Abstract—Stochasticity in the gene-expression process can create variability in the level of a given mRNA/protein across a homogenous population of living cells. Random fluctuations between different promoter states have been implicated as a major source of noise in the expression of many genes. These fluctuations are typically modeled through a two-state promoter architecture, where the promoter of a gene transitions between an active (ON) and inactive (OFF) state, spending an exponentially distributed time-interval in each state. High levels of mRNA production occur from the active state, while the inactive state allows for a low basal rate of production. Recent data has shown the existence of three-state promoter architectures with memory, where the time spent in the inactive state is gamma distributed. Here we analyze stochastic models of both two-state and three-state promoter architectures and identify key differences in their stochastic dynamics.

Quantifying distance between probability distributions using standard metrics reveals that the difference in the mRNA copy number distributions for both promter architectures is maximum when the stability of the active promoter state is comparable to the stability of the mRNA transcript. Our results further show that mRNA auto-correlations decay more rapidly for a three-state promoter architecture compared to a two-state promoter architecture. Interestingly, we find that in certain parameter regimes the three-state promoter architecture can yield negative mRNA auto-correlations. Finally, we discuss how these results can be useful for identifying genetic promoter architectures from single-cell mRNA data.

I. INTRODUCTION

Gene-expression is the process by which information contained within genes is used to synthesize mRNAs and proteins through transcription and translation, respectively. Experimental work has show that the thermal molecular motion and low-copy numbers of biochemical species inside living cells make gene-expression an inherently stochastic process [1]–[6]. An important consequence of stochastic gene-expression is that the level of a given mRNA or protein can show variations from cell-to-cell, in spite of the fact that gene-expression is that the level of a given mRNA or protein is comparable to the stability of the mRNA transcript. Our work focuses on stochastic models of both two-state and three-state promoter architectures and identify key differences in their stochastic dynamics.

In this case, the mRNA distribution is predicted to be a Poisson distribution [7]. However, experimentally obtained mRNA histograms show considerable deviations from a Poisson distribution indicating that more complex models are needed to explain the data [8]–[11]. A commonly used model for the transcription process is a two-state promoter architecture, where the promoter fluctuates between an active and inactive state with rates \( k_{on} \) and \( k_{off} \) (see Fig. 1: [12]–[14]). Here mRNA transcripts are synthesized from the active (inactive) state at rate \( k_{b} \). Since transcription from the inactive state is inefficient \( k_{b} \ll k_{m} \). Synthesized mRNAs degrade at a rate \( \gamma_{m} \). An implicit assumption in the stochastic formulation of this model is that promoter transitions are memoryless: the promoter resides in the ON and OFF state for an exponentially distributed time-interval with mean \( 1/k_{off} \) and \( 1/k_{on} \), respectively. For this case, the steady-state probability of observing \( m \) mRNA copy numbers is

\[
p_{m}(m) = e^{-k_{b}} \sum_{n=0}^{m} \frac{(k_{m} - k_{b})^{m-n}}{(m-n)!} \frac{(k_{on} + n)!}{(k_{on} + k_{off} + n)!} F_{1}\left(\alpha = k_{on} + n; \beta = k_{on} + k_{off} + n; \frac{k_{b}}{k_{b} - k_{m}}\right)
\]

where \( \Gamma(.) \) is a gamma function, \( F_{1}(\alpha, \beta, \gamma) \) is the confluent hypergeometric function of the first kind and

\[
\alpha = k_{on}/\gamma_{m}, \beta = k_{on} + k_{off}/\gamma_{m} = k_{on}/\gamma_{m}, k_{b} = k_{b}/\gamma_{m}
\]

are model parameters normalized by the mRNA degradation rate \( \gamma_{m} \) [8], [13], [15], [16]. Although the above model has been widely used to capture stochasticity in the gene-expression process, recent experimental work has argued that the time spent in the OFF state is better approximated by a gamma distribution, in particular, a sum of two independent, identical exponential random variables [17]. Towards that end we model the transcription process as a three-state promoter architecture, where the transition to an ON state occurs through a two-step mechanism (Figure 1). For this architecture, an exact analytical expression for the steady-state mRNA distribution is derived. To determine when the stochasticity variability predicted by a two-state and a three-state architecture is most distinct, we compute the Kolmogorov distance between their mRNA probability distributions. Our analysis reveals that this distance is largest when the stability of the active promoter state is comparable to the stability of the mRNA transcript.

