CSIAM Trans. Appl. Math.
doi: 10.4208/csiam-am.SO-2023-0049

Vol. **x**, No. x, pp. 1-25
xx 2024

A Novel Computational Method for Two-State Transcription Model with Non-Exponential ON and OFF Durations

Manyi Zheng, Zhishan Qiu, Feng Jiao* and Qiwen Sun*

Guangzhou Center for Applied Mathematics, Guangzhou University, Guangzhou 510006, P.R. China

Received 19 October 2023; Accepted 27 December 2023

Abstract. The fluctuation of mRNA molecule numbers within an isogenic cell population is primarily attributed to randomly switching between active (ON) and inactive (OFF) periods of gene transcription. In most studies the waiting-times for ON or OFF states are modeled as exponential distributions. However, increasing data suggest that the residence durations at ON or OFF are non-exponential distributed for which the traditional master equations cannot be presented. By combining Kolmogorov forward equations with alternating renewal processes, we present a novel method to compute the average transcription level and its noise by circumventing the bottleneck of master equations under gene ON and OFF switch. As an application, we consider lifetimes of OFF and ON states having Erlang distributions. We show that: (i) multiple steps from OFF to ON force the oscillating transcription while multiple steps from ON to OFF accelerate the transcription, (ii) the increase of steps between ON and OFF rapidly reduces the transcription noise to approach its minimum value. This suggests that a large number of steps between ON and OFF are not needed in the model to capture the stochastic transcription data. Our computation approach can be further used to treat a series of transcription cycles which are non-lattice distributed.

AMS subject classifications: 60K05, 34A35, 34A38, 92B05

Key words: Stochastic gene transcription, two-state transcription model, master equations, non-Markov process, alternating renewal processes.

1 Introduction

In both prokaryotes and eukaryotes, gene transcription is the core process in the transmission of genetic information, which flows from DNA to RNA to protein in single cells [22]. The new in vivo RNA detection technique, such as single-cell fluorescence

*Corresponding author. *Email addresses:* jiaof@gzhu.edu.cn (F. Jiao), qwsun@gzhu.edu.cn (Q. Sun)

microscopy and synthetic genetic constructs, has allowed real-time monitoring of transcription events in individual living cells [1, 24, 34]. In experiments, mRNA synthesis is monitored to be random and discontinuous [17, 37]. Randomness in transcription leads to highly variable mRNA distributions, resulting in phenotypic heterogeneity in a cell population [20, 32].

Mathematical models have been built to characterize gene transcription and explore its stochastic regulation [6,30,39,44,47]. In the classical two-state transcription model [30], the promoter is thought to switch randomly between two fundamental states: active and inactive, manifested by the observed transcriptional burst occurring in short-lived active states interspersed by long-lived inactive intervals [27, 37]. To explore the mechanisms that regulate stochastic mRNA production in response to environmental changes, the three-state transcription model instead of a two-state model was proposed to describe the transcription process [4, 39]. In this model, the OFF state is composed of two sub-states connected in series, and both obey the exponential distribution. The assertion was then validated in experiments [11, 37], and was generalized to different models with multiple sub-OFF or sub-ON states [28, 42, 47], and multiple signaling pathways [21, 35, 36]. Also, some complex models and methods were established to estimate system parameters or calculate the probability distribution functions of transcripts [3, 5, 14, 15, 28, 31, 45].

In most models mentioned above, it was assumed that the switching rates between different transcriptional states are constants. In other words, the state switching can be governed by a first-order linear differential equation with a constant coefficient. In the two-state transcription model, the promoter is assumed to switch stochastically between an OFF state and an ON state at rates k_{OFF} and k_{ON}, respectively. The two rates k_{OFF} and k_{ON} are usually defined as

$$k_{OFF} = \frac{1}{T_{ON}}, \quad k_{ON} = \frac{1}{T_{OFF}}, \tag{1.1}$$

where T_{ON} is the average duration of a burst and T_{OFF} is the average time between two consecutive bursts. Since the rates k_{OFF} and k_{ON} are constants, the lifetimes of the OFF and the ON states should have the exponential distribution. From the memoryless property of the exponential distributions, it follows that the two-state transcription process in fact is a continuous-time Markov chain. Other models built in this way also have the Markov property. It is the Markov property that allows us to obtain the chemical master equations of gene transcription.

With the help of real-time monitoring, increasing experiments indicate that the lifetimes of the ON and/or the OFF states are not always exponentially distributed [12, 27, 37, 40]. In recent years, non-Markovian processes have attracted increasing interest [7, 18, 23, 43]. Using single-cell time-lapse bioluminescence imaging, Suter *et al.* [37] monitored transcription kinetics of some genes in mouse fibroblasts and found that the duration of the OFF state should be described by summing two sequential exponential processes. How to bypass the molecular memory in the reaction process to establish an appropriate transcriptional model is crucial for the study of stochastic models.

To overcome such difficulties, several new approaches have been presented. For example, using Markovian approximation method, one could study intracellular reaction processes with molecular memory [46]. Moreover, several transcription models with time-dependent kinetic rates have been established which allows the calculation of the average levels and distributions of transcripts and proteins [3,10,16,19,29,38].

By introducing an alternating renewal process, we will propose a novel method for computing the transcription level and the corresponding noise. In the model, we only assume that the lifetimes of the transcription cycles are independent identically distributed but nonlattice. The method we propose here allows the length of OFF (ON) time to depend on the previous ON (OFF) time, and is applicable to most two-state transcription models for which gene switches between the ON and OFF states with non-exponential distributed durations. The rest of this paper is organized as follows. In Section 2 we introduce the principle and process of the modeling in detail and gives the differential equations for the mean and second moments when the promoter is in the ON and OFF states. By employing the alternating renewal process, we give the stationary forms of the transcription frequency, the average transcription levels and the second moments in Sections 3 and 4. We perform numerical simulations to explore the contribution of the different lifetimes on transcription in Section 5.

2 The model

2.1 The description of the model

In this paper, we discuss the occurrence of a discontinuous transcription process, with individual genes shifting between active and inactive periods, resulting in transcriptional outputs oscillating periodically when the two states alternate randomly. The inactive state is characterized by a lack of specific binding of the transcription factors to the promoter and no RNA polymerase elongating the coding region. Thus, no RNA transcript is synthesized in such state. Once the promoter is recognized and a bubble is created, the RNA synthesis begins. During elongation, the transcription bubble moves along DNA, and the RNA chain is extended by adding nucleotides to the growing chain. The state in which the transcription event occurs is called the gene ON state. The active state's exit is defined as the instant that the transcript is released and the bubble closes. Then the promoter enters an inactive state again.

The two-state model has been widely used to describe transcriptional fluctuations in bacteria, yeast, and mammalian cells [9,23,30]. In this model, as depicted in Fig. 1, a gene promoter is assumed to fluctuate randomly between an inactive state and an active state. Initially, the promoter resides at the inactive state at time T_0, and remains OFF for a time Y_1, then it is turned ON at time T_1 once the RNA polymerase binds to the TATA box of the gene and moves along the template to synthesize RNA. The promoter will remain ON for a time Z_1 until the last RNA polymerase reaches a terminator sequence, where

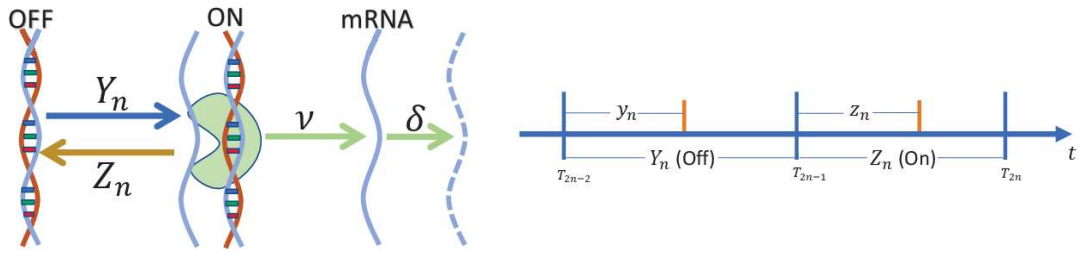

Figure 1: Discontinuous transcription with the gene switches randomly between the OFF and the ON states. RNA synthesis is catalyzed by the enzyme RNA polymerase. The gene is activated by binding RNA polymerases to TATA-box and resides at the active state for a time Z_n. RNA polymerases move along the template, synthesizing RNA with a constant rate ν, which will be turned over with a rate δ. The gene is inactivated when the last RNA polymerase reaches a terminator sequence, then resides at the inactive state for a time Y_n. In this state, there is no RNA transcript production occurring, but the existing transcripts are turned over with a constant rate δ.

the transcription ends. Then it goes OFF at time T_2 for a time Y_2, then ON at time T_3 for a time Z_2 and so forth. Let $T_0=0$. For our model depicted in Fig. 1,

$$T_{2n-1}=T_{2n-2}+Y_n, \quad T_{2n}=T_{2n-1}+Z_n, \quad n\geq 1.$$

Thus, T_{2n-1} is the waiting time that the promoter just leaves the OFF state for n-th time, and T_{2n} the waiting time that the system leaves the ON state for n-th time. If the promoter resides at the OFF state for n-th time at t, then $y_n=t-T_{2n-2}$ is the time from t since last state transition occurring from ON to OFF, and $\bar{y}_n=T_{2n-1}-t$ is the time from t until the next state transition occurring. y_n is called the age at time t that the promoter is in OFF for n-th time, and $\bar{y}_n$ is called the residual life at time t. Thus, $Y_n=y_n+\bar{y}_n$ represents the lifetime of the n-th OFF state. Similarly, if the promoter resides at ON state the n-th time at t, then $z_n=t-T_{2n-1}$ is the time from t since last state transition from OFF to ON. We make the following assumptions to complete the description of the model:

(1) The transition between OFF and ON states is instantaneous, such that the distribution of transcripts does not change during the process.

(2) The lifetimes Y_n and Z_n are independent and identically distributed random variables that follow continuous positively-valued density functions.

(3) During each OFF state, there is no production of transcripts, and mRNA molecules are turned over with a rate parameter δ.

(4) During each ON state, the expression level of transcripts is controlled by two independent random events, simple birth and death with rates ν and δ.

Assumption (1) is reasonable. The reaction rate for the bacterial RNA polymerase is about 40 to 50 nucleotides per second for most transcripts [22], which helps the RNA polymerase be bound to or released from DNA sequence in seconds. Assumption (1) also indicate that each moment of transcripts is maintained during states transitions, especially

the first and the second moments. Assumption (2) indicates that the gene transcription process is an alternating renewal process between OFF and ON states for different periods of time. Assumptions (3) and (4) indicate that the transcripts are synthesized with a constant rate ν only when a stable DNA/pol-II binding is formed, and are turned over dependent on the transcript number at time t. The turned over rate δ is derived by the half-life of transcripts.

