Theory of transcription bursting: Stochasticity in the transcription rates

Rajamanickam Murugan

Department of Biotechnology, Indian Institute of Technology Madras

Chennai, Tamil Nadu India. Email: rmurugan@gmail.com

ABSTRACT

Transcription bursting creates variation among the individuals of a given population. Bursting emerges as the consequence of turning on and off the transcription process randomly. There are at least three sub-processes involved in the bursting phenomenon with different timescale regimes viz. flipping across the on-off state channels, microscopic transcription elongation events and the mesoscopic transcription dynamics along with the mRNA recycling. We demonstrate that when the flipping dynamics is coupled with the microscopic elongation events, then the distribution of the resultant transcription rates will be over-dispersed. This in turn reflects as the transcription bursting with over-dispersed non-Poisson type distribution of mRNA numbers. We further show that there exist optimum flipping rates (α_C, β_C) at which the stationary state Fano factor and variance associated with the mRNA numbers attain maxima. These optimum points are connected via $\alpha_C = \sqrt{\beta_C\left(\beta_C + \gamma_r\right)}$. Here α is the rate of flipping from the on-state to the off-state, β is the rate of flipping from the off-state to the on-state and γ_r is the decay rate of mRNA. When $\alpha = \beta = \chi$ with zero rate in the off-state channel, then there exist optimum flipping rates at which the non-stationary Fano factor and variance attain maxima. Here $\chi_{C,v} \simeq 3k_r^+ / 2\left(1 + k_r^+ t\right)$ (here k_r^+ is the rate of transcription purely through the on-state elongation channel) is the optimum flipping rate at which the variance of mRNA attains a maximum and $\chi_{C,\kappa} \simeq 1.72/t$ is the optimum flipping rate at which the Fano factor attains a maximum. Close observation of the transcription mechanism reveals that the RNA polymerase performs several rounds of stall-continue type dynamics before generating a complete mRNA. Based on this observation, we model the transcription event as a stochastic trajectory of the transcription machinery across these on-off state elongation channels. Each mRNA transcript follows different trajectory. The total time taken by a given trajectory is the first passage time (FPT). Inverse of this FPT is the resultant transcription rate associated with the particular mRNA. Therefore, the time required to generate a given mRNA transcript will be a random variable. For a stall-continue type dynamics of RNA polymerase, we show that the overall average transcription rate can be expressed as $k_r \simeq h_\infty^+ k_r^+$ where $k_r^+ \simeq \lambda_r^+ / L$, λ_r^+ is the microscopic transcription elongation rate in the on-state channel and L is the length of a complete mRNA transcript and $h_\infty^+ = \left[\beta / \left(\alpha + \beta\right)\right]$ is the stationary state probability of finding the transcription machinery in the on-state.

KEY WORDS

Transcription bursting, on-off state flipping, over-dispersion, Fano factor, mRNA dynamics, non-Poisson distribution

INTRODUCTION

Subcellular process such as transcription and translation of various genes are essential for the survival of an organism [1, 2]. Since these processes are mesoscopic in nature, their outcomes such as mRNAs and protein molecules are prone to great number fluctuations. Particularly, cells with identical genetic materials, produce different levels of mRNAs and proteins corresponding to various genes at a given time point. Such variations in the expression of different genes across the population of individual cells is essential for the survival of an organism against various extreme environmental conditions. The molecular number fluctuations across various cells of a population, can be influenced by both intrinsic and extrinsic factors. The intrinsic noise in gene expression arises well within the system itself which is characterized by a set of kinetic parameters such as transcription and translation rate constants. On the other hand, external environmental factors such as temperature and nutrient fluctuations can perturb these intrinsic kinetic parameters which ultimately emerges as the extrinsic noise component [3].

The stochasticity of the constitutive gene expression has been extensively investigated both theoretically and experimentally [4-10]. The fluctuations in the number of mRNAs and protein molecules can be well characterized by the corresponding population *mean*, *variance*, *coefficient of variation* and the *Fano factor*. Here *coefficient of variation* = (*standard deviation* / *mean*) and *Fano factor* = (*variance* / *mean*). The Fano factor measures the extent of deviation of the molecular number fluctuations from the standard Poisson process for which the Fano factor = 1. Populations which exhibit Fano factor < 1 are under-dispersed and those populations which exhibit Fano factor > 1 are over-dispersed. Detailed theoretical calculations and subsequent experimental studies on the unregulated gene expression have shown that the Fano factors associated with the protein number fluctuations is more than one and its deviation from the Poisson increases linearly with the translational efficiency [8, 11]. Here the *translational efficiency = translation rate* / *decay rate of mRNA molecules*. On the other hand, the molecular number fluctuations in mRNAs follows a typical Poisson process with Fano factor = 1.

Gene expression was initially thought as a continuous process and the probability density functions associated with the mRNAs and protein number fluctuations are assumed to be a monomodal type. Later experiments revealed the interrupted and bursting nature of mRNAs and protein numbers along the temporal axis [12, 13]. Such transcriptional bursting can result in a bimodal or multimodal type density functions associated with the number of mRNAs and proteins [12, 14, 15]. By definition, two stage gene expression involves only transcription and translation processes and three stage gene expression incudes on-off state dynamics of the promoter along with the transcription and translation. The main sources of transcription bursting can be attributed to the initiation and elongation steps [14]. In the process of transcription initiation, the promoter seems to be turned on and off in a random manner via the binding-unbinding of the regulatory transcription factors (TFs). On the other hand, the RNA polymerase enzyme complex (RNAP) can undergo stall-continue type dynamics in the transcription elongation process. Both these types of dynamics ultimately introduce a time dependent stochasticity in the overall transcription rate constant. Transcription bursting phenomenon has been studied in detail [16-23]. The effects of various factors such as negative feedback [17], presence of enhancer elements [24, 25] on the frequency of bursting and burst size have also been unraveled in detail.

