A Novel Approach for Calculating Exact Forms of mRNA Distribution in Single-Cell Measurements

Jiaxin Chen and Feng Jiao *

Guangzhou Center for Applied Mathematics, Guangzhou University, Guangzhou 510006, China; jiaxin19982021@163.com
* Correspondence: jiaof@gzhu.edu.cn

Abstract: Gene transcription is a stochastic process manifested by fluctuations in mRNA copy numbers in individual isogenic cells. Together with mathematical models of stochastic transcription, the massive mRNA distribution data that can be used to quantify fluctuations in mRNA levels can be fitted by \( P_m(t) \), which is the probability of producing \( m \) mRNA molecules at time \( t \) in a single cell. Tremendous efforts have been made to derive analytical forms of \( P_m(t) \), which rely on solving infinite arrays of the master equations of models. However, current approaches focus on the steady-state (\( t \to \infty \)) or require several parameters to be zero or infinity. Here, we present an approach for calculating \( P_m(t) \) with time, where all parameters are positive and finite. Our approach was successfully implemented for the classical two-state model and the widely used three-state model and may be further developed for different models with constant kinetic rates of transcription. Furthermore, the direct computations of \( P_m(t) \) for the two-state model and three-state model showed that the different regulations of gene activation can generate discriminated dynamical bimodal features of mRNA distribution under the same kinetic rates and similar steady-state mRNA distribution.

Keywords: gene transcription; master equations; generating function; time-dependent mRNA distribution

1. Introduction

Gene transcription is a central biological process that transcribes genetic information stored in DNA into messenger RNA (mRNA) molecules. This process undergoes a series of complex interplays among transcription factors, RNA polymerases, chromatin modifying enzymes, and so on to control the complement of mRNA in the cell [1,2]. mRNAs are often randomly produced because of low copy numbers of biochemical molecules participating in their random collisions. The stochasticity of gene transcription has been manifested by observed pulsatile bursts, with each episode of transcriptional activity interrupted by irregular gene-off periods, which in turn leads to mRNA level fluctuations in an isogenic cell population [3–5].

The transcriptional fluctuation is conventionally quantified by measuring mRNA distribution histogram data for the gene of interest at the single-cell level [3,6] (Figure 1), which can be performed by counting the number of mRNA molecules in individual cells using single-molecule RNA fluorescence in situ hybridization (smFISH) for selected genes with single-transcript sensitivity [7,8] or exploiting the single-cell RNA sequencing (scRNA-seq) technology for abundances of genes of the whole transcriptome [4,9]. The histogram data provide a statistics base for approximations of mass function \( P_m(t) \) that denotes the probability of producing \( m \) copies of mRNA molecules in one cell at time \( t \), and the transcription noise defined as the variance of mRNA copy numbers divided by mean squared [7,10–12]. The histogram and mass function \( P_m(t) \) depict a panoramic view of mRNA distribution, whereas the noise quantifies the deviation of mRNA numbers in individual cells from the mean. When combined, these concepts may also signify a phenotypic trend in the cell population [1,13–15].
Most mRNA distributions observed in single-cell RNA-sequencing (seq) data are of three types [6,16]: Decaying distribution (Figure 1a), as shown by the tetO promoter in mammalian cells [10], for which $P_m$ decreases in $m$; unimodal distribution (Figure 1b), as shown by the Mbln2 gene in mouse fibroblast cells [4], for which $P_m$ peaks uniquely exist at $m > 0$; bimodal distribution (Figure 1c), as shown by E. coli genes [17], for which $P_m$ has two peaks: there is a peak at $m = 0$ and the another peak at $m > 1$. A decaying or unimodal distribution with low noise may suggest a phenotypic homogeneity, while a bimodal distribution with high noise supports a binary process that steers cells into sub-populations with distinct cell identities to implement the cellular “bet-hedging” strategy in response to environmental stresses [6,18–20].

![Figure 1. Three modes of mRNA distribution. (a) Decaying distribution for which the mass function $P_m(t)$ gradually decreases in mRNA copy number $m = 0, 1, \ldots$. (b) Unimodal distribution that $P_m$ has a unique peak at $m > 0$. (c) Bimodal distribution that $P_m$ takes two peaks: one at $m = 0$ and the other at $m > 0$. mRNA distribution data (blue bars) and their fits by $P_m$ (red lines) of the two-state model (2) are shown for the (a) tetO promoter in CHO cells [10], (b) Mbln2 gene in mouse fibroblast cells [4], and (c) promoter P lac/ara in E. coli cells [17].](image)

When combined with mathematical models, the computation and analysis of mass function $P_m(t)$ facilitate the understanding of how the different gene regulation mechanisms and system parameters influence mRNA distribution profiles for both prokaryotic and eukaryotic genes [7,21–23]. In stochastic transcription models, the calculation for the exact forms of $P_m(t)$ is transformed to solve the corresponding chemical master equations that involve continuous variable time $t$ and discrete random variable $m$ of mRNA copy number [24–26]. By introducing the generating function $V(z,t)$ with two variables, $z$ and $t$,

$$V(z,t) = \sum_{m=0}^{\infty} (1+z)^m P_m(t),$$

and multiplying the master equations by $(1+z)^m$, the master equations can be transformed into a system of partial differential equations of $V(z,t)$ [2,24]. Thereafter, if $V(z,t)$ can be solved from the partial differential equations, then $P_m(t)$ is obtained by the reverse calculation

$$P_m(t) = \frac{1}{m!} \left. \frac{\partial^m V(z,t)}{\partial z^m} \right|_{z=1}. \quad (1)$$

Tremendous efforts have been made to understand the steady-state cases of mathematical models for which the system of partial differential equations of $V(z,t)$ is reduced to a system of ordinary differential equations as $t \to \infty$. The expressions of $\lim_{t \to \infty} V(z,t)$ and $\lim_{t \to \infty} P_m(t)$ can be solved in the form of hypergeometric functions or singular integrals [2,10,16], which have facilitated good fits to the three distribution modes in Figure 1. When time $t$ is finite, the scenario becomes more interesting as inducible genes in E. coli, yeast, and mammalian cells can exhibit all the decaying, bimodal, and unimodal distributions over time [8,17,21,27]. However, the dynamic transition patterns among dif-
different distribution modes and the duration and significance of each dynamic distribution mode are tightly regulated by different regulation mechanisms [23] and system parameter regions [24].

In this study, our aim was to calculate the exact expressions for the dynamic mRNA distribution $P_m(t)$ of mathematical models. In this case, we can still transform partial differential equations of $V(z,t)$ into a system of ordinary differential equations through the method of characteristics [24,28]. However, a standard method to obtain $V(z,t)$ is lacking owing to the presence of variable coefficients in the transformed ordinary differential equations. This dilemma limits the calculation of $V(z,t)$ only within a few of the limiting parameter regions, such as zero values of the mRNA degradation rate [29] and gene inactivation rate [30], or fast switching between two promoter states such that the two states can be mathematically approximated as a single state [23,31].

In this study, we sought to develop a new approach to solve $V(z,t)$ for larger system parameter regions of mathematical models where the parameters are positive and finite. We can transform the calculation of $V(z,t)$ by solving a classical higher-order ordinary differential equation. This breakthrough allows us to express $V(z,t)$ as a closed form of hypergeometric functions, which facilitates the computation of $P_m(t)$ at each time point through the reverse formula (1) by numerical simulation [32]. We illustrate the approach for the widely used three-state model and the corresponding master equations in Section 2. We provide a detailed derivation of $V(z,t)$ in Section 3. Several simplifications of $V(z,t)$ are discussed in Section 4, for which the exact forms of $P_m(t)$ within special parameter regions can be derived. Finally, we provide a brief discussion of our results and their potential biophysical implications in Section 5.

2. Mathematical Models and Master Equations

One of the most classical models in the study of gene transcription is the so-called two-state model that has been utilized to capture the stochastic bursty transcription [4–6,10,33].

In the two-state model

$$\text{gene off} \xrightarrow{\lambda_2} \text{gene on} \xrightarrow{\gamma} \text{mRNA} \xrightarrow{\delta} \emptyset,$$

the genes are suggested to transition randomly between off (inactive) and on (active) states. The mRNA molecules are newly produced only when genes are in the on state and are degraded following certain process. Each biochemical process in the two-state model (2) is determined by a single rate-limiting steps.