In this paper, we assume that the two lifetimes Y_n and Z_n are Erlang distributed, and the probability density functions for them are

$$f_{Y_n}(\tau)=\frac{\lambda^{k_1}\tau^{k_1-1}e^{-\lambda\tau}}{(k_1-1)!},\quad \tau>0, \tag{2.1}$$

$$f_{Z_n}(\tau)=\frac{\gamma^{k_2}\tau^{k_2-1}e^{-\gamma\tau}}{(k_2-1)!},\quad \tau>0, \tag{2.2}$$

where k_1,k_2 are two positive integers and $\lambda>0,\gamma>0$. The Erlang distributed probability distribution functions are

$$F_{Y_n}(\tau,k_1,\lambda)=\int_0^\tau f_{Y_n}(u)du=1-\sum_{m=0}^{k_1-1}\frac{\lambda^m\tau^m e^{-\lambda\tau}}{m!},$$

$$F_{Z_n}(\tau,k_2,\gamma)=\int_0^\tau f_{Z_n}(u)du=1-\sum_{m=0}^{k_2-1}\frac{\gamma^m\tau^m e^{-\gamma\tau}}{m!}.$$

When $k_1=1,k_2=1$, the model is the classical two-state transcription model which has been widely studied [30]. When $k_1=2,k_2=1$, this model describes a transcription process having three functional states established by Tang [39].

2.2 The differential equations

For a given time $t\geq 0$, we let $X=X(t)$ denote the discrete random variable that specifies the system state. We write $X(t)=O$ if the promoter resides at the inactive state, and $X(t)=E$ if the transcription is activated at time t. Let $M(t)$ denote the mRNA number for the gene of our interest in single cells at time t. We define two joint probabilities to quantify the transcription system states. Let $P_O(m,t)$ be the probability that there are m mRNA copies and the promoter resides at the inactive state, that is,

$$P_O(m,t)=\text{Prob}\{X(t)=O,M(t)=m\}.$$

Similarly, we define

$$P_E(m,t)=\text{Prob}\{X(t)=E,M(t)=m\}$$

to be the probability that there are m mRNA copies in the cell and the promoter resides at the active state. To determine the mean transcript level and the noise when the promoter resides at each transcription state, we introduce two conditional probabilities

$$p_O(m,t)=\text{Prob}\{M(t)=m\,|\,X(t)=O\}=P_O(m,t)/P_O(t), \tag{2.3}$$

$$p_E(m,t) = \text{Prob}\{M(t)=m \,|\, X(t)=E\} = P_E(m,t)/P_E(t), \tag{2.4}$$

where the two denominators

$$P_O(t) = \text{Prob}\{X(t)=O\} = \sum_{m=0}^{\infty} P_O(m,t),$$
$$P_E(t) = \text{Prob}\{X(t)=E\} = \sum_{m=0}^{\infty} P_E(m,t)$$

define the transcriptional inefficiency and efficiency. They also denote the respective probabilities that the promoter resides at the inactive and active states. By (2.3) and (2.4), the two conditional probabilities $p_O(m,t)$ and $p_E(m,t)$ are the probabilities that there are m mRNA copies when the promoter resides at the two respective states.

In above assumption (1), we have assumed that the state transition is instantaneous. If the promoter resides at the OFF or the ON state, then the time evolution of $p_O(m,t)$ or $p_E(m,t)$ is only determined by the birth and death process of transcripts.

By using the Kolmogorov forward equations, we calculate the time evolutions of these probabilities (2.3) and (2.4). We first suppose that the promoter remains at the inactive state in time interval $[T_{2n-2}, T_{2n-1})$ $(n \geq 1)$ and will leave this state at time T_{2n-1}. For any time $t \in [T_{2n-2}, T_{2n-1})$, Δt is an infinitesimal time increment such that $t+\Delta t \in [T_{2n-2}, T_{2n-1})$. Suppose that there are m mRNA copies at time $t+\Delta t$. Then the basic model assumptions (1)-(4) imply that one of the following events must occur at time t:

(i) There is no elimination of transcripts taking place during the time interval $(t, t+\Delta t)$ with a probability $p_O(m,t)(1-m\delta\Delta t)$.

(ii) There is one transcript being eliminated during the time interval $(t, t+\Delta t)$ with a probability $p_O(m,t)m\delta\Delta t$.

Adding the probabilities in (i)-(ii) together gives $p_O(m, t+\Delta t)$. Dividing the resulting equality by Δt and then letting $\Delta t \to 0$, we obtain

$$p_O'(m,t) = -m\delta p_O(m,t) + (m+1)\delta p_O(m+1,t). \tag{2.5}$$

When the promoter resides at the active state in time interval $[T_{2n-1}, T_{2n})$ $(n \geq 1)$, one of the following events may occur at time t:

(i) There is no production or elimination of transcripts taking place during $(t, t+\Delta t)$ with a probability $p_E(m,t)(1-\nu\Delta t)(1-m\delta\Delta t)$.

(ii) There is one transcript being eliminated during the time interval $(t, t+\Delta t)$ with a probability $p_E(m,t)m\delta\Delta t$.

(iii) There is one transcript being produced during the time interval $(t, t+\Delta t)$ with a probability $p_E(m,t)\nu\Delta t$.

By using a similar discussion as above, we obtain the time evolution of $p_E(m,t)$ as

$$p'_E(m,t)=\nu p_E(m-1,t)+(m+1)\delta p_E(m+1,t)-\nu p_E(m,t)-m\delta p_E(m,t). \tag{2.6}$$

The Eqs. (2.5) and (2.6) give the basic differential equations. When the gene promoter transits randomly between the OFF and the ON states, then $p_O(m,t)$ and $p_E(m,t)$ switch randomly according to state transition.

Without loss of generality, we assume that the transcription starts from the OFF state with none transcripts existence in the cell at time $t=0$. It gives the initial condition

$$\begin{aligned} &P_O(0,0)=1, \quad P_O(m,0)=0, \quad m>0,\\ &P_E(m,0)=0, \qquad\qquad\quad m\geq 0. \end{aligned} \tag{2.7}$$

Furthermore, we have

$$\begin{aligned} &p_O(t)=p_{On}(\tau)\equiv 1, \quad p_E(t)=p_{En}(\tau)\equiv 0, \quad T_{2n-2}\leq t<T_{2n-1}, \quad 0\leq\tau<Y_n,\\ &p_O(t)=p_{On}(\tau)\equiv 0, \quad p_E(t)=p_{En}(\tau)\equiv 1, \quad T_{2n-1}\leq t<T_{2n}, \qquad 0\leq\tau<Z_n. \end{aligned}$$

In the model, T_n is the exact time that the state transition occurs. To simplify calculation, we introduce a new timer τ such that the transcription system reclocks at time T_n, and stamp each variable $V(t)$ by a subscript such as $V_{On}(\tau)$ or $V_{En}(\tau)$ to denote this variable resides at the inactive or active states the n-th time at time t. The time evolution of $V(t)$ is governed by the following ordinary differential equations:

$$\begin{aligned} &V_{On}(\tau)=V_{On}(t), \quad T_{2n-2}\leq t<T_{2n-1}, \quad 0\leq\tau<Y_n, \quad \frac{d\tau}{dt}=1,\\ &V_{En}(\tau)=V_{En}(t), \quad T_{2n-1}\leq t<T_{2n}, \qquad 0\leq\tau<Z_n, \quad \frac{d\tau}{dt}=1. \end{aligned}$$

3 The stationary frequency of elongation and the average transcript levels

3.1 The transcription frequency

To give the probabilities that the promoter resides at the inactive and active states, we need the following definition and lemmas.

Definition 3.1. *A nonnegative random variable X is said to be lattice if there exists $d\geq 0$ such that*

$$\sum_{n=0}^{\infty} P(X=nd)=1.$$

For example, if X is a random variable with Bernoulli distribution, then its distribution function F is lattice. In our model, the lifetimes $\{Y_n\}$ and $\{Z_n\}$ are continuous positive-valued distributed, thus they are nonlattice. Let F_{Y_n} be the distribution of Y_n, F_{Z_n} the distribution of Z_n, and F the distribution of Y_n+Z_n, then we have

Lemma 3.1. *If* $\mathbf{E}[Y_n+Z_n]<\infty$ *and* F *is nonlattice, then*

$$\begin{aligned} P_O^* &= \lim_{t\to\infty} P_O(t) = \frac{\mathbf{E}[Y_n]}{\mathbf{E}[Y_n]+\mathbf{E}[Z_n]}, \\ P_E^* &= \lim_{t\to\infty} P_E(t) = \frac{\mathbf{E}[Z_n]}{\mathbf{E}[Y_n]+\mathbf{E}[Z_n]}. \end{aligned} \tag{3.1}$$

The proof of this lemma can be found in [33]. We call the limit $P_E^* = \lim_{t\to\infty} P_E(t)$ the stationary frequency of elongation. When the transcription system reaches a steady state, the transcription frequency P_E^* is approximately the ratio of the average sojourn time in ON state to the average lifetime of one transcription cycle. The formula P_E^* also indicates the percentage of expressing cells among a population of isogenic cells.

In the transcription cycle, the residence time vectors (Y_n, Z_n) are independent and identically distributed. Lemma 3.1 indicates, as we expect intuitively, that the percentage that the promoter resides at the active state equals the average ON time over the total average transcription cycle time.

Lemma 3.2. *If the state durations* $\{Y_n\}$ *and* $\{Z_n\}$ *are Erlang-distributed with density functions* (2.1) *and* (2.2)*, then*

$$P_O^* = \frac{k_1\gamma}{k_2\lambda+k_1\gamma}, \quad P_E^* = \frac{k_2\lambda}{k_2\lambda+k_1\gamma}. \tag{3.2}$$

In the n-th transcription cycle, we re-clock the time when the state transition from $(n-1)$-th ON state to n-th OFF state occurs. Let y_n be the time such that the promoter remains OFF during $[0, y_n]$. Then y_n would represent the age of the promoter in n-th OFF state. By using the renewal theory, we have that

$$F_{y_n}(\tau) = \text{Prob}\{y_n \le \tau\} = E[\min(Y_n,\tau)]/E[Y_n] = \int_0^\tau \overline{F}_{Y_n}(y)dy/E[Y_n],$$

where $\overline{F}_{Y_n}(y) = 1 - F_{Y_n}(y)$.

Lemma 3.3. *If* $\{Y_n\}$ *is Erlang-distributed, then the distribution function of* y_n *is*

$$F_{y_n}(\tau) = \text{Prob}\{y_n \le \tau\} = \frac{1}{k_1}[F(\tau,1,\lambda)+F(\tau,2,\lambda)+\cdots+F(\tau,k_1,\lambda)], \tag{3.3}$$

where $F(\tau,k,\lambda)$ *is Erlang distribution function with shape parameter* k *and scale parameter* $1/\lambda$.