Since probability distributions do not provide any information on the time-scale of fluctuations, we compute the mRNA auto-correlation function for a three-state promoter architecture. Previous work had shown that mRNA auto-correlations in a two-state architecture monotonically de-
increase to zero \cite{15,16,18}. Interestingly, we find that the three-state architecture can yield non-monotonic auto-correlation functions in certain parameter regimes. Thus, by appropriately designing these systems it may be possible to use signatures in the auto-correlation function to discriminate between alternative genetic promoter architectures.

The paper is organized as follows. In Section II, we introduce the three-state promoter model and derive an analytical expression for the mRNA probability distribution. Difference between distributions corresponding to the two-state and three-state promoter architectures are computed in Section III. Auto-correlation function of mRNA copy numbers is derived in Section IV. Finally, conclusions are presented in Section V.

## II. Stochastic Analysis of the Three-State Promoter Architecture

One way to incorporate memory in the promoter transitions is by introducing intermediate states. Consider the three-state promoter architecture illustrated in Figure 1. Here the promoter can reside in three possible states: \( P_0 \), \( P_1 \) and \( P_2 \). State \( P_2 \) is a transcriptionally active (ON) state where mRNAs are made at a high transcription rate \( k_m \). From \( P_2 \), the promoter can transition to a transcriptionally inactive (OFF) state \( P_0 \). For the promoter to now transition back to an ON state, it has pass through a transcriptionally inactive intermediate state \( P_1 \). Assuming the rate of transitions from \( P_0 \) to \( P_1 \) and \( P_1 \) to \( P_2 \) are given by \( k_1 \) and \( k_2 \), respectively, the promoter stays OFF for

\[
E(k_1) + E(k_2),
\]

where \( E(k_1) \) denotes an exponentially distributed random variable with mean \( 1/k_1 \). If \( k_1 = k_2 = 2k_m \), this would correspond to a gamma distributed dwell time in the inactive state with mean \( 1/k_m \) and variance \( 1/2k_m^2 \). In principal, further intermediate states can be incorporated to model a gamma distributed dwell time with any arbitrary variance. However, to be consistent with experimental work that proposed a two-step mechanism for promoter switching ON \cite{17}, we only consider a single intermediate state. As in the two-state promoter model, we assume then when the promoter is OFF (i.e., in states \( P_0 \) and \( P_1 \)), mRNAs are synthesized at a low basal rate \( k_b \ll k_m \). In the remainder of this section we provide a stochastic description of the three-state promoter model and derive an exact analytical solution for the steady-state probability distribution for the mRNA copy number.

We represent the state of the promoter at time \( t \) by \( g_i(t) \) with \( g_i(t) = 1 \) \((g_i(t) = 0)\) denoting that the promoter is (is not) in the state \( P_i, i \in \{0,1,2\} \). Since there is only one copy

