Framework estimation of stochastic gene activation using transcription average level

Liang Chen^{1,2,†}, Genghong Lin^{1,2,†} and Feng Jiao^{1,2,*}

¹Center for Applied Mathematics, Guangzhou University, Guangzhou, P. R. China.

²School of Mathematics and Information Sciences, Guangzhou University, Guangzhou, P. R. China.

Abstract Gene activation is usually a non-Markovian process that has been modeled as various frameworks that consist of multiple rate-limiting steps. Understanding the exact activation framework for a gene of interest is a central problem for single-cell studies. In this paper, we focus on the dynamical data of the average transcription level $M(t)$, which is typically neglected when deciphering gene activation. Firstly, the smooth trend lines of $M(t)$ data present rich, visually dynamic features. Secondly, tractable analysis of $M(t)$ allows the establishment of bijections between $M(t)$ dynamics and system parameter regions. Because of these two clear advantages, we can rule out frameworks that fail to capture $M(t)$ features and we can further test potential competent frameworks by fitting $M(t)$ data. We implemented this procedure to determine an exact activation framework for a large number of mouse fibroblast genes under tumor necrosis factor induction; the cross-talk between the signaling and basal pathways is crucial to trigger the first peak of $M(t)$, while the following damped gentle $M(t)$ oscillation is regulated by the multi-step basal pathway. Moreover, the fitted parameters for the mouse genes tested revealed two distinct regulation scenarios for transcription dynamics. Taken together, we were able to develop an efficient procedure for using traditional $M(t)$ data to estimate the gene activation frameworks and system parameters. This procedure, together with sophisticated single-cell transcription data, may facilitate a more accurate understanding of stochastic gene activation.

Author Summary It has been suggested that genes randomly transit between inactive and active states, with mRNA produced only when a gene is active. The gene activation process has been modeled as a framework of multiple rate-limiting steps listed sequentially, parallel, or in combination. The system step numbers and parameters can be predicted

[†]L.C. and G.L. contributed equally.

^{*}Corresponding author: F.J. (jf1704@163.com).

by computationally fitting sophisticated single-cell transcription data. However, current algorithms require a prior hypothetical framework of gene activation. We found that the prior estimation of the framework can be achieved using the traditional dynamical data of mRNA average level $M(t)$ which present easily discriminated dynamical features. The theory regarding $M(t)$ profiles allows us to confidently rule out other frameworks and to determine optimal frameworks by fitting $M(t)$ data. We successfully applied this procedure to a large number of mouse fibroblast genes and confirmed that $M(t)$ is capable of providing a reliable estimation of gene activation frameworks and system parameters.

1 Introduction

Gene transcription is a random process in virtually all genomic loci, for which messenger RNA (mRNA) molecules for active genes are produced in a bursting fashion in which an episode of transcriptional activity is interrupted by irregular gene inactivation periods [1–3]. A central problem in the study of stochastic gene transcription has been understanding regulation scenarios that control random gene activation (*on*) and inactivation (*off*) in response to environmental changes [4–7]. The tremendous effort expended in this endeavor over the last few decades has generated massive amounts of data at the single-cell level and has produced many important observations [8–10].

Real-time imaging of transcriptional bursting makes it possible to count the durations of each gene *on* and *off* period along the entire timeline, which generates duration distributions for both gene *on* and *off* periods, respectively. For instances of the *Escherichia coli* $P_{lac/ara}$ promoter [11] and yeast *FLO11* genes [12], their *on* and *off* periods are all well fitted by single exponential distributions. These observations support the classical two-state model shown in Fig. 1a, that the gene for turning genes *on* and *off* are all controlled by single rate-limiting biochemical steps, with synthesis of mRNA when the gene is *on* and degradation of the gene all being controlled by single rate-limiting steps [1, 3]. The exponentially distributed *on* period is one of the few universal features of transcription present in both prokaryotic and eukaryotic genes [6, 8, 13]. However, the duration of the *off* state is highly gene-specific, and is manifested by the observed gamma distribution with a unique peak for mouse fibroblast genes [8] and *E. coli tetA* promoters [14]. Gamma distribution or more general non-Markovian processes can be mathematically explained by assuming that the gene for activating the *on* period is directed by a single pathway consisting of multiple sequential rate-limiting steps (e.g., the three-state model in Fig. 1b) [15], the parallel competitive rate-limiting pathways (e.g., the cross-talking pathways model in Fig. 1c) [16], or a

combination of both (e.g., Fig. 1d) [17]. However, these theoretical approaches are unable to determine the best framework to describe the observed non-Markovian gene activation, although they have provided good approximations to the downstream distribution of gene transcription [18, 19].

Figure 1: Different frameworks that direct stochastic gene activation (*on*). Other processes of gene inactivation (*off*), mRNA synthesis when the gene is *on*, and mRNA degradation are all determined by single rate-limiting steps at constant rates. (a) Two-state model. The gene is activated through a single rate-limiting step at a constant rate. (b) Three-state model. Gene activation is regulated by two sequential rate-limiting steps at a constant rate. (c) Cross-talking pathway model. The gene is activated by two different competitive rate-limiting pathways with selection probabilities q_1 and q_2 of the two pathways satisfying $q_1 + q_2 = 1$. (d) Cross-talking three-state model. The gene is activated either by a pathway consisting of two sequential rate-limiting steps, or alternatively, by a single rate-limiting pathway, with constant rates.

The snapshot data for the distribution histogram of mRNA copy numbers in an isogenic cell population at different time points carry rich dynamic information on fluctuations in transcription [20]. When combined with mathematical models, the fit of mRNA (or other RNA types) distribution data has served as a powerful tool for revealing the multi-step regulation of the activation of different genes in bacteria, yeast, and human cells [9, 20, 21]. However, the calculation of exact forms of dynamical mRNA distribution requires solving infinite arrays of chemical master equations under the whole parameter region of the models, which is beyond the scope of standard theoretical methods, even for the simplest two-state model (Fig. 1a) [22–24]. Fitting of mRNA distribution data must integrate several computational tools to determine the rate-limiting step numbers in the activation framework

and search for suitable system parameters [9,20]. However, current computation algorithms focus only on a class of prior hypothetical multi-step gene activation; therefore, additional competent frameworks may be ignored. Moreover, some typical dynamical transition patterns among the different mRNA distribution modes can be well exhibited by all the models in Fig. 1 [24–26], preventing a direct way to rule out models that do not panoramically match the transition patterns of dynamical mRNA distribution.

The steady-state measurement of gene transcription under different cellular conditions has generated a large dataset of mRNA distribution and its mean level M , the Fano factor ϕ (the variance over M), and noise CV^2 (ϕ over M) [1,10]. By virtue of mathematical models, fitting steady-state data has revealed a large spectrum of regulation scenarios that cells utilize in response to environmental changes [3,10]. The steady-state mRNA distributions observed so far are often shown to be decaying, unimodal, or bimodal [1,3]. However, the models in Fig. 1 can only generate the three distribution modes shown at steady-state [15,23,27,28], suggesting that the limited mRNA distribution modalities are insufficient to map reversely onto the diversified frameworks of gene activation. The steady-state data of noise CV^2 , Fano factor ϕ , and mean level M , when mapped as scattered points onto $M-CV^2$ and $M-\phi$ planes, provide a diagram of trend lines of CV^2 and ϕ against M under varying environments [2,10]. For a given gene of interest in *E. coli*, yeast, or mammalian cells, the trend lines fitted by different models have revealed distinct regulation scenarios [29]. However, the scenario that plays a dominant role in gene regulation remains elusive.

In contrast to the time-consuming single-cell measurements that require RNA labeling and imaging with high sensitivity and resolution [11,21], the dynamical mRNA average level $M(t)$ can be relatively easily captured by conventional methods at the cell population level [21,30]. Previous studies have revealed rich temporal profiles of $M(t)$ for different genes and cellular conditions, such as monotonic increases in the *E. coli* promoter $P_{lac/ara}$ [11], up-and-down behavior in the *c-Fos* gene in human osteosarcoma [21], multiple peaks in mouse fibroblast genes [30,31], and even oscillations in yeast stress-induced genes [32]. These observations give rise to the problem of whether such rich dynamical behaviors of $M(t)$ can be mapped back to the diversified frameworks of gene activation. To achieve this goal, the key objective is to establish bijections between the dynamical features of $M(t)$ and the parameter regions for certain mathematical models. This allows us to rule out models that do not capture the exhibited dynamical features of $M(t)$ and to test the simplest of the remaining models on the basis of their fit to $M(t)$ data. In this study, we assumed that gene activation is regulated by a combination of sequential and parallel pathways, as shown in Fig. 1, and we illustrated the way that dynamic mRNA average level data can be utilized to help estimate the gene activation frameworks.

2 Results

2.1 Cross-talking three-state model

To make the paper easier to follow, we focused on analyzing the mRNA average level $M(t)$ data from a large group of mouse fibroblast genes under cytokine tumor necrosis factor (TNF) stimulation conditions [30, 31]. With the exception of the simple monotonic growth of $M(t)$ generated by late response genes, the rich non-monotonic behaviors of $M(t)$ have also been determined, such as up-and-down for the *Fos* gene, up-down-up for the *Cxcl1* gene, and damped oscillation with multiple peaks for the *Nfkb1a* gene.

To determine the framework that can best capture the rich transcription dynamics of mouse fibroblast genes, theoretical bijections between the dynamical features of $M(t)$ and three mathematical models were established (Table 1). The two-state model shown in Fig. 1a can only generate monotonic increasing dynamics of $M(t)$ [29], and thus is not suitable for the discussion of non-monotonic dynamical behaviors. The three-state model shown in Fig. 1b is proven to display damped oscillatory dynamics of $M(t)$ under a certain parameter region [33, 34]. However, such oscillation behavior is almost invisible owing to its rapid exponential decay, and only slightly slows down the dynamic increase in $M(t)$ [34]. The frameworks with two or more parallel pathways can capture the up-and-down dynamics of $M(t)$, but fails to generate more complex transcription dynamics [29, 35]. Moreover, the cross-talking pathways model (Fig. 1c) generates up-and-down $M(t)$ only when the stronger pathway is frequently selected to active the gene [29], which is incompatible with the robust up-and-down dynamics of $M(t)$, even if the TNF induction level is extremely low [31]. Taken together, the activation frameworks of a single pathway or parallel pathways alone are insufficient to capture the rich dynamics of $M(t)$ from mouse fibroblast genes (Table 1).

By combining the three-state model (Fig. 1b) and cross-talking pathways model (Fig. 1c), it is possible to generate new dynamic $M(t)$ features. The simplest combination is shown in Fig. 1d, which we call the cross-talking three-state model. This model can be viewed as adding a parallel pathway in the three-state model or decomposing a pathway of the

Mathematical models	$M(t)$ dynamical profiles	$\iff$	Parameter regions
(a) Two-state model	Increase		All parameters
(b) Three-state model	Increase		$\alpha^2 \geq \xi$
	Almost increase ^a		$\alpha^2 < \xi$
(c) Cross-talking pathways model	Increase		$\Lambda \geq \min\{\delta, \gamma\}$
	Up-and-down		$\Lambda < \min\{\delta, \gamma\}$
(d) Multiple pathways model	Increase		$x_1 \geq \min\{\delta, \alpha_1\}$
	Up-and-down		$x_1 < \min\{\delta, \alpha_1\}$

Table 1: Bijection theory between $M(t)$ dynamical features and system parameter regions for the (a) two-state [29], (b) three-state [33, 34], (c) cross-talking pathways [29] and (d) multiple pathways models [35]. $\alpha, \xi, \Lambda, \alpha_1$, and x_1 are the auxiliary numbers associated with the system parameters, and γ and δ are the gene inactivation rate and mRNA degradation rate, respectively. ^aThe three-state model generates a damped oscillatory $M(t)$ when $\alpha^2 < \xi$ [33, 34]. However, such oscillation decays exponentially and displays visually increasing dynamics [34].

cross-talking pathways model into two sequential steps, as shown in the following scheme:

We assumed that the gene is activated by two competitive pathways. These are the weak basal pathway, which has a selection probability q_1 and consists of two sequential rate-limiting steps with strength rates κ_1 and κ_2 , or alternatively the strong rate-limiting signaling pathway, which has a selection probability q_2 and strength rate λ , satisfying:

$$0 < q_1, q_2 < 1, \quad q_1 + q_2 = 1, \quad \text{and} \quad 0 < \kappa_1, \kappa_2 < \lambda < \infty.$$

The basal pathway is regulated independently by a spontaneous mechanism to maintain basal transcription levels under normal cellular growth conditions [36, 37]. The assumption of two sequential steps and small strength rates for the basal pathway is in close agreement with the real-time imaging data of the *off* period of 16 mouse fibroblast genes [8]. Moreover, if κ_1 or κ_2 is relatively large, then the basal pathway can be mathematically approximated by a single rate-limiting step [33], and the framework (1) reduces to the cross-talking pathways model (Fig. 1c). A stronger signaling pathway is triggered when cells receive external cues,

and downstream transcription factors (TFs) are activated by special signal transduction pathways to upregulate gene transcription [31, 38]. For each target gene, its activation is ultimately mediated through the binding of downstream TFs in the basal or signaling pathways at the cognate DNA sites in the gene promoter or enhancer domains [5, 38]. The selection probabilities q_1 and q_2 may then quantify the concentration and availability of activated TFs in each pathway to competitively form TF/DNA binding configurations, while the inducible activation rate λ of the signaling pathway quantifies the binding accessibility and strength between the corresponding TFs and DNA sites [5, 17, 35, 38].

