Nascent RNA kinetics with complex promoter architecture: Analytic results and parameter inference

Changhong Shi[1], Xiyan Yang[2], Tianshou Zhou[3,4,*], and Jiajun Zhang[3,4,*]

1. School of Public Health, Guangzhou Medical University, Guangzhou, China
2. School of Financial Mathematics and Statistics, Guangdong University of Finance, Guangzhou, China
3. School of Mathematics, Sun Yat-sen University, Guangzhou, China
4. Guangdong Province Key Laboratory of Computational Science, Sun Yat-sen University, Guangzhou, China

[*]The correspondence author: mcszhtsh@mail.sysu.edu.cn, zhjiajun@mail.sysu.edu.cn

Abstract: Transcription is a stochastic process that involves several downstream operations, which make it difficult to model and infer transcription kinetics from mature RNA numbers in individual cell. However, recent advances in single-cell technologies have enabled a more precise measurement of the fluctuations of nascent RNA that closely reflect transcription kinetics. In this paper, we introduce a general stochastic model to mimic nascent RNA kinetics with complex promoter architecture. We derive the exact distribution and moments of nascent RNA using queuing theory techniques, which provide valuable insights into the effect of the molecular memory created by the multistep activation and deactivation on the stochastic kinetics of nascent RNA. Moreover, based on the analytical results, we develop a statistical method to infer the promoter memory from stationary nascent RNA distributions. Data analysis of synthetic data and a realistic example, the *HIV-1* gene, verifies the validity of this inference method.

Introduction

Transcription is a complex biochemical process that leads to stochastic fluctuations in mRNA and subsequent protein production. The mature mRNA distribution data can be modeled using a two-state gene expression model, which has successfully revealed the mechanism of bursting and the non-Poisson nature of transcription [1-7]. However, the expression levels of mature mRNA are strongly influenced by various post-transcriptional processes such as splicing, nuclear export, and miRNA mediation [8-10]. Therefore, conventional gene-expression models focusing on mature mRNA may not be suitable for interpreting and inferring transcriptional processes, and more elaborate models need to be developed.

An alternative to measuring transcriptional products is to count the number of nascent RNAs (or the number of RNA polymerase molecules, RNAP) at the single-cell level. Recent experimental methods such as molecular fluorescence in situ hybridization (smFISH) can provide a steady-state distribution of nascent RNA [11-13]. Fitting experimental data of nascent RNA using stochastic models of nascent RNA allows for a more precise understanding of the regulatory steps involved in transcriptional processes [14, 15]. Therefore, numerous stochastic models have

been recently developed to simulate and infer nascent RNA kinetics [16-23]. For instance, Choubey, et al., [22] proposed a two-state model for the stochastic kinetics of nascent RNA, which assumed that a gene switches between an active and an inactive state, transcription initiation occurs only in the active state, and polymerase moves from one base pair to the next until reaching the end of the gene. The authors used queuing theory to derive the analytical expressions of transient and stationary expectations, as well as the variance of nascent RNA in the case of deterministic elongation time. Additionally, Heng, et al., [14] introduced a coarse-grained version of that model by neglecting RNAP movement between base pairs and derived the analytical expression of steady-state nascent RNA distribution.

It is worth noting that the two-state model assumes that the time spent in the active or inactive state follows an exponential distribution. While this assumption may be reasonable for bacteria [24], recent experimental studies have shown that the inactive periods of many genes in mammalian cells follow non-exponential distributions [25]. This suggests that the transition from an inactive to an active state is a multi-step process containing multiple inactive states. More generally, multiple transcriptional active and inactive states may exist in both prokaryotic and eukaryotic cells. For example, in *E. coli*, the Promoter for Repressor Maintenance (PRM) promoter of phage lambda is regulated by two different transcription factors (TFs) binding to two sets of operators, which can be brought together through DNA looping. Consequently, the number of regulatory states of the PRM promoter is up to 128 [26]. In contrast, eukaryotic promoters would be more complex since they involve nucleosomes competing with or being removed by TFs [27]. In addition to the conventional regulation by TFs, eukaryotic promoters can undergo epigenetic regulation through histone modifications [28-30], leading to complex promoter structures. A multi-step process can create the memory between molecular events, implying that waiting-time distributions between these events are non-exponential and lead to non-Markovian dynamics of the underlying system.

In recent years, gene expression models with complex promoter multi-state switching mechanisms have been widely studied theoretically [31-41]. However, most of these studies focus on the distributions of mature RNA and protein in a steady-state or transient state, and how promoter structure affects nascent RNA kinetics remains incompletely understood. Recently, several stochastic models of nascent RNA with multi-state promoter structure have been studied. Holehouse et al. [19,57] proposed a mechanistic model with 3-state promoter structure and obtained exact steady state distribution for nascent RNA using the delayed chemical master equation (dCME). Juraj et al. [18] developed a general framework that allowed the derivation of nascent RNA distributions for gene models with stochastic and multistep initiation but with deterministic elongation and termination. This framework is based on the assumption that the transcription initiation is a renewal process, which is true when there is only one active gene state and multiple inactive gene states. Also, we note that previous studies primarily examined the steady-state

or time-dependent behavior of mature mRNA or protein. Thus far, there has been still a lack of detailed studies on stochastic nascent RNA kinetics in the case of arbitrarily complex promoter structures. Therefore, the aim of this study is twofold: to develop a general stochastic model of nascent RNA with complex promoter structure and to derive the nascent RNA distribution to show how molecular memory impacts the stochastic kinetics of nascent RNA.

A general stochastic model of nascent RNA kinetics

We introduce a stochastic gene-expression model with complex promoter switching followed by a deterministic elongation process. Assume that the gene promoter has L distinct states, denoted by $S_1, S_2, ..., S_L$. Biophysically, these states correspond to different conformational states during chromatin remodeling or to different binding states for TFs. We use the matrix $\mathbf{W} = \left[\omega_{ij}\right]$ to describe the switching rates between promoter states, where elements ω_{ij} represent the transition rate from state S_i to state S_j, satisfying the relationship $\omega_{ii} = -\sum_{j \neq i} \omega_{ij}$. We also use the diagonal matrix $\mathbf{\Lambda} = \mathrm{diag}\left(k_1, k_2, ..., k_L\right)$ to describe the rates of transcription initiation, where k_i is the transcription initiation rate when the promoter state is at S_i state. Also assume that the elongation of a single transcript is deterministic and the elongation speed is a constant k_{elo} kb/min. Consequently, the elongation time is $t_{\mathrm{elo}} = l_{\mathrm{gene}} / k_{\mathrm{elo}}$, where l_{gene} is the gene length. Although a transcriptional elongation process would be very complex, e.g., it would involve pausing and backtracking of polymerases along the gene body, our assumption of deterministic elongation is justified for a vast number of genes [22,42]. We set the survival function of the nascent RNA as

$$H(t) = \begin{cases} 1, & t \leq t_{\mathrm{elo}}, \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

Let $P_i(n,t)$ denote the probability that the number of nascent RNAs is n at time t, given that promoter state is i at time 0 and the number of nascent RNA is 0 at time 0, i.e.,

$$P_i\left(n,t\right) = \Pr\left\{N(t) = n \middle| N(0) = 0, U(0) = i\right\}, \quad i = 1, 2, ..., L, \tag{2}$$

where $U(t)$ and $N(t)$ represent the promoter state and the nascent RNA number at time t, respectively. For convenience, we denote $\boldsymbol{P}\left(n,t\right) = \left(P_1\left(n,t\right), ..., P_L\left(n,t\right)\right)^{\mathrm{T}}$, where the symbol 'T' represents transpose. Using the method proposed in our recent study [43], we can derive the following so-called generalized master equation for $\boldsymbol{P}(n,t)$

$$\frac{\partial}{\partial t}\boldsymbol{P}\left(n,t\right) = \left[\mathbf{W} - H(t)\mathbf{\Lambda}\right]\boldsymbol{P}\left(n,t\right) + H\left(t\right)\mathbf{\Lambda}\boldsymbol{P}\left(n-1,t\right). \tag{3}$$

Let $\boldsymbol{G}\left(z,t\right) = \left(G_1\left(z,t\right), G_2\left(z,t\right), ..., G_L\left(z,t\right)\right)^{\mathrm{T}}$ be the probability generating function corresponding to $\boldsymbol{P}\left(n,t\right)$, i.e., $\boldsymbol{G}\left(z,t\right) = \sum_{n=0}^{\infty} z^n \boldsymbol{P}\left(n,t\right)$. Then Eq. (3) can be transformed to

$$\frac{\partial}{\partial t}\boldsymbol{G}\left(z,t\right) = \left[\mathbf{W} + \left(z-1\right)H\left(t\right)\mathbf{\Lambda}\right]\boldsymbol{G}\left(z,t\right), \tag{4}$$

with the initial condition $\boldsymbol{G}\left(z,0\right) = \boldsymbol{e}$, where $\boldsymbol{e} = \left(1,1,...,1\right)^{\mathrm{T}}$.