<table>
<thead>
<tr>
<th>Event description</th>
<th>Reset in population count</th>
<th>Propensity function ( f(g_0,g_1,g_2,m) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Promoter transition from ( P_0 ) to ( P_1 ) state</td>
<td>( g_0(t) \rightarrow g_0(t) + 0, g_1(t) \rightarrow g_1(t) + 1 )</td>
<td>( k_1 g_0(t) )</td>
</tr>
<tr>
<td>Promoter transition from ( P_1 ) to ( P_2 ) state</td>
<td>( g_1(t) \rightarrow g_1(t) + 0, g_2(t) \rightarrow g_2(t) + 1 )</td>
<td>( k_2 g_1(t) )</td>
</tr>
<tr>
<td>Promoter transition from ( P_2 ) to ( P_0 ) state</td>
<td>( g_2(t) \rightarrow g_2(t) + 0, g_0(t) \rightarrow g_0(t) + 1 )</td>
<td>( k_{off} g_2(t) )</td>
</tr>
<tr>
<td>mRNA transcription from ON state</td>
<td>( m(t) \rightarrow m(t) + 1 )</td>
<td>( k_m g_2(t) )</td>
</tr>
<tr>
<td>mRNA transcription from OFF state</td>
<td>( m(t) \rightarrow m(t) + 1 )</td>
<td>( k_b (g_0(t) + g_1(t)) )</td>
</tr>
<tr>
<td>mRNA degradation</td>
<td>( m(t) \rightarrow m(t) - 1 )</td>
<td>( \gamma m(m(t)) )</td>
</tr>
</tbody>
</table>
Let $m(t)$ denote the mRNA copy number inside the cell at time $t$. Then, the two-state promoter model can be described by a set of discrete events listed in Table I. As in the stochastic formulation of chemical kinetics [19]–[21], events “fire” at exponentially distributed time intervals, and whenever a specific event occurs the state of the system is updated accordingly to a reset map (see second column in Table I). The third column lists the propensity functions $f(g_0, g_1, g_2, m)$, which determine how often an event occurs. More specifically, the probability that an event will occur in the next infinitesimal time interval $(t, t + dt)$ is given by $f(g_0, g_1, g_2, m)dt$. Probabilistic events described in Table I define a stochastic model where dwell time in the ON state is exponentially distributed but time spent in the OFF state is given by (3). Moreover, when the promoter is ON(OFF) then mRNA production is a Poisson process with rate $k_m(k_b)$. Finally, each mRNA lives for an exponentially distributed time interval with mean $1/\gamma_m$, where $\gamma_m$ is the mRNA degradation rate.

Let $p_i(m, t)$ be the probability that at time $t$ the promoter is in state $P_i$, $i \in \{0, 1, 2\}$ with $m$ number of mRNA molecules in the system. The time evolution of these probabilities is given by the following Chemical Master Equation (CME)

$$
\frac{\partial p_0(m, t)}{\partial t} = k_{off} p_2(m, t) - k_1 p_0(m, t) + k_b [p_0(m-1, t) - p_0(m, t)] + \gamma_m[(m+1)p_0(m+1, t) - m p_0(m, t)]
$$

$$
\frac{\partial p_1(m, t)}{\partial t} = k_{off} p_2(m, t) - k_1 p_1(m, t) + k_b [p_0(m-1, t) - p_1(m, t)] + \gamma_m[(m+1)p_1(m+1, t) - m p_1(m, t)]
$$

$$
\frac{\partial p_2(m, t)}{\partial t} = k_2 p_1(m, t) - k_{off} p_2(m, t) + k_m [p_2(m-1, t) - p_2(m, t)] + \gamma_m[(m+1)p_2(m+1, t) - m p_2(m, t)]
$$

[20]–[22]. The generating functions are defined as

$$
F_i(z,t) := \sum_{m=0}^{\infty} z^m p_i(m, t), \quad F(z,t) := \sum_{m=0}^{\infty} z^m p(m, t)
$$

$$
F(z,t) := \sum_{i=0}^{2} F_i(z,t), \quad p(m, t) := \sum_{i=0}^{2} p_i(m, t).
$$

Here $F_i(z,t)$ and $p(m, t)$ denote the generating function and probability distribution function of $m(t)$, respectively. In terms of the generating functions (6), the CME (5) can be written as

$$
\frac{\partial F_0(z,t)}{\partial t} = k_{off} F_2(z,t) - k_1 F_0(z,t) + k_b (z-1) F_0(z,t) + \gamma_m(1-z) \frac{dF_0(z,t)}{dz}
$$

$$
\frac{\partial F_1(z,t)}{\partial t} = k_2 F_1(z,t) - k_{off} F_2(z,t) + k_m (z-1) F_1(z,t) + \gamma_m(1-z) \frac{dF_1(z,t)}{dz}
$$

$$
\frac{\partial F_2(z,t)}{\partial t} = k_2 F_1(z,t) - k_{off} F_2(z,t) + k_m (z-1) F_2(z,t) + \gamma_m(1-z) \frac{dF_2(z,t)}{dz}.
$$