A central problem in studying stochastic gene transcription is understanding the regulation of gene random switching between off and on states under varying external conditions [3,22,33–35]. The technique of transcriptional real-time imaging allows the measurements of off and on periods for each gene within the whole timescale, which has generated massive distribution data for the duration periods of off and on states. This revealed the universal exponentially distributed on period for both prokaryotic and eukaryotic genes [1,36,37], which supported the control of a single rate-limiting step in gene inactivation in the two-state model (2). However, modeling a single rate-limiting step directing gene on may not be sufficient to capture the ordered multiple biochemical reactions that mediate the recruitment of transcriptional activators, chromatin modification enzymes, RNA polymerases and so on [1,5]. Many other models has been developed based on the two-state model (2) to capture the multiple sequential steps in the gene activation [2,31,35,36,38].

We illustrate our approach using the simplest framework of multi-step gene activation for which the gene is activated by two ordered rate-limiting steps. This two-step activation has been utilized to capture the data of non-exponential distribution with a unique peak for gene off period, as observed for tetA gene and $P_{lac/ara}$ promoter in E. coli cells under certain temperatures and stresses [35,39,40], and 16 mouse fibroblast genes under Trichostatin A (TSA) treatments [36]. By introducing two-step gene activation, the two-state model (2) is generalized to the three-state model [9,36,41,42]. As shown in Figure 2, each process in the
three-state model is described by a single rate-limiting step, with the duration for each state being independent and exponentially distributed with constant rates. The model consists of random rotation between two sequential gene off states (states 1 and 2) and one on state (state 3) with activation rates $\lambda_1, \lambda_2$ and inactivation rate $\gamma$, in coupling with the simple birth and death processes of mRNAs with synthesis rate $v$ and degradation rate $\delta$.

Figure 2. Three-state model of gene transcription [41,42].

The analytical expressions of the steady-state mRNA distribution and its mean and noise for the three-state model were derived [2,41,42]. However, it still lacks the exact forms for the dynamical mRNA distribution $P_m(t)$. Let $P_{m,i}(t), i = 1, 2, 3$ denote the probabilities that the gene resides at state $i$ with $m$ mRNA copy numbers being produced at time $t$. Then

$$P_m(t) = P_{m,1}(t) + P_{m,2}(t) + P_{m,3}(t).$$

Following the standard procedure, the calculation of $P_{m,i}(t)$ can be transformed to solve the chemical master equations with inactive genes and zero transcripts at the initial time [41]:

$$\frac{dP_{m,1}(t)}{dt} = \gamma P_{m,3}(t) - (m\delta + \lambda_1)P_{m,1}(t) + (m + 1)\delta P_{m+1,1}(t),$$

(3)

$$\frac{dP_{m,2}(t)}{dt} = \lambda_1 P_{m,1}(t) - (m\delta + \lambda_2)P_{m,2}(t) + (m + 1)\delta P_{m+1,2}(t),$$

(4)

$$\frac{dP_{m,3}(t)}{dt} = \lambda_2 P_{m,2}(t) - (v + m\delta + \gamma)P_{m,3}(t) + (m + 1)\delta P_{m+1,3}(t) + vP_{m-1,3}(t),$$

(5)

$$P_{0,1}(0) = 1, P_{0,2}(0) = P_{0,3}(0) = 0, \text{ and } P_{m,i} = 0 \text{ for } m \geq 1 \text{ and } i = 1, 2, 3.$$

(6)

where, by convention, $P_{-1,3}(t) = 0$.

The system (3)–(6) cannot be solved by iteration as the evolution of $P_{m,i}(t)$ is related to $P_{m+1,i}(t), P_{m-1,i}(t)$, and itself. The traditional method introduces generating functions

$$V_i(z,t) = \sum_{m=0}^{\infty} (1+z)^m P_{m,i}(t), \ i = 1, 2, 3,$$

(7)

which transform the system (3)–(6) to partial differential equations [2]:

$$\frac{\partial V_1}{\partial t}(z,t) = -\lambda_1 V_1(z,t) + \gamma V_3(z,t) - \delta z \frac{\partial V_1}{\partial z}(z,t),$$

(8)

$$\frac{\partial V_2}{\partial t}(z,t) = -\lambda_2 V_2(z,t) + \lambda_1 V_1(z,t) - \delta z \frac{\partial V_2}{\partial z}(z,t),$$

(9)

$$\frac{\partial V_3}{\partial t}(z,t) = -\gamma V_3(z,t) + \lambda V_2(z,t) + v z V_3(z,t) - \delta z \frac{\partial V_3}{\partial z}(z,t),$$

(10)

$V_1(z,0) = 1, \ V_2(z,0) = V_3(z,0) = 0.$

(11)
System (8)–(11) possesses a unique solution according to the theory of partial differential equations [28]. However, the exact expression of the solution remains elusive.

3. Expressing Generating Function \( V(z,t) \)

When time \( t \to \infty \), the expression of \( \lim_{t \to \infty} V(z,t) \) can be expressed by a single generalized hypergeometric function \( 2F_2 \) [2], denoted by

\[
2F_2(x_1, x_2; y_1, y_2; x) = 1 + \sum_{k=1}^{\infty} \frac{(x_1)_k (x_2)_k}{k! (y_1)_k (y_2)_k} x^k,
\]

with \( (a)_k = a(a+1) \cdots (a+k-1) \), \( k \geq 1 \), in the form of rescaled parameters [2,41]

\[
\lambda_1 = \frac{\lambda_1}{\delta}, \quad \lambda_2 = \frac{\lambda_2}{\delta}, \quad \gamma = \frac{\gamma}{\delta}, \quad \delta = \frac{\nu}{\delta}, \quad \delta_1 = \bar{\alpha} + \bar{\beta}, \quad \delta_s = \bar{\alpha} - \bar{\beta}, \tag{12}
\]

with

\[
\bar{\alpha} = \frac{\lambda_1 + \lambda_2 + \gamma}{2} \quad \text{and} \quad \bar{\beta} = \sqrt{\delta^2 - (\lambda_1 \lambda_2 + \lambda_1 \gamma + \lambda_2 \gamma)}, \quad \delta^2 > \bar{\beta}.
\]

When the time is finite, the next result shows that \( V(z,t) \) takes more complicated expression compared to the steady-state case.

**Theorem 1.** The generating function \( V(z,t) \) given by (7) can be expressed in a closed form

\[
V(z,t) = \frac{A(vze^{-\delta t})}{2\beta_0 e^{\delta t} - \delta_z} 2F_2(\lambda_1, \lambda_2; \delta_z, \delta_s; v\bar{z})
- \frac{\nu \lambda_1 \lambda_2 B(vze^{-\delta t})e^{-\delta t}}{2\beta_0 \delta_1 (\delta_1 - 1) e^{\delta t}} 2F_2(\lambda_1 - \delta_1 + 1, \lambda_2 - \delta_1 + 1; 2 - \delta_1, -2\beta + 1; v\bar{z})
+ \frac{\nu \lambda_1 \lambda_2 C(vze^{-\delta t})e^{-\delta t}}{2\beta_0 \delta_1 (\delta_1 - 1) e^{\delta t}} 2F_2(\lambda_1 - \delta_1 + 1, \lambda_2 - \delta_1 + 1; 2 - \delta_1, 2\beta + 1; \bar{v}z).
\tag{13}
\]

Here \( A, B \) and \( C \) are given by

\[
A(s) = \delta_1 x_2 F_2(\lambda_1 - \delta_1 + 1, \lambda_2 - \delta_1 + 1; -\delta_1, -2\beta + 1; s)
- \delta_1 x_2 F_2(\lambda_1 - \delta_1 + 1, \lambda_2 - \delta_1 + 1; -\delta_1, -2\beta + 1; s)
+ \delta_1 x_2 F_2(\lambda_1 - \delta_1 + 1, \lambda_2 - \delta_1 + 1; -\delta_1, -2\beta + 1; s)
\]

\[
B(s) = \frac{(\lambda_1 + 1)(\lambda_2 + 1)s}{(\delta_1 + 1)(\delta_2 + 1)} 2F_2(\lambda_1 + 2, \lambda_2 + 2; \delta_1 + 2, \delta_2 + 2; s)
- \delta_1 x_2 F_2(\lambda_1 - \delta_1 + 1, \lambda_2 - \delta_1 + 1; 1 - \delta_1, 2\beta + 1; s)
+ \delta_1 x_2 F_2(\lambda_1 + 1, \lambda_2 + 1; \delta_1 + 1, \delta_2 + 1; s)
\]

\[
C(s) = \frac{(\lambda_1 + 1)(\lambda_2 + 1)s}{(\delta_1 + 1)(\delta_2 + 1)} 2F_2(\lambda_1 + 2, \lambda_2 + 2; \delta_1 + 2, \delta_2 + 2; s)
- \delta_1 x_2 F_2(\lambda_1 - \delta_1 + 1, \lambda_2 - \delta_1 + 1; 1 - \delta_1, -2\beta + 1; s)
+ \delta_1 x_2 F_2(\lambda_1 + 1, \lambda_2 + 1; \delta_1 + 1, \delta_2 + 1; s)
\]

\[
\times x_2 F_2(\lambda_1 - \delta_1 + 1, \lambda_2 - \delta_1 + 1; -\delta_1, 2\beta + 1; s).
\]
\textbf{Proof.} By the method of characteristics [24,28], we let \( z_0 \) be a parameter, and
\[
z = z_0 e^{it}, \quad u_i(t) = V_i\left(z_0 e^{it}, t\right), \quad i = 1, 2, 3, \quad \text{and} \quad u(t) = u_1(t) + u_2(t) + u_3(t). \tag{14}
\]