2.2 Dynamics of $M(t)$ and the bijection with parameter regions

To establish the bijection between the $M(t)$ profiles and the parameter regions of the model (1), we first need to calculate the exact forms of $M(t)$ in terms of system parameters. At time $t \geq 0$, let random variable $X(t) = X = o_{11}, o_{12}, o_2, e$, specify the states of gene *off* 1 for basal pathway, gene *off* 1 for signaling pathway, gene *off* 2, and gene *on*, respectively. Then define:

$$P_{m,X}(t) = \text{Prob}\{\text{the system is residing at state } X \text{ with } m \text{ mRNA molecules at time } t\},$$

and the mass function:

$$P_m(t) = P_{m,o_{11}}(t) + P_{m,o_{12}}(t) + P_{m,o_2}(t) + P_{m,e}(t), \quad m = 0, 1, \dots$$

that quantifies the probability of m mRNA transcripts at time t in a single cell.

Following the standard procedure, we can obtain an infinite array of master equations with respect to $P_{m,X}(t)$ [15, 33, 39]:

$$P'_{m,o_{11}}(t) = -(\kappa_1 + m\delta)P_{m,o_{11}}(t) + (m+1)\delta P_{m+1,o_{11}}(t) + q_1\gamma P_{m,e}(t), \quad (2)$$

$$P'_{m,o_{12}}(t) = -(\lambda + m\delta)P_{m,o_{12}}(t) + (m+1)\delta P_{m+1,o_{12}}(t) + q_2\gamma P_{m,e}(t), \quad (3)$$

$$P'_{m,o_2}(t) = -(\kappa_2 + m\delta)P_{m,o_2}(t) + (m+1)\delta P_{m+1,o_2}(t) + \kappa_1 P_{m,o_{11}}(t), \quad (4)$$

$$P'_{m,e}(t) = -(\gamma + v + m\delta)P_{m,e}(t) + (m+1)\delta P_{m+1,e}(t) + vP_{m-1,e}(t) + \kappa_2 P_{m,o_2}(t) + \lambda P_{m,o_{12}}(t). \quad (5)$$

Summing up (2)-(5) gives the master equation of $P_m(t)$:

$$P'_m(t) = (m+1)\delta P_{m+1}(t) - vP_{m,e}(t) - m\delta P_m(t) + vP_{m-1,e}(t). \quad (6)$$

We did not focus on solving (2)-(6), which are beyond the scope of current mathematical methods within all parameter regions [22–26]. However, the master equations set a basis for calculating analytical forms for the gene state probabilities $P_{o_{11}}, P_{o_{12}}, P_{o_2}, P_e$, and mean transcript level $M(t)$, defined as:

$$P_X(t) = \sum_{m=0}^{\infty} P_{m,X}(t), \quad X = o_{11}, o_{12}, o_2, o_e, \quad \text{and} \quad M(t) = \sum_{m=0}^{\infty} m P_m(t).$$

Using these definitions and (2)-(6), we derived the equations for the four state probabilities and the mean transcription level:

$$\begin{cases} P'_{o_{11}}(t) = -\kappa_1 P_{o_{11}}(t) + q_1 \gamma P_e(t), & P'_{o_{12}}(t) = -\lambda P_{o_{12}}(t) + q_2 \gamma P_e(t), \\ P'_{o_2}(t) = \kappa_1 P_{o_{11}}(t) - \kappa_2 P_{o_2}(t), & P'_e(t) = \lambda P_{o_2}(t) + \kappa_2 P_{o_2}(t) - \gamma P_e(t), \\ M'(t) = v P_e(t) - \delta M(t). \end{cases} \quad (7)$$

Because the system must reside on exactly one gene state at any time, we set an arbitrary initial condition for (7):

$$P_{o_{11}}(0) + P_{o_{12}}(0) + P_{o_2}(0) + P_e(0) = 1 \quad \text{and} \quad M(0) \geq 0. \quad (8)$$

We utilized the Laplace transform method to solve the first-order differential system (7)-(8) (*Materials and Methods*). We defined a polynomial function as:

$$\begin{aligned} h(x) = & [(\kappa_1 \lambda + \kappa_2 \lambda + \kappa_1 \kappa_2) P_e(0) + \kappa_1 \kappa_2 P_{o_{11}}(0) + \lambda(\kappa_1 + \kappa_2) P_{o_{12}}(0) + \kappa_2(\lambda + \kappa_1) P_{o_2}(0)]x \\ & + P_e(0)x^3 + [(\kappa_1 + \lambda + \kappa_2) P_e(0) + \lambda P_{o_{12}}(0) + \kappa_2 P_{o_2}(0)]x^2 + \kappa_1 \kappa_2 \lambda, \end{aligned}$$

and two auxiliary numbers c_1 and c_2 in terms of the system parameters (*Materials and Methods* (18)-(19)). It can be verified that zero is a simple eigenvalue of the coefficient matrix for the system of the first four equations in (7). The other non-zero eigenvalues a_1, a_2 , and a_3 are calculated in terms of c_1, c_2 , and system parameters (*Materials and Methods* (15)-(19)). Under the arbitrary initial condition (8), the average mRNA level $M(t)$ is found to be as follows:

(1) If $c_1^2 < c_2$, then a_1, a_2, a_3 are real numbers with $0 < a_1 < a_2 < a_3$ (*Materials and*

Methods (22)), and $M(t)$ takes the form of:

$$M(t) = \frac{v \kappa_1 \kappa_2 \lambda}{\delta a_1 a_2 a_3} - \frac{vh(-a_1)e^{-a_1 t}}{a_1(a_2 - a_1)(a_3 - a_1)(\delta - a_1)} + \frac{vh(-a_2)e^{-a_2 t}}{a_2(a_2 - a_1)(a_3 - a_2)(\delta - a_2)} - \frac{vh(-a_3)e^{-a_3 t}}{a_3(a_3 - a_1)(a_3 - a_2)(\delta - a_3)} + \frac{vh(-\delta)e^{-\delta t}}{\delta(\delta - a_1)(\delta - a_2)(\delta - a_3)} + M(0)e^{-\delta t}. \quad (9)$$

(2) If $c_1^2 = c_2$, then $a_1, a_2, a_3 > 0$ are real numbers with $a_1 \neq a_2 = a_3$ (*Materials and Methods* (24)), and $M(t)$ takes the form of:

$$M(t) = \frac{v \kappa_1 \kappa_2 \lambda}{\delta a_1 a_2^2} (1 - e^{-a_2 t}) - \frac{vh(-a_1)(e^{-a_1 t} - e^{-a_2 t})}{a_1(\delta - a_1)(a_2 - a_1)^2} + \frac{vh(-a_2)te^{-a_2 t}}{a_2(a_2 - a_1)(\delta - a_2)} + \frac{vh(-\delta)(e^{-\delta t} - e^{-a_2 t})}{\delta(\delta - a_1)(\delta - a_2)^2} + M(0)e^{-\delta t}. \quad (10)$$

(3) If $c_1^2 > c_2$, then $a_1 > 0$ is a real number, whereas a_2 and a_3 are conjugate complexes, then let $a_r = \text{Re}(a_2)$ and $a_i = \text{Im}(a_2)$ (*Materials and Methods* (26)). Then, $M(t)$ takes the form of:

$$M(t) = \frac{v \kappa_1 \kappa_2 \lambda}{\delta a_1(a_r^2 + a_i^2)} - \frac{vh(-a_1)e^{-a_1 t}}{a_1(\delta - a_1)[(a_1 - a_r)^2 + a_i^2]} + \frac{vh(-\delta)e^{-\delta t}}{\delta(\delta - a_1)[(\delta - a_r)^2 + a_i^2]} + \bar{A}\sqrt{1 + A^2} \cos(a_i t + \theta) e^{-a_r t} + M(0)e^{-\delta t}, \quad (11)$$

where the constants $\bar{A}$, A , and θ correlate with the parameters and initial conditions (*Materials and Methods* (28)).

We characterized the dynamic profiles of $M(t)$ with almost no expression products at the initial time t [11, 21, 30]. We assumed that transcription starts from the gene *off* 1 state and counts only the newly produced mRNA molecules. This gives the following initial values:

$$(P_{o_{11}}(0), P_{o_{12}}(0), P_{o_2}(0), P_e(0), M(0)) = (q_1, q_2, 0, 0, 0). \quad (12)$$

We started with an interesting case, $c_1^2 > c_2$. Then, $M(t)$ was expressed in exact form (11), which contains a cosine function, and suggested possible oscillatory dynamics of $M(t)$. However, the coefficient of the cosine function damps exponentially, which dramatically weakened the oscillation visually, as manifested by our numerical examples and observations from the three-state model [34]. The exact form (11) can only be viewed as a steady-state value, adding three exponential functions at most. If the exact form of $M(t)$ with $M(0) = 0$ has such a structure, then it can be proved that the most complex dynamics of $M(t)$ will

only take a unique peak such as up-and-down or up-down-up profiles (*Materials and Methods* Theorem 1).

For the other case of $c_1^2 \leq c_2$, the eigenvalues a_1, a_2 , and a_3 are real, and the exact forms (9)-(10) of $M(t)$ do not contain oscillatory functions but contain multiple exponential functions. The case $c_1^2 = c_2$ rarely occurs in biology and can be viewed mathematically as a limiting case of $c_1^2 < c_2$. For $c_1^2 < c_2$, a rigorous statement of the $M(t)$ profiles is inevitably technical. Let x_1 and x_2 denote the two roots of:

$$H(x) = vq_2\lambda x^2 + v(q_1\kappa_1\kappa_2 + q_2\lambda\kappa_1 + q_2\lambda\kappa_2)x + v\kappa_1\kappa_2\lambda.$$

When both x_1 and x_2 are complex numbers, there are a total of four parameter correlations, and we showed that $M(t)$, expressed by (9), either increases monotonically for all $t > 0$, or develops an up-down-up profile. If x_1 and x_2 are real values, we can classify all 60 correlations among x_1, x_2, a_1, a_2, a_3 , and δ into three categories that correspond to three distinct dynamical behaviors: the increasing, up-and-down, and up-down-up profiles of $M(t)$. We illustrated detailed mathematical results and their proof in Theorem 2 of the *Materials and Methods*. In summary, for the cross-talking three-state model (1), even if $M(t)$ is expressed in different exact forms (9)-(11) and influenced by various parameter correlations, $M(t)$ can exhibit, but exhibits only three features: increasing, up-and-down, and up-down-up dynamics.

2.3 Fitting dynamical transcription data of mouse fibroblast genes

We demonstrated three dynamical profiles of $M(t)$ generated by model (1). These distinct behaviors are in agreement with the observed dynamic trends of average mRNA levels in mouse fibroblast genes in response to TNF [30, 31]. For instance, Hao and Baltimore [30] divided 180 activated mouse fibroblast genes under TNF induction into three groups, separately characterized by the short, median, and long half-lives of the transcripts. As shown in Fig. 2, they found that group I genes responded quickly by forming a sharp dynamical peak at average transcription levels, group II genes did not respond quickly but most still formed a gentle transcription peak along the timeline, and group III mRNAs accumulated rather slowly and gradually increased in abundance during the observation window. The transcription data from the 12 representative mouse fibroblast genes shown in Fig. 2 contains four genes displaying the up-down-up trend of transcription dynamics (*Edn1*, *Cxcl1*, *Ccl2*, *Icam1*), as well as others displaying either dynamical up-and-down or monotonic increasing transcription. Using exact forms (9)-(11) of $M(t)$, the model (1) provides good theoretical

fits to all 12 datasets (coefficient of determination $R^2 > 0.9$).

Figure 2: Fit of transcription data by the cross-talking three-state model (1). Black circles represent the dynamical data of mRNA average levels for 12 mouse fibroblast genes under TNF induction [30]. The genes are divided into groups I, II, and III, which are separately characterized by the short, median, and long half-lives of the transcripts [30]. The red lines are generated using exact forms (9)-(11) of the model (1), which provide a good fit to the data points ($R^2 > 0.9$). The fitted system parameters are listed in Table S1 (*Supporting Information*).

When fitting the data shown in Fig. 2, we assigned extremely small values to the activation rates κ_1 and κ_2 of weak basal pathway, and restricted the mRNA degradation rate δ of each gene within the δ region of the gene group to which they belong [30] (*Supporting Information*, Table S1). Therefore, the fitted δ values show a significant negative correlation with gene groups I, II, and III because the gene groups themselves are classified by their transcript half-lives (Fig. 3a). Intriguingly, the freely fitted inactivation rate γ and activation rate λ of the signaling pathway also exhibited negative correlation with gene groups (Fig. 3a). This observation suggests that the simultaneous large λ and γ may separately help enhance the height of the peak and suppress the stationary values of $M(t)$ in group I. In contrast, the

small λ and γ implemented contrary functions that destroy the dynamical peak of $M(t)$ and lift up the stationary mRNA numbers in group III. However, the probability q_2 of the signaling pathway does not clearly correlate with the gene groups but varies for different genes (Fig. 3a). Note that the dynamics of $M(t)$ in the three gene groups are mainly discriminated by its first peak. Our observations suggest that, once cells receive external cues, the frequency of the signaling pathway directing gene activation may not play a crucial role in regulating the dynamical peak of transcription level.