At the steady state we obtain

$$
0 = k_{off} \bar{F}_2(z) - k_1 \bar{F}_0(z) + k_b (z-1) \bar{F}_0(z) + \gamma_m(1-z) \frac{d\bar{F}_0(z)}{dz}
$$

$$
0 = k_2 \bar{F}_1(z) - k_{off} \bar{F}_2(z) + k_m (z-1) \bar{F}_1(z) + \gamma_m(1-z) \frac{d\bar{F}_1(z)}{dz}
$$

$$
0 = k_2 \bar{F}_1(z) - k_{off} \bar{F}_2(z) + k_m (z-1) \bar{F}_2(z) + \gamma_m(1-z) \frac{d\bar{F}_2(z)}{dz}
$$

where

$$
\bar{F}_i(z) := \lim_{t \to \infty} F_i(z,t), \quad \bar{F}(z) := \lim_{t \to \infty} F(z,t).
$$

Defining

$$
\bar{G}_i(z) := \bar{F}_i(z) \exp\left(-\frac{k_b(z-1)}{\gamma_m}\right), \quad i \in \{0, 1, 2\}
$$

$$
\bar{G}(z) := \sum_{i=0}^{2} \bar{G}_i(z) = \bar{F}(z) \exp\left(-\frac{k_b(z-1)}{\gamma_m}\right)
$$

one can write equation (8) as

$$
0 = k_{off} \bar{G}_2(z) - k_1 \bar{G}_0(z) + \gamma_m(1-z) \frac{d\bar{G}_0(z)}{dz}
$$

$$
0 = k_2 \bar{G}_1(z) - k_{off} \bar{G}_2(z) + \gamma_m(1-z) \frac{d\bar{G}_1(z)}{dz}
$$

$$
0 = k_2 \bar{G}_1(z) - k_{off} \bar{G}_2(z) + (k_m - k_b)(z-1) \bar{G}_2(z) + \gamma_m(1-z) \frac{d\bar{G}_2(z)}{dz}.
$$

Addition of equations (11a)-(11c) yields

$$
(k_m - k_b) \bar{G}_2(z) = \gamma_m \frac{d\bar{G}(z)}{dz}.
$$

Now we have from equations (11c) and (12)

$$
\bar{G}_1(z) = \frac{\gamma_m}{(k_m - k_b)k_2} [(k_{off} - (k_m - k_b)(z-1)] \frac{d\bar{G}(z)}{dz}
$$

$$
+ \gamma_m(z-1) \frac{d^2\bar{G}(z)}{dz^2}.
$$

Solving equations (10b), (12) and (13) we get

$$
\bar{G}_0(z) = \frac{\gamma_m}{(k_m - k_b)k_2} [(k_{off} - (k_m - k_b)(z-1)] \frac{d\bar{G}(z)}{dz}
$$

$$
+ (\gamma_m(z-1) + k_2) \frac{d^2\bar{G}(z)}{dz^2}.
$$

Then, with the help of equations (12), (13) and (14), either
\[ p_{\text{three}}(m) = e^{-\hat{k}_b} \sum_{n=0}^{m} \frac{\hat{k}_b^m}{n!} \left( \hat{k}_m - \hat{k}_b \right)^n \frac{\Gamma(\hat{k}_1 + n) \Gamma(\hat{k}_2 + n) \Gamma(h_1) \Gamma(h_2)}{\Gamma(\hat{k}_1 + \hat{k}_2 + n + m) \Gamma(h_1 + n) \Gamma(h_2 + n)} \cdot p_{F_q}(\hat{k}_2 + n, \hat{k}_1 + n; h_1 + n, h_2 + n; -(\hat{k}_m - \hat{k}_b)). \] (18)

\[ p_{\text{two}}(m) = \frac{\hat{k}_m \Gamma(\hat{k}_1 + m) \Gamma(\hat{k}_2 + m) \Gamma(\hat{k}_1) \Gamma(\hat{k}_2) \Gamma(h_1) \Gamma(h_2)}{m! \Gamma(\hat{k}_1) \Gamma(\hat{k}_2) \Gamma(h_1 + m) \Gamma(h_2 + m)} \cdot p_{F_q}(\hat{k}_2 + m, \hat{k}_1 + m; h_1 + m, h_2 + m; -(\hat{k}_m - \hat{k}_b)). \] (19)

\[ p_{\text{three}}(m) = \frac{\hat{k}_m \Gamma(\hat{k}_1 + m) \Gamma(\hat{k}_2 + m) \Gamma(\hat{k}_1) \Gamma(\hat{k}_2) \Gamma(h_1) \Gamma(h_2)}{m! \Gamma(\hat{k}_1) \Gamma(\hat{k}_2) \Gamma(h_1 + m) \Gamma(h_2 + m)} \cdot p_{F_q}(\hat{k}_2 + m, \hat{k}_1 + m; h_1 + m, h_2 + m; -(\hat{k}_m - \hat{k}_b)). \] (20)