Then Equations (8)–(10) can be transformed into the initial value problems
\[
u'_1(t) = -\lambda_1 u_1(t) + \gamma u_3(t), \tag{15}
\]
\[
u'_2(t) = -\lambda_2 u_2(t) + \lambda_1 u_1(t), \tag{16}
\]
\[
u'_3(t) = \lambda_2 u_2(t) - \gamma u_3(t) + \nu_0 e^{it} u_3(t), \tag{17}
\]
\[
u_1(0) = 1, \quad \nu_2(0) = 0. \tag{18}
\]

To transform this system further, we define
\[
\nu_i(x) = u_i(t), \quad i = 1, 2, 3, \quad \text{and} \quad \nu(x) = u(t), \quad \text{where} \quad x = \sigma z_0 e^{it}. \tag{19}
\]

By using the chain rule, we find \( \nu'_i(t) = \delta x \nu'_i(x) \) for \( i = 1, 2, 3 \), and transform (15)–(18) to a more tractable problem
\[
x \nu'_i(x) = -\lambda_1 \nu_1(x) + \gamma \nu_3(x), \tag{20}
\]
\[
x \nu'_2(x) = -\lambda_2 \nu_2(x) + \lambda_1 \nu_1(x), \tag{21}
\]
\[
x \nu'_3(x) = \lambda_2 \nu_2(x) - \gamma \nu_3(x) + x \nu_3(x), \tag{22}
\]
\[
\nu_1(0) = 1, \quad \nu_2(0) = 0. \tag{23}
\]

Next, we proceed by extracting third-order linear equations for \( \nu_3 \) and \( \nu = \nu_1 + \nu_2 + \nu_3 \) from (20)–(22). Let \( a, b, c \) and \( d \) be real constants, and \( y = y(x) \) be a smooth function of \( x \). We introduce the linear operator
\[
L_{c,d}(y) = x^2 y'''(x) + x(1 - x + c + d)y''(x) + (cd - x(a + b + 1))y'(x) - aby(x). \tag{24}
\]

We claim that \( \nu_3 \) and \( \nu \) are separately the unique solutions for the following initial value problems
\[
L_{\beta_1 + 1, \delta_2 + 1}^{\lambda_1 + 1, \lambda_2 + 1}(\nu_3) = 0, \quad \nu_3(\sigma \nu_0) = \nu_3'(\sigma \nu_0) = 0 \quad \text{and} \quad \nu_3''(\sigma \nu_0) = \lambda_1 \lambda_2 / (\nu \nu_0)^2, \tag{25}
\]
and
\[
L_{\delta_1, \delta_2}^{\lambda_1, \lambda_2}(\nu) = 0, \quad \nu(\sigma \nu_0) = 1 \quad \text{and} \quad \nu'(\sigma \nu_0) = \nu''(\sigma \nu_0) = 0. \tag{26}
\]

The verification of the initial conditions is trivial. To confirm \( L_{\beta_1 + 1, \delta_2 + 1}^{\lambda_1 + 1, \lambda_2 + 1}(\nu_3) = 0 \), we first multiply Equation (22) by \( x^{\lambda_2} \) and then take derivatives. It gives
\[
x^{\lambda_2 + 1} \nu_3''(x) + (\lambda_2 + 1) x^{\lambda_2} \nu_3'(x) = \lambda_2 x^{\lambda_2} w_2'(x) + \lambda_2^2 x^{\lambda_2 - 1} w_2(x) - (\sigma x^{\lambda_2} - x^{\lambda_2 + 1}) \nu_3'(x)
\]
\[-(\lambda_2 \sigma x^{\lambda_2 - 1} - (\lambda_2 + 1) x^{\lambda_2}) w_3(x).
\]

A substitution of Equation (21) cancels out the first two terms in the right hand side and simplifies the identity to
\[
x^2 \nu_3''(x) + x((\lambda_2 + \gamma + 1) - x) \nu_3'(x) + (\lambda_2 \gamma - (\lambda_2 + 1) x) \nu_3(x) - \lambda_1 \lambda_2 w_1(x) = 0. \tag{27}
\]
In turn, we multiply (27) by $x^{\hat{\lambda}_1}$ and then take derivatives. We get
\[
x^2w'''(x) + x((2\delta + 3) - x)w''(x) + (\xi + 2\lambda + 1 - x(\lambda_1 + \gamma + 2))w'(x)
\]
\[
+ (\lambda_1\lambda_2\gamma x^{-1} - (\lambda_1 + 1)(\lambda_2 + 1))w_3(x) - \lambda_1\lambda_2w'_1(x) - \lambda_1^2\lambda_2 x^{-1}w_1(x) = 0.
\] (28)

Substitution of Equation (20) into (28) cancels out the last two terms in the left hand side and simplifies the identity to
\[
x^2w'''(x) + x((2\delta + 3) - x)w''(x)
\]
\[
+ [\xi + 2\lambda + 1 - x(\lambda_1 + \lambda_2 + 3)]w_3(x) - (\lambda_1 + 1)(\lambda_2 + 1)w_3(x) = 0.
\] (29)

According to the definition (12), we have
\[
(\delta_5 + 1) + (\delta_1 + 1) = 2\lambda + 2 \quad \text{and} \quad (\delta_5 + 1)(\delta_1 + 1) = \xi + 2\lambda + 1.
\] (30)

From (29) and (30) it follows immediately that
\[
L_{\delta_1 + 1, \delta_5 + 1}^{\lambda_1, \lambda_2}(w_3) = x^2y'''(x) + x[1 - x + (\delta_1 + 1) + (\delta_5 + 1)]y''(x)
\]
\[
- (\lambda_1 + 1)(\lambda_2 + 1)y'(x) + [(\delta_1 + 1)(\delta_5 + 1) - x((\lambda_1 + 1) + (\lambda_2 + 1) + 1)]y'(x)
\]
\[
- x((\lambda_1 + 1) + (\lambda_2 + 1) + 1)]y'(x) = 0.
\]

To confirm $L_{\delta_1, \delta_5}^{\lambda_1, \lambda_2}(w) = 0$, we first derive $w'(x) = w_3(x)$ by adding Equations (26)–(22), and then use $L_{\delta_1 + 1, \delta_5 + 1}^{\lambda_1, \lambda_2 + 1}(w_3) = 0$ to find
\[
\left[ L_{\delta_1, \delta_5}^{\lambda_1, \lambda_2}(w) \right]' = [x^2w'''(x) + x[1 - x + \delta_1 + \delta_5]w''(x)
\]
\[
+ [\delta_1\delta_5 - x(\lambda_1 + \lambda_2 + 1)]w'(x) - \lambda_1\lambda_2w(x)]'
\]
\[= x^2w'''(x) + x[1 - x + (\delta_1 + 1) + (\delta_5 + 1)]w''(x)
\]
\[+ [(\delta_1 + 1)(\delta_5 + 1) - x((\lambda_1 + 1) + (\lambda_2 + 1) + 1)]w'(x)
\]
\[- (\lambda_1 + 1)(\lambda_2 + 1)w'(x)
\]
\[= x^2w'''(x) + x[1 - x + (\delta_1 + 1) + (\delta_5 + 1)]w''(x)
\]
\[+ [(\delta_1 + 1)(\delta_5 + 1) - x((\lambda_1 + 1) + (\lambda_2 + 1) + 1)]w'(x)
\]
\[- (\lambda_1 + 1)(\lambda_2 + 1)w_3(x)
\]
\[= L_{\delta_1 + 1, \delta_5 + 1}^{\lambda_1, \lambda_2 + 1}(w_3) = 0.
\]