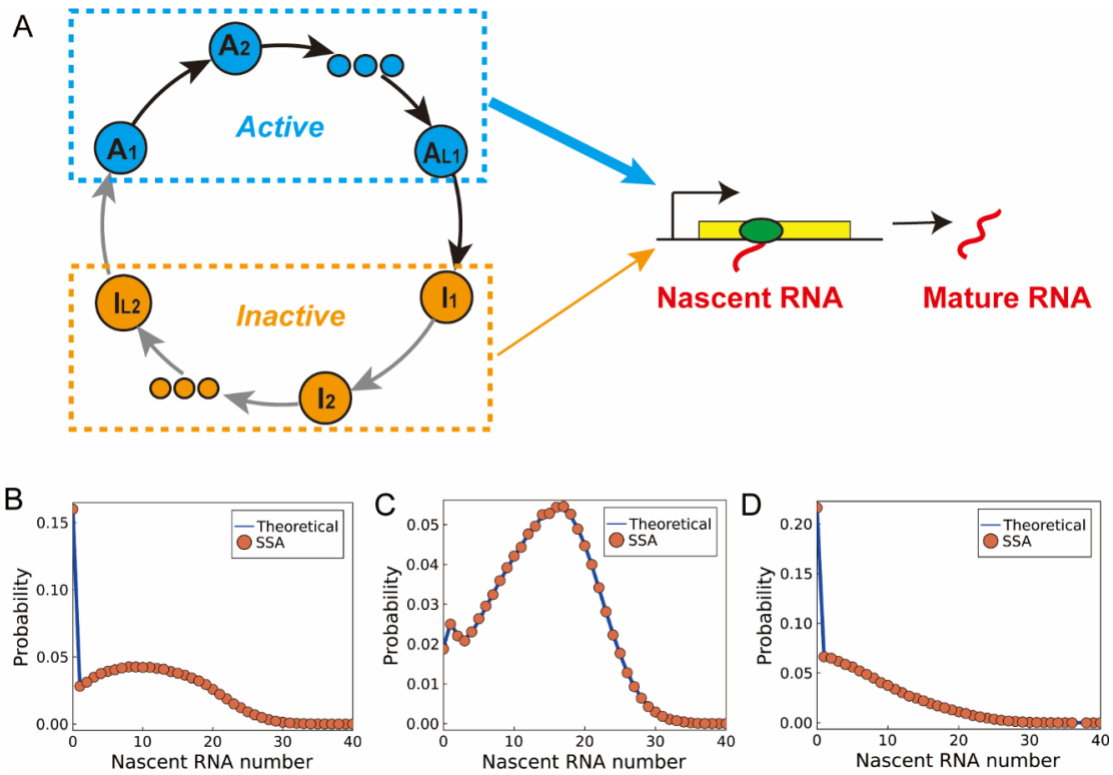

Fig. 1. Schematic diagram for a model of nascent RNA with the cyclic promoter structure, where the blue and red cycles represent transcription active (ON) and inactive (OFF) states, respectively. (B)(C)(D) The steady-state nascent RNA distributions obtained by the matrix-form solution and the stochastic simulation algorithm (SSA). The parameter values are set as: (A) $u=1, v=1.5, L_1=L_2=2, k_1=20, k_0=0$. (B) $u=0.7,\ v=1.5, L_1=L_2=2, k_1=20, k_0=1$. (C) $u=2, v=1.5, L_1=1, L_2=3, k_1=20, k_0=0$.

As a special case, we consider the model with a cyclic promoter structure, i.e., the promoter proceeds sequentially through several irreversible active (ON) state and inactive (OFF) states, which altogether form a cycle (Fig. 1 (A)). Note that this cyclic model can be used to describe many processes involved in transcription such as recruitment of various polymerases and transcription factors, assembly of the pre-initiation complexes, chromatin remodeling, and histone modification [36]. For simplicity, we assume that the transcription initiation rate is k_1 at each active state and k_0 at each inactive state, where $k_1 > k_0$. Specifically, we assume the gene promoter has L states, which are distinguished into L_1 active states and L_2 inactive states. Denote by $A_1,...,A_{L_1}$ ON states and by $I_1,...,I_{L_2}$ OFF states. The corresponding biochemical reactions are listed in Table 1. Note that if $L_1=1$ and $L_2>1$, the cyclic model becomes the multi-OFF mechanism (i.e., the promoter has multiple OFF states but only has one ON state). If $L_1>1$ and $L_2=1$, the cyclic model becomes the multi-ON mechanism (i.e., the promoter has multiple ON states but only has one OFF state). In particular, if $L_1=1$ and $L_2=1$, the cyclic model reduces to the classical two-state model. In a word, the cyclic model includes many transcription models being previously studied.

Table 1 Biochemical reactions for a transcription model to be studied.

Reaction	**Description**
$\mathrm{A}_i \xrightarrow{\omega_{i,i+1}^{(\mathrm{on})}} \mathrm{A}_{i+1}$, $i=1,2,\ldots,L_1-1$	Transitions between active states
$\mathrm{I}_j \xrightarrow{\omega_{j,j+1}^{(\mathrm{off})}} \mathrm{I}_{j+1}$, $j=1,2,\ldots,L_2-1$	Transitions between inactive states
$\mathrm{A}_{L_1} \xrightarrow{\omega_{L_1,1}^{(\mathrm{on})}} \mathrm{I}_1$, $\mathrm{I}_{L_2} \xrightarrow{\omega_{L_2,1}^{(\mathrm{off})}} \mathrm{A}_1$	Transitions between active and inactive states
$\mathrm{A}_i \xrightarrow{k_1} \mathrm{A}_i + \mathrm{nascentRNA}$	Transcription initiation at the active state
$\mathrm{I}_j \xrightarrow{k_0} \mathrm{I}_j + \mathrm{nascentRNA}$	Transcription initiation at the inactive state
$\mathrm{nascent\ RNA} \underset{t_{\mathrm{elo}}}{\Rightarrow} \varnothing$	Nascent RNA elongation

To make an equitable comparison between models with different values of L, we assume that the transition rates at the ON (OFF) states are the same, i.e., $\omega_{1,2}^{(\mathrm{on})}=\ldots=\omega_{L_1-1,L_1}^{(\mathrm{on})}=\omega_{L_1,1}^{(\mathrm{on})}=L_1u$, $\omega_{1,2}^{(\mathrm{off})}=\ldots=\omega_{L_2-1,L_2}^{(\mathrm{off})}=\omega_{L_2,1}^{(\mathrm{off})}=L_2v$, where u and v are two constants. With these assumptions, the dwell time at the ON state, τ_{on}, obeys an Erlang distribution with shape parameter L_1 and mean $1/u$, and the dwell time at the OFF state, τ_{off}, also follows an Erlang distribution with shape parameter L_2 and mean $1/v$. Note that if the noise of the dwell time is characterized by the coefficient of variation, it is $1/L_1$ for the dwell time τ_{on}, and $1/L_2$ for the dwell time τ_{off}. As $L_i\to\infty$, this noise tends to zero and the dwell time converges to a deterministic value and the molecular memory is the strongest in each case. In the following, L_1 and L_2 are called as memory index of promoter activation and deactivation, respectively.

Distribution of nascent RNA in steady state

In a previous study [43], we used the finite state projection (FSP) method to calculate a numerical nascent RNA distribution $\boldsymbol{P}(n,t)$ based on Eq. (3). Although this method can ensure that the approximation error is within a range of some pre-specified tolerance level, it is inefficient to obtain an accurate nascent RNA distribution in the steady state since it often takes a long time when a system reaches the steady state. In this section, we present two forms of the exact solutions for the stationary distribution of nascent RNA: Laplace-form solution and matrix-form solution.