Lemma 3.3 gives the distribution of the age y_n when the duration Y_n is Erlang-distributed. Similarly, let z_n be the age that the promoter remains ON during $[0, z_n]$ in n-th transcription cycle, then the distribution of the age z_n can be obtained by replacing λ, k_1 by γ, k_2 in (3.3), that is,

$$F_{z_n}(\tau) = \text{Prob}\{z_n \le \tau\} = \frac{1}{k_2}[F(\tau,1,\gamma)+F(\tau,2,\gamma)+\cdots+F(\tau,k_2,\gamma)]. \tag{3.4}$$

However, experimental observations show that the cell cycle time distribution (CCTD) is typically non-monotonic and differs substantially from an exponential distribution.

3.2 The average transcript level in each state

To begin, recall that the expected value of the random variable $M(t)$ is defined by

$$m(t)=\mathbf{E}[M(t)]=\sum_{m=0}^{\infty} mP(m,t), \tag{3.5}$$

where the probability mass function

$$P(m,t)=\text{Prob}\{M(t)=m\}=P_O(m,t)+P_E(m,t)$$

is the probability that there are m transcripts at time t in single cells. By using (2.3), (2.4) and the above probability mass function, we rewrite $m(t)$ as

$$m(t)=\sum_{m=0}^{\infty} m[P_O(m,t)+P_E(m,t)]=m_O(t)\cdot P_O(t)+m_E(t)\cdot P_E(t), \tag{3.6}$$

where

$$m_O(t)=\sum_{m=0}^{\infty} mp_O(m,t), \quad m_E(t)=\sum_{m=0}^{\infty} mp_E(m,t) \tag{3.7}$$

are the average transcript numbers in single cells at time t when the promoter resides at the two states.

Thus, the average transcript number $m(t)$ only depends on the mean levels $m_O(t)$, $m_E(t)$ and the two probabilities $P_O(t), P_E(t)$ that the promoter remains at the OFF and the ON states. As stated in assumptions (3) and (4), mRNA molecules are produced with a constant rate ν in ON state, and there is no production taking place in OFF state. Intuitively, $m_O(t)$ decreases if the duration of inactive state increases and $m_E(t)$ increases if the duration of active state increases. In one case, $P_O(t)$ will increase when inactive state duration increases and active state duration remains stable. In another case, $P_E(t)$ increases when active state duration increases and inactive state duration remains stable. Thus, the mean transcript level may increase with the probability that the promoter remains on active state and the duration of this state.

For clarity, we give four notations, that is,

$$\langle m_{On}\rangle=\int_0^{\infty} m_{On}(\tau)dF_{Y_n}(\tau), \quad \langle m_{En}\rangle=\int_0^{\infty} m_{En}(\tau)dF_{Z_n}(\tau), \tag{3.8}$$

$$\overline{m_{On}}=\int_0^{\infty} m_{On}(\tau)dF_{y_n}(\tau), \qquad \overline{m_{En}}=\int_0^{\infty} m_{En}(\tau)dF_{z_n}(\tau), \tag{3.9}$$

where $\langle m_{On}\rangle$ and $\langle m_{En}\rangle$ are the average transcript numbers at the moments that the promoter leaves the OFF and ON states in the n-th transcription cycle, $\overline{m_{On}}$ and $\overline{m_{En}}$ are the average numbers during the two states in the n-th transcription cycle. Then we have

$$\lim_{t\to\infty} m_O(t)=\lim_{n\to\infty}\overline{m_{On}}, \quad \lim_{t\to\infty} m_E(t)=\lim_{n\to\infty}\overline{m_{En}}. \tag{3.10}$$

Above limitation implies that we only need to calculate $\lim_{n\to\infty}\overline{m_{On}}$ and $\lim_{n\to\infty}\overline{m_{En}}$, then derive the stationary transcript level from (3.6) by taking limits.

Theorem 3.1. *Suppose that the state lifetimes $\{Y_n\}$ and $\{Z_n\}$ are Erlang-distributed, then the average transcript numbers $\langle m_{On}\rangle$ and $\langle m_{En}\rangle$ when the promoter leaves the n-th OFF and the n-th ON states are given as*

$$\langle m_{On}\rangle = M_{On}\cdot\frac{\lambda^{k_1}}{(\delta+\lambda)^{k_1}},\quad \langle m_{En}\rangle = \frac{\nu}{\delta}+\Big(M_{En}-\frac{\nu}{\delta}\Big)\cdot\frac{\gamma^{k_2}}{(\delta+\gamma)^{k_2}},$$

where M_{On} and M_{En} are the (average) initial transcript numbers at the beginning of the OFF and the ON states in n-th transcription cycle. And $\langle m_{On}\rangle$ and $\langle m_{En}\rangle$ satisfy following recursion formulas:

$$\langle m_{On+1}\rangle = \langle m_{On}\rangle\cdot\frac{\lambda^{k_1}\gamma^{k_2}}{(\delta+\lambda)^{k_1}(\delta+\gamma)^{k_2}}+\frac{\nu}{\delta}\cdot\frac{\lambda^{k_1}\big[(\delta+\gamma)^{k_2}-\gamma^{k_2}\big]}{(\delta+\lambda)^{k_1}(\delta+\gamma)^{k_2}},$$
$$\langle m_{En+1}\rangle = \langle m_{En}\rangle\cdot\frac{\lambda^{k_1}\gamma^{k_2}}{(\delta+\lambda)^{k_1}(\delta+\gamma)^{k_2}}+\frac{\nu}{\delta}\cdot\frac{(\delta+\gamma)^{k_2}-\gamma^{k_2}}{(\delta+\gamma)^{k_2}}.$$

When $n\to\infty$, the limits of these two average numbers are given as

$$\langle m_O\rangle = \frac{\nu}{\delta}\cdot\frac{\lambda^{k_1}\big[(\delta+\gamma)^{k_2}-\gamma^{k_2}\big]}{(\delta+\lambda)^{k_1}(\delta+\gamma)^{k_2}-\lambda^{k_1}\gamma^{k_2}},\tag{3.11}$$

$$\langle m_E\rangle = \frac{\nu}{\delta}\cdot\frac{(\delta+\lambda)^{k_1}\big[(\delta+\gamma)^{k_2}-\gamma^{k_2}\big]}{(\delta+\lambda)^{k_1}(\delta+\gamma)^{k_2}-\lambda^{k_1}\gamma^{k_2}}.\tag{3.12}$$

In Theorem 3.1, $\langle m_O\rangle$ gives the average transcript number at the moment that the promoter is activated, and $\langle m_E\rangle$ gives the number at the moment that the promoter is inactivated.

Proof. We will give the detailed process to obtain the average transcript levels at steady state when the promoter resides at the OFF and the ON states respectively. Firstly, we consider the case that the promoter resides at the OFF state during the n-th transcription cycle at time t and remains OFF for a time Y_n. Let τ be the time from t since the last state transition occurs, that is $\tau = t - T_{2n-2}$. Then $\tau = 0$ at the moment that the promoter is inactivated at $t = T_{2n-2}$ and $\tau = Y_n$ when the transcription is activated at time $t = T_{2n-1}$. The average transcript number $m_{On}(\tau)$ during this state is defined to be the sum of $m\cdot p_{On}(m,\tau), 0\le\tau<Y_n$. Multiplying (2.5) by m and summing up these products lead to

$$\frac{dm_{On}(\tau)}{d\tau} = -\delta m_{On}(\tau),\quad 0\le\tau<Y_s.\tag{3.13}$$

We will solve this equation by using the Laplace transform. Without loss of generality, we assume that there are M_{On} transcripts at the beginning of the OFF state. By applying the Laplace transform to (3.13) and noticing that the initial condition is $m_{On}(0) = M_{On}$, we transform the differential equation (3.13) into an algebraic equation

$$s\mathcal{L}\big(m_{On}(\tau)\big) - M_{On} = -\delta\mathcal{L}\big(m_{On}(\tau)\big),$$

which can be written as

$$\mathcal{L}\big(m_{On}(\tau)\big)=\frac{M_{On}}{s+\delta}. \tag{3.14}$$

By applying the inverse Laplace transform to (3.14), we obtain the average transcript level at time τ when the promoter resides at the OFF state with the initial condition M_{On},

$$m_{On}(\tau)=M_{On}\cdot e^{-\delta\tau},\quad 0\le\tau<Y_n. \tag{3.15}$$

Then the transcript number at time Y_n when the promoter leaves the OFF state is given by

$$m_{On}(Y_n)=M_{On}\cdot e^{-\delta Y_n}. \tag{3.16}$$

Since the sojourn time Y_n in the OFF state is Erlang-distributed with rates k_1 and λ in the time interval $[0,\infty)$ and the density function $f_{Y_n}(\tau)$ has been given in (2.1), then we integrate (3.16) with respect to Y_n over the time interval $[0,\infty)$ and derive the average transcript number $\langle m_{On}\rangle$ at the moment that the promoter leaves the OFF state, that is,

$$\langle m_{On}\rangle=\int_0^\infty m_{On}(\tau)f_{Y_n}(\tau)d\tau=\int_0^\infty m_{On}(\tau)\cdot\frac{\lambda^{k_1}\tau^{k_1-1}e^{-\lambda\tau}}{(k_1-1)!}d\tau=M_{On}\cdot\frac{\lambda^{k_1}}{(\delta+\lambda)^{k_1}}. \tag{3.17}$$

Eq. (3.17) implies that the average transcript number $\langle m_{On}\rangle$ is linearly dependent on the initial value M_{On}. Thus, (3.17) also holds when we compute the expected value of M_{On}.

Next, we consider another case that the promoter resides at the n-th ON state. In fact, once the transcription site of the promoter is occupied by RNA polymerase, the gene transcription is activated and the polymerase moves along the encoding region and elongates RNA. After turned on, the promoter remains ON for a time Z_n. Let τ be the time from t since the last state transition occurs, that is $\tau=t-T_{2n-1}$. Then $\tau=0$ at the moment that the promoter is activated at $t=T_{2n-1}$, and $\tau=Z_n$ when the promoter is inactivated at $t=T_{2n}$. In this new state, the average transcript level $m_{En}(\tau)$ is the sum of $m\cdot p_{En}(m,\tau)$. Multiplying (2.6) by m and summing them up give the time evolution of $m_{En}(\tau)$, that is,

$$\frac{dm_{En}(\tau)}{d\tau}=\nu-\delta m_{En}(\tau),\quad 0\le\tau<Z_n. \tag{3.18}$$

Similarly, we apply the Laplace transform to (3.18) and then derive

$$s\mathcal{L}\big(m_{En}(\tau)\big)-M_{En}=\mathcal{L}(\nu)-\delta\mathcal{L}\big(m_{En}(\tau)\big),$$

which can be rewritten and decomposed as

$$\mathcal{L}\big(m_{En}(\tau)\big)=\frac{\nu+sM_{En}}{s(s+\delta)}=\frac{\nu/\delta}{s}+\frac{M_{En}-\nu/\delta}{s+\delta}. \tag{3.19}$$

By applying the inverse Laplace transform to (3.19), we get the form of $m_{En}(\tau)$ as

$$m_{En}(\tau)=\frac{\nu}{\delta}+\Big(M_{En}-\frac{\nu}{\delta}\Big)e^{-\delta\tau},\quad 0\le\tau<Z_n. \tag{3.20}$$

Then the average transcript number at time Z_n is given by

$$m_{En}(Z_n)=\frac{\nu}{\delta}+\left(M_{En}-\frac{\nu}{\delta}\right)e^{-\delta Z_n}. \tag{3.21}$$

Since the sojourn time Z_n is also Erlang-distributed with rates k_2 and γ in time interval $[0,\infty)$ and its density function has been given in (2.2), then the average transcript number at the end of the n-th ON state is

$$\langle m_{En}\rangle=\int_0^\infty\left[\frac{\nu}{\delta}+\left(M_{En}-\frac{\nu}{\delta}\right)e^{-\delta\tau}\right]f_{Z_n}(\tau)d\tau=\frac{\nu}{\delta}+\left(M_{En}-\frac{\nu}{\delta}\right)\frac{\gamma^{k_2}}{(\delta+\gamma)^{k_2}}. \tag{3.22}$$

Eq. (3.22) implies that $\langle m_{En}\rangle$ is linearly dependent on M_{En}, and it holds when we take the expected value to M_{En}.

