above two cases when the transcription factor is an activator. In the case of the “slow switch” limit, we first introduce the so-called probability-generating functions (Zhang and Zhou [51]; Shahrezaei and Swain [52]) to transfer the CME into a coupled set of ordinary differential equations with respect to different transfer functions. If we set all the parameters to be normalized by the degradation rate δ , that is, $k_{\text{on}}/\delta \rightarrow k_{\text{on}}$, $k_{\text{off}}/\delta \rightarrow k_{\text{off}}$, $\lambda_1/\delta \rightarrow \lambda_1$, $\lambda_0/\delta \rightarrow \lambda_0$, and

$\gamma_1 = k_{\text{on}}TF_1(u)$, $\gamma_0 = k_{\text{off}}$, the analytic solution for steady-state probability distribution can be expressed as (Appendix 2)

$$P(m) = \frac{gA}{m!} \sum_{n=0}^m \binom{m}{n} \lambda_0^{m-n} [Q]^n \frac{(a-1)_n}{(b-1)_n} {}_1F_1 \cdot (a+n-1, b+n-1; -Q), \quad (5)$$

where ${}_1F_1$ is the confluent hypergeometric function, with $\lambda = \lambda_1 - \lambda_0$, $Q = \lambda$, $g = \lambda + \gamma_1 + \gamma_0/\gamma_1 - \lambda/\gamma_1$, $a = 1 + \gamma_1$, $b = 1 + (\gamma_0 + \gamma_1)$, and $A = e^{-\lambda_0} [g {}_1F_1(a-1, b-1; 0)]^{-1}$; $(c)_n$ is the Pochhammer symbol defined by $(c)_n = \Gamma(c+n)/\Gamma(c)$; and $\binom{m}{n}$ is a combination number of choosing n molecules from m molecules.

The above analytical distribution contains nearly all information of the downstream expression system, including the mean and variance, as well as the noise strength defined the same as before. The mean and the variance under the case of noiseless input are calculated according to

$$\langle m \rangle_c = G'(1), \sigma_{mc}^2 = G''(1) + G'(1) - [G'(1)]^2, \quad (6)$$

where $G(z) = G_0(z) + G_1(z)$ represents the total generating function (its analytical formulas and two factorial functions $G_0(z)$ and $G_1(z)$ are given in Appendix 2). Hence, we have

$$\begin{aligned} G'(1) &= Ae^{\lambda_0} [g\lambda_0 {}_1F_1(a-1, b-1, 0) + \lambda_1 {}_1F_1(a, b, 0)], \\ G''(1) &= Ae^{\lambda_0} [g\lambda_0^2 {}_1F_1(a-1, b-1, 0) + 2\lambda\lambda_0 {}_1F_1(a, b, 0) \\ &\quad + \frac{\lambda^2(\gamma_1+1)}{\gamma_1+\gamma_0+1} {}_1F_1(a+1, b+1, 0)], \end{aligned} \quad (7)$$

where A , g , Q , a , and b are all compounded parameters given as above. Thus, the square of the noise intensity under the case of noiseless input is calculated according to the following formula:

$$\eta_{mc}^2 = \frac{\sigma_{mc}^2}{\langle m \rangle_c^2} = \frac{G''(1) + G'(1) - [G'(1)]^2}{[G'(1)]^2}. \quad (8)$$

In contrast to the “slow switching” limit, we consider the “fast switching” limit based on the conditional probability (Friedman et al. [49]; Thomas et al. [25]; Ochab-Marcinek and Tabaka [31]). For eukaryotic cells, the promoter switches more slowly than the cell cycle (defined by degradation rate scale) (Raj et al. [20]; Munsky [21]; Raser [28]), and then, the stationary solution of the discrete CME can be decomposed into two decoupled equations for the condition probability of two active states at the steady state, which are identical modulo, a factor of λ_1 and λ_0 :

$$\begin{aligned} \lambda_0(E^{-1} - I)P_0(m, t) + \delta(E - I)[mP_0(m, t)] &= 0, \\ \lambda_1(E^{-1} - I)P_1(m, t) + \delta(E - I)[mP_1(m, t)] &= 0. \end{aligned} \quad (9)$$

The above equations have been solved exactly in Appendix (A24), which means that the distribution of mRNA is Poissonian with the parameters λ_0/δ and λ_1/δ . Note that the

ON-OFF transition of the promoter is independent of mRNA levels, even if the transition rate is controlled by the extracellular random input. Therefore, we must deduce the time evolution of the two promoter states, that is, the ON-OFF dwell time distribution.

Generally, we consider the distribution of dwell time of the OFF state when the transcription factor is an activator under the k_{on} modulation paradigm. Considering the first passage time problem in the rigour of mathematics (Ge et al. [26]; Hu et al. [35]; Van Kampen [48]), we suppose the initial system is in the OFF state with input signal $X(0) = x$. Denoting τ as the first escape time from the OFF state (Figure 1), the survival probability $Q^{(0)}(\tau, x)$ that the system never switches to the ON state up to time t is given by

$$\begin{aligned} Q^{(0)}(\tau, x) &= P(\tau > t | X(0) = x) = \int_0^\infty Q^{(0)}(\tau | x) p_{ss}(x) dx \\ &= \int_0^\infty \exp(-k'_{\text{on}} \tau) P_{ss}(x) dx \\ &= E^x [\exp(-k'_{\text{on}} \tau)] \geq \exp(E^x(-k'_{\text{on}} \tau)), \end{aligned} \quad (10)$$

where $k'_{\text{on}} = k_{\text{on}} TF_1(x) = k_{\text{on}} x^n / k'_d + x^n$, the last inequality is obtained by Jensen's inequality, and the above inequality tells us that the extracellular input signal will induce the dwell time in the OFF state to be a nonexponential distribution (which is different from the classical Markov assumption), independent of other factors, including noise strength or the correlation time. For our model, if $\lambda_1 \gg \lambda_0$, then the downstream gene expression has the following approximation:

$$\text{burst frequency} = \frac{1}{\tau_{\text{off}}} \text{ and burst size} = \lambda_1 \tau_{\text{on}}, \quad (11)$$

where τ_{on} and τ_{off} are the dwell times of the ON-OFF state (Larson [18]; Dar et al. [22]; Huang et al. [32]). These results illustrate that the downstream gene will only adjust the OFF-state time or burst frequency to adapt to the upstream stimulus, which is the frequency modulation (FM) strategy (Hao and O'Shea [4]; Micali et al. [5]), meaning that the k_{on} modulation paradigm is essentially the frequency modulation (FM) paradigm. Similarly, we can say that the k_{off} modulation paradigm is essentially the amplitude modulation (AM) paradigm because the extracellular signal will only change the ON-state dwell time and the burst size of downstream gene expression.

We will elaborate on the regulation effect of random input in an analytical way. In particular, if the switching rate of the promoter is slower than the cell cycle, the gene expression is equal to the combination of two on-state gene expression models; that is, the distribution of mRNA is the combination of two Poisson distributions weighted by the mean dwell time of the ON-OFF state (Hao [4]; Thomas et al. [25]; Qian et al. [27]):

$$\begin{aligned} \Pi(M = m) &= \frac{\langle \tau_{\text{on}} \rangle}{\langle \tau_{\text{on}} \rangle + \langle \tau_{\text{off}} \rangle} P\left(\frac{\lambda_1}{\delta}\right) + \frac{\langle \tau_{\text{off}} \rangle}{\langle \tau_{\text{on}} \rangle + \langle \tau_{\text{off}} \rangle} P\left(\frac{\lambda_0}{\delta}\right), \\ &= \frac{\langle \tau_{\text{on}} \rangle}{\langle \tau_{\text{on}} \rangle + \langle \tau_{\text{off}} \rangle} \frac{\lambda_1^m}{m! \delta^m} e^{-\lambda_1/\delta} + \frac{\langle \tau_{\text{off}} \rangle}{\langle \tau_{\text{on}} \rangle + \langle \tau_{\text{off}} \rangle} \frac{\lambda_0^m}{m! \delta^m} e^{-\lambda_0/\delta}. \end{aligned} \quad (12)$$

Here, $P(\bullet)$ is a Poisson distribution, and $\langle \tau_{\text{on}} \rangle$ and $\langle \tau_{\text{off}} \rangle$ represent the mean dwell time of the ON-OFF states; for more details, see Appendix 3.

Thereby, we can also calculate the mean and variance under the case of random input according to equation (12):

$$\langle m \rangle_s = \frac{\langle \tau_{\text{on}} \rangle}{\langle \tau_{\text{on}} \rangle + \langle \tau_{\text{off}} \rangle} \frac{\lambda_1}{\delta} + \frac{\langle \tau_{\text{off}} \rangle}{\langle \tau_{\text{on}} \rangle + \langle \tau_{\text{off}} \rangle} \frac{\lambda_0}{\delta}. \quad (13)$$

And,

$$\begin{aligned} \sigma_{ms}^2 &= \int_m m^2 \Pi(m) dm - \langle m \rangle_s^2 = (c - c^2) \lambda_1^2 + (d - d^2) \lambda_0^2 \\ &\quad + (c \lambda_1 + d \lambda_0 - 2cd \lambda_0 \lambda_1), \end{aligned} \quad (14)$$

where $c = \langle \tau_{\text{on}} \rangle / (\langle \tau_{\text{on}} \rangle + \langle \tau_{\text{off}} \rangle)$, $d = \langle \tau_{\text{off}} \rangle / (\langle \tau_{\text{on}} \rangle + \langle \tau_{\text{off}} \rangle)$ and $\lambda_1 = \lambda_1 / \delta$, $\lambda_0 = \lambda_0 / \delta$.