This implies that $L_{\delta_1, \delta_5}^{\lambda_1, \lambda_2}(w)$ is a constant. Be applying initial values (25)–(26) and the relation $w'(x) = w_3(x)$, we find
\[
L_{\delta_1, \delta_5}^{\lambda_1, \lambda_2}(w) = L_{\delta_1, \delta_5}^{\lambda_1, \lambda_2}(w(\tilde{\nu}z_0)) = (\tilde{\nu}z_0)z_0^2w''(\tilde{\nu}z_0) - \lambda_1\lambda_2 = (\tilde{\nu}z_0)^2w''(\tilde{\nu}z_0) - \lambda_1\lambda_2 = 0.
\]

To proceed, it is crucial to solve the initial value problem (26). Once the solution $w(x)$ is obtained, $V(z,t)$ follows immediately by substituting $x = \nu z$ and $z_0 = ze^{-\nu t}$ into $w(x)$.
Noting that \( L_{\delta_2}^{\tilde{\lambda}_2}(w) = 0 \) in (26) is the generalized hypergeometric equation. If \( \delta_1 - \delta_2 \neq 1, 2, \ldots \), then \( L_{\delta_1}^{\tilde{\lambda}_1}(w) = 0 \) possesses three independent particular solutions [43]

\[
\begin{align*}
\tilde{W}_{1}(x) &= 2F_2(\tilde{\lambda}_1, \tilde{\lambda}_2; \delta_1, \delta_1; x), \\
\tilde{W}_{2}(x) &= x^{1-\delta_2}2F_2(\tilde{\lambda}_1 - \delta_1 + 1, \tilde{\lambda}_2 - \delta_1 + 1; 2 - \delta_1; -2\delta_1 + 1; x), \\
\tilde{W}_{3}(x) &= x^{1-\delta_2}2F_2(\tilde{\lambda}_1 - \delta_2 + 1, \tilde{\lambda}_2 - \delta_2 + 1; 2 - \delta_2; 2\delta_2 + 1; x).
\end{align*}
\]

Therefore, we derive the general solution for \( L_{\delta_1, \delta_2}(w) = 0 \):

\[
w(x) = Aw_1(x) + Bw_2(x) + Cw_3(x),
\]

where \( A, B \) and \( C \) are undetermined constants. Substituting initial value condition \( w(\bar{v}z_0) = 1 \) and \( w'(\bar{v}z_0) = w''(\bar{v}z_0) = 0 \) into (34) leads to the following system of equations

\[
\begin{align*}
A[w_1(\bar{v}z_0)w_3'(\bar{v}z_0) - w_1'(\bar{v}z_0)w_3(\bar{v}z_0)] + B[w_2(\bar{v}z_0)w_3'(\bar{v}z_0) - w_2'(\bar{v}z_0)w_3(\bar{v}z_0)] &= w_3''(\bar{v}z_0). \\
A[w_1'(\bar{v}z_0)w_3''(\bar{v}z_0) - w_1''(\bar{v}z_0)w_3'(\bar{v}z_0)] + B[w_2'(\bar{v}z_0)w_3''(\bar{v}z_0) - w_2''(\bar{v}z_0)w_3'(\bar{v}z_0)] &= w_3'''(\bar{v}z_0).
\end{align*}
\]

In order to solve (35)–(37), we multiply (37) by \(-w_3'(\bar{v}z_0)/w_3''(\bar{v}z_0)\) and add (35), we obtain

\[
A[w_1'((\bar{v}z_0)w_3''(\bar{v}z_0)) - w_1''((\bar{v}z_0)w_3'(\bar{v}z_0))] + B[w_2'((\bar{v}z_0)w_3''(\bar{v}z_0)) - w_2''((\bar{v}z_0)w_3'(\bar{v}z_0))] = w_3'''(\bar{v}z_0).
\]

Similarly, multiplying (37) by \(-w_3'(\bar{v}z_0)/w_3''(\bar{v}z_0)\) and adding (36) lead to

\[
A[w_1'((\bar{v}z_0)w_3''(\bar{v}z_0)) - w_1''((\bar{v}z_0)w_3'(\bar{v}z_0))] + B[w_2'((\bar{v}z_0)w_3''(\bar{v}z_0)) - w_2''((\bar{v}z_0)w_3'(\bar{v}z_0))] = w_3'''(\bar{v}z_0).
\]

Solving (38) and (39) through straightforward calculation, we obtain

\[
\begin{align*}
A &= \frac{w_1'(\bar{v}z_0)w_3''(\bar{v}z_0) - w_1''(\bar{v}z_0)w_3'(\bar{v}z_0)}{\bar{W}(\bar{v}z_0)}, \\
B &= \frac{w_2'(\bar{v}z_0)w_3''(\bar{v}z_0) - w_2''(\bar{v}z_0)w_3'(\bar{v}z_0)}{\bar{W}(\bar{v}z_0)}.
\end{align*}
\]

where \( \bar{W}(s) \) is the Wronskian of \( w_1, w_2 \) and \( w_3 [32, 43] \), namely,

\[
\begin{align*}
\bar{W}(s) &= w_1(s)[w_2'(s)w_3''(s) - w_2''(s)w_3'(s)] - w_1'(s)[w_2(s)w_3''(s) - w_2''(s)w_3(s)] \\
&\quad + w_1''(s)[w_2(s)w_3'(s) - w_2'(s)w_3(s)].
\end{align*}
\]

Substituting (40) and (41) into (37) leads to

\[
C = \frac{w_1'(\bar{v}z_0)w_3''(\bar{v}z_0) - w_1''(\bar{v}z_0)w_3'(\bar{v}z_0)}{\bar{W}(\bar{v}z_0)}.
\]

In order to express \( A, B \) and \( C \) in the form of \( 2F_2 \), we need to calculate \( w_i(s), w''_i(s), i = 1, 2, 3 \), and \( \bar{W}(s) \). Noting the theory on the differentiation of the generalized hypergeometric function that [43]

\[
\frac{d^k}{dx^k}2F_2(x_1, x_2; y_1, y_2; x) = \frac{(x_1)_k(x_2)_k}{(y_1)_k(y_2)_k}2F_2(x_1 + k, x_2 + k; y_1 + k, y_2 + k; x), k = 1, \ldots
\]
and
\[
\frac{d^k x^{y_1-1} F_2(x_1, x_2; y_1, y_2; x)}{dx^k} = (y_1 - k) x^{y_1-k-1} F_2(x_1, x_2; y_1 - k, y_2; x), k = 1, \ldots. \tag{45}
\]

Applying (31)–(33), (44) and (45), we obtain

\[
w'_1(s) = \frac{\lambda_1 \lambda_2}{\delta_1 \delta_2} 2F_2(\lambda_1 + 1, \lambda_2 + 1; \delta_1 + 1, \delta_2 + 1; s), \tag{46}
\]

\[
w''_1(s) = \frac{\lambda_1 \lambda_2 (\lambda_1 + 1) (\lambda_2 + 1)}{\delta_1 \delta_2 (\delta_1 + 1) (\delta_2 + 1)} 2F_2(\lambda_1 + 2, \lambda_2 + 2; \delta_1 + 1, \delta_2 + 2; s), \tag{47}
\]

\[
w'_2(s) = \frac{1 - \delta_1}{s \delta_1 + 1} 2F_2(\lambda_1 - \delta_1 + 1, \lambda_2 - \delta_1 + 1; 1 - \delta_1, -2\beta + 1; s), \tag{48}
\]

\[
w''_2(s) = \frac{\delta_1 (1 - \delta_1)}{s \delta_1 + 1} 2F_2(\lambda_1 - \delta_1 + 1, \lambda_2 - \delta_1 + 1; -\delta_1, -2\beta + 1; s), \tag{49}
\]

\[
w'_3(s) = \frac{1 - \delta_2}{s \delta_2 + 1} 2F_2(\lambda_1 - \delta_2 + 1, \lambda_2 - \delta_2 + 1; 1 - \delta_2, 2\beta + 1; s), \tag{50}
\]

\[
w''_3(s) = -\frac{\delta_2 (1 - \delta_2)}{s \delta_2 + 1} 2F_2(\lambda_1 - \delta_2 + 1, \lambda_2 - \delta_2 + 1; -\delta_2, 2\beta + 1; s). \tag{51}
\]

We claim
\[
\check{W}(s) = 2\beta (\delta_1 - 1) (\delta_2 - 1) e^{s \delta_1 + \delta_2 + 1}. \tag{52}
\]

To verify (52), we notice that \(w_1, w_2\) and \(w_3\) are the independent solutions of the third order ordinary differential equation \(L_{\delta, \lambda, \beta}^x(w) = 0\) given in (26). Therefore, for any given real number \(s_0\), \(\check{W}(s)\) takes the following form [44]

\[
\check{W}(s) = \check{W}(s_0) \exp\left\{ - \int_{s_0}^s \frac{\delta_1 + 1 - s}{s} ds \right\} = \check{W}(s_0) e^{s_0 (s_0 / s) \delta_1 + \delta_2 + 1}. \tag{53}
\]

Since \(s_0\) is arbitrary, we choose \(s_0 = 0\). Noting that \(2F_2(x_1, x_2; y_1, y_2; 0) = 1\). Substituting (31)–(33), (42) and (46)–(51) into (53) leads to (52) immediately.