Figure 3: Regulation of transcription dynamics by system parameters. (a) The fitted parameters for 12 mouse fibroblast genes are listed in units of gene groups (*Supporting Information*, Table S1). Different gene groups exhibit distinct temporal transcription modes [30]. The values of mRNA degradation rate δ , inactivation rate γ , and activation rate λ of the signaling pathway are all negatively correlated with gene groups I, II, and III. The probability q_2 of the signaling pathway does not follow a clear correlation with the gene groups. (b) Based on the fitted parameter set for the *Cxcl1* gene, increasing δ , γ , and λ transit transcription dynamics from monotonic to up-down-up mode. An increase in q_2 generates switches among multiple transcription dynamical modes, where the non-monotonic modes are displayed within most of the q_2 variation region (0, 1).

To further understand the regulation of $M(t)$ profiles, we varied the parameters δ, γ, λ , and q_2 under the fitted parameter sets of all 12 genes in Fig. 2. This procedure reveals a uniform regulation mode for each system parameter. As shown in Fig. 3b for *Cxcl1* gene, the variation in δ behaves as a bilateral switch to regulate $M(t)$: there is a threshold value such that $M(t)$ increases monotonically while δ stays below the threshold, but switches to a non-monotonic profile once δ exceeds the threshold. Such bilateral regulation of δ has been observed to play an important role in controlling the temporal transcription mode in mouse fibroblasts [30, 31]. In addition, both γ and λ play the same bilateral roles in regulating $M(t)$ dynamics (Fig. 3b), reinforcing previous observation of smaller δ, γ, λ for group III genes that generate increasing $M(t)$ with larger δ, γ, λ for groups I and II genes that exhibit non-monotonic transcription dynamics (Fig. 3a). The regulation scenario of q_2 is different because it generates multiple switches among distinct $M(t)$ profiles, when q_2 increases from 0 to 1 (Fig. 3b). Exceptions need to be made for cases where q_2 approaches 0 or 1, which generates increasing transcription dynamics, because the dynamical peak of $M(t)$ seems to be robust in almost all the variation regions of $q_2 \in (0, 1)$. Note that q_2 is closely related to signal strength. Our observations fit with the ubiquitous transcription dynamical peak of mouse fibroblast genes under TNF induction from the lowest to the highest levels [31].

2.4 Cross-talking n -state model for oscillatory transcription dynamics

Our bijection theory shows that the cross-talking three-state model (1) cannot generate multiple dynamical peaks of $M(t)$. Therefore, the model (1) can be ruled out when $M(t)$ exhibits oscillation [30, 32]. We focused on transcription with constant kinetic rates under stable inductions to avoid complicated $M(t)$ dynamics regulated by time-dependent rates in response to time-varying signals [40]. The bijection theory (Table 1) shows that multiple parallel pathways induce at most one dynamical peak, and therefore introducing more parallel pathways in model (1) may not capture oscillatory $M(t)$. We then considered decomposing the basal pathway of model (1) into multiple sequential steps, as shown in Fig. 4a for the cross-talking n -state model. Calculation of $M(t)$ can follow the same procedure for model (1), which relies on solving a system of differential equations for which the coefficient matrix may have multiple pairs of eigenvalues expressed by conjugate complexes (*Materials and Methods*). According to the classical theory of ordinary differential equations, the exact forms of $M(t)$ may contain multiple periodic cosine functions, and it is therefore plausible to visualize oscillatory dynamics.

To test $M(t)$ oscillation induced by the sequential multi-step gene activation, we com-

Figure 4: Framework for generating oscillatory transcription dynamics. (a) The cross-talking n -state model for which the gene is activated either by the strong signaling pathway with probability q_2 , or alternatively by the weak basal pathway that consists of $n - 1$ rate-limiting steps with probability $q_1 = 1 - q_2$. (b) When gene activation is directed by a single multi-step pathway ($q_2 = 0$), the increase in the step number prolongs the initial response lag and enhances the damped oscillation of the transcription level $M(t)$. (c) When the gene is activated by the cross-talk between pathways ($q_2 > 0$), the increase in the step number in the basal pathway has almost no impact on the initial quick up-and-down dynamics of $M(t)$, but significantly enhances the following damped oscillatory behavior. (d) The curve of $M(t)$ (red lines) generated by the framework in (a) captures the oscillatory trend of transcription data (black circles) for the mouse fibroblast *Nfkb1a* gene under TNF induction [30]. The parameters in (b) and (c) are the fitted rates of the *Cxcl1* gene in Table S1 with $\kappa_1 = \dots = \kappa_{n-1}$ to maintain a constant T_{off} . Parameters in (d) are listed in Table S1.

pared the $M(t)$ profiles for different step numbers. This procedure requires that all comparisons are restricted to a constant average duration T_{off} of the gene *off* state [15, 34]. For the cross-talk n -state model (Fig. 4a), T_{off} is given by:

$$T_{off} = q_1 \sum_{i=1}^{n-1} \frac{1}{\kappa_i} + \frac{q_2}{\lambda}, \quad n \geq 2.$$

To guarantee the unchanged T_{off} , we set three parameter scaling conditions where activation rates $\kappa_1, \dots, \kappa_{n-1}$ are separately scaled identically [14, 41], differently [8, 14], and alterna-

tively [14]:

$$\begin{aligned} \text{(C1)} : \kappa_1 &= \kappa_2 = \dots = \kappa_{n-1}, \text{ with } \kappa_1 = \frac{q_1(n-1)}{T_{\text{off}} - q_2/\lambda}; \\ \text{(C2)} : \kappa_2 &= 2\kappa_1, \kappa_3 = 3\kappa_1, \dots, \kappa_{n-1} = (n-1)\kappa_1, \text{ with } \kappa_1 = q_1 \frac{1 + \frac{1}{2} + \dots + \frac{1}{n-1}}{T_{\text{off}} - q_2/\lambda}; \\ \text{(C3)} : \kappa_1 &= \kappa_3 = \dots, \text{ and } \kappa_2 = \kappa_4 = \dots = 3\kappa_1, \text{ with } \kappa_1 = \begin{cases} \frac{2q_1(n-1)}{3(T_{\text{off}} - q_2/\lambda)}, & n \text{ is odd,} \\ \frac{q_1(2n-1)}{3(T_{\text{off}} - q_2/\lambda)}, & n \text{ is even.} \end{cases} \end{aligned}$$

We initially examined the case of $q_2 = 0$, for which gene activation is directed by a single multi-step pathway [14,15,26]. Under parameter scaling condition (C1), we generated several $M(t)$ curves under different activation step numbers, using numerical simulations from the corresponding system of differential equations (*Materials and Methods* (43)). As shown in Fig. 4b under the fitted parameters of the *Cxcl1* gene in Fig. 2, multi-step gene activation for large step numbers triggers significant damped oscillatory $M(t)$, where the significance of the oscillation is positively correlated with the step number. However, the system displays lag times of more than 8 h to reach the first peak of $M(t)$. This slow transcription response contradicts the rapid peak of $M(t)$ within $0.5 \sim 2$ h for mouse fibroblast genes [30, 31]. Moreover, the damped oscillation and response lag of $M(t)$ were robust against the parameter scaling conditions (C2) and (C3) (*Supporting Information*, Fig. S1). Taken together, the large number of sequential gene activation steps facilitates the oscillatory dynamics of $M(t)$, while a single multi-step pathway was not able to induce quickly-peaked transcription dynamics of mouse fibroblast genes.

We then quantified $q_2 \neq 0$ to introduce the cross-talking regulation of pathways on $M(t)$ dynamics. We generated $M(t)$ curves (*Materials and Methods* (42)) under the fitted parameters of the *Cxcl1* gene in Fig. 2 and conditions (C1)-(C3). As shown in Fig. 4c and Fig. S2 in *Supporting Information*, the system generates oscillatory $M(t)$ with two major features. Firstly, $M(t)$ displays a quick and sharp first peak within the initial time region, while the height and sharpness of the first peak are nearly independent of the step number and parameter scaling conditions of the basal pathway. Secondly, $M(t)$ displays a damped and gentle second and following peaks that are tightly correlated with the step number and parameter conditions, similarly to the oscillation of $M(t)$ induced by a single multi-step pathway (Fig. 4b). These two features capture multiple dynamical peaks for the transcription of the mouse fibroblast *Nfkb1a* gene [30] (Fig. 4d). Taken together, the cross-talk between signaling and basal pathways plays a dominant role in generating the first quick up-and-down transcription dynamics, while the subsequent gentle and damped transcription oscillation is

induced by the multi-step regulation in the weak basal pathway.

2.5 Cross-verification for monotonic transcription dynamics

Our method requires rich $M(t)$ data dynamics to rule out frameworks that cannot sufficiently capture the dynamic features. When $M(t)$ displays simple monotonicity, all frameworks (Fig. 1) may provide good fits to the $M(t)$ data. However, we cannot claim that the two-state model is the best model to explain the data, even if it is the simplest compared to the other frameworks. This is because monotonic $M(t)$ may be generated by the manipulated small degradation rate δ (Fig. 3b) for accurately counting transcript numbers in single-cell measurements [11, 14]. One way to avoid the masking effect of small δ on $M(t)$ -rich dynamics is to consider the average level of nascent RNA that is independent of mRNA degradation [4, 42]. In contrast, some single-cell data for transcription of mRNA molecules display smooth trend lines along the timeline, such as the noise $CV^2(t)$, Fano factor $\phi(t)$, and probability $P_0(t)$ of the gene producing zero transcripts. These indexes may be used in conjunction with $M(t)$ to cross-verify the gene activation frameworks.

The calculation of $\phi(t)$ and $CV^2(t)$ are standard, and require solving the second moment $\mu_2(t)$ of the system of ordinary differential equations derived from the corresponding master equations [33, 39]. We calculated the exact forms of $\mu_2(t)$ for the cross-talking three-state model (1) (*Materials and Methods* (48)), and $\phi(t) = \mu_2(t)/M(t) - M(t)$ and $CV^2(t) = \phi(t)/M(t)$ were then readily obtained. The probability $P_0(t)$ is one of the solutions of master equations. Calculation of $P_0(t)$ involves introducing the generating function $V(z, t)$ to transform the master equation into a system of partial differential equations for $V(z, t)$. However, while solving $V(z, t)$ is somewhat difficult, we can still express $V(z, t)$ in closed forms of hypergeometric functions for different models [23, 26]. Consequently, $P_0(t)$ can be obtained by $P_0(t) = V(-1, t)$. The expressions of $\phi(t)$, $CV^2(t)$, and $P_0(t)$ are not as neat as that of $M(t)$, which prevents us from establishing theoretical bijections between their dynamical profiles and parameter regions. Nevertheless, the computation of $\phi(t)$, $CV^2(t)$, and $P_0(t)$ through their expressions or the corresponding differential equations (*Materials and Methods*) can generate dynamical curves for fitting data [11, 40].

We hypothesized that different gene activation frameworks can generate distinct dynamics of $\phi(t)$, $CV^2(t)$, and $P_0(t)$ under the same monotonic $M(t)$. To verify this, we first fitted a group of increasing $M(t)$ data to all four frameworks in Fig. 1 (see Fig. 5a). Under the fitted parameters we separately generated dynamical curves of $\phi(t)$, $CV^2(t)$, and $P_0(t)$ corresponding to each framework. The curves of noise $CV^2(t)$ clustered within a low-value region, probably because of their large denominator $M(t)$ (Fig. 5b), suggesting that the noise data

may not help discriminate gene activation frameworks. The curves of the Fano factor $\phi(t)$ are clearly separated into two categories: the curves are very high under cross-talking pathways, but are relatively low under a single pathway (Fig. 5c). This separation reinforces the conclusion that the parallel pathways enhance ϕ [39], while sequential steps suppress ϕ [15] under the same mean level M at steady-state. The curves of $P_0(t)$ follow the same separation as that of $\phi(t)$ (Fig. 5d), which reinforces the observed suppressed $P_0(t)$ under multi-step gene activation [26]. Taken together, the dynamics of $\phi(t)$ and $P_0(t)$ may help discriminate the activation frameworks regulated by cross-talking pathways or a single pathway.

Figure 5: Discrimination of gene activation frameworks under monotonic dynamical mRNA average level. (a) The increasing transcription mean data (black circles, *Mmp13* gene [30]) are fitted by different models. (b) Dynamical curves of transcription noise for each model are clustered with relatively small values most of the time. For (c) transcription Fano factor and (d) probability of genes not producing any transcript, the dynamical curves generated by a single pathway deviate from the curves generated by cross-talking pathways. Parameters: $v = 2630 \text{ h}^{-1}$, $q_2 = 0.5$; two-state model, $(\lambda, \gamma, \delta) = (2.21, 4, 0.5) \text{ h}^{-1}$; three-state model, $(\kappa_1, \kappa_2, \gamma, \delta) = (5.09, 5.09, 4, 0.55) \text{ h}^{-1}$; cross-talking pathways model, $(\kappa, \lambda, \gamma, \delta) = (0.07, 1.2, 1.2, 0.15) \text{ h}^{-1}$; cross-talking three-state model, $(\kappa_1, \kappa_2, \lambda, \gamma, \delta) = (0.24, 0.1, 1.2, 1, 0.15) \text{ h}^{-1}$.