The noise strength is calculated as follows:

$$\eta_{ms}^2 = \frac{\sigma_{ms}^2}{\langle m \rangle_s^2}. \quad (15)$$

Comparing the noiseless with the random input, we can see the difference in the downstream expression noise or variance:

$$\Delta \sigma^2 = \sigma_{ms}^2 - \sigma_{mc}^2 = (c - c^2) \lambda_1^2 + (d - d^2) \lambda_0^2 + (c \lambda_1 + d \lambda_0 - 2cd \lambda_0 \lambda_1),$$

$$- \frac{\lambda_1^2 \gamma_0 \gamma_1 + \lambda_1 \gamma_1 (\gamma_0 + \gamma_1) (1 + \gamma_0 + \gamma_1)}{(\gamma_0 + \gamma_1)^2 (1 + \gamma_0 + \gamma_1)}. \quad (17)$$

This is a function of the burst frequency (α) and burst size (β) of the upstream input signal, as well as the relevant parameters of downstream gene expression.

The above results and explicit formulas make it possible to systemically analyze the regulatory mechanism of how the interaction of extracellular random input controls and regulates both gene expression and bursting kinetics.

3. Results

Stochastic switching between the different promoter states, as an effective survival strategy, often leads to the burst expression of gene products, which is characterized by two kinetic indices, burst size and burst frequency, which are also two different strategies of the gene to cope with extracellular stimuli, that is, AM and FM. We will demonstrate how extracellular random input controls the correlation between gene expression and bursting kinetics.

3.1. The Extracellular Fluctuations Prompt the Dwell Time Drift. When extracellular random input affects gene expression by the transfer function under different modulation paradigms (see Table 1), it will affect the ON-OFF dwell time. Essentially, it only changes the OFF-state dwell time under the k_{on} modulation paradigm; that is, the k_{on}

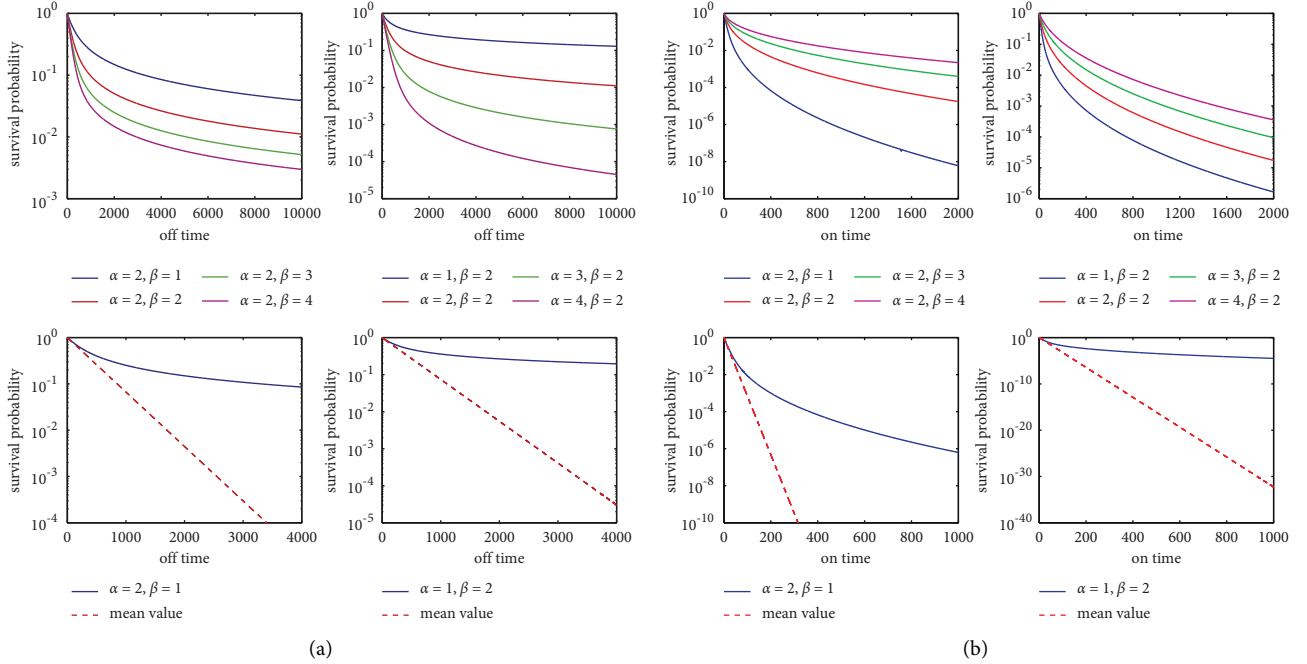


FIGURE 2: The input signal induces dwell time drift with an activator. The first row represents the upstream input as noise, whereas the second row represents the difference in the dwell time between the noise and noiseless signals. (a) The OFF-state dwell time under the FM paradigm with the extracellular stimulus increasing its burst frequency or burst size; (b) the ON-state dwell time under the AM paradigm when the extracellular stimulus strength increases, and the means of the coloured lines are similar to (a). Other parameters are set as $k_{\text{on}} = 0.01$, $k_{\text{off}} = 0.1$, $k_d = 10$, $n = 2$, $\lambda_0 = 1$, $\lambda_1 = 50$, $\alpha = 2$, and $\beta = 2$ (as the standard of comparison).

modulation paradigm is a FM paradigm in terms of burst kinetics (Hao and O'Shea [4]; Micali et al. [5]; Dar et al. [22]). Similarly, we say that the k_{off} modulation paradigm is essentially the AM paradigm. In this case, equation (10) tells us that the extracellular random input causes the survival probability or dwell time distribution not to satisfy the assumption of exponential distribution, that is, the ON-OFF-state dwell time has some memory (Pedraza and Paulsson [16]; Lu et al. [33]). Indeed, memory is a strategy that cells have evolved and is also beneficial for the gene to cope with extracellular environment fluctuations (Perkins and Swain [53]). Specifically, when the transcription factor is an activator, it will reduce the OFF-state dwell time or amplify the ON-state dwell time to activate the downstream expression, and the influence of extracellular random input under different modulation paradigms can be found in Figure 2.

Of note, the extracellular stimulus strength is always the same in the corresponding level, but to regulate the upstream signal burst size (left of the first row, burst frequency $\alpha = 2$, the burst size varies) or burst frequency (right of the first row, burst size $\beta = 2$, the burst frequency varies) to achieve the effective control for downstream expression (the coloured lines from blue to purple in Figure 2(a)), and the situations later are the exact same as Figure 2(a)). Exactly, Figure 2(a) illustrates the OFF-state survival probability or dwell time distribution changes under the FM paradigm with increasing extracellular stimulus strength (from the blue to purple lines). The OFF-state dwell time will extracellular stimulus increasing, with

the burst size (left) or burst frequency (right) of the extracellular stimulus signal varying. However, the reduced amplitude has a significant difference. Comparing the first and second columns of the first row and considering the different characters of the upstream noise, we can see that the downstream gene expression is more sensitive to the upstream frequency signal than the amplitude signal; that is, it will make the OFF-state dwell time decrease faster when the burst frequency varies. Second, comparing the blue line (noise input) with red dotted line (noiseless input) in the second line also shows that the OFF-state dwell time distribution does not obey an exponential distribution when the extracellular input is noise, revealing that the OFF-state dwell time distribution has some memory induced by random input and extracellular noise will make the OFF-state dwell time drift positively or focus.

Similarly, Figure 2(b) represents the ON-state dwell time distribution changes under the AM paradigm. The obvious difference between (A) and (B) is that the ON-state dwell time will increase when the extracellular stimulus strength keeps the same value and increases from 2 to 8 by increasing its burst size (left) or burst frequency (right) (see the first row of (B), from blue line to purple line). It indicates that the downstream gene expression is more sensitive to the burst size of the upstream signal than its burst frequency. In other words, the extracellular fluctuations will prompt the OFF-state dwell time drift when comparing the noiseless input, and there will be an obvious difference between the AM and FM paradigms, which also demonstrates that gene expression is indeed affected by extracellular stimulus signals.