Both the binding-unbinding of regulatory TFs proteins with the promoter and stall-continue type dynamics of RNAP in the elongation process ultimately switch on or off the transcription event in a time dependent and random manner. Generally, transcription elongation generates positive supercoiling ahead of the RNAP complex and leaves negative supercoil trail behind the RNAP [26-28]. For a smooth transcription elongation process, these positive and negative supercoils need to be relaxed by the gyrase and topoisomerase I enzymes respectively [14]. When there is a lack in the number of gyrase molecules to handle the accumulation of positive supercoils ahead of the transcribing RNAP, then RNAP complex cannot progress further and subsequently transcription event will be stalled. When enough gyrase molecules arrive and relax the accumulated positive supercoils, then the transcription event continues and so on. Recent studies revealed the possibility of co-transcription of several RNAPs in a convoy with finite interval among them [29, 30]. Within a convoy of RNAPs with appropriate spacing among them, there is a possibility that the negative supercoil trail of the first RNAP will be cancelled by the positive supercoil ahead of the second RNAP and so on. This is the fluid model of transcription elongation which requires generally short distances among the RNAPs [29]. The cooperative interactions among the RNAPs of a convoy can be either collaborative or antagonistic depending on the spacing among them which can be well modelled by the push and push-pull mechanisms [29-31]. The overall interactions among the RNAPs of a convoy ultimately result in the random speed up or slowdown of the transcription elongation dynamics of individual RNAPs apart from the stall-continue type dynamics of the convoy in a time dependent manner.

Clearly, there are three different timescale regimes involved in the transcription bursting viz. (**a**) the timescale associated with the microscopic elongation transitions in the generation of a complete mRNA transcript, (**b**) the timescale associated with the on-off flipping dynamics and (**c**) the timescale associated with the mesoscopic dynamics of mRNA along with its recycling. Interplay of these processes (**a**), (**b**), and (**c**) results in the over-dispersion of the distribution of mRNA numbers. Almost all the theoretical and experimental studies on the transcription bursting 1) assumed slower timescale for the on-off state flipping than the timescale of transcription and decay and 2) concentrated on the steady state of the gene expression. In this paper, using a combination of theoretical and simulation tools we will show that the flipping across the on-off transcription elongation channels introduce stochasticity in the transcription rate and there exists an optimum flipping rate at which the variance and Fano factor of mRNA numbers attain the maxima.

THEORY

Let us consider the typical transcription event along with the recycling of mRNAs well within the cellular environment as depicted in **Fig. 1**. When the transcription happens in a "stall-continue" or "on-off" mode of RNA polymerase enzyme complex (RNAP in prokaryotes and RNA Pol II in eukaryotes), then it can be described by the following set of coupled master equations.

$$\begin{aligned} \partial_t p^+_{m,t} &= k^+_r p^+_{m-1,t} + \gamma_r (m+1) p^+_{m+1,t} - \left(k^+_r + \gamma_r m\right) p^+_{m,t} - \alpha p^+_{m,t} + \beta p^-_{m,t} \\ \partial_t p^-_{m,t} &= k^-_r p^-_{m-1,t} + \gamma_r (m+1) p^-_{m+1,t} - \left(k^-_r + \gamma_m m\right) p^-_{m,t} + \alpha p^+_{m,t} - \beta p^-_{m,t} \end{aligned} \tag{1}$$

Let us denote $p_{m,t}$ as the overall probability of finding m number of mRNA molecules at time t starting from zero number of mRNAs at time $t = 0$. The probabilities of finding the transcription machinery at the on-off states are described by $p^{\pm}_{m,t}$ where the superscript (+) denotes the on-state

and (-) denotes the off-state and the corresponding transcription rates are $k_r^{\pm}$ (s^{-1} when mRNA concentration is measured in numbers or M/s when mRNA concentration is measured in M). Let us assume that the length of the complete mRNA transcript is L bp where 1 bp = 3.4 x 10^{-10} m. When there is a flipping across the on-off state channels, then the synthesis of a given mRNA transcript of interest will eventually follow a random trajectory as described in **Fig. 1**. We denote the resultant transcription rate of this random trajectory as ξ which is the inverse of the total time taken by the particular trajectory and it varies from trajectory to trajectory. We define the *average effective or resultant transcription rate* as k_r which is the average of ξ across several trajectories of the transcription event as $k_r = \langle \xi \rangle$. In general, one finds that $k_r^- \ll k_r^+$ and $k_r^- \ll k_r \leq k_r^+$.