Laplace-form solution

The closed-form solution of Eq. (4) is generally difficult to obtain. Under the assumption that the elongation time is deterministic, we find that the solution of Eq. (4) takes the form

$$\boldsymbol{G}(z,t)=\begin{cases}\exp\left[\left(\mathbf{W}+(z-1)\boldsymbol{\Lambda}\right)t\right]\boldsymbol{e} & t\le t_{\mathrm{elo}}\\ \exp\left[\mathbf{W}\left(t-t_{\mathrm{elo}}\right)\right]\exp\left[\left(\mathbf{W}+(z-1)\boldsymbol{\Lambda}\right)t_{\mathrm{elo}}\right]\boldsymbol{e} & t>t_{\mathrm{elo}}\end{cases}. \tag{5}$$

Note that matrix exponential function $\exp\left[\mathbf{W}\left(t-t_{\mathrm{elo}}\right)\right]$ tend to $\boldsymbol{e\pi}$ when $t\to\infty$, where the row vector $\boldsymbol{\pi}$ denotes the

steady state of the promoter, satisfying $\boldsymbol{\pi}\mathbf{W}=\mathbf{0}$ and $\boldsymbol{\pi}\boldsymbol{e}=1$. Since the steady-state distribution of nascent RNA is independent of the initial state: $P(n)\equiv P_1(n,\infty)=...=P_L(n,\infty)$, we have $G(z)\equiv G_1(z,\infty)=G_2(z,\infty)=...=G_L(z,\infty)$. Combining these results with Eq. (5), we can derive the analytical expression of the stationary nascent RNA distribution as follows

$$G(z)=\boldsymbol{\pi}\exp\left[\left(\mathbf{W}+(z-1)\boldsymbol{\Lambda}\right)t_{\text{elo}}\right]\boldsymbol{e}, \tag{6}$$

which is a function of t_{elo} in principle, re-denoted by $G(z,t_{\text{elo}})$. The Laplace transform of $G(z,t_{\text{elo}})$ with respect to t_{elo} is given by

$$G^*(z,s)=\boldsymbol{\pi}\left(s\mathbf{I}-\mathbf{U}-\boldsymbol{\Lambda}z\right)^{-1}\boldsymbol{e}, \tag{7}$$

where $\mathbf{U}=\mathbf{W}-\boldsymbol{\Lambda}$. Note that the function $G^*(z,s)$ has the following Taylor expansion

$$G^*(z,s)=\sum_{n=0}^{\infty}\boldsymbol{\pi}\left[\left(s\mathbf{I}-\mathbf{U}-\boldsymbol{\Lambda}z\right)^{-1}\boldsymbol{\Lambda}\right]^n\left(s\mathbf{I}-\mathbf{U}\right)^{-1}\boldsymbol{e}z^n. \tag{8}$$

On the other hand, according to the definition of probability generating function, we know that function $G^*(z,s)$ can be written as $G^*(z,s)=\sum_{n=0}^{\infty}P^*(n,s)z^n$, where $P^*(n,s)$ is the Laplace transform of $P(n)$ (which is also a function of t_{elo} in principle) with respect to t_{elo}. Therefore, we obtain the Laplace-form solution of stationary nascent RNA distribution

$$P^*(n,s)=\boldsymbol{\pi}\left[\left(s\mathbf{I}-\mathbf{U}\right)^{-1}\boldsymbol{\Lambda}\right]^n\left(s\mathbf{I}-\mathbf{U}\right)^{-1}\boldsymbol{e}. \tag{9}$$

Then, the inverse Laplace transform gives the explicit expression of the probability distribution $P(n)$. For $n=0$, we have

$$P(n=0)=\boldsymbol{\pi}\exp\left(\mathbf{U}t_{\text{elo}}\right)\boldsymbol{e}. \tag{10}$$

For $n>0$, it is possible, albeit tedious, to derive the analytical expression of $P(n)$. In any case, we can numerically calculate the probability distribution with the above formula.

We can also derive the analytical expression for the stationary moments of the nascent RNA with a general promoter switching. Let $u^{(k)}$ denotes the *k*-th factorial moments of nascent RNA number N in the steady state, i.e., $u^{(k)}=E\left[N(N-1)...(N-k+1)\right]$. Applying the formula $u^{(k)}=\partial_z^k G(z)|_{z=1}$ to Eq. (6), we obtain the first two factorial moments (see refs. [44,45] for details)

$$u^{(1)}=\boldsymbol{\pi}\boldsymbol{\Lambda}\boldsymbol{e}t_{\text{elo}}, \tag{11}$$

and

$$u^{(2)}=t_{\text{elo}}^2\left(\boldsymbol{\pi}\boldsymbol{\Lambda}\boldsymbol{e}\right)^2+2t_{\text{elo}}\left[\boldsymbol{\pi}\boldsymbol{\Lambda}\left(\tau\boldsymbol{e}\boldsymbol{\pi}-\mathbf{W}\right)^{-1}\boldsymbol{\Lambda}\boldsymbol{e}-\tau^{-1}\left(\boldsymbol{\pi}\boldsymbol{\Lambda}\boldsymbol{e}\right)^2\right]+2\boldsymbol{\pi}\boldsymbol{\Lambda}\left(\exp(\mathbf{W}t_{\text{elo}})-\mathbf{I}\right)\left(\tau\boldsymbol{e}\boldsymbol{\pi}-\mathbf{W}\right)^{-2}\boldsymbol{\Lambda}\boldsymbol{e}, \tag{12}$$

where $\tau \neq 0$. Therefore, the Fano factor, which is defined as the ratio of the variance over the mean, is given by

$$FF = 1 + 2\frac{\boldsymbol{\pi\Lambda}(\tau\boldsymbol{e\pi} - \mathbf{W})^{-1}\boldsymbol{\Lambda e} - \tau^{-1}(\boldsymbol{\pi\Lambda e})^2}{\boldsymbol{\pi\Lambda e}} + \frac{2}{t_{\text{elo}}}\frac{\boldsymbol{\pi\Lambda}(\exp(\mathbf{W}t_{\text{elo}}) - \mathbf{I})(\tau\boldsymbol{e\pi} - \mathbf{W})^{-2}\boldsymbol{\Lambda e}}{\boldsymbol{\pi\Lambda e}}. \tag{13}$$

Note that the value of FF does not depend on the choice of τ. Setting $\tau = -1$, we have

$$FF = 1 + 2\frac{(\boldsymbol{\pi\Lambda e})^2 - \boldsymbol{\pi\Lambda}(\boldsymbol{e\pi} + \mathbf{W})^{-1}\boldsymbol{\Lambda e}}{\boldsymbol{\pi\Lambda e}} + \frac{2}{t_{\text{elo}}}\frac{\boldsymbol{\pi\Lambda}(\exp(\mathbf{W}t_{\text{elo}}) - \mathbf{I})(\boldsymbol{e\pi} + \mathbf{W})^{-2}\boldsymbol{\Lambda e}}{\boldsymbol{\pi\Lambda e}}. \tag{14}$$

In particular, as $t_{\text{elo}} \to \infty$, we have

$$FF = 1 + 2\frac{(\boldsymbol{\pi\Lambda e})^2 - \boldsymbol{\pi\Lambda}(\boldsymbol{e\pi} + \mathbf{W})^{-1}\boldsymbol{\Lambda e}}{\boldsymbol{\pi\Lambda e}}. \tag{15}$$

Matrix-form solution

Although Eq. (9) allows us to calculate the closed-form solution of stationary nascent RNA distribution for some special cases, we find that it is inefficient to calculate the stationary distribution numerically in general cases. Therefore, we try to derive the matrix-form solution of the stationary nascent RNA distribution to facilitate numerical computation. Assuming that there exists a large N such that $P_i(n,t) = 0$ for $i = 1, 2, \ldots, K$ and $n > N + 1$. We denote

$$\boldsymbol{Q}(t) = \left(P_1(0,t), \ldots, P_L(0,t), P_1(1,t), \ldots, P_L(1,t), \ldots, P_1(N,t), \ldots, P_L(N,t)\right)^{\mathrm{T}}. \tag{16}$$

For $t \le t_{\text{elo}}$, Eq. (3) can be solved by the following ordinary differential equation (ODE) up to time t

$$\frac{\partial}{\partial t}\boldsymbol{Q}(t) = \mathbf{A}\boldsymbol{Q}(t), \tag{17}$$

where matrix $\mathbf{A}$ takes the form

$$\mathbf{A} = \begin{bmatrix} \mathbf{W}-\boldsymbol{\Lambda} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \boldsymbol{\Lambda} & \mathbf{W}-\boldsymbol{\Lambda} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{W}-\boldsymbol{\Lambda} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{W}-\boldsymbol{\Lambda} \end{bmatrix}_{L(N+1)\times L(N+1)}, \tag{18}$$

and the initial condition is

$$\boldsymbol{Q}(0) = \left(\underbrace{1, 1, \ldots, 1}_{L}, \underbrace{0, 0, \ldots, 0}_{L}, \ldots, \underbrace{0, 0, \ldots, 0}_{L}\right)^{\mathrm{T}}. \tag{19}$$

For $t \le t_{\text{elo}}$, the solution of the above equation is

$$\boldsymbol{Q}(t) = \exp(\mathbf{A}t)\boldsymbol{Q}(0). \tag{20}$$

Therefore, we have

$$\boldsymbol{P}(n,t) = \mathbf{E}_{n+1}\exp(\mathbf{A}t)\boldsymbol{Q}(0), \tag{21}$$

where $\mathbf{E}_n = \boldsymbol{e}_n \otimes \boldsymbol{e}$, $\otimes$ represents the Kronecker production, $\boldsymbol{e}_n$ denotes the N-dimension vector where the n-th coordinate is 1 and the others are zero, i.e., $\boldsymbol{e}_n = (0,...,1,...,0)_{1\times(N+1)}$, and $\boldsymbol{e} = (1,1,...,1)_{1\times L}$.