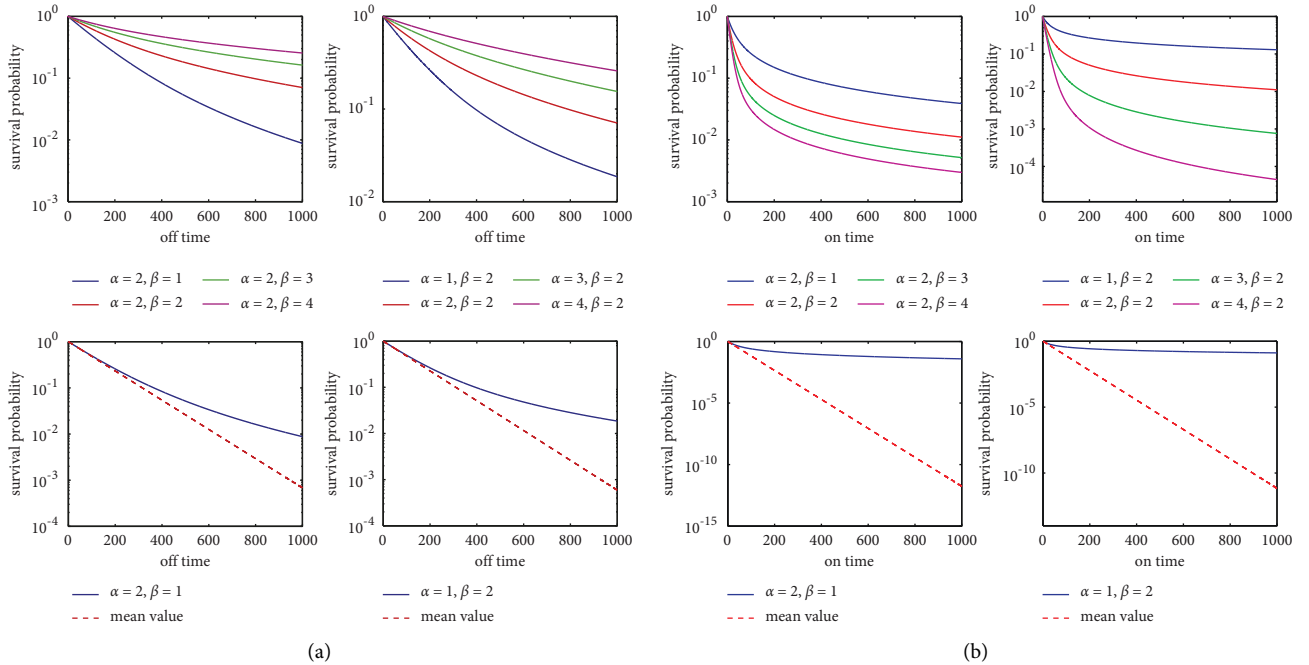


FIGURE 3: The extracellular fluctuations make the dwell time drift with a repressor, and the first row represents the upstream input as noise, whereas the second row represents the difference in the dwell time between the noise and noiseless signal. (a) The OFF-state dwell time under the FM paradigm with the extracellular stimulus increasing its burst frequency or burst size; (b) the ON-state dwell time under the AM paradigm when the extracellular stimulus strength increases. The means of the coloured lines are similar to (a). The other parameters are the same as in Figure 2.

Figure 3 illustrates the difference between random input and noiseless input under different modulation paradigms when the transcription factor is a repressor.

In contrast to the activator, the repressor always increases the OFF-state dwell time or decrease the ON-state dwell time to achieve its biology function under the different decoding paradigms. Figure 3(a) demonstrates that the OFF-state dwell time increases with extracellular stimulus strengthening under the FM paradigm in the first row. However, the characteristics of the upstream signal change the downstream gene OFF-state dwell time, as shown in the first row of Figure 3(a). Specifically, the left of the first row shows the OFF-state dwell time will have the typical focusing effect stimulated by the random input with the upstream signal burst size varying (left) or burst frequency varying (right). Moreover, the second row shows us the difference between the noise input (blue line) and noiseless input (red dotted line), illustrating that the extracellular noise makes the OFF-state dwell time drift positively compared with the noiseless input. However, the amplitude of drift has an obvious difference even if the stimulus strength remains the same. In contrast to (A), Figure 3(b) illustrates that the ON-state dwell time decreases with extracellular stimulus strengthening.

Comparing Figure 2 with Figure 3, we can see that extracellular noise will induce the ON- or OFF-state dwell time to be nonexponentially distributed, indicating that the ON or OFF state will have some memory and that the gene state involves multiple steps instead of one-step biochemical reactions (Pedraza and Paulsson [16]; Lu et al. [33]). This is a

common phenomenon in the process of biochemical reactions (Perkins and Swain [54]). Although the modulation paradigm and extracellular stimulus strength are the same, the different characters of the upstream signal will lead the ON or OFF dwell times to be significantly different (compared with Figure 2(a) with Figure 3(a), and Figure 2(b) with Figure 3(b)). This difference tells us that gene expression not only relies on the modulation paradigm but is also dependent on specific signalling pathways.

3.2. Extracellular Fluctuations Affect Burst Kinetics. To illustrate the characteristics of the downstream gene burst kinetics, we investigate two indices, burst frequency and burst size, under two different modulation paradigms, that is, the FM and AM paradigms. In general, the burst frequency is defined by the OFF-state dwell time, while the burst size is defined by the ON-state time (Larson [18]; Dar et al. [22]). Additionally, equation (12) indicates the relationship between burst kinetics and dwell time.

We first investigate the burst kinetics with extracellular random input under the FM paradigm. Obviously, when the transcription factors are activators, the extracellular stimulus will reduce the OFF-state dwell time, as shown in Figure 2(a), resulting in the increase in the burst frequency of downstream gene expression as the stimulus strength increases (see Figure 4).

Figure 4(a) shows that the burst frequency of downstream gene expression monotonically increases with an

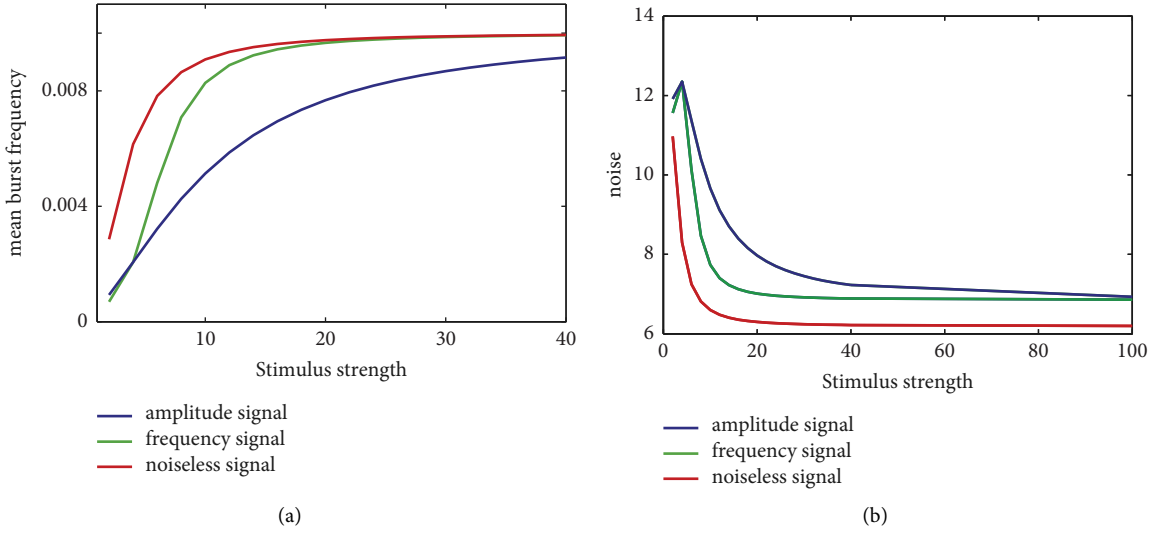


FIGURE 4: The mean burst frequency (a) and the noise (b) of downstream gene expression are under the FM paradigm, and the mean burst size of downstream gene expression is constant ($\langle \text{burst size} \rangle = 500$) when the strengthening extracellular stimulus increases and activates downstream gene expression. The red line represents the upstream stimulus signal as noiseless, and the green line represents the upstream stimulus signal as random with the burst frequency (α) changing and the burst size (β) being equal to 2, while the blue line represents the upstream stimulus signal as also random with the burst frequency (α) being equal to 2 and the burst size (β) changing. Other parameters are set as $k_{\text{on}} = 0.01$, $k_{\text{off}} = 0.1$, $k_d = 10$, $n = 2$, $\lambda_0 = 1$, and $\lambda_1 = 50$.

increase in extracellular strength and then reaches the stable limit value. However, there is a significant difference between the different stimulus signals. Considering that the upstream signal has different characteristics, that is, noiseless signal (red line), frequency signal (green line), and amplitude signal (blue line), we can evaluate their advantage on regulating burst kinetics by the paired comparison method.

First, comparing the green with blue lines, we can see that the burst frequency of downstream gene expression will increase faster when the upstream signal is a frequency signal and the burst size is equal to 2; that is, regulating upstream frequency signal is the most effective way to modify downstream gene expression. Second, when the upstream stimulus is random input, the burst frequency of downstream gene expression increases more slowly than the case that the upstream signal is noiseless (comparing the red, green, and blue lines), indicating that the lower the extracellular noise strength is, the greater the mean burst frequency is. This means that the extracellular noise will make the burst frequency of downstream gene expression drift in the negative direction. In fact, the burst frequency and burst size can control the downstream expression product, that is, the mean of gene products is equal to the mean burst frequency multiplied by the mean burst size, that is, $\langle M \rangle \approx \langle BF \rangle \langle BS \rangle$ (Dar et al. [22]; Huang et al. [53]). Therefore, if the mean burst size of downstream gene expression is a constant, the downstream gene product numbers will increase with the same trend; that is, the mean mRNA will also increase monotonically (refer to Figure S1 in Appendix 4). Figure S1 tells us that the noiseless input will make the gene expression process more efficient than the random input, and the frequency signal stimulus is also

more efficient than the amplitude signal stimulus for downstream gene expression.