We have assumed a non-zero transcription rate $k_r^- \neq 0$ for the off-state channel mainly due to the fact that the RNA polymerase complex completely stalls only when the level of positive supercoiling accumulating ahead of transcription is strong enough to completely stop its further progression. Otherwise, the progression of the RNA polymerase over transcription will be slower than the normal speed (but not equal to zero). Here α (s^{-1}) is the rate of transition from the on-state to the off-state and β (s^{-1}) is the rate of transition from the off-state to the on-state of transcription. The ratio $\sigma = k_r / \alpha$ is generally defined as the *transcription efficiency* [3] or the burst size. Here σ is the average number of mRNA molecules generated in the on-state of transcription. The first order decay rate of mRNAs is described by γ_r (s^{-1}). The main assumption in **Eqs. 1** is that the timescales associated with the flipping rates (α, β) are comparable with or higher than the timescales associated with the transcription rates $k_r^{\pm}$. We will show in the later sections that $k_r^{\pm}$ will be rescaled with (α, β) when the timescales associated with (α, β) are shorter than the timescales associated with the generation a full mRNA transcript. When the average transcription speed of the RNAP is $\lambda_r^{\pm}$ (bp/s) and the length of the final mRNA transcript is L bp, then the average transcription rate through the respective on and off channels will be $k_r^{\pm} \simeq \lambda_r^{\pm} / L$ (1/s). In other words, $\tau_r^{\pm} = 1/k_r^{\pm}$ (s) is the average time required to generate a full mRNA transcript via the respective on-off channels. Here one should note that the maximum elongation of speed of RNA pol II seems to be close to ~100 bp/s [32, 33]. Typical transcription elongation speed of the prokaryotic RNAP complex is $\lambda_r^+ \sim$ 50-60 bp/s [2]. The average resultant transcription rate k_r will be a function of the variables $\lambda_r^{\pm}$, L, α and β. The initial and normalization conditions associated with **Eqs. 1** can be given as follows.

$$p_{m,0}^+ = \delta(m);\ p_{m,0}^- = 0;\ \sum_m p_{m,t}^+ + \sum_m p_{m,t}^- = 1;\ p_{m,t} = p_{m,t}^+ + p_{m,t}^- \tag{2}$$

Eqs. 1 and **2** can be solved using the standard method of generating function formalism. Upon defining the generating functions as $G_{s,t}^{\pm} = \sum_{m=0}^{\infty} s^m p_{m,t}^{\pm}$, **Eq. 1** can written as follows.

$$\begin{aligned} \partial_t G_{s,t}^+ &= k_r^+ (s-1) G_{s,t}^+ + \gamma_r (1-s) \partial_s G_{s,t}^+ - \alpha G_{s,t}^+ + \beta G_{s,t}^- \\ \partial_t G_{s,t}^- &= k_r^- (s-1) G_{s,t}^- + \gamma_r (1-s) \partial_s G_{s,t}^- + \alpha G_{s,t}^+ - \beta G_{s,t}^- \end{aligned} \tag{3}$$

General solution to **Eqs. 3** was obtained in terms of confluent hypergeometric functions in Ref. [34] using Galilean type transformation of variables (Eqs. 13 and 14 in Ref. 31). However, these solutions in terms of confluent hypergeometric functions are very difficult to handle. Particularly, the derivation of the expressions for the variance and the Fano factor associated with the mRNA number fluctuations will not be a difficult task. To avoid this hurdle, we follow the route of case-splitting technique to derive all the possible approximations under various conditions which are simple and also easy to handle. Upon performing few basic operations, this system can be rewritten as the following set of two uncoupled partial differential equations.

$$\begin{aligned}
&G_{s,t}^{+} = \left[\gamma_r (s-1)\partial_s G_{s,t}^{-} + \partial_t G_{s,t}^{-} + \left((1-s)k_r^{-} + \beta\right) G_{s,t}^{-}\right] \Big/ \alpha \\
&\gamma_r^2 (s-1)^2 \partial_s^2 G_{s,t}^{-} = -2\gamma_r (s-1)\partial_s \partial_t G_{s,t}^{-} - \partial_t^2 G_{s,t}^{-} + A_1 \partial_s G_{s,t}^{-} + A_2 \partial_t G_{s,t}^{-} + A_3 G_{s,t}^{-} \\
&A_1 = \gamma_r (1-s)\left[\gamma_r^2 + \left(k_r^{-} + k_r^{+}\right)(1-s) + \alpha + \beta\right] \\
&A_2 = -\left[\left(k_r^{-} + k_r^{+}\right)(1-s) + \alpha + \beta\right];\ A_3 = (s-1)\left[k_r^{-}\gamma_r - k_r^{-}k_r^{+}s + \left(k_r^{+} + \alpha\right)k_r^{-} + \beta k_r^{+}\right]
\end{aligned} \tag{4}$$

Here we have the initial conditions as $G_{s,0}^{+} = 1, G_{s,0}^{-} = 0$ [35]. **Eqs. 3-4** are the central equations of this paper from which we derive several interesting results. Upon approximately solving these partial differential equations, one can recover the expressions for the probability density functions associated with the mRNA fluctuations $p_{m,t}^{\pm}$ as follows.

$$G_{1,t}^{-} + G_{1,t}^{+} = 1;\ p_{m,t}^{\pm} = \lim_{s\to 1}\left[\partial_s^m G_{s,t}^{\pm} / m!\right];\ p_{m,t} = \lim_{s\to 1}\left[\partial_s^m \left(G_{s,t}^{+} + G_{s,t}^{-}\right) / m!\right] \tag{5}$$

Further, using these generating functions, one can straightforwardly derive the expressions for the various time dependent statistical quantities associated with the population of mRNAs such as mean ($\eta_{m,t}$), variance ($\nu_{m,t}$), coefficient of variation ($\mu_{m,t}$) and Fano factor ($\kappa_{m,t}$) as follows.

$$\begin{aligned}
&\eta_{m,t} = \langle m \rangle = \lim_{s\to 1}\left[\partial_s \left(G_{s,t}^{+} + G_{s,t}^{-}\right)\right] \\
&\nu_{m,t} = \langle m^2 \rangle - \langle m \rangle^2 = \lim_{s\to 1}\left[\partial_s^2 \left(G_{s,t}^{+} + G_{s,t}^{-}\right)\right] + \eta_{m,t} - \eta_{m,t}^2;\ \mu_{m,t} = \nu_{m,t} / \eta_{m,t}^2;\ \kappa_{m,t} = \nu_{m,t} / \eta_{m,t}
\end{aligned} \tag{6}$$

With this background, we will consider the following exactly solvable cases and other interesting approximations of **Eqs. 3-4**.