For $t > t_{\text{elo}}$, the solution of Eq. (3) is

$$\boldsymbol{P}(n,t) = \exp\left(\mathbf{W}(t - t_{\text{elo}})\right)\boldsymbol{P}(n, t_{\text{elo}}). \tag{22}$$

As $t \to \infty$, we obtain the following matrix-form solution for the stationary nascent RNA distribution

$$P(n) = \boldsymbol{\pi}\mathbf{E}_{n+1}\exp(\mathbf{A}t_{\text{elo}})\boldsymbol{Q}(0). \tag{23}$$

To verify the validity of the theoretical solution given by Eq. (23), we consider a four-state cyclic model of gene expression with deterministic elongation time. Fig. 1 (B)(C)(D) exhibits numerical [46] and theoretical results for the steady-state distribution of nascent RNA for three different sets of parameter values. From these diagrams, we observe that both numerical and theoretical results are well consistent.

The two-state gene expression model: A case study

As a special case, we consider a general two-state model of gene expression with a deterministic elongation time, where we assume that the promoter switches between two states: an inactive state A and an active state I. The transcription initial rate is k_0 at the I state and k_1 at the A state. The two-state gene expression model is a special case of the cyclic model with $L_1 = 1$ and $L_2 = 1$. The reaction scheme for this process is listed below

$$\begin{aligned}
&\mathrm{I} \underset{\omega_{21}}{\overset{\omega_{12}}{\rightleftharpoons}} \mathrm{A},\\
&\mathrm{A} \xrightarrow{k_1} \text{nascent RNA},\\
&\mathrm{I} \xrightarrow{k_0} \text{nascent RNA},\\
&\text{nascent RNA} \underset{t_{\text{elo}}}{\Rightarrow} \varnothing
\end{aligned} \tag{24}$$

Previous studies have derived the analytical solution of nascent RNA distribution under the assumption that the transcription initiation only occurs at the active states [14,18], i.e., $k_1 > 0$ and $k_0 = 0$. In this section, we derive the steady state distribution of nascent RNA for a general two-state gene expression model with leaky transcription, i.e., $k_1 > k_0 \geq 0$.

Applying Eq. (9), and using Sylvster's matrix theorem to calculate $\left[(s\mathbf{I} - \mathbf{U})^{-1}\boldsymbol{\Lambda}\right]^n$, we can finally obtain the Laplace-form solution of nascent RNA distribution for the general two-state model as following

$$P^*(0,s) = \frac{s + \lambda_1 + \lambda_2 - (k_1\omega_{12} + k_0\omega_{21})/(\omega_{12} + \omega_{21})}{(s + \lambda_1)(s + \lambda_2)}, \tag{25}$$

$$P^*(n,s)=\frac{\gamma_1^n-\gamma_2^n}{\gamma_1-\gamma_2}\frac{k_0\omega_{21}(s+k_1+\omega_{12}+\omega_{21})^2+k_1\omega_{12}(s+k_0+\omega_{12}+\omega_{21})^2}{(\omega_{12}+\omega_{21})\left[(s+\lambda_1)(s+\lambda_2)\right]^{n+1}}$$
$$-\frac{\gamma_1\gamma_2(\gamma_1^{n-1}-\gamma_2^{n-1})}{\gamma_1-\gamma_2}\frac{s+\lambda_1+\lambda_2-(k_1\omega_{12}+k_0\omega_{21})/(\omega_{12}+\omega_{21})}{\left[(s+\lambda_1)(s+\lambda_2)\right]^{n+1}},\quad \text{for } n>0, \tag{26}$$

where $-\lambda_1$ and $-\lambda_2$ are roots of the quadratic equation

$$(s+\lambda_1)(s+\lambda_2)=(s+k_0+\omega_{12})(s+k_1+\omega_{21})-\omega_{12}\omega_{21}=0, \tag{27}$$

and γ_1 and γ_2 are roots of the quadratic equation

$$(x-\gamma_1)(x-\gamma_2)=\left[x-k_0(s+k_1+\omega_{21})\right]\left[x-k_1(s+k_0+\omega_{12})\right]-k_0k_1\omega_{12}\omega_{21}=0. \tag{28}$$

When $k_0=0$, the roots of the above equation are $\gamma_1=0$, $\gamma_2=k_1(s+\omega_{12})$. Therefore,

$$P^*(0,s)=\frac{s+\lambda_1+\lambda_2-k_1\omega_{12}/(\omega_{12}+\omega_{21})}{(s+\lambda_1)(s+\lambda_2)}, \tag{29}$$

$$P^*(n,s)=\frac{k_1\omega_{12}(s+\omega_{12}+\omega_{21})^2\left[k_1(s+\omega_{12})\right]^{n-1}}{(\omega_{12}+\omega_{21})\left[(s+\lambda_1)(s+\lambda_2)\right]^{n+1}}, \tag{30}$$

which reproduce the result in a previous study [18].

Using Eq.(13) and setting $\tau=\omega_{12}+\omega_{21}$, we obtain the Fanor factor of the nascent RNA number in the steady state

$$FF=1+\frac{2\omega_{12}\omega_{21}(k_0-k_1)^2}{(\omega_{12}+\omega_{21})^2(\omega_{12}k_0+\omega_{21}k_1)}+\frac{2\omega_{12}\omega_{21}(k_0-k_1)^2}{(\omega_{12}+\omega_{21})^3(\omega_{12}k_0+\omega_{21}k_1)}\frac{e^{-(\omega_{12}+\omega_{21})t_{elo}}-1}{t_{elo}}. \tag{31}$$

When $k_1>0$ and $k_0=0$, we have

$$FF=1+\frac{2\omega_{12}k_1}{(\omega_{12}+\omega_{21})^2}+\frac{2\omega_{12}k_1}{(\omega_{12}+\omega_{21})^3}\frac{e^{-(\omega_{12}+\omega_{21})t_{elo}}-1}{t_{elo}}, \tag{32}$$

which reproduces the result in a previous study [23].

Effect of molecular memory on nascent RNA distribution

Previous studies have demonstrated that molecular memory can modulate the distribution of mature RNA in both steady state and transient state [31-34]. However, the effect of molecular memory on the stationary distribution of nascent RNA remains incompletely understood. In this section, we investigate the influence of molecular memory on the stationary nascent RNA distribution using the above cyclic model. Note that this model considers two types of molecular memory: the memory during the active period (characterized by index L_1 and called activation memory) and the memory during the inactive period (characterized by index L_2 and called deactivation memory).

To understand how activation and deactivation memory affect the stationary nascent RNA distribution, we first

focus on the cyclic model with fast or slow promoter switching. Under the slow-switching condition, i.e., $u \to 0$, $v \to 0$, it follows from Eq. (6) that $G(z) = \pi_0 \exp\left(k_0 t_{\text{elo}}(z-1)\right) + \pi_1 \exp\left(k_1 t_{\text{elo}}(z-1)\right)$ with $\pi_0 = v/(u+v)$ and $\pi_1 = u/(u+v)$, indicating that the stationary nascent RNA distribution is a mixture Poisson distribution. In contrast, under the fast-switching condition, i.e., $u \to \infty$, $v \to \infty$, Eq. (6) becomes $G(z) = \exp\left((\pi_0 k_0 + \pi_1 k_1) t_{\text{elo}}(z-1)\right)$, suggesting that the stationary nascent RNA distribution is a Poisson distribution. Therefore, both activation and deactivation memory do not alter the stationary distribution of nascent RNA in the limiting cases of fast and switching regimes.