Third, similar to the mean burst frequency, the downstream gene expression noise also depends heavily on upstream stimulus, as shown in Figure 4(b). Comparing the red line with the blue and green lines, we can see that the expression noise in the noiseless input paradigm is less than that in the random input paradigm, meaning that the extracellular noise will increase the expression noise. In particular, amplitude signal stimulus will make the downstream expression noise maximal among the three signal stimulus paradigms when the stimulus strength is the same, as represented by the blue line in Figure 4(b). Fourth, the extracellular random input prompts the expression noise to have a weak peak value, indicating that there may exist a critical value for upstream noise strength to regulate bidirectionally the expression stability of downstream gene. From Figures 4(a) and 4(b), we can observe that an increasing upstream random signal will reduce the mean burst frequency or the mean number of mRNAs but amplify the expression noise. This shows that the randomness of upstream active signal is destructive for downstream gene expression under the FM paradigm.

Similarly, when the transcription factors are repressors under the FM paradigm, the burst frequency of downstream gene expression has some differences, as shown in Figure 5.

Of note, the stimulus strength is the same for the three different signals. We can see that the downstream burst frequency will monotonically decrease as the upstream stimulus increases, as shown in Figure 5(a). Specifically, the extracellular noise will induce the downstream burst frequency to be less than that in the noiseless input (comparing the red, blue with green lines); that is, the noise input will

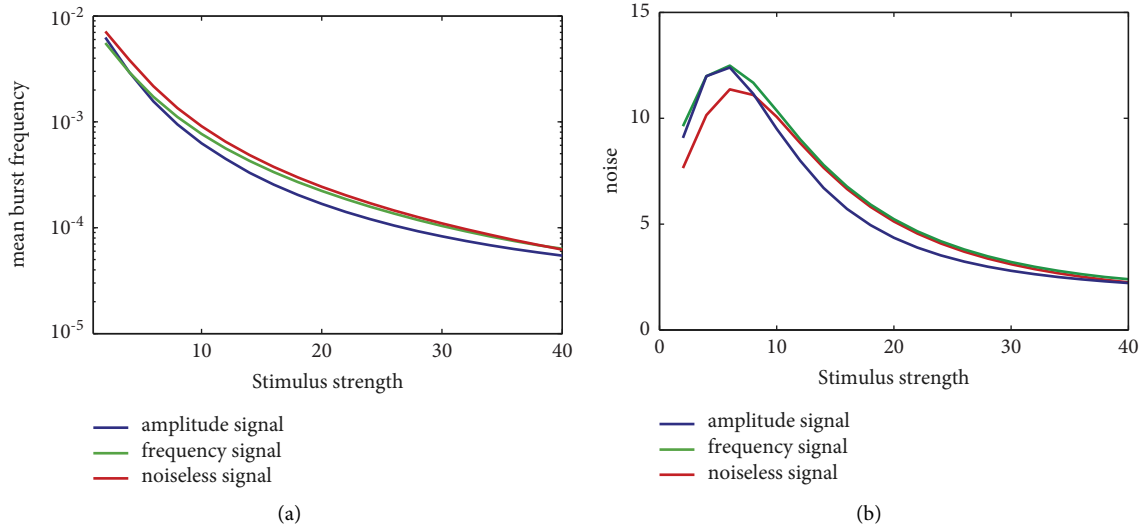


FIGURE 5: The mean burst frequency (a) and the noise (b) of downstream gene expression are under the FM modulation paradigm, and the mean burst size of downstream gene expression is constant ($\langle \text{burst size} \rangle = 500$) when the extracellular stimulus strengthens and represses downstream gene expression. The red line represents the upstream stimulus signal as noiseless, and the green line represents the upstream stimulus signal as random with the burst frequency (α) changing and the burst size (β) being equal to 2, while the blue line represents the upstream stimulus signal as also random with the burst frequency (α) being equal to 2 and the burst size (β) changing. The other parameters are the same as in Figure 4.

promote the functions of the gene components to be maximal, for example, the repressors will repress the downstream gene expression as much as possible. This indicates that extracellular noise is beneficial for gene expression. Meanwhile, the difference between the blue and green lines tells us that downstream expression is more sensitive to amplitude signals than to frequency signals. Figure 5(b) illustrates that the downstream gene expression noise first increases and then monotonically decreases, and there is a threshold value of stimulus strength for expression stability. Obviously, the amplitude signal minimizes the gene expression noise when the extracellular signal strength is greater than a fixed value (see the blue line in Figure 5(b)). From Figures 5(a) and 5(b), we can determine that the mean burst frequency or mRNA number (Figure S2 in Appendix 4) will be the lowest and the expression noise will also be the lowest when the upstream signal is the amplitude signal and represses the downstream expression under the FM paradigm (refer to the blue line in Figures 5(a) and 5(b)). This is counterintuitive because the traditional theory indicates that a smaller burst frequency will result in higher expression noise or that a smaller mean mRNA number will lead to larger expression noise (Paulsson [13]; Raser and O’Shea [14]); that is, the extracellular stimulus will reshape the gene expression mode because the gene expression system will form a coherent resonance under the interference of an extracellular stimulus (Qian et al. [27]; Hu et al. [34]). From the viewpoint of biology, this means that the upstream amplitude signal will be optimal for repressing downstream gene expression to achieve its cellular function.

In contrast to the FM paradigm, there is an obvious difference in the AM paradigm. In fact, the AM paradigm is modulated to achieve the ON-to-OFF transition rate or the

ON-state time dwell time. However, if the upstream signal is a repressor and the stimulus strength is very large or greater than some fixed value, it will lead to the gene state approximately dwelling on the ON state; that is, the gene expression is approximately constitutive gene expression. In this case, the burst expression will disappear; that is, the extracellular stimulus will repress the burst gene expression. Considering that the OFF-state dwell time is a constant in the AM paradigm, we will directly investigate the downstream mRNA expression mean and the expression noise, as shown in Figure 6.

Figures 6(a) and 6(c) show that the activator prompts gene expression, while the repressor suppresses gene expression as stimulus strength increases. The case in which the upstream signal is in amplitude makes the downstream transcription more efficient than the other two cases (the frequency signal and the noiseless signal) (comparing the blue line with the green line and red line in Figure 6(a)). If the extracellular stimulus suppresses downstream gene expression, the upstream noiseless signal is the most efficient (comparing the red line with the green and blue lines in Figure 6(c)), revealing that the upstream signal will have a different function for the downstream expression. Specifically, if the upstream signal activates gene expression, genes will use extracellular noise to optimize the efficiency of downstream expression, meaning that noise is beneficial for gene expression. However, when the upstream signal is a repressor, extracellular noise is harmful to downstream gene expression, which will reduce the efficiency of gene expression.

Furthermore, the expression noise also exhibits the tight relationship with the upstream stimulus. For the upstream active signal, the expression noise decreases monotonically