3 Conclusion and Discussion

The classical two-state model used in single-cell studies posits that a gene will randomly transition between *on* (active) and *off* (inactive) states, with mRNA molecules being produced only when the gene is *on* (Fig. 1a). Compared to the universal feature of a single rate-limiting step turning gene *off* [6, 8, 13], the process of turning a gene *on* is usually non-Markovian, and is influenced by multiple rate-limiting fluctuations [6, 7, 13]. The framework of gene activation has been modeled by listing rate-limiting steps sequentially [8, 14] (Fig. 1b), parallelly [16, 35] (Fig. 1c), or in the form of their combinations [17] (Fig. 1d). Recent studies have facilitated efficient computation of downstream transcription distribution under non-Markovian gene activation [18, 19]. However, the best method for mapping the

transcription distribution data back to the exact picture of the activation framework for the gene of interest remains elusive.

Distinct gene activation frameworks cannot be discriminated by the steady-state single-cell data on mRNA distribution or its noise, Fano factor, and mean, because the limited steady-state diagrams can be reliably illustrated by different frameworks [15, 23, 27–29]. When the time dimension is included, the distribution histograms of both gene *off* periods and mRNA copy numbers provide a complete static basis to computationally search the optimal number of rate-limiting steps in gene activation and their kinetic parameters [9, 14, 20]. However, these methods require a prior hypothetical framework of gene activation that cannot be easily deduced from limited gene *off* distribution modes [15–17] or from similar dynamical transition patterns of mRNA distribution profiles [24–26]. Therefore, an efficient method for determining confident gene activation frameworks is urgently required prior to any attempt to fit single-cell dynamical transcription data.

In this paper, we demonstrated that the dynamics of transcript average level $M(t)$ can serve as a competent candidate to facilitate the efficient estimation of the gene activation framework and system parameters. Firstly, compared with time-consuming single-cell measurements [11, 21], the dynamic features of $M(t)$ for a large number of genes can be captured relatively quickly by traditional cell population methods. Secondly, compared to the multiple uneven distribution profiles at discrete time points [9, 20], a single smooth curve of $M(t)$ along the timeline presents easily discriminated dynamic features and provides a more efficient theoretical fit to the data. Thirdly, compared to the lack of a method to calculate dynamical mRNA distribution [22, 23], the exact forms of $M(t)$ are rather neat and can be achieved by the standard theory of ordinary differential equations.

By taking advantage of the $M(t)$ expressions and their simple forms (Eqs.(9)-(11)), we were able to establish theoretical bijections between the $M(t)$ dynamics and parameter regions for different gene activation frameworks. Subsequently, frameworks that cannot capture the exhibited $M(t)$ dynamical features can be ruled out, while the optimal forms of the other potential frameworks are further determined by fitting $M(t)$ data. We illustrated this idea by analyzing activation frameworks for a large number of TNF-induced mouse fibroblast genes. These genes display rich transcription dynamics that can be categorized into three main features: increasing, up-and-down, and up-down-up profiles of $M(t)$ [30, 31]. Our bijection theories (Table 1 and *Materials and Methods* Theorems 1-2) show that these three distinct dynamics cannot be achieved by the sequential or parallel rate-limiting steps alone, but can be captured by the simplest form of combined sequential and parallel steps (Fig. 1d). We call this framework the cross-talking three-state model, as depicted in (1), for which the gene is activated either by the weak basal pathway consisting of two sequential

steps, or alternatively by the strong signaling pathway.

The cross-talking three-state model (1) can serve as a reliable model because it provides a good theoretical fit for all the representative datasets on the transcription dynamics of mouse fibroblast genes [30] (Fig. 2). Moreover, analysis of the freely fitted system parameters (Table S1) reveals two regulation scenarios for $M(t)$ dynamics (Fig. 3). For the mRNA degradation rate δ , inactivation rate γ , and activation rate λ of the signaling pathway, $M(t)$ maintains a monotonic increase when δ, γ , and λ remain below their thresholds while switching non-monotonically once δ, γ , and λ exceed the threshold values (Fig. 3b). This bilateral regulation results in an interesting phenomenon: δ, γ , and λ are simultaneously large for the genes displaying non-monotonic $M(t)$, whereas they are simultaneously small for genes displaying monotonic $M(t)$ (Fig. 3a). In contrast to the gene-specific δ, γ , and λ , the environment influenced the probability q_2 that the signaling pathway does not play a crucial role in regulating the first peak of $M(t)$ (Fig. 3). This observation is consistent with the robust transcription dynamical peak of mouse fibroblast genes, regardless of how the TNF induction level varies [31].

We note that a small number of mouse fibroblast genes display damped transcription oscillation, with the first peak forming rapidly within the initial period [30, 31]. We suggest that multiple sequential rate-limiting steps may play a role in triggering oscillatory behavior. Therefore, we developed the cross-talking n -state model by decomposing the basal pathway of model (1) into multiple steps (Fig. 4a). We first ruled out the framework of a single multi-step pathway as it triggers a long lag reaching the first peak of $M(t)$ (Fig. 4b, Fig. S1), which contradicts the observed rapid peak of $M(t)$. However, when we recovered the cross-talk between the signaling pathway and the multi-step basal pathway, the initial lag time disappeared and the first peak of $M(t)$ formed quickly (Fig. 4c, Fig. S2). Intriguingly, the step number and parameter scaling condition in the basal pathway have almost no effect on the first peak of $M(t)$, whereas they significantly influence the other dynamical peaks. Together with the transcription data of the mouse fibroblast genes (Fig. 4d), we found that cross-talking regulation between pathways is crucial to trigger the first rapid, sharp peak of $M(t)$, while the multi-step regulation facilitates the following damped and gentle oscillatory dynamics.

Thus, we developed a procedure to estimate frameworks of gene activation using transcription level $M(t)$ dynamical data at the cell population level. This procedure readily disqualifies frameworks that do not perform satisfactorily, and determines potential frameworks that provide the best data fit. When more sophisticated single-cell distribution data of the gene *off* period [8, 14] or mRNA copy numbers [9, 20, 21] are used, our procedure provides a prior estimation of activation frameworks and parameter rates to facilitate more compu-

tationally efficient search for optimal step numbers and parameters. This procedure relies on the rich dynamics of $M(t)$. When $M(t)$ behaves monotonically, the other transcription indexes are required to cross-verify the activation frameworks. For instance, the dynamical Fano factor or the probability of no transcript being produced may help discriminate frameworks regulated by a single pathway or cross-talking pathways (Fig. 5).

Future work may utilize additional data for different genes and transcription indexes to test and develop our procedure. In addition, we anticipate the inclusion of nascent RNA data that are not influenced by the mRNA degradation rate and may therefore provide more direct information on gene activation frameworks [4,42]. Finally, a cell cycle description [22,41,43] may be introduced into gene activation frameworks to eliminate estimation errors caused by disregarding cell cycle stochasticity [41].

Materials and Methods

Exact forms of $M(t)$ for cross-talking three-state model (1)

To calculate mRNA average level $M(t)$ for the model (1), we need solve the first four differential equations in system (7) for gene state probabilities P_{o11} , P_{o12} , P_{o2} and P_e under the arbitrary initial condition (8):

$$\begin{cases} P'_{o11}(t) = -\kappa_1 P_{o11}(t) + q_1 \gamma P_e(t), & P'_{o12}(t) = -\lambda P_{o12}(t) + q_2 \gamma P_e(t), \\ P'_{o2}(t) = \kappa_1 P_{o11}(t) - \kappa_2 P_{o2}(t), & P'_e(t) = \lambda P_{o12}(t) + \kappa_2 P_{o2}(t) - \gamma P_e(t), \\ P_{o11}(0) + P_{o12}(0) + P_{o2}(0) + P_e(0) = 1. \end{cases} \quad (13)$$

To solve the above first-order differential system, we denote by

$$\begin{cases} L_{1j}(x) = \mathcal{L}(P_{o1j}(t)) = \int_0^\infty e^{-xt} P_{o1j}(t) dt, & j = 1, 2, \\ L_2(x) = \mathcal{L}(P_{o2}(t)) = \int_0^\infty e^{-xt} P_{o2}(t) dt & \text{and} & L_e(x) = \mathcal{L}(P_e(t)) = \int_0^\infty e^{-xt} P_e(t) dt. \end{cases}$$

Then using Laplace transform in the theory of ordinary differential equation to system (13) yields

$$\begin{cases} P_{o11}(0) = (x + \kappa_1) L_{11}(x) - q_1 \gamma L_e(x), & P_{o12}(0) = (x + \lambda) L_{12}(x) - q_2 \gamma L_e(x), \\ P_{o2}(0) = -\kappa_1 L_{11}(x) + (x + \kappa_2) L_2(x), & P_e(0) = -\lambda L_{12}(x) - \kappa_2 L_2(x) + (x + \gamma) L_e(x). \end{cases} \quad (14)$$

The determinant of coefficient matrix for the above system can be directly calculated as

$$x(x + a_1)(x + a_2)(x + a_3) = x[x^3 + (\kappa_1 + \lambda + \kappa_2 + \gamma)x^2 + (\kappa_1\lambda + \kappa_1\kappa_2 + \kappa_2\lambda + \kappa_1\gamma + q_1\lambda\gamma + \kappa_2\gamma)x + \kappa_1\lambda\kappa_2 + q_1\kappa_1\lambda\gamma + q_2\kappa_1\kappa_2\gamma + q_1\kappa_2\lambda\gamma],$$

where a_1, a_2 and a_3 are non-zero eigenvalues of the coefficient matrix for the system (13). To calculate a_1, a_2 and a_3 , we let the above formula equal to 0, and then solve the cubic polynomial equation using Cardano formula. This gives

$$a_1 = \frac{\kappa_1 + \kappa_2 + \lambda + \gamma}{3} - \left(\sqrt[3]{-c_1 + \sqrt{c_1^2 - c_2}} + \sqrt[3]{-c_1 - \sqrt{c_1^2 - c_2}} \right), \quad (15)$$

$$a_2 = \frac{\kappa_1 + \kappa_2 + \lambda + \gamma}{3} - \left(\frac{-1 + \sqrt{3}i}{2} \sqrt[3]{-c_1 + \sqrt{c_1^2 - c_2}} - \frac{1 + \sqrt{3}i}{2} \sqrt[3]{-c_1 - \sqrt{c_1^2 - c_2}} \right), \quad (16)$$

$$a_3 = \frac{\kappa_1 + \kappa_2 + \lambda + \gamma}{3} - \left(\frac{-1 - \sqrt{3}i}{2} \sqrt[3]{-c_1 + \sqrt{c_1^2 - c_2}} - \frac{1 - \sqrt{3}i}{2} \sqrt[3]{-c_1 - \sqrt{c_1^2 - c_2}} \right), \quad (17)$$

where

$$c_1 = \frac{\kappa_1\kappa_2\lambda + q_1\kappa_1\lambda\gamma + q_2\kappa_1\kappa_2\gamma + q_1\kappa_2\lambda\gamma}{2} + \frac{(\kappa_1 + \kappa_2 + \lambda + \gamma)^3}{27} - \frac{(\kappa_1 + \kappa_2 + \lambda + \gamma)(\kappa_1\lambda + \kappa_1\kappa_2 + \lambda\kappa_2 + \kappa_1\gamma + q_1\lambda\gamma + \lambda\gamma)}{6} \neq 0, \quad (18)$$

$$c_2 = -\frac{1}{27} \left[(\kappa_1\lambda + \kappa_1\kappa_2 + \lambda\kappa_2 + \kappa_1\gamma + q_1\lambda\gamma + \lambda\gamma) - \frac{(\kappa_1 + \kappa_2 + \lambda + \gamma)^2}{3} \right]^3 \neq 0. \quad (19)$$

Applying Cramer's rule in the theory of linear algebra on the system (14) readily gives

$$\begin{aligned} L_e(x) &= \frac{h(x)}{x(x + a_1)(x + a_2)(x + a_3)}, \quad \text{with} \\ h(x) &= [(\kappa_1\lambda + \kappa_2\lambda + \kappa_1\kappa_2)P_e(0) + \lambda(\kappa_1 + \kappa_2)P_{o_{12}}(0) + \kappa_2(\lambda + \kappa_1)P_{o_2}(0) + \kappa_1\kappa_2P_{o_{11}}(0)]x \\ &\quad + P_e(0)x^3 + [(\kappa_1 + \lambda + \kappa_2)P_e(0) + \lambda P_{o_{12}}(0) + \kappa_2P_{o_2}(0)]x^2 + \kappa_1\kappa_2\lambda. \end{aligned} \quad (20)$$