We further investigate the impact of activation and deactivation memory on the stationary nascent RNA distribution in the case that promoter switching is neither too fast nor too slow. To illustrate this, we present the $u-v$ phase diagram for the two-state model with $L_1 = 1$ and $L_2 = 1$ (Fig. 2(A)), the multi-ON model with $L_1 = 5$ and $L_2 = 1$ (Fig. 2(B)) and the multi-OFF model with $L_1 = 1$ and $L_2 = 5$ (Fig. 2(C)). In all cases, the stationary nascent RNA distribution exhibits three main features: unimodal distribution with zero peak (Region I or Fig. 2(D)), unimodal distribution with non-zero peak (Region II or Fig. 2(E)), and bimodal distribution with a zero peak and a non-zero peak (Region III or Fig. 2(F)). Comparing these phase diagrams across the different models, we observe molecular memory can influence the size of each region: (1) increasing the activation memory L_1 leads to the transition from unimodal distribution of type I to bimodal distribution, and results in a significant reduction of region I ; (2) increasing deactivation memory L_2 causes the transition from bimodal distribution to the unimodal distribution of type II, and leads to an enlargement of Region II. In addition, we discover that increasing the deactivation memory L_2 in the mult-OFF model induces a trimodal distribution with a zero peak and two non-zero peaks ((type IV or Fig. 2(G)), which has not been reported in previous studies.

In addition, we demonstrate stationary distributions of nascent RNA under different choices of L_1 and L_2, referring to Fig. 2(H). From this figure, we observe that increasing deactivation memory L_2 tends to eliminate the zero peak, thus transforming the bimodal distribution into a unimodal distribution. When $L_1 = 1$, increasing L_2 can induce a new non-zero peak (the upper row in Fig. 2(H)), leading to the trimodal distribution in Region IV. When L_2 is sufficient large, i.e., $L_2 = 15$, the peak at zero almost disappears and the trimodal distribution becomes a bimodal distribution with two non-zero peaks. When both activation and deactivation memory are larger than 1, i.e., $L_1 > 1$ and $L_2 > 1$, two non-zero peaks of the bimodal distribution tend to merge together, leading to a unimodal distribution.

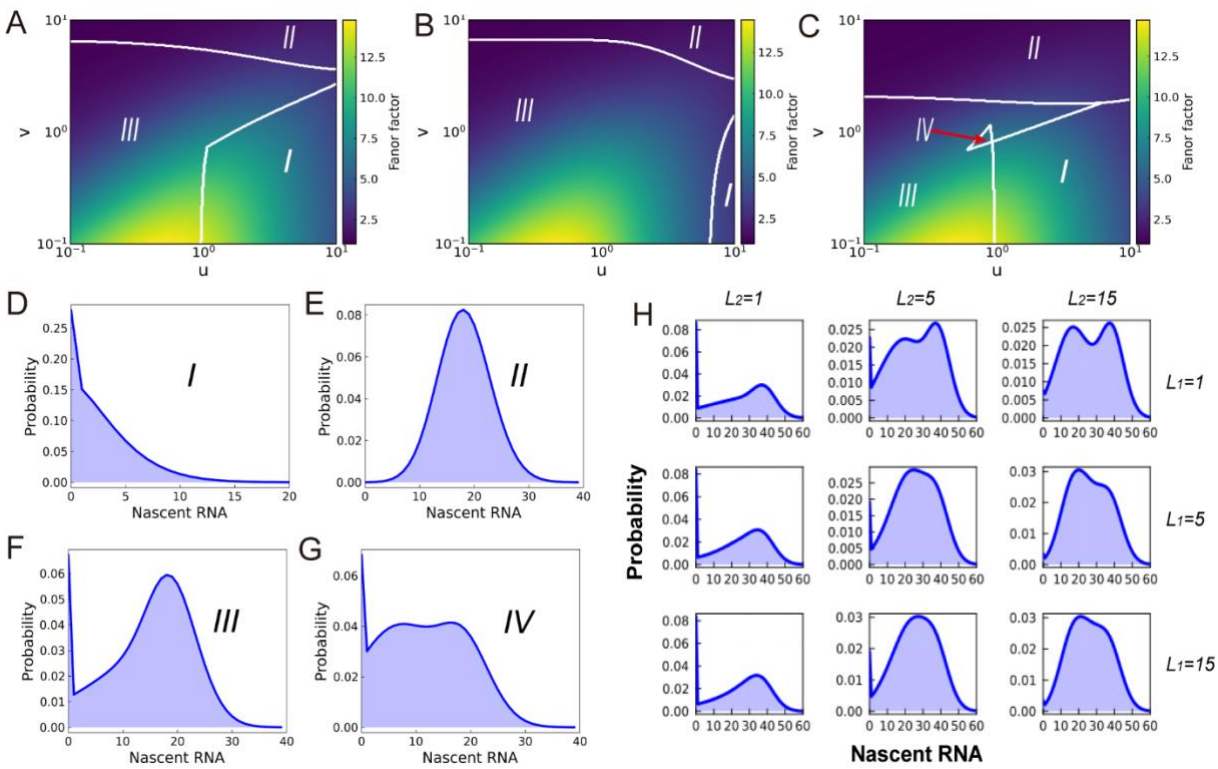

Fig. 2. The effect of molecular memory on the nascent RNA distribution in the steady state. (A) The *u*-*v* phase diagram for the two-state model $(L_1 = 1, L_2 = 1)$; (B) The *u*-*v* phase diagram for the multi-ON model $(L_1 = 5, L_2 = 1)$; (C) The *u*-*v* phase diagram for the multi-OFF model $(L_1 = 1, L_2 = 5)$; other parameter values are set as $k_1 = 20, k_0 = 0$; (D)(E)(F)(G) The distributions of different types, where the parameter values for each type are set as those for Type I $u = 12, v = 2$; Type II $u = 1, v = 10$; Type III $u = 0.5, v = 1.5$; Type IV $u = 0.8, v = 1.2, L_2 = 8$. The other parameter values are the same as those in (A); (D) the steady state of nascent RNA distribution under different choices of L_1 and L_2. The other parameters values are set as: $u = 0.8, v = 1.5, k_1 = 40, k_0 = 0$.

To understand why increasing the activation memory L_1 can transform the unimodal distribution of type I to the bimodal distribution, we investigate the stationary nascent RNA distribution in the burst regime ($u \ll v$). From Fig. 3(C)(D), we observe that increasing activation memory L_1 diminishes the probability $P(n=0)$ and $P(n=1)$ simultaneously, concurrently giving rise to a valley between two peaks. This implies that activation memory L_1 can induce bimodal distribution in the burst regime. However, within the burst regime, increasing L_2 significantly reduces the probability $P(n=0)$ while increase the probability $P(n=1)$ (Fig. 3(A)(B)), and the unimodal distribution remains unchanged. Hence, the activation and deactivation memories exhibit distinct modulation effects on nascent RNA distribution. It should be noted that this characteristic may serve as an indicator for distinguishing activation memory form station nascent RNA distribution, i.e., when the stationary nascent RNA displays bimodal distribution in the bursting regime, it suggests that the activation memory may be larger than 1.