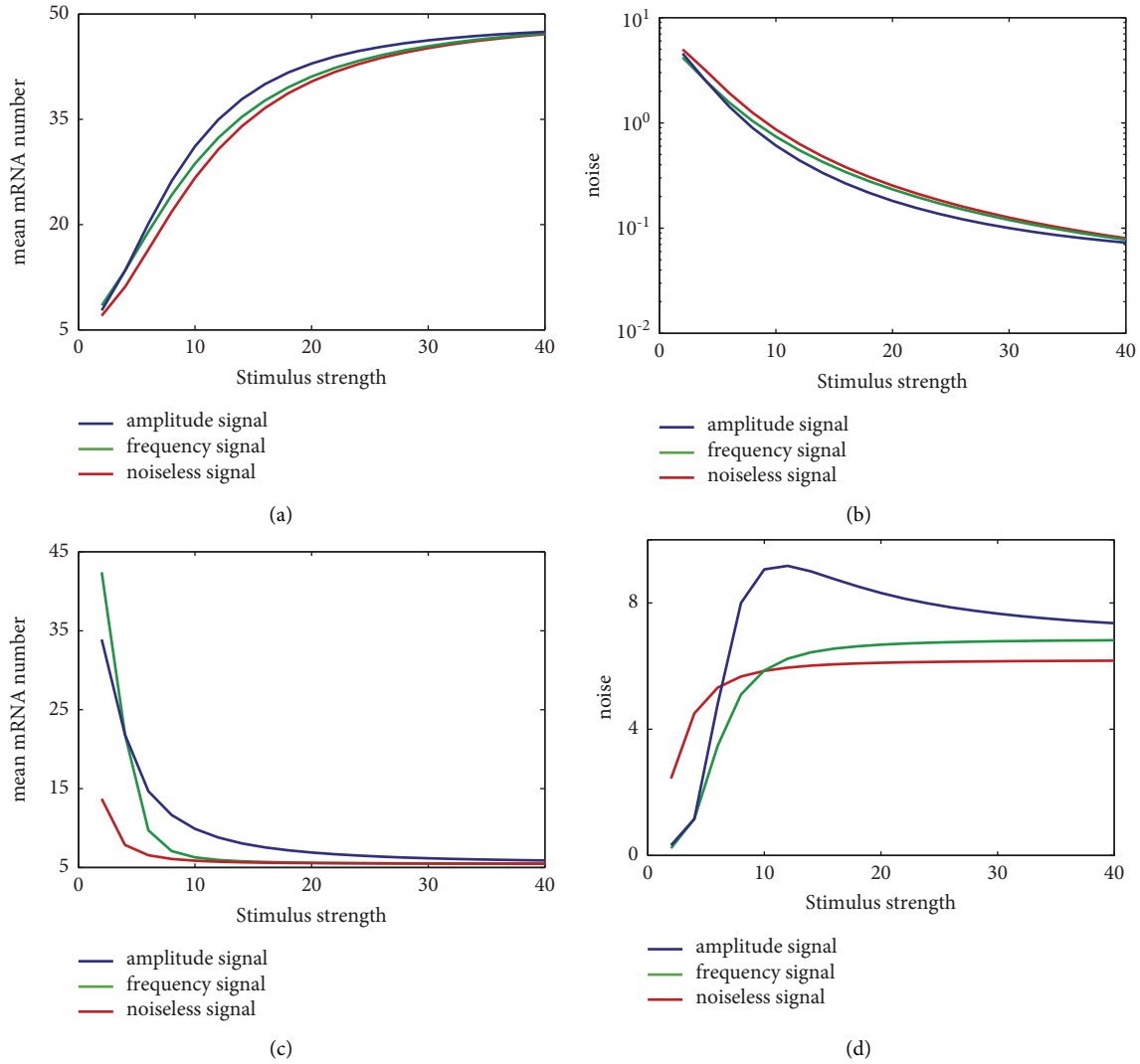


FIGURE 6: The mean mRNA number (a, c) and expression noise (b, d) of downstream gene expression under the AM paradigm. (a, b) indicates that the upstream signal is an activator, while (c, d) indicates that the upstream signal is a repressor. The red line represents the upstream stimulus signal as noiseless, and the green line represents the upstream stimulus signal as random with the burst frequency (α) changing and the burst size (β) being equal to 2, while the blue line represents the upstream stimulus signal as also random with the burst frequency (α) being equal to 2 and the burst size (β) changing. The other parameters are the same as in Figure 2.

with stimulus increasing (refer to Figure 6(b)); that is, the greater the extracellular noise, the less the gene expression noise (see the blue or green line in Figure 6(b)), and the upstream amplitude signal will make the expression noise be least (see the blue line in Figure 6(b)), which further reveals that the extracellular noise is beneficial for the gene expression because the extracellular noise makes the mean mRNA number maximal and the expression noise minimal. But the expression noise has different characteristics, as shown in Figure 6(d). It will monotonically increase as the stimulus increases under noiseless input (see the red line in Figure 6(d)) and frequency signal stimulus (see the green line in Figure 6(d)), and it will increase monotonically first and then decrease monotonically under the amplitude signal stimulus mode (see the blue line in Figure 6(d)). Comparing the coloured lines, we can see that the noiseless input will

minimize the expression noise when the stimulus strength is greater than some value. And, the expression noise under the frequency signal mode is less than that under the amplitude signal mode (comparing the green line with the blue line in Figure 6(d)), indicating that the frequency signal stimulus is better than the amplitude signal stimulus mode for cell stability. Figures 6(c) and 6(d) indicate that the noiseless input will minimize the mean and noise of mRNA when the stimulus strength is greater than some value. That is, when the extracellular stimulus affects downstream gene expression in the AM paradigm, noiseless input will make the gene achieve its function as soon as possible. In this case, noise is harmful to gene expression.

Therefore, Figures 5 and 6 show that extracellular noise has a selective function for downstream gene expression. This phenomenon has been discovered in the yeast

Saccharomyces cerevisiae (Sanchez and Golding [55]) and the transcription protein AP-2 in tumorigenesis (Allouche et al. [39]; Delacroix et al. [40]; Hilger-Eversheim et al. [41]; Matsumoto et al. [42]), which indicates that the gene sometimes uses noise to activate gene expression but sometimes uses noise to suppress gene expression (Paulsson [13]; Wang et al. [32]; Elowitz et al. [56]; Paulsson [57]), depending on the specific extracellular environment.

3.3. Extracellular Fluctuations Induce Phenotypic Diversity. It has been proven by many biological experiments that phenotype switching is an effective method for cells to adapt fluctuating environments, and it is also specific forms for cell decisions to survive better in complex environments (Acar et al. [2]), in which noise may be a potential factor for inducing phenotypic heterogeneity (McAdams [10]; Kaern et al. [58]; Sanchez et al. [59]). Currently, cell phenotype switching is the core issue of epigenetics that determines cell fitness by many factors, including cell adaptation, robustness and growth rate, and noise (also called phenotypic variation). Among these factors, noise is the focus of attention in biology (Wang et al. [32]; Kussell and Leibler [60]). It may be beneficial for the cells to live in the fluctuation environment, at least for population diversity, and it can supply more alternate ways to adjust its survival strategy when the cell encounters sudden changes, such as changes in temperature, illumination by sensing, or the concentrations of ligand molecules (Hung et al. [17]; Lehner [61]). In addition, the number of peaks of steady-state product distribution and its change process are important measures of phenotypic diversity in constant or fluctuating environments (Kaern et al. [58]).

The bimodal distribution as a common result has been verified in many single-cell experiments in vivo (Thomas et al. [25]; To and Maheshri [62]), demonstrating that the cell indeed can encode a distinct signal at different expression levels even in isogenic cells. In fact, many researchers have reported that if promoter switching is a slow process relative to the expression process, it will result in a distribution characterized by different expression levels, such as bimodal distributions (Thomas et al. [25]; Qian et al. [27]) or some mixed exponential distributions (Raser [28]) and multimodal distributions (Thomas et al. [25]). Despite these findings, it is unclear how extracellular random signal input regulates the steady-state distribution of gene products.

Therefore, we first investigate the mechanism of how extracellular random input can regulate the dynamic features of the mRNA steady-state distribution, as shown in Figure 7. Figure 7 indicates that the activator will prompt the bimodal distribution, while the repressor will make the distribution change from bimodal to unimodal; that is, the activator (repressor) increases (decreases) the number of gene products by modulating the peak value of the expression distribution. Specifically, when the transcription factors are activators, the high expression peak will be higher to increase the mRNA number under the FM paradigm (see Figure 7(a)), meaning that the extracellular stimulus will not change the bimodal but change the high expression peak.

However, under the AM paradigm, the extracellular stimulus will not only change the high expression peak but also make the distribution change into bimodal (refer to Figure 7(c)); that is, the AM paradigm affects downstream gene expression by binary channels. The reason behind this situation may be that the AM paradigm focuses on the information of promoter dwell time at ON or OFF states, while the FM paradigm only emphasizes when the promoter switches from the OFF states into ON states; that is, the AM paradigm is the realization of classic results of Berg and Purcell that the ligand concentration is estimated by time averaging (Micali et al. [5]). Second, when the transcription factors are repressors, the high expression peak will be smaller to decrease the mRNA number (see Figures 7(b) and 7(d)), indicating that the extracellular noise will suppress the bimodal as the stimulus increasing. Specifically, the stimulus signal will induce the expression distribution change from bimodal to unimodal under the FM paradigm, but the case is nearly opposite under the AM paradigm (see the green and blue drawing in Figure 7(d)). Moreover, the distribution will almost not change to unimodal under the noiseless paradigm (see the red drawing in Figure 7(d)). This difference indicates the effect of extracellular noise on gene expression. Although the stimulus strength is the same, the extracellular noise source and the decoding paradigm will affect the downstream gene expression together.

Finally, we investigate the effect of extracellular noise on the peak of stationary distribution, especially on the high expression peak, to uncover the sensitivity of noise regulation, as shown in Figure 8. The upstream frequency signal as activators promotes the high expression to be larger than amplitude signal stimulus under the FM paradigm (comparing the green and blue drawings in Figure 8(a)); that is, the coherent resonance indeed has advantages. However, the noiseless input induces the higher expression peak to the stable value fastest, indicating that the extracellular noise is harmful for gene expression under the FM paradigm. Similarly, the amplitude signal will suppress the higher peak value more severely than the frequency signal and noiseless input under the FM paradigm (see Figure 8(b)); that is, the gene uses extracellular noise to survive and that the amplitude fluctuation is more sensitive to downstream gene expression.

In contrast to the FM paradigm, the situation will be different under the AM paradigm. Specifically, when the transcription factors are activators, the amplitude signal is more sensitive to the higher expression peak (see the second column in Figure 8(c)), indicating the superiority of coherent resonance. However, when the transcription factors are repressors, the noiseless input will make the gene achieve its goal fastest (see the red histogram in Figure 8(d)).