To calculate analytical formula for mRNA average level $M(t)$, we first notice that it satisfies

$$M'(t) = -\delta M(t) + vP_e(t).$$

Then applying Laplace transform to $M(t)$ gives

$$xL_M(x) - M(0) = vL_e(x) - \delta L_M(x), \quad \text{with} \quad L_M(x) = \mathcal{L}(M(t)) = \int_0^\infty e^{-xt} M(t) dt.$$

The substitution of $L_e(x)$ in (20) further leads to

$$L_M(x) = \frac{vL_e(x) + M(0)}{x + \delta} = \frac{vh(x) + M(0)x(x + a_1)(x + a_2)(x + a_3)}{x(x + a_1)(x + a_2)(x + a_3)(x + \delta)}. \quad (21)$$

(1) If $c_1^2 < c_2$, then a_1, a_2 and a_3 given in (15)-(19) are real numbers satisfying $0 < a_1 < a_2 < a_3$. Also, the direct calculation can simplify a_1, a_2 and a_3 to be

$$\begin{cases} a_1 = \frac{\kappa_1 + \kappa_2 + \lambda + \gamma}{3} - 2\sqrt[6]{c_2} \cos \theta_0, & a_2 = \frac{\kappa_1 + \kappa_2 + \lambda + \gamma}{3} - 2\sqrt[6]{c_2} \cos(\theta_0 + 4\pi/3), \\ a_3 = \frac{\kappa_1 + \kappa_2 + \lambda + \gamma}{3} - 2\sqrt[6]{c_2} \cos(\theta_0 + 2\pi/3), & \text{where } \theta_0 = \frac{\arccos(c_1/\sqrt{c_2})}{3} \in (0, \pi/3). \end{cases} \quad (22)$$

We decompose $L_M(x)$ given by (21) into the sum of partial fractions as

$$L_M(x) = \frac{\alpha_0}{x} + \frac{\alpha_1}{x + a_1} + \frac{\alpha_2}{x + a_2} + \frac{\alpha_3}{x + a_3} + \frac{\alpha_\delta}{x + \delta}, \quad (23)$$

where the coefficients $\alpha_0, \alpha_1, \alpha_2, \alpha_3$ and α_δ satisfy

$$\begin{aligned} & \alpha_0(x + a_1)(x + a_2)(x + a_3)(x + \delta) + \alpha_1x(x + a_2)(x + a_3)(x + \delta) + \alpha_2x(x + a_1)(x + a_3)(x + \delta) \\ & + \alpha_3x(x + a_1)(x + a_2)(x + \delta) + \alpha_\delta x(x + a_1)(x + a_2)(x + a_3) \\ & = vh(x) + M(0)x(x + a_1)(x + a_2)(x + a_3). \end{aligned}$$

The separate substitution of $x = 0, -a_1, -a_2, -a_3, -\delta$ into above equation yields

$$\begin{aligned} \alpha_0 &= \frac{v}{\delta} \frac{\kappa_1 \kappa_2 \lambda}{a_1 a_2 a_3}, \quad \alpha_1 = -\frac{vh(-a_1)}{a_1(a_2 - a_1)(a_3 - a_1)(\delta - a_1)}, \quad \alpha_2 = -\frac{vh(-a_2)}{a_2(a_1 - a_2)(a_3 - a_2)(\delta - a_2)}, \\ \alpha_3 &= -\frac{vh(-a_3)}{a_3(a_1 - a_3)(a_2 - a_3)(\delta - a_3)}, \quad \text{and} \quad \alpha_\delta = -\frac{vh(-\delta)}{\delta(a_1 - \delta)(a_2 - \delta)(a_3 - \delta)} + M(0). \end{aligned}$$

Then the application of the inverse Laplace transform to $L_M(x)$ given in (23) readily leads to

$$M(t) = \alpha_0 + \alpha_1 e^{-a_1 t} + \alpha_2 e^{-a_2 t} + \alpha_3 e^{-a_3 t} + \alpha_\delta e^{-\delta t},$$

which is exactly the expression (9).

(2) If $c_1^2 = c_2$, then it can be verified that a_1, a_2 and a_3 given by (15)-(19) are positive real numbers taking the form of

$$a_1 = \frac{\kappa_1 + \kappa_2 + \lambda + \gamma}{3} + 2\sqrt[3]{c_1}, \quad \text{and} \quad a_2 = a_3 = \frac{\kappa_1 + \kappa_2 + \lambda + \gamma}{3} - \sqrt[3]{c_1}. \quad (24)$$

We decompose $L_M(x)$ given in (21) into the sum of partial fractions as

$$L_M(x) = \frac{\bar{\alpha}_0}{x} + \frac{\bar{\alpha}_1}{x + a_1} + \frac{\bar{\alpha}_\delta}{x + \delta} + \frac{\bar{\alpha}_2}{x + a_2} + \frac{\bar{\alpha}_3}{(x + a_2)^2}, \quad (25)$$

where the coefficients $\bar{\alpha}_0, \bar{\alpha}_1, \bar{\alpha}_2, \bar{\alpha}_3$ and $\bar{\alpha}_\delta$ satisfy

$$\begin{aligned} & \bar{\alpha}_0(x + a_1)(x + a_2)^2(x + \delta) + \bar{\alpha}_1x(x + a_2)^2(x + \delta) + \bar{\alpha}_2x(x + a_1)(x + a_2)(x + \delta) \\ & + \bar{\alpha}_3x(x + a_1)(x + \delta) + \bar{\alpha}_\delta x(x + a_1)(x + a_2)^2 = vh(x) + M(0)x(x + a_1)(x + a_2)(x + a_3). \end{aligned}$$

Then the separate substitution of $x = 0, -a_1, -a_2, -\delta$ into this equation yields

$$\begin{aligned} \bar{\alpha}_0 &= \frac{v}{\delta} \frac{\kappa_1 \kappa_2 \lambda}{a_1 a_2^2}, \quad \bar{\alpha}_1 = -\frac{vh(-a_1)}{a_1(\delta - a_1)(a_2 - a_1)^2}, \quad \bar{\alpha}_2 = M(0) - (\bar{\alpha}_0 + \bar{\alpha}_1 + \bar{\alpha}_\delta), \\ \bar{\alpha}_3 &= -\frac{vh(-a_2)}{a_2(a_1 - a_2)(\delta - a_2)}, \quad \text{and} \quad \bar{\alpha}_\delta = -\frac{vh(-\delta)}{\delta(a_1 - \delta)(a_2 - \delta)^2} + M(0). \end{aligned}$$

Then the application of the inverse Laplace transform to $L_M(x)$ given in (25) readily leads to

$$M(t) = \bar{\alpha}_0 + \bar{\alpha}_1 e^{-a_1 t} + \bar{\alpha}_2 e^{-a_2 t} + \bar{\alpha}_3 t e^{-a_2 t} + \bar{\alpha}_\delta e^{-\delta t},$$

which verifies the expression (10).

(3) If $c_1^2 > c_2$, then a_1 is a positive real number while a_2 and a_3 are conjugate complexes. Let $a_r = \text{Re}(a_2)$ and $a_i = \text{Im}(a_2)$. Then expressions (15)-(19) suggest

$$\begin{cases} a_1 = \frac{\kappa_1 + \kappa_2 + \lambda + \gamma}{3} - \left(\sqrt[3]{-c_1 + \sqrt{c_1^2 - c_2}} + \sqrt[3]{-c_1 - \sqrt{c_1^2 - c_2}} \right), \\ a_r = \frac{\kappa_1 + \kappa_2 + \lambda + \gamma}{3} + \frac{1}{2} \left(\sqrt[3]{-c_1 + \sqrt{c_1^2 - c_2}} + \sqrt[3]{-c_1 - \sqrt{c_1^2 - c_2}} \right), \\ a_i = -\frac{\sqrt{3}}{2} \left(\sqrt[3]{-c_1 + \sqrt{c_1^2 - c_2}} - \sqrt[3]{-c_1 - \sqrt{c_1^2 - c_2}} \right). \end{cases} \quad (26)$$

We decompose $L_M(x)$ of (21) into the sum of partial fractions as

$$L_M(x) = \frac{\tilde{\alpha}_0}{x} + \frac{\tilde{\alpha}_1}{x + a_1} + \frac{\tilde{\alpha}_\delta}{x + \delta} + \frac{\tilde{\alpha}_2 x + \tilde{\alpha}_3}{(x + a_r)^2 + a_i^2}, \quad (27)$$

with $\tilde{\alpha}_0, \tilde{\alpha}_1, \tilde{\alpha}_2, \tilde{\alpha}_3$ and $\tilde{\alpha}_\delta$ satisfying

$$\begin{aligned} & \tilde{\alpha}_0(x + a_1)(x + \delta)[(x + a_r)^2 + a_i^2] + \tilde{\alpha}_1x(x + \delta)[(x + a_r)^2 + a_i^2] + (\tilde{\alpha}_2x + \tilde{\alpha}_3)x(x + a_1)(x + \delta) \\ & + \tilde{\alpha}_\delta x(x + a_1)[(x + a_r)^2 + a_i^2] = vh(x) + M(0)x(x + a_1)[(x + a_r)^2 + a_i^2]. \end{aligned}$$

Then the separate substitution of $x = 0, -a_1, -\delta$ into this equation yields

$$\tilde{\alpha}_0 = \frac{v}{\delta} \frac{\kappa_1 \kappa_2 \lambda}{a_1(a_r^2 + a_i^2)}, \quad \tilde{\alpha}_1 = \frac{-vh(-a_1)}{a_1(\delta - a_1)[(a_1 - a_r)^2 + a_i^2]}, \quad \text{and} \quad \tilde{\alpha}_\delta = \frac{-vh(-\delta)}{\delta(a_1 - \delta)[(\delta - a_r)^2 + a_i^2]} + M(0).$$

The other two $\tilde{\alpha}_2$ and $\tilde{\alpha}_3$ are obtained by comparing the coefficients:

$$\begin{aligned}\tilde{\alpha}_2 &= M(0) - (\tilde{\alpha}_0 + \tilde{\alpha}_1 + \tilde{\alpha}_\delta), \\ \tilde{\alpha}_3 &= vP_e(0) + (a_1 + 2a_r)M(0) - [\alpha_0(2a_r + a_1 + \delta) + \tilde{\alpha}_1(2a_r + \delta) + \tilde{\alpha}_2(a_1 + \delta) + \tilde{\alpha}_\delta(2a_r + a_1)].\end{aligned}$$

Using above $\tilde{\alpha}_2$ and $\tilde{\alpha}_3$ we further introduce the symbols

$$\bar{A} = \tilde{\alpha}_2, \quad A = \frac{\tilde{\alpha}_3/\tilde{\alpha}_2 - a_r}{a_i}, \quad \text{and} \quad \theta = \arccos \frac{1}{\sqrt{1 + A^2}}. \quad (28)$$

Then the application of inverse Laplace transform to $L_M(x)$ given in (27) readily leads to

$$\begin{aligned}M(t) &= \tilde{\alpha}_0 + \tilde{\alpha}_1 e^{-a_1 t} + \tilde{\alpha}_\delta e^{-\delta t} + \bar{A}[\cos(a_i t) + A \sin(a_i t)]e^{-a_r t} \\ &= \tilde{\alpha}_0 + \tilde{\alpha}_1 e^{-a_1 t} + \tilde{\alpha}_\delta e^{-\delta t} + \bar{A}\sqrt{1 + A^2} \cos(a_i t + \theta)e^{-a_r t},\end{aligned}$$

which is exactly the expression (11).

Dynamical profiles of $M(t)$ for cross-talking three-state model (1)

In this section we utilize exact forms of $M(t)$ to understand its dynamical features. We consider the initial condition (12) for which the gene is silence and there is almost no transcripts at the initial time [11, 21, 30]. The following two theorems illustrate the cases of $c_1^2 > c_2$ and $c_1^2 < c_2$, respectively.

Theorem 1. *For arbitrary non-zero real numbers a, b, c and positive real numbers M^*, α, β, r , if $M(t)$ is expressed as*

$$M(t) = M^* + ae^{-\alpha t} + be^{-\beta t} + ce^{-rt} \geq 0, \quad 0 < \alpha < \beta < r, \quad M(0) = 0, \quad \text{and} \quad M'(0) \geq 0,$$

then $M(t)$ displays only one of increasing, up-and-down and up-down-up dynamical profiles.

Proof. The derivative of $M(t)$ gives

$$M'(t) = -\alpha ae^{-\alpha t} - \beta be^{-\beta t} - rce^{-rt} \implies M'(t)e^{rt} = -[rc + \alpha ae^{(r-\alpha)t} + \beta be^{(r-\beta)t}]. \quad (29)$$

We divided the discussion into two cases of $a > 0$ and $a < 0$.