The equation (11a) or (11b) or (11c) can be expressed as

\[ \frac{(z-1)^2 dG(z)}{dz^2} + \left( (1+k_s)(z-1) - k_d(z-1)^2 \right) \frac{dG(z)}{dz} + \left( k_p - k_d(1+\hat{k}_1 + \hat{k}_2)(z-1) \right) G(z) = 0 \] (15)

where \( k_d = \hat{k}_m - \hat{k}_b \) and

\[ k_s = \hat{k}_1 + \hat{k}_2 + \hat{k}_b \text{off}, \quad k_p = \hat{k}_1 k_2 + \hat{k}_b \text{off} + \hat{k}_2 \text{off}. \] (16)

The solution of the above equation is a generalized hypergeometric function and is given by (using Mathematica)

\[ G(z) = \text{C}_p F_q[\hat{k}_2, \hat{k}_1; h_1, h_2; (\hat{k}_m - \hat{k}_b)(z-1)], \] (17)

where

\[ h_1 = \frac{1}{2} \left( k_s - \sqrt{k_s^2 - 4k_p} \right) \] (18a)

\[ h_2 = \frac{1}{2} \left( k_s + \sqrt{k_s^2 - 4k_p} \right) \] (18b)

\( C \) is the normalization constant and \( p_{F_q}(a,b,c,d,z) \) is the generalized hypergeometric function (GHF). Thus we have from (10b)

\[ F(z) = \exp\left( \frac{k_s(z-1)}{\gamma_m} \right) G(z) \]

\[ = C \exp\left( \frac{k_b(z-1)}{\gamma_m} \right) p_{F_q}[\hat{k}_2, \hat{k}_1; h_1, h_2; (\hat{k}_m - \hat{k}_b)(z-1)]. \] (19)

The normalization constant can be determined from the condition \( F(1) = 1 \) and is given by \( C = 1 \). From equation (19) one can obtain the expression for the mRNA probability distribution function as (18).

### III. QUANTIFYING DISTANCE BETWEEN PROBABILITY DISTRIBUTIONS

The steady-state mRNA level distribution for a two-state and a three-state promoter architecture are given by \( p_{\text{two}}(m) \) (Eq. (1)) and \( p_{\text{three}}(m) \) (Eq. (18)), respectively. In this section we define a distance between these distributions and explore how this distance varies with model parameters. To simply analytical expressions we assume \( k_b = 0 \), i.e., when the promoter resides in the inactive state mRNA production in completely shut OFF. When \( k_b = 0 \), distributions (1) and (18) reduce to (19) and (20), respectively.

For a fair comparison between both distributions it is essential that they both have the mean. Towards that end we assume \( k_1 = k_2 = 2k_{\text{ON}} \). This ensures that the average time promoter spends in the OFF state is same for both architectures and given by \( 1/k_{\text{ON}} \). As other quantities (mean promoter dwell time in the ON state, transcription rate from the ON state, etc.) are also identical between them, both architectures yield the same steady-state average mRNA copy number

\[ \langle m \rangle = f_{\text{on}} \frac{k_m}{\gamma_m} \] (21)

where \( \langle \cdot \rangle \) denotes the expected value at steady-state,

\[ f_{\text{on}} = \frac{k_m}{k_{\text{off}} + k_m} \] (22)

is the fraction of time the promoter is in the ON state, and \( k_m \) is the mRNA production rate from the ON state.