As stated in assumption (1), the state transition from OFF to ON is instantaneous, which implies that there is no production or elimination of transcripts taking place during such an infinitesimal time. Thus, the transcript distribution at the beginning of the n-th ON state is maintained at the same level as that at the end of the n-th OFF state. Without loss of generality, we could assume that the initial condition for $m_{En}(\tau)$ is

$$m_{En}(0)=M_{En}=\langle m_{On}\rangle. \tag{3.23}$$

Also the distribution of transcripts does not change during an infinitesimal time increment when the promoter transfers from the ON state in the n-th transcription cycle to the OFF state in the $(n+1)$-th transcription cycle, thus we take the initial value of $m_{On+1}(\tau)$ as

$$m_{On+1}(0)=M_{On+1}=\langle m_{En}\rangle, \tag{3.24}$$

where $m_{On+1}(\tau)$ is the temporal mRNA expression level when the promoter resides at the $(n+1)$-th OFF state and M_{On+1} is its initial number. From Eqs. (3.17), (3.22)-(3.24), we obtain recurrence formulas for $\langle m_{On}\rangle$ and $\langle m_{En}\rangle$, that is,

$$\begin{aligned}\langle m_{On+1}\rangle&=\langle m_{On}\rangle\cdot\frac{\lambda^{k_1}\gamma^{k_2}}{(\delta+\lambda)^{k_1}(\delta+\gamma)^{k_2}}+\frac{\nu\lambda^{k_1}\left[(\delta+\gamma)^{k_2}-\gamma^{k_2}\right]}{\delta(\delta+\lambda)^{k_1}(\delta+\gamma)^{k_2}},\\ \langle m_{En+1}\rangle&=\langle m_{En}\rangle\cdot\frac{\lambda^{k_1}\gamma^{k_2}}{(\delta+\lambda)^{k_1}(\delta+\gamma)^{k_2}}+\frac{\nu\left[(\delta+\gamma)^{k_2}-\gamma^{k_2}\right]}{\delta(\delta+\gamma)^{k_2}}.\end{aligned}$$

Taking limits to above two equations with respect to n gives

$$\begin{aligned}\langle m_O\rangle&=\frac{\nu\lambda^{k_1}\left[(\delta+\gamma)^{k_2}-\gamma^{k_2}\right]}{\delta\left[(\delta+\lambda)^{k_1}(\delta+\gamma)^{k_2}-\lambda^{k_1}\gamma^{k_2}\right]},\\ \langle m_E\rangle&=\frac{\nu(\delta+\lambda)^{k_1}\left[(\delta+\gamma)^{k_2}-\gamma^{k_2}\right]}{\delta\left[(\delta+\lambda)^{k_1}(\delta+\gamma)^{k_2}-\lambda^{k_1}\gamma^{k_2}\right]}.\end{aligned}$$

The two transcript levels $\langle m_O\rangle$ and $\langle m_E\rangle$ in above equations give the average numbers at steady state when the promoter just leaves inactive and active states respectively. □

Theorem 3.2. *Under the conditions of Theorem* 3.1*, the average transcript numbers* $\overline{m_O}$ *and* $\overline{m_E}$ *when the promoter resides at the OFF and the ON states are given as*

$$\overline{m_O}=\frac{\nu}{\delta}\cdot\frac{\lambda\left[(\delta+\lambda)^{k_1}-\lambda^{k_1}\right]\left[(\delta+\gamma)^{k_2}-\gamma^{k_2}\right]}{k_1\delta\left[(\delta+\lambda)^{k_1}(\delta+\gamma)^{k_2}-\lambda^{k_1}\gamma^{k_2}\right]}, \tag{3.25}$$

$$\overline{m_E}=\frac{\nu}{\delta}\cdot\left[1-\frac{\gamma\left[(\delta+\lambda)^{k_1}-\lambda^{k_1}\right]\left[(\delta+\gamma)^{k_2}-\gamma^{k_2}\right]}{k_2\delta\left[(\delta+\lambda)^{k_1}(\delta+\gamma)^{k_2}-\lambda^{k_1}\gamma^{k_2}\right]}\right]. \tag{3.26}$$

Then the stationary average transcript number in single cells is

$$m^*=\overline{m_O}\cdot P_O^*+\overline{m_E}\cdot P_E^*=\frac{\nu}{\delta}\cdot\frac{k_2\lambda}{k_2\lambda+k_1\gamma}. \tag{3.27}$$

Proof. We have given the probability distribution function of the time y_n as shown in Lemma 3.3, then the its density function is

$$f_{y_n}(\tau)=\frac{d}{d\tau}F_{y_n}(\tau)=\frac{1}{k_1}\left[\lambda e^{-\lambda\tau}+\frac{\lambda^2\tau e^{-\lambda\tau}}{1!}+\cdots+\frac{\lambda^{k_1}\tau^{k_1-1}e^{-\lambda\tau}}{(k_1-1)!}\right]. \tag{3.28}$$

When the promoter resides at the n-th OFF state, we have derived the temporal expression of transcripts as shown in (3.15). Then the mean transcription level over the whole OFF state is

$$\overline{m_{On}}=\int_0^\infty m_{On}(\tau)f_{y_n}(\tau)d\tau=M_{On}\cdot\frac{1}{k_1}\sum_{i=1}^{k_1}\left(\frac{\lambda}{\delta+\lambda}\right)^i. \tag{3.29}$$

Similarly, the probability density function of the time z_n in the n-th ON state is

$$f_{z_n}(\tau)=\frac{1}{k_2}\left[\gamma e^{-\gamma\tau}+\frac{\gamma^2\tau e^{-\gamma\tau}}{1!}+\cdots+\frac{\gamma^{k_2}\tau^{k_2-1}e^{-\gamma\tau}}{(k_2-1)!}\right]. \tag{3.30}$$

And, we have derived the temporal transcript level as shown in (3.20) when the promoter is in the ON state. Then the mean transcript level over this ON state is

$$\overline{m_{En}}=\int_0^\infty m_{En}(\tau)f_{z_n}(\tau)d\tau=\frac{\nu}{\delta}+\left(M_{En}-\frac{\nu}{\delta}\right)\cdot\frac{1}{k_2}\sum_{i=1}^{k_2}\left(\frac{\gamma}{\delta+\gamma}\right)^i. \tag{3.31}$$

Taking limits to $\overline{m_{On}}$ and $\overline{m_{En}}$ with respect to n, we have

$$\overline{m_O}=M_O\cdot\frac{1}{k_1}\sum_{i=1}^{k_1}\left(\frac{\lambda}{\delta+\lambda}\right)^i, \tag{3.32}$$

$$\overline{m_E}=\frac{\nu}{\delta}+\left(M_E-\frac{\nu}{\delta}\right)\cdot\frac{1}{k_2}\sum_{i=1}^{k_2}\left(\frac{\gamma}{\delta+\gamma}\right)^i. \tag{3.33}$$

Since $M_O=\langle m_E\rangle$ and $M_E=\langle m_O\rangle$ have been given in Theorem 3.1, then we substitute them into above equations and derive (3.25) and (3.26).

Taking limits to $m(t)$ with respect to t and substitute (3.2), (3.25) and (3.26) into (3.6), we get the mean transcript level at steady state in single cells as shown in (3.27). □

4 The stationary noise of transcripts

To characterize the fluctuations of transcripts in cell populations, we introduce the noise $\eta^2(t)$, the variance $\sigma^2(t)$ normalized by the square of the average $m(t)$ and the noise strength $\Phi(t)$, the ratio between the variance and the average, that is,

$$\eta^2(t)=\frac{\sigma^2(t)}{m^2(t)},\quad \Phi(t)=\frac{\sigma^2(t)}{m(t)}, \tag{4.1}$$

where $\sigma^2(t)=\mu(t)-m^2(t)$. As the average transcript levels have been given in Theorems 3.1 and 3.2, it suffices to evaluate the second moment functions in this section.

By definition, the second moment of the transcript number $M(t)$ is

$$\mu(t)=\mathrm{E}[M^2(t)]=\sum_{m=0}^{\infty}m^2P(m,t). \tag{4.2}$$

In order to derive the form of the second moment, it is helpful to consider two values

$$\mu_O(t)=\sum_{m=0}^{\infty}m^2p_O(m,t),\quad \mu_E(t)=\sum_{m=0}^{\infty}m^2p_E(m,t). \tag{4.3}$$

The two values give the second moments of transcripts in the OFF and ON states, respectively. Then $\mu(t)$ can be split into

$$\mu(t)=\mu_O(t)\cdot P_O(t)+\mu_E(t)\cdot P_E(t). \tag{4.4}$$

Similarly, we define four notations, that is,

$$\langle\mu_{On}\rangle=\int_0^{\infty}\mu_{On}(\tau)dF_{Y_n}(\tau),\quad \langle\mu_{En}\rangle=\int_0^{\infty}\mu_{En}(\tau)dF_{Z_n}(\tau), \tag{4.5}$$

$$\overline{\mu_{On}}=\int_0^{\infty}\mu_{On}(\tau)dF_{y_n}(\tau),\qquad \overline{\mu_{En}}=\int_0^{\infty}\mu_{En}(\tau)dF_{z_n}(\tau), \tag{4.6}$$

where $\langle\mu_{On}\rangle$ and $\langle\mu_{En}\rangle$ give the second moments of transcripts at the moment that the promoter leaves the n-th OFF or the n-th ON states, $\overline{\mu_{On}}$ and $\overline{\mu_{En}}$ give the second moments of transcripts when the promoter resides at the n-th OFF or the n-th ON states. And we have

$$\lim_{t\to\infty}\mu_O(t)=\lim_{n\to\infty}\overline{\mu_{On}},\quad \lim_{t\to\infty}\mu_E(t)=\lim_{n\to\infty}\overline{\mu_{En}}. \tag{4.7}$$