Case I. $\alpha \neq \beta;\ k_r^{-} = 0;\ \gamma_r = 0$

This case has been investigated earlier in Refs. [35, 36] (Eqs. 11 and 12 in Ref. [35]). Under these conditions, **Eqs. 4** can be reduced to the following form which is exactly solvable.

$$\partial_t^2 G_{s,t}^{-} - \left(k_r^{+}(s-1) - (\alpha+\beta)\right)\partial_t G_{s,t}^{-} - k_r^{+}\beta(s-1) G_{s,t}^{-} = 0;\ G_{s,t}^{+} = \left[\beta G_{s,t}^{-} + \partial_t G_{s,t}^{-}\right] \Big/ \alpha \tag{7}$$

The general solution to this set of partial differential equations can be written as follows.

$$\begin{aligned}
&G_{s,t}^{+} = \left[\exp\left((w-2\beta)t/2\right)/2\alpha\right]\left(F_1(s)(w+u)\exp(ut/2) + F_2(s)(w-u)\exp(-ut/2)\right) \\
&G_{s,t}^{-} = \exp\left((w-2\beta)t/2\right)\left(F_1(s)\exp(ut/2) + F_2(s)\exp(-ut/2)\right)
\end{aligned} \tag{8}$$

Here $u=\sqrt{w^2+4\alpha\beta}$ where $w=k_r^+\left(s-1\right)+\left(\beta-\alpha\right)$. Noting that $G_{1,t}^-+G_{1,t}^+=1$, the functions $F_1(s)$ and $F_2(s)$ and the particular solution to **Eqs. 8** can be derived as follows.

$$F_1\left(s\right)+F_2\left(s\right)=0;\ F_1\left(s\right)\left(w+u\right)+F_2\left(s\right)\left(w-u\right)=2\alpha;\ \therefore F_1\left(s\right)=\alpha/u;\ F_2\left(s\right)=-\alpha/u \tag{9}$$

Upon substituting these functions into **Eqs. 8** one finally obtains the following expression for the generating functions.

$$\begin{aligned} G_{s,t}^+&=\left[\exp\left(\left(w-2\beta\right)t/2\right)/2u\right]\left(\left(w+u\right)\exp\left(ut/2\right)-\left(w-u\right)\exp\left(-ut/2\right)\right)\\ G_{s,t}^-&=\left[\exp\left(\left(w-2\beta\right)t/2\right)\alpha/u\right]\left(\exp\left(ut/2\right)-\exp\left(-ut/2\right)\right)\end{aligned} \tag{10}$$

Using the overall generating function, one can derive the mean number of mRNAs as follows.

$$\eta_{m,t}=\left\langle m\right\rangle=\lim_{s\to1}\partial_s\left(G_{s,t}^++G_{s,t}^-\right)=k_r^+\left(\alpha\left[1-\exp\left(-\left(\alpha+\beta\right)t\right)\right]+\beta\left(\alpha+\beta\right)t\right)\Big/\left(\beta+\alpha\right)^2 \tag{11}$$

For short timescales, $\eta_{m,t}\simeq k_r^+t$ which is similar to the deterministic result and for long timescale it behaves as $\eta_{m,t}\simeq\left[k_r^+\beta/\left(\beta+\alpha\right)\right]t$. Noting that $\lim_{s\to1}\partial_s^2\left(G_{s,t}^++G_{s,t}^-\right)=\left\langle m^2\right\rangle-\left\langle m\right\rangle$, which can be written explicitly as follows.

$$\left\langle m^2\right\rangle-\left\langle m\right\rangle=\left(k_r^+\Big/\left(\alpha+\beta\right)^2\right)^2\left\{\begin{aligned}&\alpha^2\left(2+\beta t\left(4+\beta t\right)\right)+2\alpha\left(\beta^2t\left(2+\beta t\right)-2\beta\right)+\\&+\beta^4t^2-2\alpha\left(\alpha^2t+\alpha\beta t+\alpha-2\beta\right)\exp\left(-\left(\alpha+\beta\right)t\right)\end{aligned}\right\} \tag{12}$$

Using **Eqs. 11** and **12** one can write down the variance, coefficient of variation and Fano factor of mRNAs as follows.

$$v_{m,t}=\lim_{s\to1}\partial_s^2\left(G_{s,t}^++G_{s,t}^-\right)+\left\langle m\right\rangle-\left\langle m\right\rangle^2;\ \mu_{m,t}=v_{m,t}/\eta_{m,t}^2;\kappa_{m,t}=v_{m,t}/\eta_{m,t} \tag{13}$$

It is remarkable to note down the following limits as time tends towards infinity.

$$\lim_{t\to0}\kappa_{m,t}=1;\ \lim_{t\to\infty}\kappa_{m,t}=\kappa_{m,\infty}=1+2\alpha k_r^+\Big/\left(\alpha+\beta\right)^2 \tag{14}$$

Clearly, in **Eqs. 14** the stationary state Fano factor $\kappa_{m,\infty}$ attains a maximum value at an optimum $\alpha_C=\beta$ which can be obtained by solving $\left[\partial\kappa_{m,\infty}/\partial\alpha\right]=0$ for α.