#### A. Kolmogorov metric

We first use the Kolmogorov metric (also referred to as the Uniform metric) to define the following distance

\[ d(p_{\text{two}}, p_{\text{three}}) = \sup \left| \sum_{i=0}^{m} p_{\text{two}}(i) - \sum_{i=0}^{m} p_{\text{three}}(i) \right|, \] (23)

which can be interpreted as the maximum absolute-value difference between the cumulative distributions of \( p_{\text{two}}(m) \) and \( p_{\text{three}}(m) \) [23]. The top row in Figure 2 plots \( d(p_{\text{two}}, p_{\text{three}}) \) as a function of \( f_{\text{on}} \) and \( k_{\text{off}} \) for \( k_m = 5 \) and 80. Our results show that for \( k_{\text{off}} \gg 1 \) or \( f_{\text{on}} \) close to zero or one, the difference between the distributions is minimal. Moreover, \( d(p_{\text{two}}, p_{\text{three}}) \) is maximized at intermediate values of \( f_{\text{on}} \) and \( k_{\text{off}} \approx 1 \).

#### B. Difference in variance

We next consider an analytically tractable measure of difference between distributions. Recall that both distributions \( p_{\text{two}}(m) \) and \( p_{\text{three}}(m) \) have the same mean. Thus, one can define the distance as the difference in the variances of the two distributions. Note that this distance cannot be used to define a metric since the distance being zero does not imply
Figure 2. Distance between mRNA probability distribution functions corresponding to a two-state and three-state promoter architecture as a function of $f_{on} = k_{on}/(k_{on} + k_{off})$ and stability of the ON promoter state $k_{off} = k_{off}/\gamma_m$. The figures on the top (bottom) correspond to distance computed using Kolmogorov metric (difference in variance). Distances are computed for a normalized transcription rate of $\hat{k}_m = 5$ (left) and $\hat{k}_m = 80$ (right).

\[
Var_{two}(m) - Var_{three}(m) = \frac{(1 - f_{on})^2 f_{on}^2 k_{off}^2 k_m^2}{(1 - f_{on} + \hat{k}_{off})(1 - f_{on})^2 + (1 + 2f_{on} - 3f_{on}^2)\hat{k}_{off} + 4\hat{k}_{off}^2 f_{on}}
\]

that the distributions are identical. From (19)-(20), variance in mRNA level is given by

\[
Var_{two}(m) = \langle m \rangle \left(1 + \frac{\hat{k}_{off} \langle m \rangle}{k_{on}(1 + k_{on} + \hat{k}_{off})} \right),
\]

\[
Var_{three}(m) := \langle m \rangle \left(1 + \frac{\hat{k}_{off}(1 + 3k_{on}) \langle m \rangle}{k_{on}(1 + 4k_{on} + \hat{k}_{off} + 4k_{on}(\hat{k}_{on} + \hat{k}_{off}))} \right)
\]

for the two-state and three-state promoter architecture, respectively. It turns out that $Var_{three}(m) < Var_{two}(m)$, i.e., a three-state architecture predicts lower stochastic variability around the mean compared to a two-state promoter architecture. Taking the difference between the variances and using the fact that (see (22))

\[
\hat{k}_{on} = \frac{\hat{k}_{off}}{\frac{1}{f_{on}} - 1}
\]

we obtain the expression shown in (25). Analysis in Mathematic reveals that this distance is always maximized at $\hat{k}_{off} = 2/3$ and $f_{on} = 5/9$. Note that the point at which maxima is achieved is independent of $\hat{k}_m$. This point is illustrated in the bottom row of Figure 2 which plots (25) as a function of $f_{on}$ and $\hat{k}_{off}$ for different values of $\hat{k}_m$.

IV. Computation of mRNA Auto-correlation Time

We next focus on the time-scale of mRNA copy number fluctuations and derive an analytical formula for the steady-state auto-correlation function of $m(t)$. As in the previous section we assume that $k_0 = 0$ and $k_1 = k_2 = 2k_{on}$. To determine the auto-correlation function we use the fact that

\[
\langle m(t+s) \rangle = \langle m(s) \rangle \langle m(t) \mid m(s), g_i(s) \rangle
\]

where $\langle \rangle$ denotes the expected value and

\[
\langle m(t+s) \mid m(s), g_i(s) \rangle
\]