Then the stationary second moment is given by

$$\mu^*=\lim_{t\to\infty}\mu(t)=\lim_{n\to\infty}\overline{\mu_{On}}\cdot P_O^*+\lim_{n\to\infty}\overline{\mu_{En}}\cdot P_E^*. \tag{4.8}$$

Theorem 4.1. *Under the conditions of Theorem* 3.1*, the second moments for transcripts existing at the end of* $(n+1)$*-th OFF and* $(n+1)$*-th ON states satisfy*

$$\begin{aligned}\langle\mu_{On+1}\rangle=&\langle\mu_{On}\rangle\frac{\lambda^{k_1}\gamma^{k_2}}{(2\delta+\lambda)^{k_1}(2\delta+\gamma)^{k_2}}\\&+\frac{\lambda^{k_1}}{(2\delta+\lambda)^{k_1}}\left[\frac{\nu(\nu+\delta)}{\delta^2}-\frac{\nu(2\nu+\delta)\gamma^{k_2}}{\delta^2(\delta+\gamma)^{k_2}}+\frac{\nu^2\gamma^{k_2}}{\delta^2(2\delta+\gamma)^{k_2}}\right]\\&+\langle m_{En}\rangle\frac{\lambda^{k_1}\left[(2\delta+\lambda)^{k_1}-(\delta+\lambda)^{k_1}\right]}{(\delta+\lambda)^{k_1}(2\delta+\lambda)^{k_1}}\\&+\langle m_{On}\rangle\frac{(2\nu+\delta)\lambda^{k_1}\gamma^{k_2}\left[(2\delta+\gamma)^{k_2}-(\delta+\gamma)^{k_2}\right]}{\delta(2\delta+\lambda)^{k_1}(\delta+\gamma)^{k_2}(2\delta+\gamma)^{k_2}},\end{aligned}\tag{4.9}$$

$$\begin{aligned}\langle\mu_{En+1}\rangle=&\langle\mu_{En}\rangle\frac{\lambda^{k_1}\gamma^{k_2}}{(2\delta+\lambda)^{k_1}(2\delta+\gamma)^{k_2}}\\&+\left[\frac{\nu(\nu+\delta)}{\delta^2}-\frac{\nu(2\nu+\delta)\gamma^{k_2}}{\delta^2(\delta+\gamma)^{k_2}}+\frac{\nu^2\gamma^{k_2}}{\delta^2(2\delta+\gamma)^{k_2}}\right]\\&+\langle m_{En}\rangle\frac{\lambda^{k_1}\gamma^{k_2}\left[(2\delta+\lambda)^{k_1}-(\delta+\lambda)^{k_1}\right]}{(\delta+\lambda)^{k_1}(2\delta+\lambda)^{k_1}(2\delta+\gamma)^{k_2}}\\&+\langle m_{On+1}\rangle\frac{(2\nu+\delta)\gamma^{k_2}\left[(2\delta+\gamma)^{k_2}-(\delta+\gamma)^{k_2}\right]}{\delta(\delta+\gamma)^{k_2}(2\delta+\gamma)^{k_2}}.\end{aligned}\tag{4.10}$$

When $n\to\infty$*, the stationary second moments for transcripts at the end of the OFF and the ON states are*

$$\begin{aligned}\langle\mu_O\rangle=\langle m_O\rangle+\langle m_O\rangle\cdot\frac{\nu}{\delta}\Bigg[&\frac{(\delta+\lambda)^{k_1}(2\delta+\gamma)^{k_2}-\lambda^{k_1}\gamma^{k_2}}{(2\delta+\lambda)^{k_1}(2\delta+\gamma)^{k_2}-\lambda^{k_1}\gamma^{k_2}}\\&-\frac{\gamma^{k_2}\left[(\delta+\lambda)^{k_1}-\lambda^{k_1}\right]\left[(2\delta+\gamma)^{k_2}-(\delta+\gamma)^{k_2}\right]}{\left[(\delta+\gamma)^{k_2}-\gamma^{k_2}\right]\left[(2\delta+\lambda)^{k_1}(2\delta+\gamma)^{k_2}-\lambda^{k_1}\gamma^{k_2}\right]}\Bigg],\end{aligned}\tag{4.11}$$

$$\begin{aligned}\langle\mu_E\rangle=\langle m_E\rangle+\langle m_E\rangle\cdot\frac{\nu(2\delta+\lambda)^{k_1}}{\delta(\delta+\lambda)^{k_1}}\Bigg[&\frac{(\delta+\lambda)^{k_1}(2\delta+\gamma)^{k_2}-\lambda^{k_1}\gamma^{k_2}}{(2\delta+\lambda)^{k_1}(2\delta+\gamma)^{k_2}-\lambda^{k_1}\gamma^{k_2}}\\&-\frac{\gamma^{k_2}\left[(\delta+\lambda)^{k_1}-\lambda^{k_1}\right]\left[(2\delta+\gamma)^{k_2}-(\delta+\gamma)^{k_2}\right]}{\left[(\delta+\gamma)^{k_2}-\gamma^{k_2}\right]\left[(2\delta+\lambda)^{k_1}(2\delta+\gamma)^{k_2}-\lambda^{k_1}\gamma^{k_2}\right]}\Bigg].\end{aligned}\tag{4.12}$$

Proof. Let the promoter reside at the inactive state the n-th time. For any $0\le\tau<Y_n$, the second moment of transcripts $\mu_{On}(\tau)$ is defined to be the sum of $m^2\cdot p_{On}(m,\tau)$. Thus, multiplying (2.5) by m^2 and taking the sum lead to

$$\mu'_{On}(\tau)=-2\delta\mu_{On}(\tau)+\delta m_{On}(\tau),\quad 0\le\tau<Y_n.\tag{4.13}$$

At the beginning of this state, we assume that the initial condition of second moment is $\mu_{On}(0)=\Lambda_{On}$. By applying the Laplace transform to (4.13), we transform (4.13) into

an algebraic equation

$$s\mathcal{L}\big(\mu_{On}(\tau)\big)-\Lambda_{On}=-2\delta\mathcal{L}\big(\mu_{On}(\tau)\big)+\delta\mathcal{L}\big(m_{On}(\tau)\big), \tag{4.14}$$

where $\mathcal{L}(m_{On}(\tau))$ has been given in (3.14). Substituting (3.14) into (4.14), we obtain

$$\mathcal{L}\big(\mu_{On}(\tau)\big)=\frac{M_{On}}{s+\delta}+\frac{\Lambda_{On}-M_{On}}{s+2\delta}.$$

Then applying the inverse Laplace transform to above decomposition gives

$$\mu_{On}(\tau)=M_{On}e^{-\delta\tau}+(\Lambda_{On}-M_{On})e^{-2\delta\tau},\quad 0\le\tau<Y_n. \tag{4.15}$$

Thus, the second moment of transcripts at the time $\tau=Y_n$ is

$$\mu_{On}(Y_n)=M_{On}e^{-\delta Y_n}+(\Lambda_{On}-M_{On})e^{-2\delta Y_n}. \tag{4.16}$$

Since the duration Y_n has the Erlang distribution with parameters λ and k_1, then the second moment of transcripts at the end of the n-th OFF state is given by integrating (4.16) with respect to Y_n on the interval $[0,\infty)$, that is,

$$\langle\mu_{On}\rangle=\int_0^\infty\mu_{On}(\tau)dF_{Y_n}(\tau)=M_{On}\frac{\lambda^{k_1}}{(\delta+\lambda)^{k_1}}+(\Lambda_{On}-M_{On})\frac{\lambda^{k_1}}{(2\delta+\lambda)^{k_1}}. \tag{4.17}$$

Similarly, the second moment of transcripts in the n-th ON state is the sum of $m^2\cdot p_{En}(m,\tau)$. And its time evolution is

$$\mu'_{En}(\tau)=2\nu m_{En}(\tau)+\nu-2\delta\mu_{En}(\tau)+\delta m_{En}(\tau),\quad 0\le\tau<Z_n. \tag{4.18}$$

Assuming $\mu_{En}(0)=\Lambda_{En}$ and applying the Laplace transform to (4.18), we transfer (4.18) into

$$s\mathcal{L}\big(\mu_{En}(\tau)\big)-\Lambda_{En}=2\nu\mathcal{L}\big(m_{En}(\tau)\big)+\mathcal{L}(\nu)-2\delta\mathcal{L}\big(\mu_{En}(\tau)\big)+\delta\mathcal{L}\big(m_{En}(\tau)\big).$$

Solving this equation, we have

$$\begin{aligned}\mathcal{L}\big(\mu_{En}(\tau)\big)=&\frac{\nu(\nu+\delta)}{\delta^2 s}-\frac{(2\nu+\delta)(\nu-\delta M_{En})}{\delta^2(s+\delta)}\\&+\frac{\delta^2\Lambda_{En}+\nu^2-(2\nu+\delta)\delta M_{En}}{\delta^2(s+2\delta)}.\end{aligned} \tag{4.19}$$

Applying the inverse Laplace transform to (4.19) gives

$$\begin{aligned}\mu_{En}(\tau)=&\frac{\nu(\nu+\delta)}{\delta^2}-\frac{(2\nu+\delta)(\nu-\delta M_{En})}{\delta^2}e^{-\delta\tau}\\&+\frac{\delta^2\Lambda_{En}+\nu^2-(2\nu+\delta)\delta M_{En}}{\delta^2}e^{-2\delta\tau}.\end{aligned} \tag{4.20}$$

When $\tau = Z_n$, the second moment is

$$\mu_{En}(Z_n) = \frac{\nu(\nu+\delta)}{\delta^2} - \frac{(2\nu+\delta)(\nu-\delta M_{En})}{\delta^2} e^{-\delta Z_n} + \frac{\delta^2\Lambda_{En}+\nu^2-(2\nu+\delta)\delta M_{En}}{\delta^2} e^{-2\delta Z_n}. \tag{4.21}$$

Integrating (4.21) gives the second moment of transcripts at the end of the n-th ON state, that is,

$$\begin{aligned}\langle \mu_{En} \rangle &= \int_0^\infty \mu_{En}(\tau) dF_{Z_n}(\tau) \\ &= \frac{\nu(\nu+\delta)}{\delta^2} - \frac{(2\nu+\delta)(\nu-\delta M_{En})}{\delta^2} \cdot \frac{\gamma^{k_2}}{(\delta+\gamma)^{k_2}} \\ &\quad + \frac{\delta^2\Lambda_{En}+\nu^2-(2\nu+\delta)\delta M_{En}}{\delta^2} \cdot \frac{\gamma^{k_2}}{(2\delta+\gamma)^{k_2}}.\end{aligned} \tag{4.22}$$

Since transitions between the OFF and the ON states are completed instantaneously, then the second moments of transcripts does not change during state transitions occur. Thus, we have following identities:

$$\mu_{En}(0) = \Lambda_{En} = \langle \mu_{On} \rangle, \quad \mu_{On+1}(0) = \Lambda_{On+1} = \langle \mu_{En} \rangle. \tag{4.23}$$