Case II. $\alpha=\beta=\chi;\ k_r^-=0;\ \gamma_r\neq0$

In this case, **Eqs. 4** reduce to the following set of uncoupled partial differential equations.

$$\begin{aligned}\gamma_r^2\left(s-1\right)^2\partial_s^2G_{s,t}^-&=\left\{\begin{aligned}&-2\gamma_r\left(s-1\right)\partial_s\partial_tG_{s,t}^--\partial_t^2G_{s,t}^-+\left(k_r^+\left(s-1\right)-2\chi\right)\partial_tG_{s,t}^-+k_r^+\chi\left(s-1\right)G_{s,t}^-+\\&\gamma_r\left(1-s\right)\left[\gamma_r-2\chi+\left(s-1\right)k_r^+\right]\partial_sG_{s,t}^-\end{aligned}\right\}\\ G_{s,t}^+&=\left(\gamma_r\left(s-1\right)\partial_sG_{s,t}^-+\chi G_{s,t}^-+\partial_tG_{s,t}^-\right)\Big/\chi\end{aligned} \tag{15}$$

Similar to this case with settings $\left(\alpha \neq \beta;\ k_r^- = 0;\ \gamma_r \neq 0\right)$ was considered in Refs. [35, 36]. The first term of the series expansion was obtained using recurrence relation method in Ref. [35]. Interestingly, the solution set for these conditions was obtained in terms of confluent hypergeometric functions using Galilean type transformation of variables in Ref. [36] (Eqs. A2-A5 in Ref. [36]). Remarkably, when $\chi \to \infty$, then **Eqs. 15** reduce to the following form.

$$G_{s,t}^{+} = G_{s,t}^{-};\ \lim_{\chi\to\infty} G_{s,0}^{\pm} = 1/2;\ \partial_t G_{s,t}^{-} = \left[k_r^{+}\left(s-1\right)/2\right] G_{s,t}^{-} + \gamma_r\left(1-s\right)\partial_s G_{s,t}^{-} \tag{16}$$

These equations are exactly solvable as follows.

$$\left.\begin{aligned} G_{s,t}^{\pm} &= \left[1/2\right]\exp\left(-\left(k_r^{+}\left(1-s\right)\left[1-\exp\left(-\gamma_r t\right)\right]/2\gamma_r\right)\right) \\ G_{s,t} &= G_{s,t}^{+} + G_{s,t}^{-} = \exp\left(-\left(k_r^{+}\left(1-s\right)\left[1-\exp\left(-\gamma_r t\right)\right]/2\gamma_r\right)\right) \end{aligned}\right\} \tag{17}$$

Using these generating functions, one can derive the Poisson type expressions for the probability density function associated with the mRNA population as follows.

$$\begin{aligned} p_{m,t}^{\pm} &= \left[1/2\right]\left(\left(k_r^{+}/2\gamma_r\right)\left[1-\exp\left(-\gamma_r t\right)\right]\right)^{m}\exp\left(-\left(k_r^{+}/2\gamma_r\right)\left[1-\exp\left(-\gamma_r t\right)\right]\right)/m! \\ p_{m,t} &= p_{m,t}^{+} + p_{m,t}^{-} = \left(\left(k_r^{+}/2\gamma_r\right)\left[1-\exp\left(-\gamma_r t\right)\right]\right)^{m}\exp\left(-\left(k_r^{+}/2\gamma_r\right)\left[1-\exp\left(-\gamma_r t\right)\right]\right)/m! \end{aligned} \tag{18}$$

Various statistical properties of mRNA population can be derived as follows.

$$\eta_{m,t} = \nu_{m,t} = \left(k_r^{+}/2\gamma_r\right)\left[1-\exp\left(-\gamma_r t\right)\right];\ \therefore \kappa_{m,t} = 1;\ \mu_{m,t} = 2\gamma_r/k_r^{+}\left[1-\exp\left(-\gamma_r t\right)\right] \tag{19}$$

These results clearly suggest that as $\chi \to \infty$, the overall average or effective transcription rate k_r rescales from k_r^{+} to $k_r^{+}/2$. The condition $\chi \to \infty$ implies that the system switches infinite number of times between on and off states with equal amount of dwell times at both these states. That is to say, the timescale of flipping between on-off states are comparable with the timescales of the microscopic transcription elongation events. Therefore, the condition $\chi \to \infty$ violates the assumption of **Eqs. 1** and **Eqs. 16-18** since the transcription rates will be a function of the flipping rate χ under such conditions.