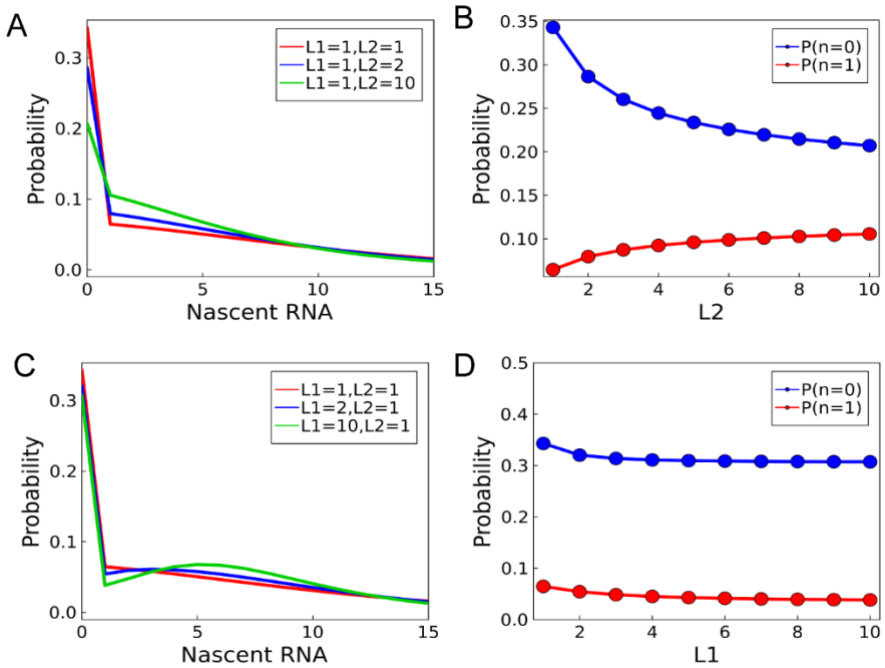

Fig. 3. The effect of molecular memory on the stationary nascent RNA distribution in the burst regime $(u \gg v)$. (A) The steady-state distributions of nascent RNA for different values of L_2; (B) The probabilities of $P(n=0)$ and $P(n=1)$ as functions of L_2; (C) The steady-state distribution of nascent RNA for different value of L_1; (D) The probabilities of $P(n=0)$ and $P(n=1)$ as functions of L_1. The other parameter values are set as: $u=3, v=1, k_1=20, k_0=0$.

By summarizing the above analysis, we know that the molecular memories during active and inactive periods can modulate the steady-state distribution of nascent RNA cooperatively. Increasing the deactivation memory L_1 tends to disrupt the zero peak in the stationary nascent RNA distribution, while enhancing the activation memory L_2 can induce a bimodal distribution in the bursting regime. Depending on the sizes of the activation and deactivation memories, distinct shapes of stationary nascent RNA distribution may emerge, including bimodal distribution with a zero peak and non-zero peak, bimodal distribution with two non-zero peaks, and trimodal distribution, unimodal distribution with zero peak, and unimodal distribution with non-zero peak.

Inferring the memory of promoter switching from a nascent RNA distribution

The nascent RNA distribution can be obtained through experimental techniques such as mRNA-FISH labeling or single-cell RNA sequencing (scRNA-seq). An intriguing question is whether the memory of promoter switching (called promoter memory for convenience) can be inferred from the stationary nascent RNA distribution. The

question of how model parameters are inferred from a given distribution has been explored in previous studies [47-50]; and the two-state non-Markov model has commonly been employed to fit the nascent RNA distribution and to elucidate the underlying mechanism of gene expression [14,15]. In this section, we propose a method based on the above analytical results to infer the promoter memory from the stationary nascent RNA distribution.

We assume that M cells are measured simultaneously at an one-time experiment and cell i has n_i copies of nascent RNA in the steady state, and the stationary distribution of nascent RNA is $P(n,\theta)$. We also assume that the promoter consists of L_1 active states and L_2 inactive states. The transition rates at the ON (OFF) states are the same, i.e., $\omega_{1,2}^{(\mathrm{on})} = \ldots = \omega_{L_1-1,L_1}^{(\mathrm{on})} = \omega_{L_1,1}^{(\mathrm{on})} = L_1 u$, $\omega_{1,2}^{(\mathrm{off})} = \ldots = \omega_{L_2-1,L_2}^{(\mathrm{off})} = \omega_{L_2,1}^{(\mathrm{off})} = L_2 v$. The initiation of nascent RNA only occurs in the active state with rate k_1. Since it is impossible to infer the elongation time and other parameters simultaneously from a stationary distribution, we fix the elongation time $t_{\mathrm{elo}} = 1$. Let $\theta = (L_1, L_2, u, v, k_1)$ be the parameter vector of the cyclic model to be inferred. We define the total likelihood function as

$$\mathrm{L}(\theta) = \prod_{i=1}^{M} P(n_i, \theta), \tag{33}$$

where $P(n_i;\theta)$ is calculated according to Eq. (23). We use an efficient gradient-free optimization algorithm (i.e., adaptive differential optimizer, referring to BlackBoxOptim.jl [51]) to find the optimal parameters by maximizing the likelihood function $\mathrm{L}(\theta)$.

To evaluate the accuracy of the method, we generate a set of synthetic data comprising 2000 cells using the stochastic simulation method [46]. It is worth noting that the parameter set for this synthetic data is sampled in the burst regime, satisfying the condition $u \gg v$. By fitting both the cyclic model and classical two-state model to this synthetic data, we observe that the former outperforms the latter in terms of goodness of fit (Fig. 4(A)). In order to assess estimation accuracy, bootstrapping resampling (i.e., random sampling with replacement) is performed to obtain a distribution of the estimated parameters, which can be utilized to estimate the standard errors and the confidence intervals of these estimated parameters [52]. As shown in Fig. 4(B), the activation memory L_1 can be accurately estimated. Furthermore, Fig.4(C) demonstrates that all 'true' parameter values fall within the 95% confidence interval estimated by the bootstrapping, indicating the precise inference of the 'true' parameters from the stationary nascent RNA distribution in the burst regime. However, extensive numerical simulations reveal limitations in accurately inferring promoter memory solely based on stationary nascent RNA distribution under certain conditions. For instance, when comparing promoter switching rates with elongation process rates—either sufficiently fast or slow—the nascent RNA distribution becomes independent of promoter memory.

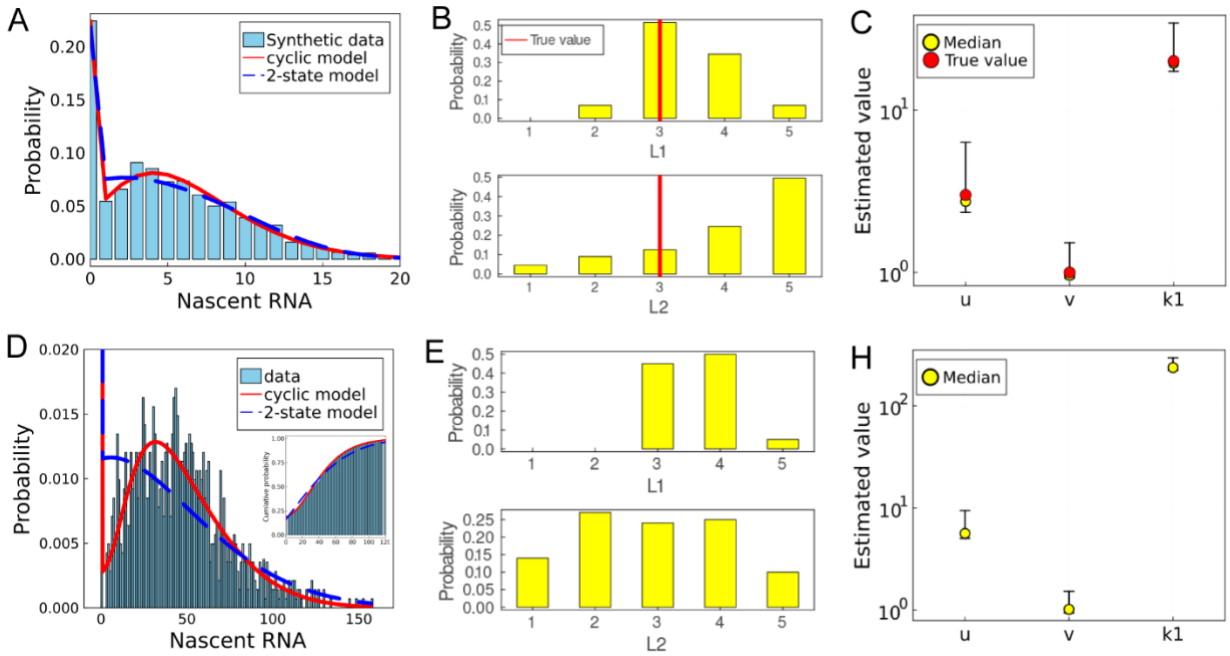

Fig. 4. Inference of promoter memory from nascent RNA distribution. (A)The fitted distributions for synthetic data using the cyclic model and the two-state model. The 'true' parameter set for the synthetic data is: $L_1 = 3$, $L_2 = 3$, $u = 3$, $v = 1$, $k_1 = 20$; (B) the distribution of promoter memory estimated by bootstrap resampling for the synthetic data. The red line represents the 'true' value; (C) the 95% confidence interval of the kinetic parameters estimated by bootstrap resampling for the synthetic data. The red circle represents the 'true' value. (D) The fitted distributions for experimental data of *HIV-1* gene using the cyclic model and the two-state model. The estimated parameters of the cyclic model are: $L_1 = 3$, $L_2 = 4$, $k_1 = 242.66$, $u = 5.12$, $v = 0.95$; the estimated parameters of the two-state model are: $k_1 = 122.21$, $u = 3.41$, $v = 1.50$. The inset is the cumulative probability function that corresponds to the distribution estimated by the cyclic and two-state models; (E) the distribution of the promoter memory estimated by bootstrap resampling for the experimental data; (H) the 95% confidence interval of kinetic parameters estimated by bootstrap resampling for the experimental data.