From (4.17), (4.22) and (4.23), we find that $\langle \mu_{On} \rangle$ and $\langle \mu_{En} \rangle$ satisfy

$$\begin{aligned}\langle \mu_{On+1} \rangle &= \langle \mu_{On} \rangle \frac{\lambda^{k_1}\gamma^{k_2}}{(2\delta+\lambda)^{k_1}(2\delta+\gamma)^{k_2}} \\ &\quad + \frac{\lambda^{k_1}}{(2\delta+\lambda)^{k_1}} \left[\frac{\nu(\nu+\delta)}{\delta^2} - \frac{\nu(2\nu+\delta)\gamma^{k_2}}{\delta^2(\delta+\gamma)^{k_2}} + \frac{\nu^2\gamma^{k_2}}{\delta^2(2\delta+\gamma)^{k_2}} \right] \\ &\quad + M_{On+1} \frac{\lambda^{k_1}\left[(2\delta+\lambda)^{k_1} - (\delta+\lambda)^{k_1}\right]}{(\delta+\lambda)^{k_1}(2\delta+\lambda)^{k_1}} \\ &\quad + M_{En} \frac{(2\nu+\delta)\lambda^{k_1}\gamma^{k_2}\left[(2\delta+\gamma)^{k_2} - (\delta+\gamma)^{k_2}\right]}{\delta(2\delta+\lambda)^{k_1}(\delta+\gamma)^{k_2}(2\delta+\gamma)^{k_2}}, \\ \langle \mu_{En+1} \rangle &= \langle \mu_{En} \rangle \frac{\lambda^{k_1}\gamma^{k_2}}{(2\delta+\lambda)^{k_1}(2\delta+\gamma)^{k_2}} \\ &\quad + \left[\frac{\nu(\nu+\delta)}{\delta^2} - \frac{\nu(2\nu+\delta)\gamma^{k_2}}{\delta^2(\delta+\gamma)^{k_2}} + \frac{\nu^2\gamma^{k_2}}{\delta^2(2\delta+\gamma)^{k_2}} \right] \\ &\quad + M_{On+1} \frac{\lambda^{k_1}\gamma^{k_2}\left[(2\delta+\lambda)^{k_1} - (\delta+\lambda)^{k_1}\right]}{(\delta+\lambda)^{k_1}(2\delta+\lambda)^{k_1}(2\delta+\gamma)^{k_2}} \\ &\quad + M_{En+1} \frac{(2\nu+\delta)\gamma^{k_2}\left[(2\delta+\gamma)^{k_2} - (\delta+\gamma)^{k_2}\right]}{\delta(\delta+\gamma)^{k_2}(2\delta+\gamma)^{k_2}}.\end{aligned}$$

From the two recurrent formulas, we could find that they are linearly dependent on M_{On} and M_{En}, then they also hold when changing the initial numbers M_{On}, M_{En} into average numbers. From (3.23) and (3.24), we get (4.9) and (4.10) by changing M_{On+1} to $\langle m_{En}\rangle$ and M_{En} to $\langle m_{On}\rangle$.

Taking limits to above two equations, we have

$$\begin{aligned}
\langle \mu_O\rangle =& \frac{\nu\lambda^{k_1}\left[(\nu+\delta)(\delta+\gamma)^{k_2}(2\delta+\gamma)^{k_2}-(2\nu+\delta)\gamma^{k_2}(2\delta+\gamma)^{k_2}+\nu\gamma^{k_2}(\delta+\gamma)^{k_2}\right]}{\delta^2(\delta+\gamma)^{k_2}\left[(2\delta+\lambda)^{k_1}(2\delta+\gamma)^{k_2}-\lambda^{k_1}\gamma^{k_2}\right]}\\
&+M_O\cdot\frac{\lambda^{k_1}(2\delta+\gamma)^{k_2}\left[(2\delta+\lambda)^{k_1}-(\delta+\lambda)^{k_1}\right]}{(\delta+\lambda)^{k_1}\left[(2\delta+\lambda)^{k_1}(2\delta+\gamma)^{k_2}-\lambda^{k_1}\gamma^{k_2}\right]}\\
&+M_E\cdot\frac{(2\nu+\delta)\lambda^{k_1}\gamma^{k_2}\left[(2\delta+\gamma)^{k_2}-(\delta+\gamma)^{k_2}\right]}{\delta(\delta+\gamma)^{k_2}\left[(2\delta+\lambda)^{k_1}(2\delta+\gamma)^{k_2}-\lambda^{k_1}\gamma^{k_2}\right]},\\
\langle \mu_E\rangle =& \frac{\nu(2\delta+\lambda)^{k_1}\left[(\nu+\delta)(\delta+\gamma)^{k_2}(2\delta+\gamma)^{k_2}-(2\nu+\delta)\gamma^{k_2}(2\delta+\gamma)^{k_2}+\nu\gamma^{k_2}(\delta+\gamma)^{k_2}\right]}{\delta^2(\delta+\gamma)^{k_2}\left[(2\delta+\lambda)^{k_1}(2\delta+\gamma)^{k_2}-\lambda^{k_1}\gamma^{k_2}\right]}\\
&+M_O\cdot\frac{\lambda^{k_1}\gamma^{k_2}\left[(2\delta+\lambda)^{k_1}-(\delta+\lambda)^{k_1}\right]}{(\delta+\lambda)^{k_1}\left[(2\delta+\lambda)^{k_1}(2\delta+\gamma)^{k_2}-\lambda^{k_1}\gamma^{k_2}\right]}\\
&+M_E\cdot\frac{(2\nu+\delta)(2\delta+\lambda)^{k_1}\gamma^{k_2}\left[(2\delta+\gamma)^{k_2}-(\delta+\gamma)^{k_2}\right]}{\delta(\delta+\gamma)^{k_2}\left[(2\delta+\lambda)^{k_1}(2\delta+\gamma)^{k_2}-\lambda^{k_1}\gamma^{k_2}\right]}.
\end{aligned}$$

Changing M_O to $\langle m_E\rangle$ and M_E to $\langle m_O\rangle$, we obtain the two stationary second moments (4.11) and (4.12). □

Theorem 4.2. *The second moments for transcripts existing in the OFF and ON states are given as*

$$\overline{\mu_O}=\overline{m_O}+[\langle\mu_E\rangle-\langle m_E\rangle]\frac{1}{k_1}\sum_{i=1}^{k_1}\left(\frac{\lambda}{2\delta+\lambda}\right)^i, \tag{4.24}$$

$$\overline{\mu_E}=\overline{m_E}+\frac{\nu}{\delta}\left[2\overline{m_E}-\frac{\nu}{\delta}\right]+\left[\langle\mu_O\rangle-\langle m_O\rangle+\frac{\nu^2}{\delta^2}-\frac{2\nu}{\delta}\langle m_O\rangle\right]\frac{1}{k_2}\sum_{i=1}^{k_2}\left(\frac{\gamma}{2\delta+\gamma}\right)^i. \tag{4.25}$$

Then the second moment of transcripts at steady state is

$$\mu^*=m^*+m^*\cdot\frac{\nu}{\delta}\cdot\left[1-\frac{\gamma\left[(\delta+\lambda)^{k_1}-\lambda^{k_1}\right]\left[(\delta+\gamma)^{k_2}-\gamma^{k_2}\right]}{k_2\delta\left[(\delta+\lambda)^{k_1}(\delta+\gamma)^{k_2}-\lambda^{k_1}\gamma^{k_2}\right]}\right]. \tag{4.26}$$

Proof. When the promoter resides at the OFF state, the analytical expression for the second moment of transcripts has been given in (4.15), and the distribution of the age y_n that the promoter remains OFF has been given in (3.28). Then the second moment for transcripts existing in the n-th OFF state is

$$\overline{\mu_{On}}=\int_0^\infty\mu_{On}(\tau)f_{y_n}(\tau)d\tau=\frac{M_{On}}{k_1}\sum_{i=1}^{k_1}\left(\frac{\lambda}{\delta+\lambda}\right)^i+\frac{\Lambda_{On}-M_{On}}{k_1}\sum_{i=1}^{k_1}\left(\frac{\lambda}{2\delta+\lambda}\right)^i.$$

When the promoter resides at the n-th ON state, the second moment of transcripts is

$$\begin{aligned}\overline{\mu_{En}}&=\int_0^\infty \mu_{En}(\tau)f_{z_n}(\tau)d\tau\\&=\frac{\nu(\nu+\delta)}{\delta^2}-\frac{(2\nu+\delta)(\nu-\delta M_{En})}{k_2\delta^2}\sum_{i=1}^{k_2}\left(\frac{\gamma}{\delta+\gamma}\right)^i\\&\quad+\frac{\delta^2\Lambda_{En}+\nu^2-(2\nu+\delta)\delta M_{En}}{k_2\delta^2}\sum_{i=1}^{k_2}\left(\frac{\gamma}{2\delta+\gamma}\right)^i.\end{aligned}$$

Taking limits to $\overline{\mu_{On}}$ and $\overline{\mu_{En}}$ with respect to n, we have

$$\overline{\mu_O}=\frac{M_O}{k_1}\sum_{i=1}^{k_1}\left(\frac{\lambda}{\delta+\lambda}\right)^i+\frac{\Lambda_O-M_O}{k_1}\sum_{i=1}^{k_1}\left(\frac{\lambda}{2\delta+\lambda}\right)^i, \tag{4.27}$$

$$\begin{aligned}\overline{\mu_E}&=\frac{\nu(\nu+\delta)}{\delta^2}-\frac{(2\nu+\delta)(\nu-\delta M_E)}{k_2\delta^2}\sum_{i=1}^{k_2}\left(\frac{\gamma}{\delta+\gamma}\right)^i\\&\quad+\frac{\delta^2\Lambda_E+\nu^2-(2\nu+\delta)\delta M_E}{k_2\delta^2}\sum_{i=1}^{k_2}\left(\frac{\gamma}{2\delta+\gamma}\right)^i.\end{aligned} \tag{4.28}$$

By changing M_O to $\langle m_E\rangle$, M_E to $\langle m_O\rangle$, Λ_O to $\langle\mu_E\rangle$ and Λ_E to $\langle\mu_O\rangle$, we get $\overline{\mu_O}$ and $\overline{\mu_E}$.