4. Discussion

The cell encodes the extracellular stimuli (normally a fast process) by sensing the concentration or frequency changes of ligand molecules and then responds to the stimulus signal by decoding the information of ligand molecules at different levels (slow process). In general, these decoding processes

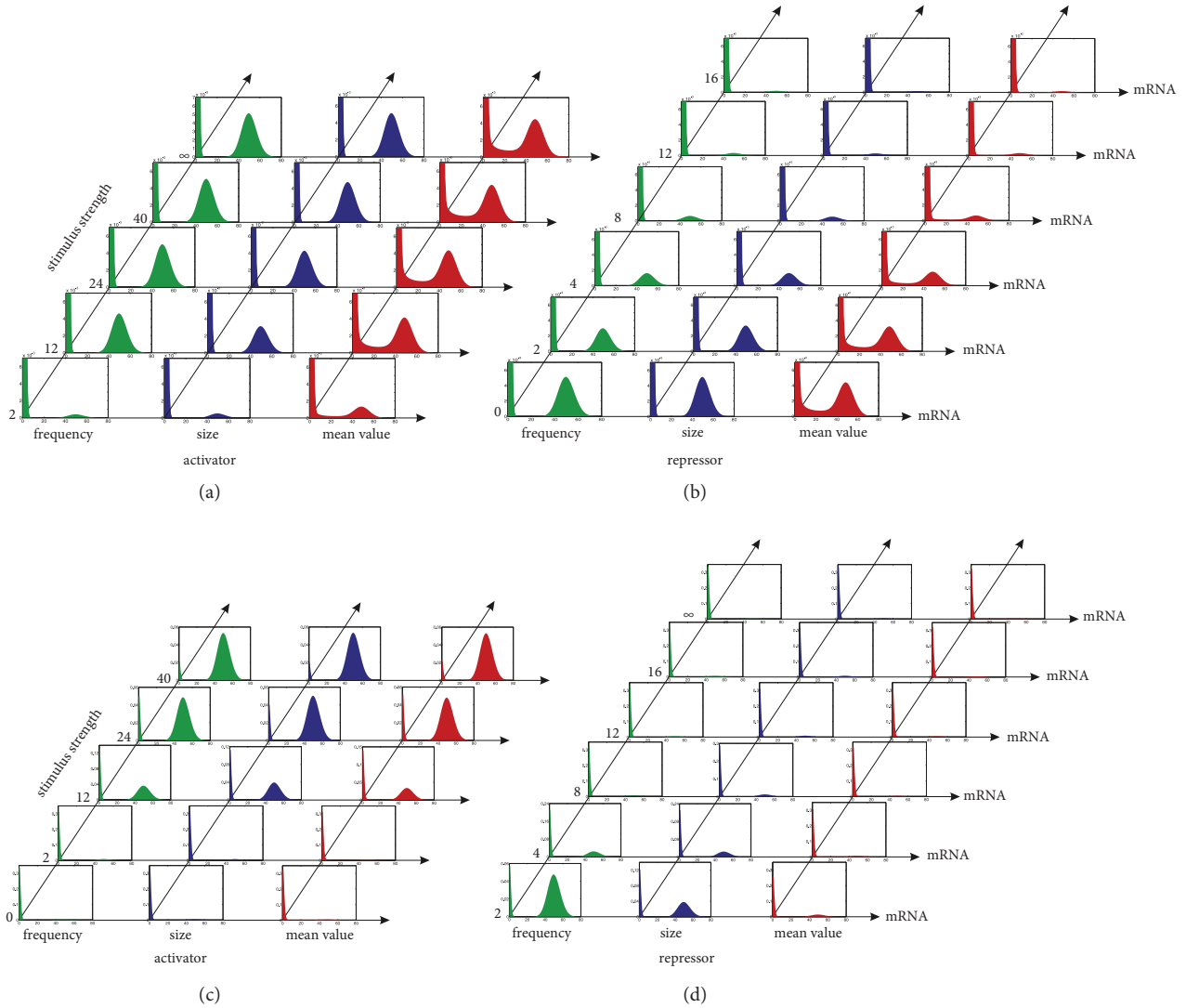


FIGURE 7: Dependence of mRNA distribution on the upstream signal of role and stimulus strength: (a, b) FM paradigm; (c, d) AM paradigm; (a, c) activator; (b, d) repressor. The green drawing represents the upstream input as the frequency signal, the blue drawing represents the upstream input as the amplitude signal, and the red drawing represents the noiseless input. Other parameter values are set as $k_{on} = 0.01$, $k_{off} = 0.1$, $k_d = 10$, $n = 2$, $\lambda_0 = 1$, and $\lambda_1 = 50$.

have two strategies, AM and FM, focusing on the different kinetic features of signals. Here, we employ a simple mechanistic gene expression model with signal stimuli, including three kinds of sources of noise: the promoter switching process, the generation and degradation process, and the signal stimulus process induced by transcription factor fluctuation. By analysis, we have shown that under different timescales and modulation paradigms, the extracellular stimulus has different functions for downstream gene expression. If the transcription factors are activators, the noiseless input will be best for downstream gene expression under the FM paradigm, while the amplitude signal is the best choice to suppress downstream burst expression. Meanwhile, under the AM paradigm, the amplitude signal will be optimal to activate downstream burst expression, and the noiseless signal that suppresses downstream gene expression will be more efficient (Figures 4–6). Furthermore,

the activator will induce bimodality, and the repressor will attenuate bimodality. The number of peaks and their value of distribution further illustrate this feature (Figures 7–8). Moreover, there often exists feedback loop module in the actual expression process due to the complexity of gene expression; however, the feedback loop would reduce into the upstream stimulus signal by an open-loop approach and then unify in the stimulus-response framework (Maleki et al. [63]; Niraj et al. [64]), which will not change the quantitative relationship in our model.

The biological function to achieve often submits to some restraints locally or globally, coming from the design principles of the network from one aspect and optimal evolutionary fitness from another (Perkins and Swain [54]; Schwanhäusser et al. [65]). First, the cell will always seek high efficiency for each process in gene expression. Here, it indicates that, under the FM paradigm, the noiseless input as

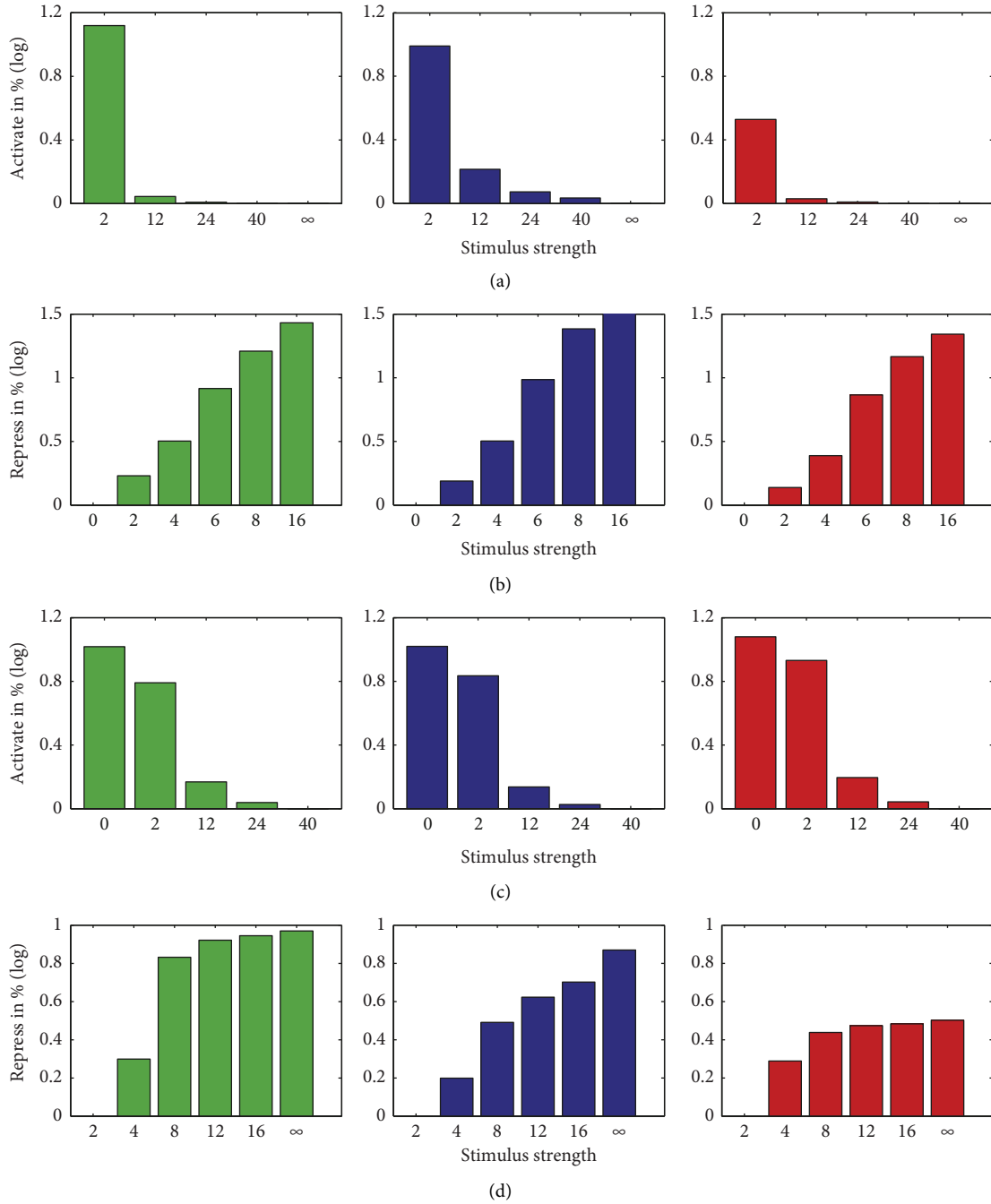


FIGURE 8: The changing trends of higher expression peaks under different paradigms: (a, b) FM paradigm; (c, d) AM paradigm; (a, c) activator; (b, d) repressor. The green drawing represents the upstream input as the frequency signal, the blue drawing represents the upstream input as the amplitude signal, and the red drawing represents the noiseless input. Other parameter values are set as $k_{\text{on}} = 0.01$, $k_{\text{off}} = 0.1$, $k_d = 10$, $n = 2$, $\lambda_0 = 1$, and $\lambda_1 = 50$.

an activator signal will prompt the downstream product to be larger than in other cases, while the gene products are the least suppressed by the amplitude signal. However, under the AM paradigm, if the upstream signal activates gene expression, then the amplitude signal will have the largest gene products, but the effect of noiseless input suppressing downstream gene expression will be best. Taken together, our results suggest that AM should be more advantageous in gene regulation, but FM is more common in the signal

sensing pathway (Hao et al. [3]; Micali et al. [5]). Another constraint is stability (Paulsson [13]; Schwanhäusser et al. [65]; Sanchez et al. [59]). We have shown that the expression noise has different effects in different contexts. The expression noise will be the lowest when the extracellular signal is a noiseless signal and an activator under the FM paradigm, while the amplitude signal will minimize the expression noise under the FM paradigm if the extracellular signal represses the downstream expression. Moreover, the

gene is more stable under the AM paradigm if the amplitude signal activates the downstream expression but will decrease the expression noise under the AM paradigm if the signal suppresses the downstream gene expression (Figures 4–6). When “efficiency” and “stability” in the regulation network are required to satisfy simultaneously, a trade-off on the fitness may exist, which is worthy of further study.