(I) If $a > 0$, then $\lim_{t \rightarrow \infty} M'(t) < 0$ by noting that the sign of $M'(t)$ is dominated by the term “ $-\alpha ae^{-\alpha t}$ ” when $t \rightarrow \infty$. Taking the derivative of $M'(t)e^{rt}$ in (29) gives

$$[M'(t)e^{rt}]' = -\alpha a(r - \alpha)e^{(r-\alpha)t}M_1(t), \quad \text{with} \quad M_1(t) = 1 + \frac{\beta b(r - \beta)}{\alpha a(r - \alpha)}e^{(\alpha-\beta)t}. \quad (30)$$

If $b \geq -\alpha a(r - \alpha)/\beta(r - \beta)$, then (30) indicates

$$M_1(t) > 0 \implies [M'(t)e^{rt}]' < 0 \implies M'(t)e^{rt} \text{ decreases, for } t > 0.$$

There are two probabilities of $M'(0) > 0$ and $M'(0) = 0$. If $M'(0) > 0$, then together with $\lim_{t \rightarrow \infty} M'(t) < 0$, there exists $t_1 > 0$ such that

$$M'(t)e^{rt} > 0, \text{ for } t \in (0, t_1), \quad M'(t_1)e^{rt_1} = 0, \quad \text{and} \quad M'(t)e^{rt} < 0, \text{ for } t > t_1.$$

Therefore, $M(t)$ increases for $t \in (0, t_1)$ and decreases for $t > t_1$ which presents up-and-down dynamics. If $M'(0) = 0$, then $M'(t)e^{rt} < 0$ for $t > 0$, and thus $M(t)$ decreases with $M(t) < 0$ all the time. This contradicts to the assumption of $M(t) \geq 0$.

If $b < -\alpha a(r - \alpha)/\beta(r - \beta)$, then (30) suggests that there is $t_2 > 0$ such that

$$M_1(t) < 0, \quad t \in (0, t_2), \quad M_1(t_2) = 0, \quad \text{and} \quad M_1(t) > 0, \quad t > t_2, \quad \text{where} \quad t_2 = \frac{1}{\alpha - \beta} \ln \left(\frac{-a\alpha(r - \alpha)}{b\beta(r - \beta)} \right).$$

Then (30) suggests that $[M'(t)e^{rt}]' > 0$ for $t \in (0, t_2)$ and $[M'(t)e^{rt}]' < 0$ for $t > t_2$. Therefore, $M'(t)e^{rt}$ increases in $(0, t_2)$ while decreases in (t_2, ∞) . Since $M'(0) \geq 0$ and $\lim_{t \rightarrow \infty} M'(t) < 0$, there exists t_3 such that $M'(t_3)e^{rt_3} = 0$. Thus $M'(t)e^{rt} > 0$ for $t \in (0, t_3)$ and $M'(t)e^{rt} < 0$ for $t > t_3$. This indicates that $M(t)$ increases for $t \in (0, t_3)$ and decreases for $t > t_3$, which is the up-and-down dynamics.

(II) If $a < 0$, the (29) indicates that $\lim_{t \rightarrow \infty} M'(t) > 0$. If $b \leq -\alpha a(r - \alpha)/\beta(r - \beta)$, then (30) gives

$$M_1(t) > 0 \implies [M'(t)e^{rt}]' > 0 \implies M'(t)e^{rt} \text{ increases, for } t > 0.$$

Since $M'(0) \geq 0$, we have $M'(t)e^{rt} \geq 0, t > 0$ which suggests that $M(t)$ increases for all the time.

On the other hand, if $b > -\alpha a(r - \alpha)/\beta(r - \beta)$, then from (30) we find that there exists $t_4 > 0$ such that

$$M_1(t) < 0, \quad t \in (0, t_4), \quad M_1(t_4) = 0, \quad \text{and} \quad M_1(t) > 0, \quad t > t_4, \quad \text{where} \quad t_4 = \frac{1}{\alpha - \beta} \ln \left(\frac{-a\alpha(r - \alpha)}{b\beta(r - \beta)} \right).$$

Together with (30), this further leads to $[M'(t)e^{rt}]' < 0$ for $t \in (0, t_4)$ and $[M'(t)e^{rt}]' > 0$ for $t > t_4$. Therefore, $M'(t)e^{rt}$ decreases in $(0, t_4)$ while increases in (t_4, ∞) . Since $M'(0) \geq 0$, there are two possibilities of $M'(t_4)e^{rt_4} \geq 0$ and $M'(t_4)e^{rt_4} < 0$. For $M'(t_4)e^{rt_4} \geq 0$, the fact of $\lim_{t \rightarrow \infty} M'(t) > 0$ indicates $M'(t)e^{rt} \geq 0$ for all $t \geq 0$, and thereby $M(t)$ increases all the time. If it is $M'(t_4)e^{rt_4} < 0$, then there exist $0 < t_5 < t_4 < t_6$ such that $M'(t_5)e^{rt_5} = M'(t_6)e^{rt_6} = 0$ and

$$M'(t)e^{rt} > 0, \quad t \in (0, t_5), \quad M'(t)e^{rt} < 0, \quad t \in (t_5, t_6), \quad \text{and} \quad M'(t)e^{rt} > 0, \quad t > t_6.$$

This indicates that $M(t)$ displays an up-down-up dynamics that $M(t)$ increases for $t \in (0, t_5) \cup (t_6, \infty)$ and decreases within $t \in (t_5, t_6)$. The proof is completed. ■

Theorem 2. *Let Condition $c_1^2 < c_2$ hold. Let*

$$H(x) = vh(x) = vq_2\lambda x^2 + v[q_1\kappa_1\kappa_2 + q_2\lambda(\kappa_1 + \kappa_2)]x + v\kappa_1\kappa_2\lambda, \quad (31)$$

and x_1 and x_2 be two roots of $H(x)$. Without loss of generality, assume $\delta \neq a_1, a_2, a_3$ given by (22). If x_1 and x_2 are complex valued, then either $M(t)$ increases monotonically for all $t > 0$, or $M(t)$ develops up-down-up dynamics. If x_1 and x_2 are real valued and $x_1 < x_2$, then we have:

(1) If one of the following occurs: (i) $-x_2 < \{a_1, a_2, a_3, \delta\} < -x_1$, (ii) $-x_2 < \{a_1, a_2, \delta\} < -x_1 < a_3$, (iii) $-x_2 < \{a_1, \delta\} < -x_1 < a_2$, (iv) $-x_2 < a_1 < a_2 < a_3 < -x_1 < \delta$, (v) $-x_2 < a_1 < a_2 < -x_1 < \min\{a_3, \delta\}$, (vi) $-x_2 < a_1 < -x_1 < \min\{a_2, \delta\}$, (vii) $-x_2 < \delta < -x_1 < a_1$, then $m(t)$ increases initially until reaching a peak and then goes down (up-and-down).

(2) If one of the following occurs: (i) $a_1 < -x_2 < a_2 < a_3 < -x_1 < \delta$, (ii) $a_1 < -x_2 < a_2 < a_3 < -x_1 < \delta$, (iii) $a_1 < -x_2 < a_2 < -x_1 < \min\{a_3, \delta\}$, (iv) $a_1 < -x_2 < \{a_2, \delta\} < -x_1 < a_3$, (v) $\max\{a_1, \delta\} < -x_2 < a_2 < -x_1 < a_3$, (vi) $\delta < -x_2 < a_1 < a_2 < -x_1 < a_3$, (vii) $a_1 < -x_2 < \delta < -x_1 < a_2$, and (viii) $\delta < -x_2 < a_1 < -x_1 < a_2$, then $M(t)$ increases monotonically for all $t > 0$.

(3) If one of the following occurs: (i) $-x_1 < \min\{a_1, \delta\}$, (ii) $a_1 < -x_2 < -x_1 < \min\{a_2, \delta\}$, (iii) $a_2 < -x_2 < -x_1 < \min\{a_3, \delta\}$, (iv) $a_3 < -x_2 < -x_1 < \delta$, (v) $\max\{a_1, \delta\} < -x_2$, (vi) $\max\{a_2, \delta\} < -x_2 < -x_1 < a_3$, (vii) $\max\{a_3, \delta\} < -x_2$, (viii) $\delta < -x_2 < -x_1 < a_1$, (ix) $a_3 < -x_2 < \delta < -x_1$, (x) $a_2 < -x_2 < a_3 < -x_1 < \delta$, (xi) $a_2 < -x_2 < \{a_3, \delta\} < -x_1$, (xii) $\{a_2, \delta\} < -x_2 < a_3 < -x_1$, (xiii) $a_1 < -x_2 < \{a_2, a_3, \delta\} < -x_1$, (xiv) $\{a_1, \delta\} < -x_2 < a_2 < a_3 < -x_1$, (xv) $\delta < -x_2 < a_1 < a_2 < a_3 < -x_1$, then either $M(t)$ increases monotonically for all $t > 0$, or $M(t)$ develops up-down-up dynamical profile.

Proof. The substitution of the initial condition (12) into the differential system (7) gives $M'(0) = 0$, $P_e'(0) > 0$, and in turn, $M''(0) = vP_e''(0) > 0$. Consequently, $M(t)$, $M'(t)$, and $M''(t)$ are all positive for $t > 0$ sufficiently small. The initial values (12) simplifies the exact form (9) of $M(t)$ in the form of

$$M(t) = \frac{v}{\delta} \cdot \frac{\kappa_1\kappa_2\lambda}{a_1a_2a_3} - \frac{\beta_1}{a_1}e^{-a_1t} - \frac{\beta_2}{a_2}e^{-a_2t} - \frac{\beta_3}{a_3}e^{-a_3t} - \frac{\beta_\delta}{\delta}e^{-\delta t},$$

where

$$\begin{aligned} \beta_1 &= \frac{H(-a_1)}{(\delta - a_1)(a_2 - a_1)(a_3 - a_1)}, & \beta_2 &= \frac{H(-a_2)}{(\delta - a_2)(a_1 - a_2)(a_3 - a_2)}, \\ \beta_3 &= \frac{H(-a_3)}{(\delta - a_3)(a_1 - a_3)(a_2 - a_3)}, & \beta_\delta &= \frac{H(-\delta)}{(a_1 - \delta)(a_2 - \delta)(a_3 - \delta)}. \end{aligned}$$

Thus

$$M'(t) = \beta_1e^{-a_1t} + \beta_2e^{-a_2t} + \beta_3e^{-a_3t} + \beta_\delta e^{-\delta t}. \quad (32)$$

We first consider the situation of complex valued x_1 and x_2 . There are total 4 parameter cases and we only present the proof for the case $\delta < a_1 < a_2 < a_3$, since the other cases $a_1 < \delta < a_2 < a_3$, $a_1 < a_2 < \delta < a_3$ and $a_1 < a_2 < a_3 < \delta$ can be treated similarly. If x_1 and x_2 are complex valued, then $H(x) > 0$ for all x . In view of $\delta < a_1 < a_2 < a_3$, we obtain

$$\beta_1 < 0, \beta_2 > 0, \beta_3 < 0 \text{ and } \beta_\delta > 0. \quad (33)$$

For $t > 0$ sufficiently large, $M'(t)$ is dominated by $\beta_\delta \exp(-\delta t)$ in (32) and is therefore positive. It follows from (32) that

$$[e^{a_3 t} M'(t)]' = e^{(a_3 - \delta)t} g(t), \text{ with } g(t) = (a_3 - \delta)\beta_\delta + (a_3 - a_1)\beta_1 e^{(\delta - a_1)t} + (a_3 - a_2)\beta_2 e^{(\delta - a_2)t}. \quad (34)$$

It is seen that

$$g(0) = M''(0) > 0 \text{ and } \lim_{t \rightarrow \infty} g(t) = (a_3 - \delta)\beta_\delta > 0. \quad (35)$$

We can see from (34) that

$$M'(t) = e^{-a_3 t} G(t), \text{ where } G(t) = \int_0^t e^{(a_3 - \delta)s} g(s) ds, \quad t > 0. \quad (36)$$

It follows from (32) and (36) that

$$G(0) = 0, \quad \lim_{t \rightarrow \infty} G(t) = +\infty, \quad \text{and} \quad G'(t) = e^{(a_3 - \delta)t} g(t). \quad (37)$$

Differentiating $g(t)$ gives

$$g'(t) = \frac{H(-a_1)}{a_2 - a_1} e^{(\delta - a_2)t} \left[e^{(a_2 - a_1)t} - \frac{H(-a_2)}{H(-a_1)} \right].$$

If $H(-a_2) \leq H(-a_1)$, then $g'(t) \geq 0$, and so $g(t) > 0$ for all $t > 0$. Hence (36) gives $M'(t) > 0$ for all $t > 0$. If $H(-a_2) > H(-a_1)$, then $g'(t) < 0$ in $(0, t_0)$ for

$$t_0 = \frac{1}{a_2 - a_1} [\ln(H(-a_2)) - \ln(H(-a_1))],$$

and becomes positive for $t > t_0$. If $g(t_0) \geq 0$, then $g(t) > 0$ for all $t > 0$ and $t \neq t_0$. Thus (36) gives again $M'(t) > 0$ for all $t > 0$. If $g(t_0) < 0$, then (35) implies that $g(t)$ has two zeros $t_1 > 0$ and $t_2 > 0$ with $g(t) > 0$ in $(0, t_1) \cup (t_2, +\infty)$, and $g(t) < 0$ in (t_1, t_2) . It follows from (37) that

$$G'(t) > 0 \text{ for } t \in (0, t_1) \cup (t_2, +\infty), \quad \text{and} \quad G'(t) < 0 \text{ for } t \in (t_1, t_2). \quad (38)$$

If $G(t_2) \geq 0$, then $G(t) > 0$ for all $t > 0$ and $t \neq t_2$, and so $M'(t) > 0$ for all $t > 0$ and $t \neq t_2$. We recall that $G(t) > 0$ for $t > 0$ both sufficiently small and large. Thus, if $G(t_2) < 0$, then (38) implies that $G(t)$ has two zeros $t_3 > 0$ and $t_4 > 0$ with $G(t) > 0$ in $(0, t_3) \cup (t_4, +\infty)$, and $G(t) < 0$

in (t_3, t_4) . It follows from (36) that $M'(t) > 0$ in $(0, t_3) \cup (t_4, +\infty)$ and $M'(t) < 0$ in (t_3, t_4) . This finishes the proof of the case when x_1 and x_2 are complex valued.