Since the stationary second moment μ^* can be split into

$$\mu^*=\overline{\mu_O}\cdot P_O^*+\overline{\mu_E}\cdot P_E^*, \tag{4.29}$$

then we only need to substitute (3.2), (4.24) and (4.25) into above equation and obtain the second moment μ^*. We have completed the proof. □

By definition of the noise strength, we can give its stationary form, that is

$$\Phi^*=1+\frac{\nu}{\delta}\cdot\left[\frac{k_1\gamma}{k_1\gamma+k_2\lambda}-\frac{\gamma\left[(\delta+\lambda)^{k_1}-\lambda^{k_1}\right]\left[(\delta+\gamma)^{k_2}-\gamma^{k_2}\right]}{k_2\delta\left[(\delta+\lambda)^{k_1}(\delta+\gamma)^{k_2}-\lambda^{k_1}\gamma^{k_2}\right]}\right]. \tag{4.30}$$

5 Discussion

The stationary transcription level and the noise strength we derived are in good agreement with existing theoretical results [30,39]. For example, if the lifetimes of the OFF and the ON states are exponentially distributed with parameters λ and γ ($k_1=1$ and $k_2=1$), then the transcription model is called telegraph model, which was firstly established by Peccoud and Ycart [30]. Based on our previous results, the stationary transcription level and the noise strength are given as

$$m^*=\frac{\nu\lambda}{\delta(\lambda+\gamma)},\quad \Phi^*=1+\frac{\nu\gamma}{(\delta+\lambda+\gamma)(\lambda+\gamma)}.$$

When $k_1=2$ and $k_2=1$, our model describes a transcription system that the promoter transits circularly among three functional states by dividing the OFF state into a ground state and an engaged state [4, 39]. By letting both the lifetimes of ground state and engaged state follow an exponential distribution with a same parameter λ, our results are consistent with the conclusions derived in [39]. At this moment, the transcription level and the noise strength are given as

$$m^*=\frac{\nu\lambda}{\delta(\lambda+2\gamma)},\quad \Phi^*=1+\frac{\nu\gamma(2\delta+3\lambda)}{[\delta^2+(2\lambda+\gamma)\delta+(\lambda^2+2\lambda\gamma)](\lambda+2\gamma)}. \tag{5.1}$$

A special case is considered by assuming the average lifetimes of the OFF and the ON states are fixed, that is, $\mathbf{E}[Y_n]=T_1$ and $\mathbf{E}[Z_n]=T_2$ are two constants. Since random variables Y_n and Z_n have Erlang distributions, then the average lifetimes of the OFF and the ON states are

$$\mathbf{E}[Y_n]=\frac{k_1}{\lambda},\quad \mathbf{E}[Z_n]=\frac{k_2}{\gamma}. \tag{5.2}$$

When k_1 and λ increase synchronously, then the stationary transcription level is maintained, but the temporal profile is different. As shown in Fig. 2, only a few mRNA molecules are synthesized during the initiation of transcription. The temporal profile displays a damped oscillation behavior. On the other hand, when k_2 and γ increase synchronously, we find that the average transcription level will reach equilibrium more quickly than in a two-state transcription system.

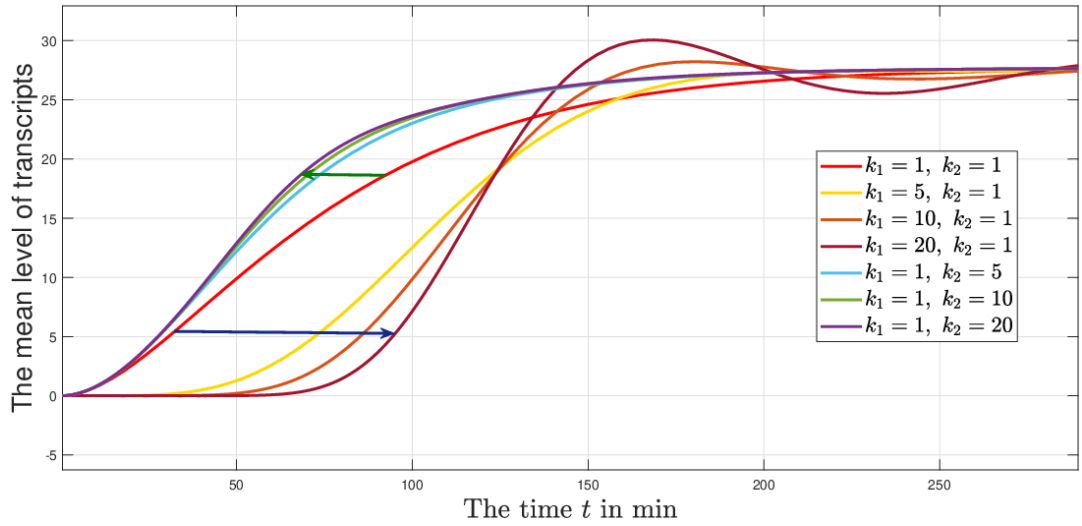

Figure 2: The temporal profile of the average transcription level. The average lifetimes of the OFF and the ON states are maintained. The red curve depicts the temporal profile of transcripts produced in the classical two-state transcription system. When k_1 increases and $k_2=1$ is fixed, the profile displays a damage oscillatory behavior. At the beginning of transcription, only a few transcripts are produced. When $k_1=1$ is fixed and k_2 increases, the profiles are simple and could enter homeostasis in a short time.

Unlike the mean transcription level, the noise strength decreases as k_1 and k_2 increase. As shown in Fig. 3, the noise strength has a maximum value when $k_1=1$ and $k_2=1$. In other words, the telegram process generates maximal noise strength. When k_1,k_2 increase to infinity and $\mathbf{E}[Y_n],\mathbf{E}[Z_n]$ are fixed, the transcription level and the noise strength at steady state are given as

$$m^*=\frac{\nu T_2}{\delta(T_1+T_2)},\quad \Phi^*=1+\frac{\nu}{\delta}\left[\frac{T_1}{T_1+T_2}-\frac{(e^{\delta T_1}-1)(e^{\delta T_2}-1)}{\delta T_2[e^{\delta(T_1+T_2)}-1]}\right]. \tag{5.3}$$

Note that the molecular memory increases while noise strength decreases as k_1 and k_2 increase, we find that the existence of molecular memory may suppress the fluctuation of mRNA molecule numbers within an isogenic cell population. Moreover, as shown in Fig. 3, transcription noise decreases to approximate the minimum value in (5.3) by slightly increasing k_1 or k_2. This suggests that the model with small number of steps between ON and OFF states may be enough to capture the observed stochastic transcription data, as shown previously the 3 steps for mouse fibroblast genes [37] and optimal 4 steps for both *E.coli* tetA promoter [48] and yeast stress response genes [28].

In fact, when T_1+T_2 is a constant, its distribution function is lattice, we could not use the alternating renewal process to analyze the transcription process. The transcription level and the noise strength could be derived by letting $k_1,k_2\to\infty$. When the shape parameters $k_1,k_2\to\infty$, we find that the lifetimes Y_n and Z_n tend to degenerate univariates. It means that the promoter spends confirmed sojourn times in the OFF and the ON states in each transcription cycle. The age y_n has a Uniform $[0,T_1]$ distribution, and z_n has a Uniform $[0,T_2]$ distribution. When $T_1\to 0$ or $T_2\to\infty$, the transcription process is usually

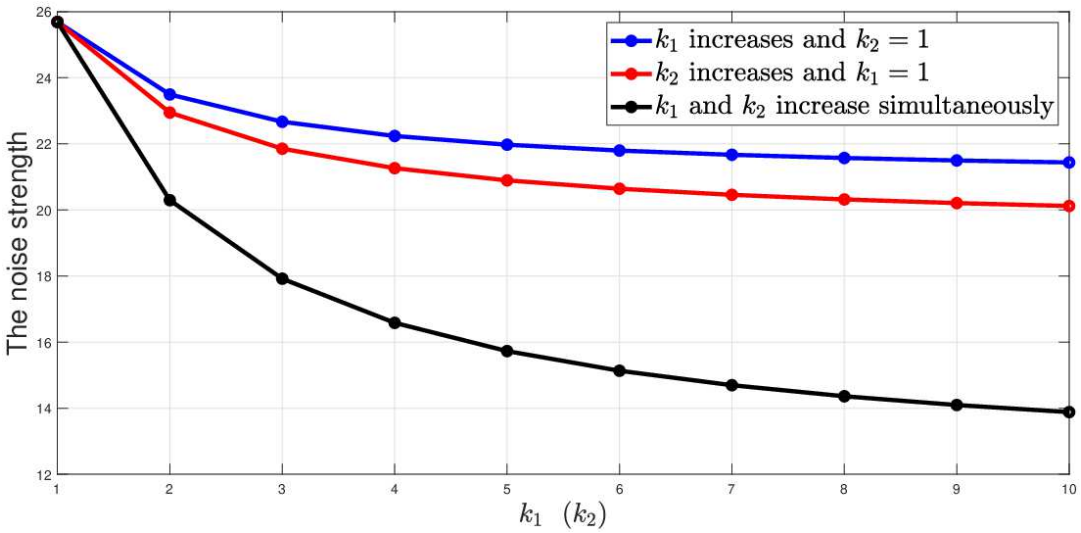

Figure 3: The noise strength decreases when k_1 and k_2 increase. The average lifetimes of the OFF and the ON states are maintained. When k_1 and/or k_2 increase, the noise strength will decrease and tend to a stationary value.

depicted by a one-state model, that is, the promoter is continuously ON. From (5.3), the mean level and the noise strength at steady state are

$$m^* = \frac{\nu}{\delta}, \quad \Phi^* = 1.$$

These results are exactly obtained in the one-state transcription model [2, 25].

Future work is required to calculate mRNA distribution $P(m,t)$. In recent years, many scholars have devoted themselves to solving the probability mass function [13, 16, 46, 47]. However, in their models, the sojourn time that the gene stays in each ON or OFF state follows an exponential distribution, but only a few of their studies have been conducted on non-exponential sojourn times [16, 46]. The method we proposed here provides a new idea for calculating the probability mass function. However, some technical difficulties need to be solved during the specific calculation process: Firstly, we need to give the time-dependent expression of the probability mass function $P(m,t|M_0)$ when the gene is always ON, where M_0 is the initial mRNA number at $t=0$. Secondly, we need to establish the iterative formulas of $P(m,Y_n|M_{On})$ and $P(m,Z_n|M_{En})$ regarding the transcription cycle n and calculate their limits. Lastly, we derive the stationary expression of $P(m)$ by averaging $P(m,\tau|m_E)$ and $P(m,\tau|m_O)$ over the sojourn times. The main difficulty is how to construct iterative sequences of $P(m,Y_n|M_{On})$ and $P(m,Z_n|M_{En})$ and find their limits.

Future work is also required to apply the method to study the cell cycle-coupled gene transcription for which mRNA synthesis rate varies with respect to the continuously increasing cell volume during each cell cycle [16, 26, 41]. Since the duration of a cell cycle is random, the waiting time for synthesizing a mRNA molecule may be assumed to follow an arbitrary distribution. We shall focus on analyzing two categories of volume growth forms: exponential and linear [26], and calculate transcription noise to study the mechanisms that maintain expression homeostasis during the cell division cycle.

6 Conclusion

The two-state gene transcription model is a conceptual framework that aims to explain the molecular mechanisms that regulate gene expression. According to this model, each gene can exist in two discrete states – an active state and an inactive state. The switch between these two states is mediated by regulatory proteins, such as transcription factors, that bind to specific DNA sequences near the gene. Usually, a set of transcription factors is needed to initiate transcription. For example, some transcription factors bind to a promoter to help form the transcription initiation complex. Other transcription factors bind to regulatory sequences to stimulate or repress transcription of the target gene. All bindings between transcription factors and promoter sequences/regulatory sequences can be deemed to be a Poisson process.