In general, dynamic disordering readjusts to the new ordering need to sense and handle the information transmitted from extracellular stimuli and consume the energy to attenuate the fluctuation from the environment (Micali et al. [5]; Mehta and Schwab [66]). The second issue is well investigated in many synthetic molecular networks, such as molecular motors, actin, and the myosin protein, while the first issue may be essential and more difficult because the global understanding of cells needs to be investigated from the viewpoint of systems theory. From the single-cell experiments and the idea of engineering principles [67–70], we believe that the collective behaviour in the large regulation network stimulated by many sources of signals can be most likely to elucidate clearly.

Data Availability

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Authors' Contributions

H. H. W. and S. J. X. proposed and designed this study, did all numerical simulations described in the article, interpreted the results, and wrote the manuscript. Z.G.W., Y.W., and H.H.W. performed an analytical treatment of a stochastic differential equation, and contributed to the discussions. Sijia Xiao and Yan Wang contributed equally to this work.

Acknowledgments

The authors would like to thank Elsevier Language Editing (<https://webshop.elsevier.com/>) for English language editing. This study was supported by the Hainan Province Science and Technology Special Fund (No. ZDYF2021SHFZ231), the National Natural Science Foundation of China (Grant nos. 12261028, 11761025, and 11961018), Natural Science Foundation of Hainan Province (Grant nos. 120RC451 and 2019RC168), Hainan Province Innovative Scientific Research Project for Graduate Students (Qhys2021-208), and the Specific Research Fund of the Innovation Platform for Academicians of Hainan Province.

Supplementary Materials

Supplementary information accompanying this article refers to the online supporting material. Figure S1: the mean mRNA number and expression noise of downstream gene

expression is under FM modulation paradigm when the transcription factors are the activator. The red line represents that the upstream stimulus signal is noiseless, and the green line represents the upstream stimulus signal is random with the burst frequency (α) changing and the burst size (β) being equal to 2, while the blue line represents that the upstream stimulus signal is also random with the burst frequency (α) being equal to 2 and the burst size (β) changing. All parameters are the same as those listed in the main text. Figure S2: the mean mRNA number (A) and the noise (B) of downstream gene expression is under FM modulation paradigm when the transcription factors are the repressor. The red line represents that the upstream stimulus signal is noiseless, and the green line represents the upstream stimulus signal is random with the burst frequency (α) changing and the burst size (β) being equal to 2, while the blue line represents that the upstream stimulus signal is also random with the burst frequency (α) being equal to 2 and the burst size (β) changing. All parameters are the same as those listed in the main text. (*Supplementary Materials*)

References

- [1] L. Cai, C. K. Dalal, and M. B. Elowitz, “Frequency-modulated nuclear localization bursts coordinate gene regulation,” *Nature*, vol. 455, no. 7212, pp. 485–490, 2008.
- [2] M. Acar, J. T. Mettetal, and A. van Oudenaarden, “Stochastic switching as a survival strategy in fluctuating environments,” *Nature Genetics*, vol. 40, no. 4, pp. 471–475, 2008.
- [3] N. Hao, B. A. Budnik, J. Gunawardena, and E. K. O’Shea, “Tunable signal processing through modular control of transcription factor translocation,” *Science*, vol. 339, no. 6118, pp. 460–464, 2013.
- [4] N. Hao and E. K. O’Shea, “Signal-dependent dynamics of transcription factor translocation controls gene expression,” *Nature Structural & Molecular Biology*, vol. 19, no. 1, pp. 31–39, 2012.
- [5] G. Micali, G. Aquino, D. M. Richards, and R. G. Endres, “Accurate encoding and decoding by single cells: amplitude versus frequency modulation,” *PLoS Computational Biology*, vol. 11, no. 6, Article ID e1004222, 2015.
- [6] S. Ben-Tabou de-Leon and E. H. Davidson, “Modeling the dynamics of transcriptional gene regulatory networks for animal development,” *Developmental Biology*, vol. 325, no. 2, pp. 317–328, 2009.
- [7] S. Komili and P. A. Silver, “Coupling and coordination in gene expression processes, a systems biology view,” *Nature Reviews Genetics*, vol. 9, no. 1, pp. 38–48, 2008.
- [8] J. C. W. Locke and M. B. Elowitz, “Using movies to analyse gene circuit dynamics in single cells,” *Nature Reviews Microbiology*, vol. 7, no. 5, pp. 383–392, 2009.
- [9] D. Muzzey and A. van Oudenaarden, “Quantitative time-lapse fluorescence microscopy in single cells,” *Annual Review of Cell and Developmental Biology*, vol. 25, no. 1, pp. 301–327, 2009.
- [10] H. H. McAdams and A. Arkin, “Stochastic mechanisms in gene expression,” *Proceedings of the National Academy of Sciences*, vol. 94, no. 3, pp. 814–819, 1997.
- [11] M. Thattai and A. van Oudenaarden, “Intrinsic noise in gene regulatory networks,” *Proceedings of the National Academy of Sciences*, vol. 98, no. 15, pp. 8614–8619, 2001.