Finally, we apply the aforementioned inference method to a realistic example. Specifically, we aim to infer the promoter memory from the stationary nascent RNA distribution for the *HIV-1* gene in cells expressing a high level of Tat which is measured by Katjana et.al. [20]. It has been reported that the expression of the *HIV-1* gene exhibits bursting kinetics, which are believed to play a crucial role in the control of latency. We fit the experimental data using the two-state model and cyclic model, and present the distribution of the parameters estimated by bootstrapping resampling in Fig. 4(E)(H). Notably, this analysis reveals that the activation memory L_1 is larger than 1. By comparing with Fig. 4(D), we find that our proposed cyclic model outperforms the two-state counterpart in terms of

fitting accuracy. Furthermore, although inferring an exact distribution of the dwell time at the active state remains challenging, our estimations strongly suggest non-exponential behavior of the dwell time during activation.

Conclusions and discussion

In this study, we have analyzed a general stochastic model of gene expression with complex promoter structure, which included many previously studied models as its special case, and derived analytical results for the stationary distribution and statistical quantities of nascent RNA, which provide insight into the stochastic kinetics of nascent RNA. Interestingly, we found that the molecular memories created by multistep activation and deactivation can modulate the steady-state distribution of nascent RNA cooperatively, e.g., increasing the deactivation memory tends to disrupt the zero peak in the stationary nascent RNA distribution, while enhancing the activation memory can induce a bimodal distribution in the bursting regime. More practically, we developed an efficient method to infer the shape parameters characterizing the non-exponential distributions of the waiting times between promoter states from stationary nascent RNA distributions, and verified the validity of our inference method by analyzing the experimental data of a realistic example.

Using the analytical solution of a cyclic model, we investigated how activation and deactivation memory affect the shape of the stationary nascent RNA distribution. Previous study [14] showed that the two-state model with deterministic elongation time can display bimodal distribution with a zero peak and non-zero peak. By contrast, our study revealed that increasing the deactivation memory can destroy the zero peak, resulting in a bimodal distribution with two non-zero peaks. Additionally, we found that increasing the initiation rate k_0 at the inactive state also eliminated the zero peak. This suggests that the zero peak in stationary nascent RNA distribution is unstable, and this fact may explain why the bimodal distribution is rarely observed in experiments. Furthermore, our study showed that by increasing activation memory, delay-induced bimodality or delay-induced zero-inflation phenomenon for nascent RNA can be further enlarged within burst switching regime as reported by Qingchao et al. [17]. However, our study has limitations since it only considers a cyclic promoter proceeding sequentially through several irreversible active and inactive state while real-life promoters may have more complex structure such as competitive cross-talking pathways [53] and reversible processes [31], all of which may affect stationary nascent RNA distributions.

Based on the analytical solution of the stationary nascent RNA, we have also developed an effective method to infer the promoter memory and transcriptional burst kinetics from a known nascent distribution. Our inferred results indicated that in the burst regime where the promoter spends most of its time in the inactive period, it is possible to infer promoter memory, in particular active memory, from the stationary nascent RNA distribution. We applied this method to the *HIV-1* gene [20], which was reported to exhibit bursting kinetics. By fitting the experimental data to the cyclic model, we found that the active memory was always larger than 1, suggesting a non-exponential dwell

time distribution during activation. However, it should be noted that accurately inferring promoter memory from stationary nascent RNA distribution may not always be feasible since different promoters with distinct active and inactive memories can generate identical stationary distribution [31,54]. Further investigation is required to address potential bias in inferring promoter memory from nascent RNA distribution. Additionally, intermediate processes of gene expression such as the partitioning of mRNA/protein during cell division [5,55,56], would contribute the noise of gene expression and potentially skew the inferred results, but these processes were not considered in our study. How details of intermediate processes affect the inferred results is worth further study.

Acknowledgments

This work was supported by grants 12001129, 12171494, 62373384, 11931019, and 12371483 from the Natural Science Foundation of P. R. China, and by grants 2019B110233002, 202007030004 from Key-Area Research and Development Program of Guangzhou, P.R. China, by grants 2022ZDZX2045 from the Special Projects in Key Fields for Colleges and Universities in Guangdong Province.

Data Availability

The nascent RNA data of *HIV-1* gene in cells expressing a high level of Tat were downloaded from ref. [20].

References

1. Raj, A., C.S. Peskin, D. Tranchina, D.Y. Vargas, and S. Tyagi. 2006. Stochastic mRNA synthesis in mammalian cells. *PLoS Biol.* 4:e309.
2. Leloup, J.C., and A. Goldbeter. 2003. Toward a detailed computational model for the mammalian circadian clock. *Proc. Natl. Acad. Sci. U. S. A.* 100:7051–7056.
3. Sepúlveda, L.A., H. Xu, J. Zhang, M. Wang, and I. Golding. 2016. Measurement of gene regulation in individual cells reveals rapid switching between promoter states. *Science. 351:1218–1222.*
4. Jones, D.L., R.C. Brewster, and R. Phillips. 2014. Promoter architecture dictates cell-to-cell variability in gene expression. Science. 346:1533–1536.
5. Cao, Z., and R. Grima. 2020. Analytical distributions for detailed models of stochastic gene expression in eukaryotic cells. *Proc. Natl. Acad. Sci.* 117:4682–4692.
6. Shahrezaei, V., and P.S. Swain. 2008. Analytical distributions for stochastic gene expression. *Proc. Natl. Acad. Sci.* 105:17256–17261.
7. Zoller, B., S.C. Little, and T. Gregor. 2018. Diverse spatial expression patterns emerge from unified kinetics of transcriptional bursting. *Cell.* 175:835–847.

8. Wan, Y., D.G. Anastasakis, J. Rodriguez, M. Palangat, P. Gudla, G. Zaki, M. Tandon, G. Pegoraro, C.C. Chow, M. Hafner, and D.R. Larson. 2021. Dynamic imaging of nascent RNA reveals general principles of transcription dynamics and stochastic splice site selection. *Cell*. 184:2878-2895.e20.
9. Filipowicz, W., S.N. Bhattacharyya, and N. Sonenberg. 2008. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat. Rev. Genet.* 9:102–114.
10. Halpern, K.B., I. Caspi, E. Elinav, I. Ulitsky, K.B. Halpern, I. Caspi, D. Lemze, M. Levy, S. Landen, E. Elinav, and I. Ulitsky. 2015. Nuclear Retention of mRNA in Mammalian Tissues Report Nuclear Retention of mRNA in Mammalian Tissues. *CellReports*. 13:2653–2662.
11. Zenklusen, D., D.R. Larson, and R.H. Singer. 2008. Single-RNA counting reveals alternative modes of gene expression in yeast. *Nat. Struct. Mol. Biol.* 15:1263–1271.
12. Larson, D.R., D. Zenklusen, B. Wu, J.A. Chao, and R.H. Singer. 2011. Real-time observation of transcription initiation and elongation on an endogenous yeast gene. *Science*. 332:475–478.
13. Halpern, K.B., S. Tanami, S. Landen, M. Chapal, L. Szlak, A. Hutzler, A. Nizhberg, and S. Itzkovitz. 2015. Bursty gene expression in the intact mammalian liver. *Mol. Cell*. 58:147–156.
14. Xu, H., S.O. Skinner, A.M. Sokac, and I. Golding. 2016. Stochastic kinetics of nascent RNA. *Phys. Rev. Lett*. 117:128101.
15. Fu, X., H.P. Patel, S. Coppola, L. Xu, Z. Cao, T.L. Lenstra, and R. Grima. 2022. Quantifying how post-transcriptional noise and gene copy number variation bias transcriptional parameter inference from mRNA distributions. *Elife*. 11:e82493
16. Szavits-Nossan, J., and R. Grima. 2022. Mean-field theory accurately captures the variation of copy number distributions across the mRNA life cycle. *Phys. Rev. E*. 105:1–15.
17. Jiang, Q., X. Fu, S. Yan, R. Li, W. Du, Z. Cao, F. Qian, and R. Grima. 2021. Neural network aided approximation and parameter inference of non-Markovian models of gene expression. *Nat. Commun.* 12:1–12.
18. Szavits-Nossan, J., and R. Grima. 2023. Steady-state distributions of nascent RNA for general initiation mechanisms. *Phys. Rev. Res*. 5:013064.
19. Braichenko, S., J. Holehouse, and R. Grima. 2021. Distinguishing between models of mammalian gene expression: telegraph-like models versus mechanistic models. *J. R. Soc. Interface*. 18:20210510.
20. Tantale, K., E. Garcia-Oliver, M.C. Robert, A. L'Hostis, Y. Yang, N. Tsanov, R. Topno, T. Gostan, A. Kozulic-Pirher, M. Basu-Shrivastava, K. Mukherjee, V. Slaninova, J.C. Andrau, F. Mueller, E. Basyuk, O. Radulescu, and E. Bertrand. 2021. Stochastic pausing at latent HIV-1 promoters generates transcriptional bursting. *Nat. Commun*. 12:1–20.
21. Filatova, T., N. Popovic, and R. Grima. 2021. Statistics of nascent and mature RNA fluctuations in a stochastic model of transcriptional initiation, elongation, pausing, and termination. *Bull. Math. Biol.* 83:1–62.
22. Choubey, S., J. Kondev, and A. Sanchez. 2015. Deciphering Transcriptional Dynamics In Vivo by Counting Nascent RNA Molecules. *PLoS Comput. Biol.* 11:1–21.
23. Choubey, S. 2018. Nascent RNA kinetics: Transient and steady state behavior of models of transcription. *Phys. Rev. E*. 97:1–11.
24. Chong, S., C. Chen, H. Ge, and X.S. Xie. 2014. Mechanism of transcriptional bursting in bacteria. *Cell*. 158:314–326.
25. Suter, D.M., N. Molina, D. Gatfield, K. Schneider, U. Schibler, and F. Naef. 2011. Mammalian genes are transcribed with widely different bursting kinetics. *Science*. 332:472–474.
26. J. M. G. Vilar and L. Saiz. CplexA: a mathematica package to study macromolecular-assembly control of gene expression. *Bioinformatics* 26:2060-2061 (2010).