We are now in a position to substitute (40), (41), (43) and (52) into (34) and obtain

\[
w(x) = \frac{w'_1(\tilde{\nu}z_0) w''_1(\tilde{\nu}z_0) - w''_1(\tilde{\nu}z_0) w'_1(\tilde{\nu}z_0)}{2\beta (\delta_1 - 1) (\delta_2 - 1)} e^{-\tilde{\nu}z_0 (\tilde{\nu}z_0) \delta_1 + \delta_2 + 1} w_1(x) \\
+ \frac{w''_1(\tilde{\nu}z_0) w'_2(\tilde{\nu}z_0) - w'_1(\tilde{\nu}z_0) w''_2(\tilde{\nu}z_0)}{2\beta (\delta_1 - 1) (\delta_2 - 1)} e^{-\tilde{\nu}z_0 (\tilde{\nu}z_0) \delta_1 + \delta_2 + 1} w_2(x) \\
+ \frac{w'_1(\tilde{\nu}z_0) w'_2(\tilde{\nu}z_0) - w''_1(\tilde{\nu}z_0) w'_2(\tilde{\nu}z_0)}{2\beta (\delta_1 - 1) (\delta_2 - 1)} e^{-\tilde{\nu}z_0 (\tilde{\nu}z_0) \delta_1 + \delta_2 + 1} w_3(x).
\]

By substituting (46)–(51) into \(w(x)\), we get

\[
w(x) = \frac{A(\tilde{\nu}z_0) w_1(x) - \lambda_1 \lambda_2 (\tilde{\nu}z_0) \delta_1 B(\tilde{\nu}z_0) + \lambda_1 \lambda_2 (\tilde{\nu}z_0) \delta_2 C(\tilde{\nu}z_0)}{2\beta \delta_1 (\delta_1 - 1) e^{\tilde{\nu}z_0}} w_2(x) + \frac{\lambda_1 \lambda_2 (\tilde{\nu}z_0) \delta_1 C(\tilde{\nu}z_0)}{2\beta \delta_1 (\delta_1 - 1) e^{\tilde{\nu}z_0}} w_3(x), \tag{54}
\]

where \(A, B\) and \(C\) are given in the theorem. By substituting (31)–(33) into (54) and replacing \(x\) with \(\tilde{\nu}z\), and \(z_0\) with \(z e^{-\delta t}\), we convert \(w(x)\) to (13). The proof is completed. \(\square\)

4. Probability Mass Function \(P_m(t)\)

The complicated expression \(V(z, t)\) given by (13) does not permit a direct utilization of (1) to derive \(P_m(t)\). However, when the parameters fall within some special regions,
the hypergeometric function $\,_{2}F_{2}$ can be dramatically simplified, which may allow the calculation of $P_m(t)$ using (1).

4.1. Inactive Gene with $\gamma = 2\delta$ and $\lambda_1 + \lambda_2 = \delta$

Denoted by $T_{\text{off}}$ and $T_{\text{on}}$ the durations of gene off and on states, respectively. When the parameters fall within the region $\gamma = 2\delta$ and $\lambda_1 + \lambda_2 = \delta$, there is

$$T_{\text{off}} = \frac{1}{\lambda_1} + \frac{1}{\lambda_2} = \frac{1}{\lambda_1} + \frac{1}{\delta - \lambda_1} \in \left(\frac{4}{\delta}, \infty\right)$$

and $T_{\text{on}} = \frac{1}{\gamma} = \frac{1}{2\delta}$

which suggests the inactive gene with $T_{\text{off}} > T_{\text{on}}$.

**Theorem 2.** If $\gamma = 2\delta$ and $\lambda_1 + \lambda_2 = \delta$, then generating function $V(z, t)$ can be expressed as

$$V(z, t) = e^{-\sigma ez^{-\delta t}} \, _{2}F_{2}(\lambda_1, \lambda_2; \lambda_1 + 1, \lambda_2 + 1; \sigma z)$$

$$- \frac{\sigma \lambda_2 e^{-(\lambda_1 + 1)t}}{(\lambda_1 - \lambda_2)(\lambda_1 + 1)e^{\sigma z - \delta t}} \, _{2}F_{1}(\lambda_1 + 1; \lambda_1 + 2; \sigma ez^{-\delta t})$$

$$+ \frac{\sigma \lambda_1 e^{-(\lambda_2 + 1)t}}{(\lambda_1 - \lambda_2)(\lambda_2 + 1)e^{\sigma z - \delta t}} \, _{2}F_{1}(\lambda_2 + 1; \lambda_2 + 2; \sigma ez^{-\delta t}).$$

and the probability mass function $P_m(t)$ can be expressed as

$$P_m(t) = \frac{\sigma^m}{m!} \left[ \lambda_1 \lambda_2 \int_{0}^{1} \int_{0}^{1} s^{\lambda_1 - 1}(s \sigma - e^{-\delta t})^m e^{\delta(\sigma - s)z} e^{2 \lambda_2 - 1} ds dz \right]$$

$$+ \frac{e^{(1-m)\delta t}}{1 - \lambda_2} \int_{0}^{1} (\lambda_1 e^{-\lambda_2 s} - \lambda_2 e^{-\lambda_1 s})^m + \sigma \delta e^{-\delta t}(1 - s) e^{\sigma z - \delta t} (1 - s) ds.$$  

We remark that if $\lambda_1 = \lambda_2$, the expression (57) still hold if we take the derivative $\lambda_1 \rightarrow \lambda_2$.

**Proof.** We suppose that $\lambda_1 \neq \lambda_2$ since the case $\lambda_1 = \lambda_2$ can be considered by taking the limit $\lambda_1 \rightarrow \lambda_2$. It can be verified that $\delta_1 = \lambda_1 + 1$ and $\delta_2 = \lambda_2 + 1$ (or $\delta_2 = \lambda_1 + 1$ and $\delta_1 = \lambda_2 + 1$) through the correlation

$$\delta_1 + \delta_2 = 2\lambda = \lambda_1 + \lambda_2 + \gamma, \quad \delta_1 \delta_2 = \delta = \lambda_1 \lambda_2 + \lambda_1 \gamma + \lambda_2 \gamma.$$  

Since $\lambda_1$ and $\lambda_2$ play exactly the same regulatory role in $\mathcal{L}^{\lambda_1, \lambda_2}(w)$, we suppose that $\delta = \lambda_1 + 1$ and $\delta_2 = \lambda_2 + 1$ are not general. In this case, $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ given in Theorem 1 can be greatly simplified. This is mainly because that in $\mathcal{A}, \mathcal{B}$ and $\mathcal{C}$, $\,_{2}F_{2}$ equals to 1 if it contains $\lambda_1 - \delta + 1$ or $\lambda_2 - \delta + 1$. Therefore, we have

$$\mathcal{A}(s) = \lambda_1 - \lambda_2, \quad \mathcal{B}(s) = \frac{(\lambda_1 + 1)(\lambda_2 + 1)s}{(\lambda_1 + 2)(\lambda_2 + 2)} \, _{2}F_{2}(\lambda_1, \lambda_2; \lambda_1 + 2, \lambda_2 + 2; \lambda_1 + 3, \lambda_2 + 3; s)$$

$$+ (\lambda_2 + 1) \, _{2}F_{2}(\lambda_1 + 1, \lambda_2 + 1; \lambda_1 + 2, \lambda_2 + 2; s),$$

$$\mathcal{C}(s) = \frac{(\lambda_1 + 1)(\lambda_2 + 1)s}{(\lambda_1 + 2)(\lambda_2 + 2)} \, _{2}F_{2}(\lambda_1 + 1, \lambda_2 + 1; \lambda_1 + 2, \lambda_2 + 2; s)$$

$$+ (\lambda_1 + 1) \, _{2}F_{2}(\lambda_1 + 1, \lambda_2 + 1; \lambda_1 + 2, \lambda_2 + 2; s).$$