is the expected number of mRNA transcripts at time $t+s$ given $m(s)$ and $g_i(s)$, $\forall i \in \{0,1,2\}$. We recall that $g_i(s) = 1$ denotes that the promoter is in state $P_i$ at time $s$. Since $g_0(s) = 1 - g_1(s) - g_2(s)$ we will only consider $g_1, g_2$ in the sequel.
To compute the conditional expectation (26) we derive the linear system of differential equations that describe the time evolution of the first and second order statistical moments of the stochastic processes \( g_1(t), g_2(t) \) and \( m(t) \). For the model described in Table I, the time derivative of the expected value of any differential function \( \varphi(g_1, g_2, m) \) is given by

\[
\frac{d\left\langle \varphi (g_1, g_2, m) \right\rangle}{dt} = \left\langle \sum_{\text{Events}} \Delta \varphi (g_1, g_2, m) f(g_1, g_2, m) \right\rangle ,
\]

(27)

where \( \Delta \varphi (g_1, g_2, m) \) is the change in \( \varphi(g_1, g_2, m) \) when an event occurs and \( f(g_1, g_2, m) \) is the event propensity function [24], [25]. For appropriate choices of \( \varphi(g_1, g_2, m) \) we obtain the following moment dynamics:

\[
\begin{align*}
\frac{d\langle g_1 \rangle}{dt} &= 2k_{on} - 2k_{on}\langle g_2 \rangle - 4k_{on}\langle g_1 \rangle \\
\frac{d\langle g_2 \rangle}{dt} &= 2k_{on}\langle g_1 \rangle - k_{off}\langle g_2 \rangle \\
\frac{d\langle m \rangle}{dt} &= m_g(m) - \gamma_m\langle m \rangle \\
\frac{d\langle m^2 \rangle}{dt} &= m_g(m) + \gamma_m(m) + 2k_{on}\langle g_2 m \rangle - 2\gamma_m\langle m^2 \rangle \\
\frac{d\langle g_1 m \rangle}{dt} &= 2k_{on}(m) - 2k_{on}\langle g_2 m \rangle + k_m\langle g_1 g_2 \rangle \\
&\quad - (\gamma_m + 4k_{on})\langle g_1 m \rangle \\
\frac{d\langle g_2 m \rangle}{dt} &= - (k_{off} + 4k_{on})\langle g_1 g_2 \rangle \\
\frac{d\langle g_2 m \rangle}{dt} &= k_m\langle g_2 \rangle + 2k_{on}\langle g_1 m \rangle - (k_{off} + \gamma_m)\langle g_2 m \rangle.
\end{align*}
\]

(28a)-(28g)

Note that since \( g_i(t) \) is a boolean variable \( \langle g_i^2(t) \rangle = \langle g_i(t) \rangle \). Steady-state analysis of the above linear equations yields

\[
\begin{align*}
\langle g_1 \rangle &= \frac{1 - f_{on}}{2}, \quad \langle g_2 \rangle = f_{on}, \quad \langle m \rangle = \frac{k_m f_{on}}{\gamma_m}, \\
\langle m^2 \rangle &= \text{Var}_{\text{three}}(m) + \langle m \rangle^2, \quad \langle g_1 g_2 \rangle = 0 \\
\langle g_1 m \rangle &= \frac{2k_{off} \gamma_m \langle m \rangle}{\gamma_n^2 + 4k_{on}(k_m + k_{off}) + \gamma_n(4k_{on} + k_{off})} \\
\langle g_2 m \rangle &= \frac{(\gamma_m + 2k_{on})^2 \langle m \rangle}{\gamma_n^2 + 4k_{on}(k_m + k_{off}) + \gamma_n(4k_{on} + k_{off})},
\end{align*}
\]

(29a)-(29d)

where variance \( \text{Var}_{\text{three}}(m) \) is given by (23) and \( f_{on} = k_{on} / (k_{on} + k_{off}) \) is the fraction of time the promoter is ON.