Case III. $k_r^- = 0;\ \alpha = \beta = \chi;\ \gamma_r = 0$

In this case, **Eqs. 15** reduce to the following form which is exactly solvable.

$$\begin{aligned} \partial_t G_{s,t}^{+} &= k_r^{+}\left(s-1\right)G_{s,t}^{+} - \chi\left(G_{s,t}^{+} - G_{s,t}^{-}\right) \\ \partial_t G_{s,t}^{-} &= -\chi\left(G_{s,t}^{-} - G_{s,t}^{+}\right) \end{aligned} \tag{20}$$

The general solution of **Eqs. 20** can be written as follows.

$$\begin{aligned} G_{s,t}^{+} &= \left[\exp\left(-\left(w-2\chi\right)t/2\right)/2\chi\right]\left(F_1\left(s\right)\left(w+u\right)\exp\left(ut/2\right) + F_2\left(s\right)\left(w-u\right)\exp\left(-ut/2\right)\right) \\ G_{s,t}^{-} &= \exp\left(-\left(w-2\chi\right)t/2\right)\left(F_1\left(s\right)\exp\left(ut/2\right) + F_2\left(s\right)\exp\left(-ut/2\right)\right) \end{aligned} \tag{21}$$

Here $w = k_r^+ (s-1)$ and $u = \sqrt{w^2 + 4\chi^2}$. Noting the initial conditions $G_{s,0}^+ = 1;\ G_{s,0}^- = 0$ for a finite χ values one finds the following expressions for the functions $F_1(s)$ and $F_2(s)$.

$$F_1(s) + F_2(s) = 0;\ F_1(s)(w+u) + F_2(s)(w-u) = 2\chi;\ \therefore\ F_1(s) = \chi/u;\ F_2(s) = -\chi/u \tag{22}$$

Upon using these, one can write down the particular solution of **Eqs. 20** as follows.

$$\begin{aligned} G_{s,t}^+ &= \left[\exp\left((w-2\chi)t/2\right)/2u\right]\left((w+u)\exp(ut/2) - (w-u)\exp(-ut/2)\right) \\ G_{s,t}^- &= \left[\exp\left((w-2\chi)t/2\right)\chi/u\right]\left(\exp(ut/2) - \exp(-ut/2)\right) \\ &\lim_{\chi\to 0} G_{s,t}^- = 0;\ \lim_{\chi\to\infty} G_{s,t}^+ = G_{s,t}^- \end{aligned} \tag{23}$$

Using **Eqs. 22** and noting that $G_{s,t} = G_{s,t}^- + G_{s,t}^+$ one can drive the following limiting conditions of the generating function.

$$\lim_{\chi\to 0} G_{s,t} = \exp\left(k_r^+ (s-1)t\right);\ \lim_{\chi\to\infty} G_{s,t} = \exp\left(k_r^+ (s-1)t/2\right) \tag{24}$$

Using **Eqs. 23** one can directly derive the probability density function associated with the mRNA population at high and low values of the flipping rate as follows.

$$\lim_{\chi\to 0} p_{m,t} = \exp\left(-k_r^+ t\right)\left(\left(k_r^+ t\right)^m / m!\right);\ \lim_{\chi\to\infty} p_{m,t} = \exp\left(-k_r^+ t/2\right)\left(\left(k_r^+ t/2\right)^m / m!\right) \tag{25}$$

Further, one can derive the mean and the variance of the mRNA population explicitly as follows.

$$\left.\begin{aligned} \eta_{m,t} &= k_r^+ \left(1 - \exp(-2\chi t) + 2\chi t\right)/4\chi \\ v_{m,t} &= \left[k_r^+/4\chi^2\right]\left(\left(k_r^+ \chi t + 2\chi^2 t - 3k_r^+/4 + \chi\right) + \left(k_r^+ - \chi\right)\exp(-2\chi t) - k_r^+ \exp(-4\chi t)/4\right) \end{aligned}\right\} \tag{26}$$

It is also interesting to note down the following limiting values.

$$\lim_{\chi\to 0} \eta_{m,t} = k_r^+ t;\ \lim_{\chi\to\infty} \eta_{m,t} = k_r^+ t/2;\ \lim_{\chi\to 0} v_{m,t} = k_r^+ t;\ \lim_{\chi\to\infty} v_{m,t} = k_r^+ t/2 \tag{27}$$

Remarkably, the functional form of the mRNA variance shows a turnover type behavior with respect to changes in χ and it has a definite maximum at the optimum flipping rate $\chi_{C,v}$. Explicit expression for this optimum χ can be derived from the following equation.

$$\left[\partial v_{m,t}/\partial\chi\right] = \left\{\begin{aligned} &2k_r^+ \chi t - 3k_r^+ + 2\chi + \exp(-2\chi t)\left[4k_r^+ \chi t - 4\chi^2 t + 4k_r^+ - 2\chi\right] \\ &-\exp(-4\chi t)\left[2k_r^+ \chi t + k_r^+\right] \end{aligned}\right\} = 0 \tag{28}$$

Upon ignoring the terms multiplying the exponentials for large values of χ and time as $t \to \infty$, **Eq. 28** can be approximated to $2k_r^+ \chi t - 3k_r^+ + 2\chi \simeq 0$ from which one finds the optimum value of χ at which the variance becomes a maximum as follows.

$$\chi_{C,v} \simeq 3k_r^+ / 2\left(1 + k_r^+ t\right) \tag{29}$$

However, **Eq. 29** suggests that the optimum value of χ will be a time dependent quantity which decreases towards zero when the time increases towards infinity. Clearly, at the steady or stationary state i.e. in the limit as $t \to \infty$, the existence of optimum flipping rate over the variance of the mRNA population will disappear. It is also interesting to note down the following limits.