27. Hornung, G., R. Bar-Ziv, D. Rosin, N. Tokuriki, D.S. Tawfik, M. Oren, and N. Barkai. 2012. Noise–mean relationship in mutated promoters. *Genome Res.* 22:2409–2417.
28. Halme, A., S. Bumgarner, C. Styles, and G.R. Fink. 2004. Genetic and epigenetic regulation of the FLO gene family generates cell-surface variation in yeast. *Cell*. 116:405–415.
29. L. M. Octavio, K. Gedeon, and N. Maheshri. Epigenetic and conventional regulation is distributed among activators of FLO11 allowing tuning of population-level heterogeneity in its expression. *PLoS Genet*. 5:e1000673 (2009);
30. Weinberger, L., Y. Voichek, I. Tirosh, G. Hornung, I. Amit, and N. Barkai. 2012. Expression noise and acetylation profiles distinguish HDAC functions. *Mol. Cell*. 47:193–202.
31. Zhang, J., L. Chen, and T. Zhou. 2012. Analytical distribution and tunability of noise in a model of promoter progress. *Biophys. J.* 102:1247–1257.
32. Zhou, T., and J. Zhang. 2012. Analytical results for a multistate gene model. *SIAM J. Appl. Math.* 72:789–818.
33. Shi, C., Y. Jiang, and T. Zhou. 2020. Queuing Models of Gene Expression: Analytical Distributions and Beyond. *Biophys. J.* 119:1606–1616.
34. Cao, Z., T. Filatova, D.A. Oyarzún, and R. Grima. 2020. A stochastic model of gene expression with polymerase recruitment and pause release. *Biophys. J.* 119:1002–1014.
35. Pedraza, J.M., and J. Paulsson. 2008. Effects of molecular memory and bursting on fluctuations in gene expression. *Science*. 319:339–343.
36. Yang, X., Z. Wang, Y. Wu, T. Zhou, and J. Zhang. 2022. Kinetic characteristics of transcriptional bursting in a complex gene model with cyclic promoter structure. *Math. Biosci. Eng.* 19:3313–3336.
37. Yang, X., S. Luo, Z. Zhang, Z. Wang, T. Zhou, and J. Zhang. 2022. Silent transcription intervals and translational bursting lead to diverse phenotypic switching. *Phys. Chem. Chem. Phys.* 24:26600–26608.
38. Jia, C., and Y. Li. 2023. Analytical time-dependent distributions for gene expression models with complex promoter switching mechanisms. *SIAM J. Appl. Math.* 83:1572–1602.
39. Chen, M., S. Luo, M. Cao, C. Guo, T. Zhou, and J. Zhang. 2022. Exact distributions for stochastic gene expression models with arbitrary promoter architecture and translational bursting. *Phys. Rev. E*. 105:14405.
40. Zhang, J., and T. Zhou. 2014. Promoter-mediated transcriptional dynamics. *Biophys. J.* 106:479–488.
41. Zhang, J., and T. Zhou. 2019. Markovian approaches to modeling intracellular reaction processes with molecular memory. *Proc. Natl. Acad. Sci.* 116:23542–23550.
42. Adelman, K., A. La Porta, T.J. Santangelo, J.T. Lis, J.W. Roberts, and M.D. Wang. 2002. Single molecule analysis of RNA polymerase elongation reveals uniform kinetic behavior. *Proc. Natl. Acad. Sci.* 99:13538–13543.
43. Shi, C., X. Yang, J. Zhang, and T. Zhou. 2023. Stochastic modeling of the mRNA life process: A generalized master equation. *Biophys. J*. 122:4023–4041.
44. Fischer, W., S. Ag, K. Meier-hellstern, and T.B. Laboratories. 1992. The Markov-modulated Poisson process (MMPP) cookbook. 18:149–171.
45. Neuts, M.F. 1979. Versatile Markovian Point Process. *J. Appl. Probab.* 16:764–779.
46. Fu, X., X. Zhou, D. Gu, Z. Cao, and R. Grima. 2022. DelaySSAToolkit. jl: stochastic simulation of reaction systems with time delays in Julia. *Bioinformatics*. 36:4243–4245.
47. Luo, S., Z. Wang, Z. Zhang, T. Zhou, and J. Zhang. 2023. Genome-wide inference reveals that feedback regulations constrain promoter-dependent transcriptional burst kinetics. *Nucleic Acids Res.* 51:68–83.
48. Cao, Z., and R. Grima. 2019. Accuracy of parameter estimation for auto-regulatory transcriptional feedback loops from noisy data. *J. R. Soc. Interface*. 16:20180967.

49. Gómez-Schiavon, M., L.F. Chen, A.E. West, and N.E. Buchler. 2017. BayFish: Bayesian inference of transcription dynamics from population snapshots of single-molecule RNA FISH in single cells. *Genome Biol.* 18:1–12.
50. Luo, S., Z. Zhang, Z. Wang, X. Yang, X. Chen, T. Zhou, and J. Zhang. 2023. Inferring transcriptional bursting kinetics from single-cell snapshot data using a generalized telegraph model. *R. Soc. Open Sci*. 10:221057.
51. Feldt, R., and A. Stukalov. 2018. Blackboxoptim. jl. *See https//github. com/robertfeldt/BlackBoxOptim. jl*.
52. Endo, T., T. Watanabe, and A. Yamamoto. 2015. Confidence interval estimation by bootstrap method for uncertainty quantification using random sampling method. *J. Nucl. Sci. Technol.* 52:993–999.
53. Jiao F, Zhu C. Regulation of gene activation by competitive cross talking pathways. *Biophys. J*, 2020, 119(6): 1204-1214.
54. Kepler, T. B., and T. C. Elston. 2001. Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations. *Biophys. J.* 81:3116–3136.
55. Beentjes, C. H. L., R. Perez-Carrasco, and R. Grima. 2020. Exact solution of stochastic gene expression models with bursting, cell cycle and replication dynamics. *Phys. Rev. E*. 101:032403.
56. Skinner, S.O., H. Xu, S. Nagarkar-Jaiswal, P.R. Freire, T.P. Zwaka, and I. Golding. 2016. Single-cell analysis of transcription kinetics across the cell cycle. *Elife*. 5:1–24.
57. Holehouse, J. 2022. Model reduction, mechanistic modelling and transience in models of stochastic chemical kinetics.