Noting that $[43]$

$$\frac{b_1 s}{c_1 c_2} \, _{2}F_{2}(b_1 + 1, b_2 + 1; c_1 + 1, c_2 + 1; s) + \, _{2}F_{2}(b_1, b_2; c_1, c_2; s) = \, _{2}F_{2}(b_1, b_2 + 1; c_1, c_2; s)$$
and
\[
\frac{b_2 s}{c_1 c_2} 2F_2(b_1 + 1, b_2 + 1; c_1 + 1, c_2 + 1; s) + 2F_2(b_1, b_2; c_1, c_2; s) = 2F_2(b_1 + 1, b_2; c_1, c_2; s).
\]

Then, \( B(s) \) in (59) and \( C(s) \) in (60) can be further simplified as
\[
B(s) = (\lambda_2 + 1) 2F_2(\lambda_1 + 1, \lambda_2 + 2; \lambda_1 + 2, \lambda_2 + 2; s) = (\lambda_2 + 1) F_1(\lambda_1 + 1, \lambda_2 + 2; s),
\]
and
\[
C(s) = (\lambda_1 + 1) 2F_2(\lambda_1 + 2, \lambda_2 + 1; \lambda_1 + 2, \lambda_2 + 2; s) = (\lambda_1 + 1) F_1(\lambda_2 + 1, \lambda_2 + 2; s),
\]
where \( F_1 \) is the confluent hypergeometric function defined by
\[
F_1(b_1; c_1; x) = \sum_{k=0}^{\infty} \frac{(b_1)_k}{(c_1)_k k!} x^k.
\]

Directly substituting \( A(s) = \lambda_1 - \lambda_2, \) (61) and (62) into (13), we obtain (56).

We are now in a position to express the Formula (56) in integrals. Noting that Euler’s integrals for \( F_1 \) and \( F_2 \) can be written as follows \[43\]
\[
1F_1(b_1; c_1; x) = \frac{\Gamma(c_1)}{\Gamma(b_1) \Gamma(c_1 - b_1)} \int_0^1 e^{sx} s^{b_1 - 1} (1 - s)^{c_1 - b_1 - 1} ds, \quad \text{for} \quad c_1 > b_1 > 0,
\]
where \( \Gamma(\cdot) \) denotes Gamma function, and
\[
2F_2(b_1, b_2; c_1, c_2; x) = M \int_0^1 \int_0^1 s^{b_1 - 1} (1 - s)^{c_1 - b_1 - 1} e^{sxe^{-\sigma}} \sigma^{b_2 - 1} (1 - \sigma)^{c_2 - b_2 - 1} d\sigma ds,
\]
when \( c_1 > b_1 > 0 \) and \( c_2 > b_2 > 0 \), where
\[
M = \frac{\Gamma(c_1) \Gamma(c_2)}{\Gamma(b_1) \Gamma(c_1 - b_1) \Gamma(c_2 - b_2)}.
\]

Therefore, using (64), (65) and noting that \( \Gamma(s + 1) = s \Gamma(s) \), we can immediately calculate
\[
\begin{align*}
2F_2(\lambda_1, \lambda_2; \lambda_1 + 1, \lambda_2 + 1; \sigma z) &= \lambda_1 \lambda_2 \int_0^1 \int_0^1 s^{\lambda_1 - 1} e^{\sigma z e^{-\sigma t}} \sigma^{\lambda_2 - 1} d\sigma ds, \\
&= \lambda_1 \lambda_2 \int_0^1 \int_0^1 s^{\lambda_1 - 1} e^{\sigma z e^{-\sigma t}} \sigma^{\lambda_2 - 1} d\sigma ds,
\end{align*}
\]
\[
1F_1(\lambda_1 + 1; \lambda_2 + 2; \sigma z e^{-\delta t}) = (\lambda_1 + 1) \int_0^1 e^{\sigma z e^{-\delta t}} s^{\lambda_1} ds,
\]
\[
1F_1(\lambda_2 + 1; \lambda_2 + 2; \sigma z e^{-\delta t}) = (\lambda_2 + 1) \int_0^1 e^{\sigma z e^{-\delta t}} s^{\lambda_2} ds.
\]

Substituting (66)–(68) into (56) leads to the integral expression of \( V(z, t) \)
\[
V(z, t) = \lambda_1 \lambda_2 \int_0^1 \int_0^1 s^{\lambda_1 - 1} e^{\sigma z (s e^{-\delta t} - e^{-\delta t})} \sigma^{\lambda_2 - 1} d\sigma ds
\]
\[
+ \frac{\delta z}{\lambda_1 - \lambda_2} \int_0^1 (\lambda_1 e^{-\lambda_2 t} s^{\lambda_2} - \lambda_2 e^{-\lambda_1 t} s^{\lambda_1}) e^{\sigma z (s e^{-\delta t} - e^{-\delta t})} ds.
\]

We extract \( P_m(t) \) from (69) by formula (1). Applying identity
\[
\partial^m (z e^{cz}) / \partial z^m = (m + cz) e^{m-1} c^z,
\]
we find
\[
\frac{\partial^m V(z,t)}{\partial z^m} = \sigma^m \lambda_1 \lambda_2 \int_0^1 \int_0^1 s^\lambda_1 - 1 (sc - e^{-\delta t})^m e^{\delta z (sc - e^{-\delta t})} e^{\lambda_2 - 1} ds \, dz
\]
\[+ \frac{\sigma^m e(1-m)\delta t}{\lambda_1 - \lambda_2} \int_0^1 (\lambda_1 e^{-\delta t} s^\lambda_2 - \lambda_2 s^\lambda_1) \frac{m + \sigma z e^{-\delta t} (s - 1)}{(s - 1)^{m+1}} e^{\delta z e^{-\delta t}(s-1)} ds.
\]

By substituting \( z = -1 \) and then dividing it by \( m! \), we obtain (57). The proof is completed.

4.2. Two-State Model with \( \lambda_1 \to \infty \)

When \( \lambda_1 \) is sufficiently large, the transition from gene off 1 to off 2 states in the three-state model (Figure 2) can be neglected, and the model collapses into the classical two-state model (2), as illustrated for yeast FLO11 gene [37]. In this subsection, we shall derive the exact form for the generating function \( V(z,t) \) of the two-state model (2). If \( \gamma = \delta \) is further assumed, reverse Formula (1) can be used to derive \( P_m(t) \).

**Theorem 3.** If \( \gamma = \delta \), then the generating function \( V(z,t) \) of the two-state model (2) can be expressed as

\[
V(z,t) = e^{-\bar{\delta}z e^{-\delta t}} F_1(\lambda; \lambda 1; 1 + \lambda 2; \bar{\delta}z) + \frac{\sigma e^{-(\lambda_2 + \gamma)t} z \gamma_1 F_1(1; 2 + \lambda_2; \gamma; 0)}{1 + \lambda_2},
\]

where

\[
P_m(t) = \frac{\lambda_2 \sigma^m}{m!} \left[ \int_0^1 e^{\delta (e^{-\delta t} - s)} (s - e^{-\delta t})^m s^\lambda_2 - 1 \, ds \right.
\]
\[\left. + (-1)^m - 1 e^{-(\lambda_2 + m)\delta t} \int_0^1 (m + \sigma z e^{-\delta t} s) e^{\delta z e^{-\delta t}(s-1)} ds \right].
\]

**Proof.** When \( \lambda_1 \to \infty \), we notice that \( \delta_1 \) and \( \delta_3 \) given by (12) satisfy

\[
\lim_{\lambda_1 \to \infty} \delta_1 = 1, \quad \text{and} \quad \lim_{\lambda_1 \to \infty} \delta_3 = \lim_{\lambda_1 \to \infty} \frac{\bar{\lambda}^2 - \bar{\beta}^2}{\delta_1} = \lim_{\lambda_1 \to \infty} \frac{\lambda_2 + \gamma + \bar{\lambda}_2 \gamma / \lambda_1}{\delta_1 / \lambda_1} = \bar{\lambda}_2 + \bar{\gamma}.
\]