$$\begin{aligned} &\mu_{m,t} = v_{m,t}/\eta_{m,t}^2;\ \lim_{t\to 0}\mu_{m,t} = \infty;\ \lim_{t\to\infty}\mu_{m,t} = 0;\ \lim_{\chi\to 0}\mu_{m,t} = 1/k_r^+ t;\ \lim_{\chi\to\infty}\mu_{m,t} = 2/k_r^+ t \\ &\kappa_{m,t} = v_{m,t}/\eta_{m,t};\ \lim_{t\to 0}\kappa_{m,t} = 1;\ \lim_{t\to\infty}\kappa_{m,t} = 1 + k_r^+/2\chi;\ \lim_{\chi\to 0}\kappa_{m,t} = 1;\ \lim_{\chi\to\infty}\kappa_{m,t} = 1 \end{aligned} \tag{30}$$

Remarkably, the Fano factor associated with the fluctuations in the number of mRNAs show a maximum deviation from the Poisson with respect to χ which can be demonstrated as follows.

$$\left[\partial\kappa_{m,t}/\partial\chi\right] = \left\{\begin{aligned} &8\chi^2t^2 - 12\chi t - 3 + \exp(-2\chi t)\left[24\chi^2t^2 + 18\chi t + 7\right] - \\ &-\exp(-4\chi t)\left[8\chi^2t^2 + 8\chi t + 5\right] + \exp(-6\chi t)\left[2\chi t + 1\right] \end{aligned}\right\} = 0 \tag{31}$$

Similar to the derivation of **Eq. 29**, one can ignore the terms multiplying the exponentials and finally one obtains the following approximation for large values of χ.

$$8\chi^2t^2 - 12\chi t - 3 \simeq 0;\ \therefore \chi_{C,\kappa} \simeq 1.72/t \tag{32}$$

Clearly, at the steady state i.e. as $t \to \infty$, the existence of optimum flipping rate over the Fano factor as well as the variance of the mRNA population will disappear. In other words, steady state theories and experiments cannot capture these important properties. Using detailed stochastic simulations, we will show in the later sections that both **Eqs. 29** and **32** are valid even when the conditions $k_r^- > 0;\ \alpha \neq \beta;\ \gamma_r > 0$ are true. In these situations, we will show in the later sections that the optimum flipping rates (α_C, β_C) which maximize the variance and the Fano factor associated with the fluctuations in the number of mRNA molecules asymptotically approach their non-zero steady state limits.

Dependency of the transcription rate on the on-off flipping dynamics

Let us consider a single transcription event. Here the length of the final mRNA transcript is *L* bp. Since the addition of each nucleotide at the ends of a growing mRNA is an energetically driven stochastic process, we can describe the entire transcription elongation as a directed walk with the microscopic transition rates $\lambda_r^{\pm}$ (bp/s). Here the superscript '+' represents the on-state and '-' represents the off-state of the promoter and $q_{n,t}^{\pm}$ are probabilities of finding of the transcription machinery with *n* number of transcribed mRNA bases in the respective on-off state channels at time *t*. Now we can consider three possible scenarios as depicted in **Figs. 2A-B** viz. (a) when all the microscopic transitions of the transcription elongation event are characterized homogeneously with the rates $\lambda_r^{\pm}$ and there is no flipping across the on-off state channels, then one finds that $k_r^{\pm} = \lambda_r^{\pm}/L$ and subsequently the average resultant transcription rate will be equal to either k_r^+ or k_r^- depending on the channel used for the transcription elongation (**Fig. 1**, **Figs. 2B1** and **2B3** respectively). On the other hand, when there is a flipping across on-off state channels, then the resultant transcription rate ξ will be somewhere in between $\left(k_r^+, k_r^-\right)$ (**Fig. 2B2**). Clearly the

transcription rate ξ associated with each mRNA transcript will be a random variable and its ensemble average across several individual transcripts is defined as $k_r = \langle \xi \rangle$. When the timescales of flipping across on-off states are comparable with the timescales of the microscopic transcription elongation events, then such randomly interrupted transcription elongation can be well described by the following set of coupled master equations.

$$\begin{aligned} \partial_t q^+_{n,t} &= \lambda^+_r \left(q^+_{n-1,t} - q^+_{n,t} \right) - \alpha q^+_{n,t} + \beta q^-_{n,t} \\ \partial_t q^-_{n,t} &= \lambda^-_r \left(q^-_{n-1,t} - q^-_{n,t} \right) + \alpha q^+_{n,t} - \beta q^-_{n,t} \end{aligned} \tag{33}$$

The initial and boundary conditions associated with **Eqs. 34** can be given as follows [37-39].

$$q_{n,t} = q^+_{n,t} + q^-_{n,t};\ q^+_{n,0} = \delta(n);\ q^-_{n,0} = 0;\ \therefore q_{n,0} = \delta(n);\ q^\pm_{L,t} = 0 \tag{34}$$

Upon defining $g^\pm_{n_0,t} = \int_0^L q^\pm_{n,t|n_0,t_0} dn$ which is the overall probability of finding the system within $n \in (0, L)$ at time t starting from $n = 0$, one can straightforwardly derive the following results.

$$g^\pm_{n_0,0} = 1;\ g^\pm_{n_0,\infty} = 0;\ T^\pm_{n_0} = -\int_0^\infty \left[t\partial_t g^+_{n_0,t} \right] dt = \int_0^\infty g^+_{n_0,t} dt;\ \int_0^\infty \left[\partial_t g^+_{n_0,t} \right] dt = g^+_{n_0,\infty} - g^+_{n_0,0} = -1 \tag{36}$$