We next use (28) to compute the condition expectation (26). By solving the linear system of differential equations (28a)-(28c) with initial conditions \( g_1(s), g_2(s) \) and \( m(s) \), respectively, we obtain

\[
\langle m(t+s) | m(s), g(s) \rangle = a_1(t)\langle m \rangle + a_2(t)g_1(s) + a_3(t)g_2(s) + a_4(t)m(s)
\]

(30)

for some time-varying coefficients \( a_1(t), a_2(t), a_3(t) \) and \( a_4(t) \) that satisfy

\[
\begin{align*}
&\quad a_1(0) = a_2(0) = a_3(0) = a_4(0) = 1 \\
&\lim_{t \to -\infty} a_1(t) = 1, \quad \lim_{t \to -\infty} a_2(t) = \lim_{t \to -\infty} a_3(t) = \lim_{t \to -\infty} a_4(t) = 0.
\end{align*}
\]

(31a)-(31b)

Substituting (30) in (26) yields

\[
\langle m(t+s)m(s) \rangle = a_1(t)\langle m \rangle^2 + a_2(t)\langle g_1 m \rangle + a_3(t)\langle g_2 m \rangle + a_4(t)\langle m \rangle^2
\]

(32)

which using values of steady-state moments computed in (29) results in the mRNA auto-correlation function defined as

\[
R(t) = \frac{\langle m(t+s)m(s) \rangle - \langle m \rangle^2}{\langle m \rangle^2}.
\]

(33)

The closed-form analytical expressions for \( R(t) \) obtained in Mathematica is too complex to be listed here but representative plots are shown in Figure 3.

Analysis of \( R(t) \) shows that the three-state promoter architecture exhibits a faster decay in mRNA auto-correlations compared to a two-state architecture (Figure 3). As is the case with mRNA probability distributions (Figure 2), both architectures yield similar auto-correlations for a wide range of parameters, in particular, for \( f_{on} \) close to zero or one. Interestingly, \( R(t) \) can be non-monotonic when \( k_{off} \ll \gamma_m \) and for intermediate values of \( f_{on} \) (Figure 3), in which case, the function undershoots to the negative region before

![Figure 3. Steady-state mRNA auto-correlation functions for a two-state (blue lines) and a three-state (red lines) promoter architecture for different values of \( k_{on} \) and \( k_{off} \). Other parameters are taken as \( k_m = 150 \) and \( k_h = 0 \). All rates have been normalized by the mRNA degradation rate, i.e., \( \gamma_m = 1 \).](image-url)
converging to zero.

V. CONCLUSION

Stochastic gene-expression is often modeled using a two-state promoter architecture with memoryless transitions between different promoter states. Here we have considered a three-state promoter architecture with memory, where the promoter resides in a transcriptionally active (inactive) state for an exponentially (gamma) distributed time-interval. For the general case, where transcription can occur from both states, we derived an analytical expression for the mRNA probability distribution (Eq. (18)).

By comparing the extent and time-scale of mRNA copy number fluctuations for both architecture we identified key differences in their stochastic dynamics. Comparison of mRNA distributions revealed that the difference between them is maximum when

\[ \frac{k_{\text{off}}}{\gamma_m} \approx 1 \& f_{\text{on}} \approx 0.5, \]  

which implies

\[ k_{\text{off}} \approx k_{\text{off}} \approx \gamma_m. \]

(34)

Thus to be able to discriminate between a two-state and a three-state promoter architecture one would require reporter systems where the stability of the promoter states are comparable to the stability of the mRNA transcript. Note that an unstable mRNA \( (\gamma_m \gg 1) \) will lead to smaller values of \( k_{\text{off}} = k_{\text{off}}/\gamma_m \ll 1 \). Similarly a stable mRNA \( (\gamma_m \ll 1) \) will have \( k_{\text{off}} = \gg 1 \), and our results show that in both cases the distance between the distributions is small (Figure 2). The mRNA half-life will have to be appropriately tuned according to (35) so as to be in a region where distributions show maximum difference.

We also derived an analytical expression for the mRNA auto-correlation function \( R(t) \). For some regions of the parameter space, \( R(t) \) can undershoot to the negative region (Figure 3). This is quite different from a two-state architecture, which always exhibits monotonically decreasing auto-correlations. Thus, the shape of the auto-correlation function may be useful in discriminating between promoter architectures. Future, work will precisely map the parameter region where non-monotonic auto-correlations are observed. Moreover, we will analyze time-series data on mRNA copy number fluctuations in single-cells to detect such signatures in \( R(t) \). Finally, we plan to perform stochastic analysis of more complex promoter architecture where the dwell time in both promoter states can be an arbitrary random variable.

REFERENCES