Therefore, the substitution of above limitations into particular solutions \( w_1(x) \), \( w_2(x) \) and \( w_3(x) \) given by (31)–(33) leads to

\[
w_1(x) \to 1 F_1(\lambda; \lambda 1; 1 + \lambda 2; \bar{\gamma}; x), \quad w_2(x) \to 0, \quad \text{and} \quad w_3(x) \to x^{1-\lambda_2 - \bar{\gamma}} 1 F_1(-\bar{\gamma} + 1; 2 - \lambda_2 - \bar{\gamma}; x),
\]

where \( 1 F_1(x; y; z) \) is a confluent hypergeometric function denoted by (63). Then, the initial values (35)–(37) reduce to

\[
A w_1(\sigma_0) + C w_3(\sigma_0) = 1, \quad \text{and} \quad A w'_1(\sigma_0) + C w'_3(\sigma_0) = 0,
\]

and the general solution \( w(x) \) given by (34) reduce to
\[
\begin{align*}
w(x) &= A w_1(x) + C w_3(x) \\
&= A \times 1 F_1(\lambda; \lambda 1; 1 + \lambda 2; \bar{\gamma}; x) + C \times x^{1-\lambda_2 - \bar{\gamma}} 1 F_1(-\bar{\gamma} + 1; 2 - \lambda_2 - \bar{\gamma}; x).
\end{align*}
\]

Here, \( A \) and \( C \) are undetermined coefficients determined by (73).
Using Wronskian-type identity \cite{32,43}

\[
1_F_1(c-d+1;1-d;z)\ 1_F_1(c; d; z) - \frac{cz}{d(1-d)} \ 1_F_1(c-d+1;2-d;z) \ 1_F_1(c+1;d+1;z) = e^z
\]

and the Kummer transformation \cite{32,43}; \( e^{-\gamma t} \ 1_F_1(c; d; z) = 1_F_1(d - c; d; -z) \), we calculate the probability generating function \( w(x) \) as

\[
w(x) = -\frac{\hat{\lambda}_t (\sigma z_0)^\lambda_1 + \gamma_2 x^{-\lambda_2 - \hat{\gamma}}}{(\lambda_2 + \gamma)(1 - \lambda_2 - \gamma)} F_1(\gamma; \hat{\lambda}_2 + \gamma + 1; -\bar{\sigma} z_0) \ F_1(1 - \gamma; \hat{\lambda}_2 - \gamma; x) + 1_F_1(-\lambda_2; 1 - \lambda_2 - \gamma; -\sigma z_0) \ F_1(\lambda_2; \hat{\lambda}_2 + \gamma; x). \tag{74}
\]

We leave the calculation of (74) to interested readers for the sake of simplicity. It remains to calculate \( V(z, t) \). In fact, the substituting \( z_0 = ze^{-dt} \) and \( x = \bar{\sigma} z_0 e^{\delta t} = \bar{\sigma} z \) into (74) yields immediately

\[
V(z, t) = -\hat{\lambda}_t e^{-\delta z e^{-dt}} \int_0^1 \ e^{\bar{\sigma} z e^{-(\lambda_2 + \gamma) t} s} \ F_1(1 - \gamma; \hat{\lambda}_2; x) \ F_1(1 - \hat{\lambda}_2; \gamma + 1; -\bar{\sigma} z_0) \ d s + 1_F_1(-\lambda_2; 1 - \lambda_2 - \gamma; -\sigma z_0) \ F_1(\lambda_2; \hat{\lambda}_2 + \gamma; x) \tag{75}
\]

If \( \gamma = \delta \) (i.e., \( \gamma = 1 \)), the definition of \( 1_F_1 \) indicates that \( 1_F_1(-\lambda_2; 1 - \lambda_2 - \gamma; -\sigma z e^{-dt}) = e^{-\sigma z e^{-dt}} \) and \( 1_F_1(1 - \gamma; \lambda_2 - \lambda_2 - \gamma; \sigma z) = 1 \). Therefore, the expression (75) reduces to (71). Noting the classic theory on the confluent hypergeometric function \cite{16,43}

\[
\Gamma (d) \Gamma (d - c) \int_0^1 e^{s z(1 - s)} \ F_1(1 - s) - \gamma - 1 \ F_1(1 - s) \ F_1(1 - \gamma; \hat{\lambda}_2; x) \ F_1(1 - \hat{\lambda}_2; \gamma + 1; -\bar{\sigma} z_0) \ d s, \tag{76}
\]

where \( \Gamma (\cdot) \) denotes Gamma function. Noticing that \( \Gamma (z + 1) = z \Gamma (z) \), we can rewrite (71) in integrals:

\[
V(z, t) = \hat{\lambda}_t e^{-\delta z e^{-dt}} \int_0^1 e^{\bar{\sigma} z e^{-(\lambda_2 + \gamma) t} s} \ F_1(1 - \gamma; \hat{\lambda}_2; x) \ F_1(1 - \hat{\lambda}_2; \gamma + 1; -\bar{\sigma} z_0) \ d s + 1_F_1(-\lambda_2; 1 - \lambda_2 - \gamma; -\sigma z e^{-dt}) \ F_1(\hat{\lambda}_2; \gamma + 1; -\bar{\sigma} z_0) \ F_1(1 - \gamma; \hat{\lambda}_2; x) + 1_F_1(-\lambda_2; 1 - \lambda_2 - \gamma; -\sigma z_0) \ F_1(\lambda_2; \hat{\lambda}_2 + \gamma; x) \tag{77}
\]

Therefore, we are now in a position express the probability distribution function \( P_m(t) \) in integrals. Applying (1), (77) and (70) we obtain

\[
P_m(t) = \frac{m^n}{m} \left[ \hat{\lambda}_t \int_0^1 e^{\bar{\sigma} z e^{-(\lambda_2 + \gamma) t} s} \ F_1(1 - \gamma; \hat{\lambda}_2; x) \ F_1(1 - \hat{\lambda}_2; \gamma + 1; -\bar{\sigma} z_0) \ d s + 1_F_1(-\lambda_2; 1 - \lambda_2 - \gamma; -\sigma z e^{-dt}) \ F_1(\hat{\lambda}_2; \gamma + 1; -\bar{\sigma} z_0) \ F_1(1 - \gamma; \hat{\lambda}_2; x) + 1_F_1(-\lambda_2; 1 - \lambda_2 - \gamma; -\sigma z_0) \ F_1(\lambda_2; \hat{\lambda}_2 + \gamma; x) \right]_{z = -1},
\]

which is exactly (72), and the proof is completed. \( \Box \)

5. Discussions and Conclusions

In this study, we developed an approach for calculating \( V(z, t) \) and \( P_m(t) \) for large parameter regions. We illustrated the approach for the three-state model (Figure 2) and successfully obtained the dynamic expressions of \( V(z, t) \) in a closed-form of hypergeometric functions (Theorem 1). In addition to the numerical simulation for computing \( P_m(t) \) using (1), we can still simplify the expression of \( V(z, t) \) within several special parameter regions to derive the integral forms of \( P_m(t) \) by elementary functions. The simplicity of those forms may facilitate us to prove in mathematical detail the dynamic profile transitions.
of mRNA distribution [24]. We discussed two special parameter regions in Theorems 2 and 3. The first parameter region presents inactive genes with a longer duration of gene off than that of gene on, which has been confirmed for thousands of eukaryotic genes [22,36] and prokaryotic genes under low induction [17,45]. The second parameter region presents the classical two-state model with arbitrary durations of the gene off state. This case has been discussed in other studies that exhibited rich dynamic $P_m(t)$ profiles [24,32]. Moreover, we noticed that $V(z,t)$ readily gives $P_0(t) = V(-1,t)$, which quantifies the probability of zero mRNA being produced. The index $P_0(t)$ quantifies the typical “none-or-all” phenomenon and signifies the cell fate decision in an isogenic cell population [17,18,45]. Our work may facilitate the computation of $P_0(t)$ to describe HIV latency [18], transcription dynamics of the $E. coli$ promoter $P_{lac/ara}$ [17], and lysogen stability from the fate-determining $cl$ gene [45].