In the following, we assume that x_1 and x_2 are real valued.

(1) We present the proof with the additional assumption that $-x_2 < a_1 < a_2 < \delta < -x_1 < a_3$, as the other cases can be dealt with by the same idea. Since

$$-x_2 < a_1 < a_2 < \delta < -x_1 < a_3 \Rightarrow \beta_1 < 0, \beta_2 > 0, \beta_3 < 0, \text{ and } \beta_\delta < 0, \quad (39)$$

by using (32) again, we obtain

$$[e^{a_2 t} M'(t)]' = e^{(a_2 - a_1)t} g_1(t), \quad g_1(t) = (a_2 - a_1)\beta_1 + (a_2 - \delta)\beta_\delta e^{(a_1 - \delta)t} + (a_2 - a_3)\beta_3 e^{(a_1 - a_3)t}, \quad (40)$$

with $f(0) = M''(0) > 0$ and $\lim_{t \rightarrow \infty} g_1(t) = (a_2 - a_1)\beta_1 < 0$. In terms of (39) and (40), we have $(a_2 - \delta)(a_1 - \delta)\beta_\delta < 0$ and $(a_2 - a_3)(a_1 - a_3)\beta_3 < 0$, and so

$$f'(t) = (\delta - a_2)(a_1 - a_2)\beta_2 e^{(\delta - a_2)t} + (\delta - a_3)(a_1 - a_3)\beta_3 e^{(\delta - a_3)t} < 0.$$

Hence, there exists a $t_5 > 0$ such that $g_1(t) > 0$ in $(0, t_5)$, $g_1(t_5) = 0$, and $g_1(t) < 0$ in $(t_5, +\infty)$. For $t > 0$ sufficiently large, $M'(t)$, dominated by the term with $\exp(-a_1)$ in (32), is negative. We also recall that $M'(t) > 0$ for $t > 0$ sufficiently small. Thus, there exists a finite $t_6 > 0$ such that $M'(t) > 0$ in $(0, t_6)$, and $M'(t_6) = 0$. Then $M''(t_6) \leq 0$, and (40) gives $f(t_6) = \exp(a_1 t_6) M''(t_6) \leq 0$. It follows that $t_6 \geq t_5$ and $f(t) < 0$ for all $t > t_6$. Hence, by using (40) again, for each $t > t_6$, we have

$$M'(t) = e^{-a_2 t} \int_{t_6}^t e^{(a_2 - \delta)s} f(s) ds < 0.$$

(2) We present the proof for the case $a_1 < a_2 < -x_2 < \delta < -x_1 < a_3$, as the other cases can be handled similarly. With this specification, we have

$$\beta_1 > 0, \beta_2 < 0, \beta_3 < 0, \text{ and } \beta_\delta < 0. \quad (41)$$

By (32) and (41), we obtain for all $t > 0$

$$[e^{a_1 t} M'(t)]' = (a_1 - \delta)\beta_\delta e^{(a_1 - \delta)t} + (a_1 - a_2)\beta_2 e^{(a_1 - a_2)t} + (a_1 - a_3)\beta_3 e^{(a_1 - a_3)t} > 0.$$

As $M'(0) = 0$, it shows that $e^{a_2 t} M'(t) > 0$, and so $M'(t) > 0$ for all $t > 0$.

(3) We give a short description for the proof with the additional condition that $x_1 < x_2 < \delta < a_1 < a_2 < a_3$, as the other cases can be proceeded in an analogous manner. Under this extra condition, (33) holds again. Thus the remaining discussion is the same as the proof given above for two complex valued x_1 and x_2 , and we omit it here. ■

Calculation of $M(t)$ for cross-talking n -state model

Following the standard procedure of [15,33,39] we obtain the master equations for the cross-talking n -state model, which further leads to the differential system of gene state probabilities and the average mRNA level $M(t)$ as follows

$$\begin{cases} P'_{o_{11}}(t) = -\kappa_1 P_{o_{11}}(t) + q_1 \gamma P_e(t), & P'_{o_{12}}(t) = -\lambda P_{o_{12}}(t) + q_2 \gamma P_e(t), \\ P'_{o_2}(t) = \kappa_1 P_{o_{11}}(t) - \kappa_2 P_{o_2}(t), & P'_{o_3}(t) = \kappa_2 P_{o_2}(t) - \kappa_3 P_{o_3}(t), \\ \dots & \\ P'_{o_{n-1}}(t) = \kappa_{n-2} P_{o_{n-2}}(t) - \kappa_{n-1} P_{o_{n-1}}(t), & P'_e(t) = \lambda P_{o_{12}}(t) + \kappa_{n-1} P_{o_{n-1}}(t) - \gamma P_e(t), \\ M'(t) = v P_e(t) - \delta M(t), \\ P_{o_{11}}(0) = q_1, & P_{o_{12}}(0) = q_2, & P_{o_j}(0) = P_e(0) = M(0) = 0, & j = 2, 3, \dots, n-1. \end{cases} \quad (42)$$

Here $P_{o_{11}}(t)$ and $P_{o_{12}}(t)$ are probabilities of gene *off* 1 state selecting weak and signaling pathways, respectively. $P_{o_j}(t), j = 2, 3, \dots, n-1$ are probabilities of gene *off* j states while the probability $P_e(t)$ is for gene *on* state. Also, we consider the initial condition for which the gene is totally *off* with no transcripts at the initial time.

For the special case of $q_2 = 0$ which suggests that the gene activation is regulated by a single multi-step pathway, the system (42) reduces to

$$\begin{cases} P'_{o_{11}}(t) = -\kappa_1 P_{o_{11}}(t) + \gamma P_e(t), & P'_{o_2}(t) = \kappa_1 P_{o_{11}}(t) - \kappa_2 P_{o_2}(t), \\ P'_{o_3}(t) = \kappa_2 P_{o_2}(t) - \kappa_3 P_{o_3}(t), & \dots & P'_{o_{n-1}}(t) = \kappa_{n-2} P_{o_{n-2}}(t) - \kappa_{n-1} P_{o_{n-1}}(t), \\ P'_e(t) = \kappa_{n-1} P_{o_{n-1}}(t) - \gamma P_e(t), & M'(t) = v P_e(t) - \delta M(t), \\ P_{o_{11}}(0) = 1, & P_{o_j}(0) = P_e(0) = M(0) = 0, & j = 2, 3, \dots, n-1. \end{cases} \quad (43)$$

The crucial step to solve the system (42) or (43) is to calculate the eigenvalues of the coefficient matrix for equations of gene state probabilities [33,39]. Here we illustrate a simple case of $\kappa_1 = \kappa_2 = \dots = \kappa_{n-1} = \kappa$ in the system (43). Then its characteristic polynomial $f(x)$ for gene state probabilities is expressed as

$$f(x) = \begin{vmatrix} x + \kappa & 0 & \dots & 0 & -\gamma \\ -\kappa & x + \kappa & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & x + \kappa & 0 \\ 0 & 0 & \dots & -\kappa & x + \gamma \end{vmatrix}_{n \times n} = (x + \kappa)^{n-1}(x + \gamma) - \kappa^{n-1}\gamma.$$

There is no standard method to solve the characteristic polynomial equation $f(x) = 0$ of degree $n \geq 5$. However, the fundamental theorem of algebra shows that $f(x) = 0$ has n roots (eigenvalues) with complex roots always exhibiting in multiple pairs of conjugate complexes.

The second moment of mRNA distribution for cross-talking three-state model (1)

By definition the second moment $\mu_2(t)$ of the distribution of mRNA copy numbers is

$$\mu_2(t) = \sum_{m=0}^{\infty} m^2 P_m(t),$$

where $P_m(t)$ is the solution of master functions (2)-(6). To derive $\mu_2(t)$, we need first calculate the partial mRNA average levels at each gene state, defined by

$$M_{1j}(t) = \sum_{m=0}^{\infty} m P_{m,o_{1j}}(t), \quad j = 1, 2, \quad M_2(t) = \sum_{m=0}^{\infty} m P_{m,o_2}(t) \quad \text{and} \quad M_e(t) = \sum_{m=0}^{\infty} m P_{m,e}(t).$$

By these definitions and (2)-(6) we can derive the equations of $M_{11}(t), M_{12}(t), M_2(t), M_e(t)$ and $\mu_2(t)$ [33, 39]

$$\begin{cases} M'_{11}(t) = -(\kappa_1 + \delta)M_{11}(t) + q_1\gamma M_e(t), & M'_{12}(t) = -(\lambda + \delta)M_{12}(t) + q_2\gamma M_e(t), \\ M'_2(t) = \kappa_1 M_{11}(t) - (\kappa_2 + \delta)M_2(t), & M'_e(t) = \lambda M_{12}(t) + \kappa_2 M_2(t) - (\gamma + \delta)M_e(t) + vP_e(t), \\ \mu'_2(t) = -2\delta\mu_2(t) + \delta M(t) + vP_e + 2vM_e(t). \end{cases} \quad (44)$$

Here we focus on the initial condition

$$(P_{o_{11}}(0), P_{o_{12}}(0), P_{o_2}(0), P_e(0), M(0), \mu_2(0)) = (q_1, q_2, 0, 0, 0, 0). \quad (45)$$

To solve the first-order differential system (44)-(45), we denote by

$$\begin{aligned} L_{M_{1j}}(x) &= \int_0^{\infty} e^{-xt} M_{1j}(t) dt, \quad j = 1, 2, & L_{M_2}(x) &= \int_0^{\infty} e^{-xt} M_2(t) dt, \\ L_{M_e}(x) &= \int_0^{\infty} e^{-xt} M_e(t) dt, & \text{and} & \quad L_{\mu_2}(x) = \int_0^{\infty} e^{-xt} \mu_2(t) dt. \end{aligned}$$

Then the application of Laplace transform to the first four equations of system (44) yields

$$\begin{cases} 0 = (x + \kappa_1 + \delta)L_{M_{11}}(x) - q_1\gamma L_{M_e}(x), \\ 0 = (x + \lambda + \delta)L_{M_{12}}(x) - q_2\gamma L_{M_e}(x), \\ 0 = -\kappa_1 L_{M_{11}}(x) + (x + \kappa_2 + \delta)L_{M_2}(x), \\ vL_e = -\lambda L_{M_{12}}(x) - \kappa_2 L_{M_2}(x) + (x + \gamma + \delta)L_{M_e}(x). \end{cases}$$

By applying Cramer's rule in the theory of linear algebra to solve the above system, we obtain

$$L_{M_e}(x) = \frac{v(x + \kappa_1 + \delta)(x + \lambda + \delta)(x + \kappa_2 + \delta)L_e(x)}{(x + \delta)(x + a_1 + \delta)(x + a_2 + \delta)(x + a_3 + \delta)}, \quad (46)$$

where $L_e(x)$ is given by (20) while a_1, a_2 and a_3 are given by (15)-(19) or by the separate (22), (24) and (26) for the corresponding cases.