In the classical two-state model, the lifetimes of the OFF and the ON states are assumed to have exponential distributions. Many transcription models have been established to explore outputs and fluctuation of transcripts among cells. Most of these mod-

els are based on Markov processes. Usually, master equations could be easily obtained under the Markov property. When multiple transcription factors are involved in gene transcription, the lifetimes of the OFF and the ON states directly extract from the data deviate from an exponential distribution [8, 37]. Then the activation rate and the inactivation rate defined by (1.1) are time-dependent, which makes the construction of the master equation impossible. To leap over such a barrier, we presented a hybrid transcription model for connecting the random switching of promoter between active and inactive states with a determined transcription fashion in each state. By employing the alternating renewal process, we derived the time evolution of the mean transcription level and the noise strength in each state. Both the transcription level and the noise strength fluctuate with state switching between OFF and ON. Taking the Erlang distribution as an example, we gave a detailed procedure for calculating the mean and the noise strength.

In the alternating renewal process, both the sequence of random variables Y_n and the sequence Z_n are independent and identically distributed, but Y_n and Z_n are allowed to be dependent. Thus, our computing method can be extended to the study of a non-Markov system [46]. For instance, when Y_n or Z_n is uniformly distributed over an interval $(0,T]$, the transcription process is non-Markovian. By employing the computational method used in this paper, we can explore the source of the variability in mRNA numbers produced in transcription systems with different lifetimes of transcription cycles.

Acknowledgments

We thank anonymous referee for insightful comments and stimulating suggestions.

This work was supported by the National Natural Science Foundation of China (Grant Nos. 12171113, 12271118) and by the Natural Science Foundation of Guangdong (Grant No. 2022A1515010242).

References

[1] A. Adan, G. Alizada, Y. Kiraz, Y. Baran, and A. Nalbant, *Flow cytometry: Basic principles and applications*, Crit. Rev. Biotechnol., 37(2):163–176, 2017.

[2] O. G. Berg, *A model for the statistical fluctuations of protein numbers in a microbial population*, J. Theor. Biol., 71(4):587–603, 1978.

[3] Z. Cao and R. Grima, *Analytical distributions for detailed models of stochastic gene expression in eukaryotic cells*, Proc. Natl. Acad. Sci. USA, 117(9):4682–4692, 2020.

[4] J. Chen and F. Jiao, *A novel approach for calculating exact forms of mRNA distribution in single-cell measurements*, Mathematics, 10(1):27, 2022.

[5] L. Chen, C. Zhu, and F. Jiao, *A generalized moment-based method for estimating parameters of stochastic gene transcription*, Math. Biosci., 345:108780, 2022.

[6] R. D. Dar, B. S. Razooky, L. S. Weinberger, C. D. Cox, and M. L. Simpson, *The low noise limit in gene expression*, PLoS ONE, 10(10):e0140969, 2015.

[7] Y. Feng, J. Dong, and X. Tang, *Non-Markovian effect on gene transcriptional systems*, Chinese Phys. Lett., 33(10):108701, 2016.

[8] C. Fritzsch, S. Baumgärtner, M. Kuban, D. Steinshorn, G. Reid, and S. Legewie, *Estrogen-dependent control and cell-to-cell variability of transcriptional bursting*, Mol. Syst. Biol., 14(2): e7678, 2018.
[9] T. Gregor, H. G. Garcia, and S. C. Little, *The embryo as a laboratory: Quantifying transcription in Drosophila*, Trends Genet., 30(8):364–375, 2014.
[10] T. Guérin, O. Bénichou, and R. Voituriez, *Non-Markovian polymer reaction kinetics*, Nat. Chem., 4(7):568–573, 2012.
[11] S. Hao and D. Baltimore, *The stability of mRNA influences the temporal order of the induction of genes encoding inflammatory molecules*, Nat. Immunol., 10(3):281–288, 2009.
[12] C. V. Harper et al., *Dynamic analysis of stochastic transcription cycles*, PLoS Biol., 9(4):e1000607, 2011.
[13] S. Iyer-Biswas, F. Hayot, and C. Jayaprakash, *Stochasticity of gene products from transcriptional pulsing*, Phys. Rev. E, 79:031911, 2009.
[14] C. Jia, *Simplification of Markov chains with infinite state space and the mathematical theory of random gene expression bursts*, Phys. Rev. E., 96:032402, 2017.
[15] C. Jia, *Kinetic foundation of the zero-inflated negative binomial model for single-cell RNA sequencing data*, SIAM J. Appl. Math., 80(3):1336–1355, 2020.
[16] C. Jia and Y. Li, *Analytical time-dependent distributions for gene expression models with complex promoter switching mechanisms*, SIAM J. Appl. Math., 83(4):1341–1653, 2023.
[17] C. Jia , L. Wang, G. Yin, and M. Zhang, *Single-cell stochastic gene expression kinetics with coupled positive-plus-negative feedback*, Phys. Rev. E., 100:052406, 2019.
[18] Q. Jiang, X. Fu, S. Yan, R. Li, W. Du, Z. Cao, F. Qian, and R. Grima, *Neural network aided approximation and parameter inference of non-Markovian models of gene expression*, Nat. Commun., 12:2618, 2021.
[19] F. Jiao, G. Lin, and J. Yu, *Approximating gene transcription dynamics using steady-state formulas*, Phys. Rev. E., 104:014401, 2021.
[20] F. Jiao and M. Tang, *Quantification of transcription noise's impact on cell fate commitment with digital resolutions*, Bioinformatics, 38(11):3062–3069, 2022.
[21] F. Jiao and C. Zhu, *Regulation of gene activation by competitive cross-talking pathways*, Biophys. J., 119(6):1204–1214, 2020.
[22] J. E. Krebs, E. S. Goldstein, and S. T. Kilpatrick, *Lewin's Genes XI*, Jones & Bartlett Learning, 2012.
[23] N. Kumar, A. Singh, and R. V. Kulkarni, *Transcriptional bursting in gene expression: Analytical results for general stochastic models*, PLoS Comput. Biol., 11(10):e1004292, 2015.
[24] J. W. Lee, A. Gyorgy, D. E. Cameron, N. Pyenson, K. R. Choi, J. C. Way, P. A. Silver, D. Del Vecchio, and J. J. Collins, *Creating single-copy genetic circuits*, Mol. Cell, 63(2):329–336, 2016.
[25] N. Maheshri and E. K. O'Shea, *Living with noisy genes: How cells function reliably with inherent variability in gene expression*, Annu. Rev. Biophys. Biomol. Struct., 36:413–434, 2007.
[26] S. Marguerat and J.Bähler, *Coordinating genome expression with cell size*, Trends Genet., 28(11): 560–565, 2012.
[27] N. Molina, D. M. Suter, R. Cannavo, B. Zoller, I. Gotic, and F. Naef, *Stimulus-induced modulation of transcriptional bursting in a single mammalian gene*, Proc. Natl. Acad. Sci. USA, 110:20563–20568, 2013.
[28] G. Neuert, B. Munsky, R. Z. Tan, L. Teytelman, M. Khammash, and A. van Oudenaarden, *Systematic identification of signal-activated stochastic gene regulation*, Science, 339:584–587, 2013.
[29] M. Opper and G. Sanguinetti, *Learning combinatorial transcriptional dynamics from gene expression data*, Bioinformatics, 26(13):1623–1629, 2010.

[30] J. Peccoud and B. Ycart, *Markovian modeling of gene-product synthesis*, Theor. Popul. Biol., 48(2):222–234, 1995.
[31] M. K. Prajapat and A. S. Ribeiro, *Added value of autoregulation and multi-step kinetics of transcription initiation*, Royal Soc. Open Sci., 5(11):181170, 2018.
[32] L. M. Reyes Ruiz, C. L. Williams, and R. Tamayo, *Enhancing bacterial survival through phenotypic heterogeneity*, PLoS Pathog., 16(5):e1008439, 2020.
[33] S. M. Ross, *Stochastic Processes*, John Wiley & Sons, 1996.
[34] M. J. Sanderson, I. Smith, I. Parker, and M. D. Bootman, *Fluorescence microscopy*, Cold Spring Harb. Protoc., pdb.top071795, 2014.
[35] Q. Sun, M. Tang, and J. Yu, *Modulation of gene transcription noise by competing transcription factors*, J. Math. Biol., 64(3):469–494, 2012.
[36] Q. Sun, M. Tang and J. Yu, *Temporal profile of gene transcription noise modulated by cross-talking signal transduction pathways*, Bull. Math. Biol., 74(2):375–398, 2012.
[37] D. M. Suter, N. Molina, D. Gatfield, K. Schneider, U. Schibler, and F. Naef, *Mammalian genes are transcribed with widely different bursting kinetics*, Science, 332:472–474, 2011.
[38] J. Szavits-Nossan and R. Grima, *Steady-state distributions of nascent RNA for general initiation mechanisms*, Phys. Rev. Res., 5:013064, 2023.
[39] M. Tang, *The mean and noise of stochastic gene transcription*, J. Theor. Biol., 253(2):271–280, 2008.
[40] T. Udomlumleart, S. Hu, and S. Garg, *Lineages of embryonic stem cells show non-Markovian state transitions*, iScience, 24(8):102879, 2021.
[41] Y. Voichek, R. Bar-Ziv, and N. Barkai, *Expression homeostasis during DNA replication*, Science, 351(6277):1087–1090, 2016.
[42] J. Wang, S. Zhang, H. Lu, and H. Xu, *Differential regulation of alternative promoters emerges from unified kinetics of enhancer-promoter interaction*, Nat. Commun., 13(1):2714, 2022.
[43] X. Yang, S. Luo, Z. Zhang, Z. Wang, T. Zhou, and J. Zhang, *Silent transcription intervals and translational bursting lead to diverse phenotypic switching*, Phys. Chem. Chem. Phys, 24:26600–26608, 2022.
[44] J. Yu, Q. Sun, and M. Tang, *The nonlinear dynamics and fluctuations of mRNA levels in cross-talking pathway activated transcription*, J. Theor. Biol., 363:223–234, 2014.
[45] C. Zhang and F. Jiao, *Using steady-state formula to estimate time-dependent parameters of stochastic gene transcription models*, BioSystems, 236:105128, 2024.
[46] J. Zhang and T. Zhou, *Markovian approaches to modeling intracellular reaction processes with molecular memory*, Proc. Natl. Acad. Sci. USA, 116(47):23542–23550, 2019.
[47] T. Zhou and J. Zhang, *Analytical results for a multistate gene model*, SIAM J. Appl. Math., 72(3):789–818, 2012.
[48] C. Zimmer, A. Häkkinen, and A. S. Ribeiro, *Estimation of kinetic parameters of transcription from temporal single-RNA measurements*, Math. Biosci., 271:146–153, 2016.