- [12] T. B. Kepler and T. C. Elston, "Stochasticity in transcriptional regulation, origins, consequences, and mathematical representations," *Biophysical Journal*, vol. 81, no. 6, pp. 3116–3136, 2001.
- [13] J. Paulsson, "Models of stochastic gene expression," *Physics of Life Reviews*, vol. 2, pp. 157–175, 2005.
- [14] J. M. Raser and E. K. O'Shea, "Noise in gene expression, origins, consequences, and control," *Science*, vol. 309, no. 5743, pp. 2010–2013, 2005.
- [15] A. Sanchez and J. Kondev, "Transcriptional control of noise in gene expression," *Proceedings of the National Academy of Sciences*, vol. 105, no. 13, pp. 5081–5086, 2008.
- [16] J. M. Pedraza and J. Paulsson, "Effects of molecular memory and bursting on fluctuations in gene expression," *Science*, vol. 319, no. 5861, pp. 339–343, 2008.
- [17] M. Hung, E. Chang, R. Hussein et al., "Modulating the frequency and bias of stochastic switching to control phenotypic variation," *Nature Communications*, vol. 5, no. 1, p. 4574, 2014.
- [18] D. R. Larson, "What do expression dynamics tell us about the mechanism of transcription?" *Current Opinion in Genetics & Development*, vol. 21, no. 5, pp. 591–599, 2011.
- [19] S. Chong, C. Chen, H. Ge, and X. S. Xie, "Mechanism of transcriptional bursting in bacteria," *Cell*, vol. 158, no. 2, pp. 314–326, 2014.
- [20] A. Raj, C. S. Peskin, D. Tranchina, D. Y. Vargas, and S. Tyagi, "Stochastic mRNA synthesis in mammalian cells," *PLoS Biology*, vol. 4, no. 10, Article ID e309, 2006.
- [21] B. Munsky and G. Neuert, "From analog to digital models of gene regulation," *Physical Biology*, vol. 12, no. 4, Article ID 045004, 2015.
- [22] R. D. Dar, B. S. Razoooky, A. Singh et al., "Transcriptional burst frequency and burst size are equally modulated across the human genome," *Proceedings of the National Academy of Sciences*, vol. 109, no. 43, pp. 17454–17459, 2012.
- [23] S. S. Dey, J. E. Foley, P. Limsirichai, D. V. Schaffer, and A. P. Arkin, "Orthogonal control of expression mean and variance by epigenetic features at different genomic loci," *Molecular Systems Biology*, vol. 11, no. 5, p. 806, 2015.
- [24] J. M. Schmiedel, S. L. Klemm, Y. Zheng et al., "MicroRNA control of protein expression noise," *Science*, vol. 348, no. 6230, pp. 128–132, 2015.
- [25] P. Thomas, N. Popovic, and R. Grima, "Phenotypic switching in gene regulatory networks," *Proceedings of the National Academy of Sciences*, vol. 111, no. 19, pp. 6994–6999, 2014.
- [26] H. Ge, H. Qian, and X. S. Xie, "Stochastic phenotype transition of a single cell in an intermediate region of gene state switching," *Physical Review Letters*, vol. 114, no. 7, Article ID 078101, 2015.
- [27] H. Qian, P. Z. Shi, and J. Xing, "Stochastic bifurcation, slow fluctuations, and bistability as an origin of biochemical complexity," *Physical Chemistry Chemical Physics*, vol. 11, no. 24, pp. 4861–4870, 2009.
- [28] J. M. Raser and E. K. O'Shea, "Control of stochasticity in eukaryotic gene expression," *Science*, vol. 304, no. 5678, pp. 1811–1814, 2004.
- [29] J. Wang, L. Xu, E. Wang, and S. Huang, "The potential landscape of genetic circuits imposes the arrow of time in stem cell differentiation," *Biophysical Journal*, vol. 99, no. 1, pp. 29–39, 2010.
- [30] X. M. Zhu, L. Yin, L. Hood, and P. Ao, "Calculating biological behaviors of epigenetic states in the phage λ life cycle," *Functional & Integrative Genomics*, vol. 4, no. 3, pp. 188–195, 2004.
- [31] A. Ochab-Marcinek and M. Tabaka, "Bimodal gene expression in noncooperative regulatory systems," *Proceedings of the National Academy of Sciences*, vol. 107, no. 51, pp. 22096–22101, 2010.
- [32] H. Wang, Z. Yuan, P. Liu, and T. Zhou, "Division time-based amplifiers for stochastic gene expression," *Molecular BioSystems*, vol. 11, no. 9, pp. 2417–2428, 2015.
- [33] H. P. Lu, L. Xun, and X. S. Xie, "Single-molecule enzymatic dynamics," *Science*, vol. 282, no. 5395, pp. 1877–1882, 1998.
- [34] B. Hu, D. A. Kessler, W. J. Rappel, and H. Levine, "Effects of input noise on a simple biochemical switch," *Physical Review Letters*, vol. 107, no. 14, Article ID 148101, 2011.
- [35] B. Hu, D. A. Kessler, W. J. Rappel, and H. Levine, "How input fluctuations reshape the dynamics of a biological switching system," *Physical Review*, vol. 86, no. 6, Article ID 061910, 2012.
- [36] D. R. Larson, D. Zenklusen, B. Wu, J. A. Chao, and R. H. Singer, "Real-time observation of transcription initiation and elongation on an endogenous yeast gene," *Science*, vol. 332, no. 6028, pp. 475–478, 2011.
- [37] A. S. Hansen and E. K. O'Shea, "Promoter decoding of transcription factor dynamics involves a trade-off between noise and control of gene expression," *Molecular Systems Biology*, vol. 9, no. 1, p. 704, 2013.
- [38] B. Munsky, Z. Fox, and G. Neuert, "Integrating single-molecule experiments and discrete stochastic models to understand heterogeneous gene transcription dynamics," *Methods*, vol. 85, pp. 12–21, 2015.
- [39] A. Allouche, G. Nolens, A. Tancredi et al., "The combined immunodetection of AP-2 α and YY1 transcription factors is associated with ERBB2 gene overexpression in primary breast tumors," *Breast Cancer Research*, vol. 10, no. 1, 2008.
- [40] L. Delacroix, D. Begon, G. Chatel, P. Jackers, and R. Winkler, "Distal ERBB2 promoter fragment displays specific transcriptional and nuclear binding activities in ERBB2 overexpressing breast cancer cells," *DNA and Cell Biology*, vol. 24, no. 9, pp. 582–594, 2005.
- [41] K. Hilger-Eversheim, M. Moser, H. Schorle, and R. Buettner, "Regulatory roles of AP-2 transcription factors in vertebrate development, apoptosis and cell-cycle control," *Gene*, vol. 260, no. 1–2, pp. 1–12, 2000.
- [42] M. Matsumoto, S. Han, T. Kitamura, and D. Accili, "Dual role of transcription factor FoxO1 in controlling hepatic insulin sensitivity and lipid metabolism," *The Journal of Clinical Investigation*, vol. 116, no. 9, pp. 2464–2472, 2006.
- [43] R. Baler, S. Covington, and D. C. Klein, "Rat arylalkylamine N-acetyltransferase gene: upstream and intronic components of a bipartite promoter," *Biology of the Cell*, vol. 91, no. 9, pp. 699–705, 1999.
- [44] L. Huang, Z. Yuan, P. Liu, and T. Zhou, "Effects of promoter leakage on dynamics of gene expression," *BMC Systems Biology*, vol. 9, no. 1, 2015.
- [45] M. Behar and A. Hoffmann, "Understanding the temporal codes of intra-cellular signals," *Current Opinion in Genetics & Development*, vol. 20, no. 6, pp. 684–693, 2010.
- [46] G. Aquino, L. Tweedy, D. Heinrich, and R. G. Endres, "Memory improves precision of cell sensing in fluctuating environments," *Scientific Reports*, vol. 4, no. 1, p. 5688, 2014.
- [47] C. W. Gardiner, *Handbook of Stochastic Methods*, Springer, Berlin, 1985.
- [48] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry*, Elsevier, 1992.
- [49] N. Friedman, L. Cai, and X. S. Xie, "Linking stochastic dynamics to population distribution: an analytical framework of

- gene expression,” *Physical Review Letters*, vol. 97, no. 16, Article ID 168302, 2006.
- [50] B. Oksendal, “Stochastic differential equations: an introduction with applications,” *SIAM Review*, vol. 29, no. 1, pp. 163–164, 2013.
- [51] J. Zhang and T. Zhou, “Promoter-mediated transcriptional dynamics,” *Biophysical Journal*, vol. 106, no. 2, pp. 479–488, 2014.
- [52] V. Shahrezaei and P. S. Swain, “Analytical distributions for stochastic gene expression,” *Proceedings of the National Academy of Sciences*, vol. 105, no. 45, pp. 17256–17261, 2008.
- [53] L. Huang, Z. Yuan, P. Liu, and T. Zhou, “Feedback-induced counterintuitive correlations of gene expression noise with bursting kinetics,” *Physical Review*, vol. 90, no. 5, Article ID 052702, 2014.
- [54] T. J. Perkins and P. S. Swain, “Strategies for cellular decision-making,” *Molecular Systems Biology*, vol. 5, no. 1, p. 326, 2009.
- [55] A. Sanchez and I. Golding, “Genetic determinants and cellular constraints in noisy gene expression,” *Science*, vol. 342, no. 6163, pp. 1188–1193, 2013.
- [56] M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain, “Stochastic gene expression in a single cell,” *Science*, vol. 297, no. 5584, pp. 1183–1186, 2002.
- [57] J. Paulsson, “Summing up the noise in gene networks,” *Nature*, vol. 427, no. 6973, pp. 415–418, 2004.
- [58] M. Kærn, T. C. Elston, W. J. Blake, and J. J. Collins, “Stochasticity in gene expression: from theories to phenotypes,” *Nature Reviews Genetics*, vol. 6, no. 6, pp. 451–464, 2005.
- [59] A. Sanchez, S. Choubey, and J. Kondev, “Regulation of noise in gene expression,” *Annual Review of Biophysics*, vol. 42, no. 1, pp. 469–491, 2013.
- [60] E. Kussell and S. Leibler, “Phenotypic diversity, population growth, and information in fluctuating environments,” *Science*, vol. 309, no. 5743, pp. 2075–2078, 2005.
- [61] B. Lehner, “Conflict between noise and plasticity in yeast,” *PLoS Genetics*, vol. 6, no. 11, Article ID e1001185, 2010.
- [62] T. L. To and N. Maheshri, “Noise can induce bimodality in positive transcriptional feedback loops without bistability,” *Science*, vol. 327, no. 5969, pp. 1142–1145, 2010.
- [63] F. Maleki and A. Becskei, “An open-loop approach to calculate noise-induced transitions,” *Journal of Theoretical Biology*, vol. 415, pp. pp145–157, 2017.
- [64] N. Kumar, T. Platini, and R. V. Kulkarni, “Exact distributions for stochastic gene expression models with bursting and feedback,” *Physical Review Letters*, vol. 113, no. 26, Article ID 268105, 2014.
- [65] B. Schwanhäusser, D. Busse, N. Li et al., “Global quantification of mammalian gene expression control,” *Nature*, vol. 473, no. 7347, pp. 337–342, 2011.
- [66] P. Mehta and D. J. Schwab, “Energetic costs of cellular computation,” *Proceedings of the National Academy of Sciences*, vol. 109, no. 44, pp. 17978–17982, 2012.
- [67] J. Zhang, Z. Yuan, and T. Zhou, “Physical limits of feedback noise-suppression in biological networks,” *Physical Biology*, vol. 6, no. 4, Article ID 046009, 2009.
- [68] L. B. Carey, D. Van Dijk, P. M. A. Sloot, J. A. Kaandorp, and E. Segal, “Promoter sequence determines the relationship between expression level and noise,” *PLoS Biology*, vol. 11, no. 4, Article ID e1001528, 2013.
- [69] C. C. Reed, A. Waterhouse, S. Kirby et al., “Decorin prevents metastatic spreading of breast cancer,” *Oncogene*, vol. 24, no. 6, pp. 1104–1110, 2005.
- [70] C. T. H. Stadtländer and H. Stadtländer, “Quantitative biology: from molecular to cellular systems,” *Journal of Biological Dynamics*, vol. 9, no. 1, pp. 336–340, 2015.