We computed $V(z,t)$ and $P_m(t)$ using the obtained analytical expressions. In Figure 3a, we generated three dynamical curves of $P_0(t) = V(-1,t)$ using $V(z,t)$ expressions (13) and (75) under three parameter groups, including the average off duration $T_{off}$ (Equation (55)), which takes $T_{off} = 25$ min with $\lambda_2 = 2\lambda_1$ for $P_{lac/ara}$ promoter in $E. coli$ cells under acidic shift [40]; $T_{off} = 70$ min with $\lambda_2 = 7\lambda_1$ for mouse fibroblast genes under TSA treatments [36]; and $T_{off} = 50$ min with $\lambda_1 \gg \lambda_2$ (two-state model) for yeast $FLO11$ gene under the addition of ethanol [37]. These dynamic curves exactly match the corresponding curves generated by the stochastic simulation algorithm (SSA), suggesting the correctness of our $V(z,t)$ expressions. In Figure 3b,c, we further generated distribution profiles of $P_m(t)$ using Formula (1) by differentiating expressions of $V(z,t)$ in Theorems 2 and 3, and confirmed their correctness using SSA. We emphasize here that the expression of $V(z,t)$ for the two-state model was obtained in [32] by another method that shows a different sign “−” compared with our expression (75). As an independent check using SSA (Figure 3a) verifies the correctness of our expression (75), the expression of $V(z,t)$ obtained in [32] needs to be corrected.

![Figure 3](image)

**Figure 3.** Dynamic profiles of distribution $P_m(t)$ generated by the derived exact formulas. (a) The curves of $P_0(t) = V(-1,t)$ against time $t$ (green lines) are generated using Formulas (13) and (75) under three parameter groups—$\delta = 0.014, \gamma = 0.167$ (min$^{-1}$) and $T_{off} = 25$ min with $\lambda_2 = 2\lambda_1$ for $E. coli$ $P_{lac/ara}$ promoter [17,40]; $\delta = 0.0109, \gamma = 0.167$ (min$^{-1}$) and $T_{off} = 70$ min with $\lambda_2 = 7\lambda_1$ for mouse fibroblast genes [22,36]; $\delta = 0.063, \gamma = 0.05$ (min$^{-1}$) and $T_{off} = 50$ min with $\lambda_1 \gg \lambda_2$ for yeast $FLO11$ gene [11,37]. (b) Dynamic decay distribution $P_m(t)$ (solid circles in lines) of the three-state model generated by the formulas in Theorem 2, with $\delta = 0.01, \lambda_1 = \lambda_2 = \delta/2, \gamma = \delta$ (min$^{-1}$). (c) Dynamic transition among the three modes of distribution $P_m(t)$ (solid circles in lines) for the two-state model generated by the formulas in Theorem 3, with $\delta = 0.01, \lambda_2 = 2\delta, \gamma = \delta$ (min$^{-1}$). The black dashed lines in (a) and hollow circles in (b,c) represent numerical results obtained by the stochastic simulation algorithm. The other parameter is $v = 20\delta$.

We further compared $P_m(t)$ dynamic profiles separately generated by different models. We first found that the steady-state mRNA distribution data of mouse $Mbln2$ gene [4] and $E. coli$ $P_{lac/ara}$ promoter [17] (Figure 1b,c) can be well fitted by both the two-state model and three-state model under the same rescaled kinetic rates $\gamma/\delta, v/\delta$ and $T_{off}\delta$ with $\lambda_1 = n\lambda_2$ for different $n \geq 1$ (Figure 4). When $n$ increases, the three-state model gradually reduces to the two-state model ($n \to \infty$). We then generated $P_m(t)$ profiles at different time points.
using both models equipped with the estimated kinetic rates and $\delta = 0.0109 \text{ min}^{-1}$ for $Mbln2$ gene [22] and $\delta = 0.014 \text{ min}^{-1}$ for $P_{lac/ara}$ promoter [17]. As shown in Figure 4a, there exhibits an ordered distribution transition pattern among decaying, bimodal, and unimodal as time develops. However, the two-state model generates the longest duration of 120 min for the intermediate bimodality compared with the decreasing bimodal duration with respect to $n$ for the three-state model. This observation suggests that the increase in gene activation step number may facilitate the suppression of heterogeneous bimodal distribution while enhancing the formation of homogeneous unimodal distribution. Such bimodal suppression was also observed in the transcription of promoter $P_{lac/ara}$ for which the dynamic mRNA distribution transition from decaying to bimodality occurs as early as 40 min for the two-state model and as late as 100 min for the three-state model with $\lambda_1 = \lambda_2$ (Figure 4b).

The approach presented here may be further implemented in other models with constant kinetic rates of transcription, especially for the multiscale model [31] and cross-talking pathway model [11], which also includes three functional gene states. Both models are different from the framework shown in Figure 2 as the multiscale model describes a change in the gene $on$ state upon production of mRNA to explicitly describe RNA polymerase dynamics; the cross-talking pathways model assumes two parallel rate-limiting gene activation steps to simulate competitive binding of transcription factors at the cognate DNA sites. By introducing the generating functions (7) and variables (14) and (19), the master equations [11,31] for the multiscale model or cross-talking pathway model can be transformed into canonical forms of third-order hypergeometric differential equations, as in (24) and (26). Then, using the method of undetermined coefficients (34), the analytical expressions for generating functions $V(z, t)$ for both models may be derived. Furthermore, the protocol of our approach may be developed for models with more than three gene states, where gene activation is regulated by multiple parallel [46,47] or sequential [2,48] rate-limiting steps, or their combinations. However, the complexity of the calculation for $P_m(t)$ is expected to increase dramatically with respect to the gene state number.

The models mentioned above may provide a universal framework to illustrate stochastic gene activation for both prokaryotic and eukaryotic genes. However, the assumption of an exponential decay of mRNA molecules in these models may fit the straight log-scale lines of degrading mRNA levels against time for $E. coli$ genes [33], but may fail to capture much more complicated mRNA degradation processes for eukaryotic genes [49]. A detailed model of eukaryotic mRNA degradation models in which the mRNA molecule decays either through the pathway of $5' \to 3'$ direction or the pathway of $3' \to 5'$ direction [50]: the $5' \to 3'$ pathway consists of multiple rate-limiting steps representing the ordered PolyA-deadenylation, decapping, and the involvement of many intracellular factors; the $3' \to 5'$ pathway is directed by a single rate-limiting step to degrade mRNA from the unprotected $3'$ ends by exosomes. Future work may include the calculation of $P_m(t)$ for models of stochastic mRNA degradation and illustrate how transcript degradation influences transcription dynamics.
Figure 4. Discriminated dynamic bimodal features of mRNA distribution generated by the two-state model and three-state model. (a) The steady-state mRNA distribution data of mouse Mbln2 gene [4] (blue bars) can be well fitted by both models with $\left( T_{\text{off}}, \gamma/\delta, v/\delta \right) = (0.71, 4.74, 38.8)$, and $\lambda_1 = n\lambda_2$ with finite $n$ for the three-state model and $n \to \infty$ for the two-state model. Dynamic distribution profiles with $\delta = 0.0109 \text{ min}^{-1}$ [22] display an intermediate dynamic bimodal distribution with the longest duration of 120 min for the two-state model, but decreasing durations with respect to $n$ for the three-state model. (b) The steady-state mRNA distribution data of E. coli P$_{\text{lac/ara}}$ promoter [17] (blue bars) can be well fitted by both models with $\left( T_{\text{off}}, \gamma/\delta, v/\delta \right) = (2.53, 0.22, 13)$ and $\lambda_1 = n\lambda_2$. Dynamic distribution profiles with $\delta = 0.014 \text{ min}^{-1}$ [17] display a dynamic distribution transition from decaying to bimodality, with the transition time as early as 40 min for the two-state model and as late as 100 min for the three-state model with $\lambda_1 = \lambda_2$.

Author Contributions: J.C. performed research; F.J. designed and wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Natural Science Foundation of China grants (Nos. 11871174, 11631005) and by the Program for Changjiang Scholars and Innovative Research Team in the University (No. IRT_16R16).

Data Availability Statement: The data that support the findings of this study are openly available at refs. [4,10,11,17,22,36,37,40].

Conflicts of Interest: The authors declare no conflict of interest.

References
7. Munsky, B.; Fox, Z.; Neuert, G. Integrating single-molecule experiments and discrete stochastic models to understand heterogeneous gene transcription dynamics. Methods 2015, 85, 12–21. [CrossRef]