To solve the second moment $\mu_2(t)$ from the system (44)-(45), we apply the Laplace transform to the last equation of (44), and then substitute (46) to obtain

$$L_{\mu_2}(x) = \frac{1}{x + 2\delta} \left(\frac{v\delta}{x + \delta} + v + \frac{2v^2(x + \kappa_1 + \delta)(x + \lambda + \delta)(x + \kappa_2 + \delta)}{(x + \delta)(x + a_1 + \delta)(x + a_2 + \delta)(x + a_3 + \delta)} \right) L_e(x). \quad (47)$$

We note that $L_e(x)$ given by (20) can be simplified under the initial condition (12) or (45). The substitution of simplified $L_e(x)$ into (47) further gives

$$L_{\mu_2}(x) = \frac{F(x)}{x(x + a_1)(x + a_2)(x + a_3)(x + \delta)(x + 2\delta)(x + a_1 + \delta)(x + a_2 + \delta)(x + a_3 + \delta)},$$

where $F(x)$ takes the form of

$$\begin{aligned} F(x) &= [q_2\lambda x^2 + [q_1\kappa_1\kappa_2 + q_2\lambda(\kappa_1 + \kappa_2)]x + \kappa_1\kappa_2\lambda] \\ &\times [v(x + 2\delta)(x + a_1 + \delta)(x + a_2 + \delta)(x + a_3 + \delta) + 2v^2(x + \lambda_1 + \delta)(x + \lambda_2 + \delta)(x + \lambda_3 + \delta)]. \end{aligned}$$

Under the condition $c_1^2 < c_2$ where a_1, a_2 and a_3 are all positive real numbers, we decompose $L_{\mu_2}(x)$ into the sum of partial fractions as

$$\begin{aligned} L_{\mu_2}(x) &= \frac{b_0}{x} + \frac{b_1}{x + a_1} + \frac{b_2}{x + a_2} + \frac{b_3}{x + a_3} + \frac{b_\delta}{x + \delta} + \frac{b_{2\delta}}{x + 2\delta} \\ &+ \frac{b_{1,\delta}}{x + a_1 + \delta} + \frac{b_{2,\delta}}{x + a_2 + \delta} + \frac{b_{3,\delta}}{x + a_3 + \delta}. \end{aligned}$$

Then the application of the inverse Laplace transform to $L_{\mu_2}(x)$ readily leads to

$$\begin{aligned} \mu_2(t) &= b_0 + b_1 e^{-a_1 t} + b_2 e^{-a_2 t} + b_3 e^{-a_3 t} + b_\delta e^{-\delta t} + b_{2\delta} e^{-2\delta t} \\ &+ b_{1,\delta} e^{-(a_1 + \delta)t} + b_{2,\delta} e^{-(a_2 + \delta)t} + b_{3,\delta} e^{-(a_3 + \delta)t}, \end{aligned} \quad (48)$$

with a_1, a_2 and a_3 being given by (22) and the coefficients taking forms of

$$\begin{aligned}
 b_0 &= \frac{F(0)}{2a_1a_2a_3\delta^2(a_1+\delta)(a_2+\delta)(a_3+\delta)}, \\
 b_1 &= -\frac{F(-a_1)}{a_1\delta(a_2-a_1)(a_3-a_1)(\delta-a_1)(2\delta-a_1)(a_2+\delta-a_1)(a_3+\delta-a_1)}, \\
 b_2 &= \frac{-F(-a_2)}{a_2\delta(a_1-a_2)(a_3-a_2)(\delta-a_2)(2\delta-a_2)(a_1+\delta-a_2)(a_3+\delta-a_2)}, \\
 b_3 &= \frac{-F(-a_3)}{a_3\delta(a_1-a_3)(a_2-a_3)(\delta-a_3)(2\delta-a_3)(a_1+\delta-a_3)(a_2+\delta-a_3)}, \\
 b_\delta &= \frac{-F(-\delta)}{a_1a_2a_3\delta^2(a_1-\delta)(a_2-\delta)(a_3-\delta)}, \\
 b_{2\delta} &= \frac{F(-2\delta)}{2\delta^2(a_1-\delta)(a_2-\delta)(a_3-\delta)(a_1-2\delta)(a_2-2\delta)(a_3-2\delta)}, \\
 b_{1,\delta} &= \frac{-F[-(a_1+\delta)]}{a_1\delta(a_1+\delta)(a_2-a_1)(a_3-a_1)(\delta-a_1)(a_2-a_1-\delta)(a_3-a_1-\delta)}, \\
 b_{2,\delta} &= \frac{-F[-(a_2+\delta)]}{a_2\delta(a_2+\delta)(a_1-a_2)(a_3-a_2)(\delta-a_2)(a_1-a_2-\delta)(a_3-a_2-\delta)}, \\
 b_{3,\delta} &= \frac{-F[-(a_3+\delta)]}{a_3\delta(a_3+\delta)(a_1-a_3)(a_2-a_3)(\delta-a_3)(a_1-a_3-\delta)(a_2-a_3-\delta)}.
 \end{aligned}$$

Supporting Information

Table S1. Fitted system parameters for dynamical transcription data of mouse fibroblast genes.

Fig. S1. Oscillatory dynamics of mRNA average level $M(t)$ generated by a single pathway with multiple sequential steps (n -state model). The increase in the step number $n-1$ prolongs the initial response lag and enhances the following damped oscillation for $M(t)$. Parameters are the fitted rates of *Cxcl1* gene in Table S1 with the average gene *off* duration T_{off} being constant under (a) $\kappa_i = i\kappa_1, i = 1, 2, \dots, n-1$; (b) $\kappa_i = \kappa_1, \kappa_j = 3\kappa_1$ for odd i and even j .

Fig. S2. Oscillatory dynamics of mRNA average level $M(t)$ generated by the cross-talking n -state model. The increase in the step number $n-1$ in the basal pathway has nearly no impact on the initial quick up-and-down dynamics of $M(t)$ but significantly enhances the following gentle and damped oscillatory behavior. Parameters are the fitted rates of *Cxcl1* gene in Table S1 with the average gene *off* duration T_{off} being constant under (a) $\kappa_i = i\kappa_1, i = 1, 2, \dots, n-1$; (b) $\kappa_i = \kappa_1, \kappa_j = 3\kappa_1$ for odd i and even j .

Author Contributions

L.C. G.L. and F.J. designed and performed the research; F.J. wrote the paper.

Acknowledgements

This work is supported by Natural Science Foundation of China grants (Nos. 11871174; 12001127; 11631005), and by Program for Changjiang Scholars and Innovative Research Team in University (No. IRT_16R16). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data Availability Statement

All relevant data are within the manuscript and its Supporting Information files.

Competing interests

The authors declare no competing financial interests.

References

- [1] Munsky, B., Neuert, G. & van Oudenaarden, A. Using gene expression noise to understand gene regulation. *Science* **336**, 183-187 (2012).
- [2] Dar, R.D. *et al.* Transcriptional burst frequency and burst size are equally modulated across the human genome. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 17454-17459 (2012).
- [3] Larsson, A.J.M *et al.* Genomic encoding of transcriptional burst kinetics. *Nature* **565**, 251-254 (2019).
- [4] Choubey, S., Kondev, J. & Sanchez, A. Deciphering transcriptional dynamics in vivo by counting nascent RNA molecules. *PLoS Comput. Biol.* **11**, e1004345 (2015).
- [5] Sepúlveda, L.A., Xu, H., Zhang, J., Wang, M. & Golding I. Measurement of gene regulation in individual cells reveals rapid switching between promoter states. *Science* **351**, 1218-1222 (2016).
- [6] Lloyd-Price, J. *et al.* Dissecting the stochastic transcription initiation process in live *Escherichia coli*. *DNA Research* **23**, 1-12 (2016).
- [7] Bhogale, P.M., Sorg, R.A., Veening, J.W. & Berg J. What makes the *lac*-pathway switch: identifying the fluctuations that trigger phenotype switching in gene regulatory systems. *Nucleic Acids Res.* **42**, 11321-11328 (2014).
- [8] Suter, D.M., Molina, N., Gatfield, D., Schneider, K., Schibler, U. & Naef, F. Mammalian genes are transcribed with widely different bursting kinetics. *Science* **332**, 472-474 (2011).
- [9] Munsky B., Fox Z. & Neuert G. Integrating single-molecule experiments and discrete stochastic models to understand heterogeneous gene transcription dynamics. *Methods* **85**, 12-21 (2015).
- [10] Sanchez, A. & Golding, I. Genetic determinants and cellular constraints in noisy gene expression. *Science* **342**, 1188-1193 (2013).
- [11] Golding, I., Paulsson, J., Zawilski, S.M. & Cox, E.C. Real-time kinetics of gene activity in individual bacteria. *Cell* **123**, 1025-1036 (2005).

- [12] Octavio, L.M., Gedeon, K. & Maheshri, N. Epigenetic and conventional regulation is distributed among activators of FLO11 allowing tuning of population-level heterogeneity in its expression *PLoS Genet.* **5**, e1000673 (2009).
- [13] Larson, D.R. What do expression dynamics tell us about the mechanism of transcription. *Curr. Opin. Genet. Dev.* **21**, 1-9 (2011).
- [14] Zimmer, C., Häkkinen, A. & Ribeiro, A.S. Estimation of kinetic parameters of transcription from temporal single-RNA measurements. *Math. Biosci.* **271**, 146-153 (2016).
- [15] Zhou, T. & Zhang, J. Analytical results for a multistate gene model. *SIAM J. Appl. Math.* **72**, 789-818 (2012).
- [16] Sun, Q., Tang, M., & Yu, J. Temporal profile of gene transcription noise modulated by cross-talking signal transduction pathways. *Bull. Math. Biol.* **74**, 375-398 (2012).
- [17] Sun, Q., Tang, M., & Yu, J. Modulation of gene transcription noise by competing transcription factors. *J. Math. Biol.* **64**, 469-494 (2012).
- [18] Jiang, Q. *et al.* Neural network aided approximation and parameter inference of non-Markovian models of gene expression. *Nat. Commun.* **12**, 2618 (2021).
- [19] Zhang, J. & Zhou, T. Markovian approaches to modeling intracellular reaction processes with molecular memory. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 23542-23550 (2019).
- [20] Neuert G. *et al.* Systematic identification of signal-activated stochastic gene regulation. *Science* **339**, 584-587 (2013).
- [21] Senecal, A. *et al.* Transcription factors modulate c-Fos transcriptional bursts. *Cell Reports* **8**, 75-83 (2014).
- [22] Cao, Z. & Grima, R. Analytical distributions for detailed models of stochastic gene expression in eukaryotic cells. *Proc. Natl. Acad. Sci. U.S.A.* **117** 4682-4692 (2020).
- [23] Iyer-Biswas, S., Hayot, F. & Jayaprakash, C. Stochasticity of gene products from transcriptional pulsing. *Phys. Rev. E* **79**, 031911 (2009).
- [24] Jiao, F., Tang, M. & Yu, J. Distribution profiles and their dynamic transition in stochastic gene transcription. *J. Differential Equations* **254**, 3307-3328 (2013).
- [25] Zhu, C., Han, G. & Jiao, F. Dynamical regulation of mRNA distribution by cross-talking signaling pathways. *Complexity* **2020**, 6402703 (2020).
- [26] Chen, Y., Gong, Q., Wu, Y. Yan, H. & Jiao, F. Dynamical mRNA distribution regulated by multi-step gene activation. *Chaos*, Under review.
- [27] Jiao, F., Sun, Q., Tang, M., Yu, J. & Zheng, B. Distribution modes and their corresponding parameter regions in stochastic gene transcription. *SIAM J. Appl. Math.* **75**, 2396-2420 (2015).
- [28] Jiao, F., Sun, Q., Lin, G. & Yu, J. Distribution profiles in gene transcription activated by the cross-talking pathway. *Discrete Contin. Dyn. Syst. B* **24**, 2799-2810 (2019).
- [29] Jiao, F. & Zhu C. Regulation of gene activation by competitive cross talking pathways. *Biophys. J.* **119**, 1204-1214 (2020).
- [30] Hao, S. & Baltimore, D. The stability of mRNA influences the temporal order of the induction of genes encoding inflammatory molecules. *Nat. Immunol.* **10**, 281-288 (2009).
- [31] Tay, S., Hughev, J.J., Lee, T.K., Lipniacki, T., Quake, S.R. & Covert, M.W. Single-cell NF- κ B dynamics reveal digital activation and analog information processing. *Nature* **466**, 267-271 (2010).
- [32] Hao, N. & O'Shea E.K. Signal-dependent dynamics of transcription factor translocation controls gene expression. *Nat. Struct. Mol. Biol.* **19**, 31-39 (2012).
- [33] Tang, M. The mean and noise of stochastic gene transcription. *J. Theor. Biol.* **253**, 271-280 (2008).
- [34] Kuang J., Tang M. & Yu J. The mean and noise of protein numbers in stochastic gene expression. *J. Math. Biol.* **67**, 261-291 (2013).
- [35] Lin, G., Jiao, F., Sun, Q., Tang, M., Yu, J. & Zhou, Z. Linking dynamical complexities from activation signals to transcription responses. *R. Soc. open sci.* **6**, 190286 (2019).
- [36] Macia, J., Regot, S., Peeters, T., Conde, N., Solé, R. & Posas, F. Dynamic signaling in the Hog1 MAPK pathway relies on high basal signal transduction. *Sci. Signal* **2**, ra13 (2009).
- [37] Jordan, A., Defechereux, P. & Verdin, E. The site of HIV-1 integration in the human genome determines basal transcriptional activity and response to Tat transactivation. *EMBO J.* **20**, 1726-1738 (2001).
- [38] Pelet, S. *et al.* Transient activation of the HOG MAPK pathway regulates bimodal Gene expression. *Science* **332**, 732-735 (2011).

- [39] Lin, G., Yu, J., Zhou, Z., Sun, Q. & Jiao, F. Fluctuations of mRNA distributions in multiple pathway activated transcription. *Discrete Contin. Dyn. Syst. B* **24**, 1543-1568 (2019).
- [40] Jiao, F., Lin, G. & Yu, J. Approximating gene transcription dynamics using steady-state formulas. *Phys. Rev. E* **104**, 014401 (2021).
- [41] Perez-Carrasco, R., Beentjes, C. & Grima, R. Effects of cell cycle variability on lineage and population measurements of messenger RNA abundance. *J. R. Soc. Interface* **17**, 20200360 (2020).
- [42] Choubey, S. Nascent RNA kinetics: Transient and steady state behavior of models of transcription. *Phys. Rev. E* **97**, 022402 (2018).
- [43] Sun, Q., Jiao, F., Lin, G., Yu, J. & Tang, M. The nonlinear dynamics and fluctuations of mRNA levels in cell cycle coupled transcription. *PLoS Comput. Biol.* **15**, e1007017 (2019).