Further, we set that $\int_0^\infty t\partial_t g^-_{n_0,t} dt \simeq 0$ especially when $\lambda^-_r \to 0$ since the probability flux entering the off-state of the system will eventually return back to the on-state. The system can exit upon generating a complete transcript only from the on-state. Here $T^\pm_n$ is the mean first passage times (MFPT) associated with the generation of a full mRNA transcript of length L starting from n number of initial bases of mRNA at time $t = 0$. One should note that the resultant transcription rate ξ is the inverse of the first passage time (FPT) which is a random variable. Clearly, the MFPTs $T^\pm_n$ obey the following set of coupled backward type master equations [38, 39].

$$\lambda^+_r T^+_{n+1} - \lambda^+_r T^+_n - \alpha T^+_n + \beta T^-_n = -1;\ \lambda^-_r T^-_{n+1} - \lambda^-_r T^-_n + \alpha T^+_n - \beta T^-_n = 0 \tag{37}$$

Here the boundary conditions directly follow from **Eqs. 34** as $T^\pm_L = 0$. Upon solving **Eqs. 37**, the overall MFPT can be calculated as the sum $T_n = T^+_n + T^-_n$. When the initial number of bases in the mRNA transcript is $n_0 = 0$, then the required MFPT to generate a complete transcript starting from zero bases will be $T_0 = T^+_0 + T^-_0$ from which one can derive the expression for the overall average transcription rate as $k_r = [1/T_0]$. The system of coupled **Eqs. 37** is exactly solvable. The solution corresponding to the first moment of the FPT $T^\pm_n$ can be written as follows.

$$\begin{aligned} T^+_n &= \left(\alpha\lambda^-_r \left[\beta(L-n) + \lambda^-_r \left(1 - W^{-L+n} \right) \right] + \beta^2 (L-n) \lambda^+_r \right) \Big/ y^2 \\ T^-_n &= \left([L-n] y - \lambda^-_r \lambda^+_r \alpha \left(1 - W^{-L+n} \right) \right) \Big/ y^2;\ y = \left[\alpha\lambda^+_r + \beta\lambda^+_r \right];\ W = \left(1 + \beta/\lambda^-_r + \alpha/\lambda^+_r \right) \end{aligned} \tag{38}$$

Here one should note the following limiting conditions.

$$\lim_{\lambda_r^- \to 0} T_n^+ = (L-n)/\lambda_r^+ ; \lim_{\lambda_r^- \to 0} T_n^- = [\alpha/\beta](L-n)/\lambda_r^+ ;\ \lim_{\lambda_r^- \to 0} T_n = [1+\alpha/\beta](L-n)/\lambda_r^+ \quad (39)$$

The second moment of the FPTs (we denote them as $R_n^{\pm}$) satisfy the following set of coupled backward type master equations [37].

$$\lambda_r^+ R_{n+1}^+ - \lambda_r^+ R_n^+ - \alpha R_n^+ + \beta R_n^- = -2T_n^+;\ \lambda_r^- R_{n+1}^- - \lambda_r^- R_n^- + \alpha R_n^+ - \beta R_n^- = -2T_n^- \quad (40)$$

Here the boundary conditions of **Eqs. 40** directly follow from **Eqs. 34** as $R_L^{\pm} = 0$. **Eqs. 40** can be derived using the following relationships.

$$R_{n_0}^{\pm} = -\int_0^{\infty} \left[t^2 \partial_t g_{n_0,t}^{\pm} \right] dt = 2\int_0^{\infty} t g_{n_0,t}^{\pm} dt;\ \int_0^{\infty} \left[t \partial_t g_{n_0,t}^{\pm} \right] dt = -\int_0^{\infty} g_{n_0,t}^{\pm} dt = -T_{n_0}^{\pm} \quad (41)$$

The set of coupled difference equations given in **Eqs. 40** along with **Eqs. 38** for the definition of $T_n^{\pm}$ is exactly solvable and the complete solution is given in **Appendix A**. One can obtain the following simplified expressions for the second moments of FPTs associated with the generation of a complete mRNA transcript especially in the limit as $\lambda_r^- \to 0$.

$$R_n^+ = [1+\alpha/\beta](L-n)(L-n+1)/\left[\lambda_r^+\right]^2 ;\ R_n^- = \alpha(L-n)\left([\alpha+\beta](L-n+1)+2\lambda_r^+\right)/\left[\beta\lambda_r^+\right]^2 \quad (42)$$

Eqs. 42 can be obtained by substituting the expressions of $T_n^{\pm}$ given by **Eqs. 39** in **Eqs. 40** subsequently solving the resultant difference equation. We denote the mean of FPTs with respect to $n = 0$ as $\eta_T^{\pm} = T_0^{\pm}$ and, $\eta_T = T_0^+ + T_0^- = T_0$ by definition. Using **Eqs. 39** and **42**, one can obtain the expressions for the variance (v_T), Fano factor (κ_T) and coefficient of variation (μ_T) of the FPTs associated with the generation of a complete mRNA transcript of length L starting from n = 0 in the limit $\lambda_r^- \to 0$ as follows.

$$\begin{aligned}
&R_n = R_n^+ + R_n^-;\ R_0 = R_0^+ + R_0^-;\ v_T = R_0 - T_0^2;\ T_0 = \eta_T = [1+\alpha/\beta]L/\lambda_r^+ \\
&v_T = L\left[\alpha^2 + 2\alpha\left(\beta+\lambda_r^+\right) + \beta^2\right]/\left(\beta\lambda_r^+\right)^2 \\
&\kappa_T = \left(v_T/\eta_T\right) = \left[\alpha^2 + 2\alpha\left(\beta+\lambda_r^+\right) + \beta^2\right]/\beta[\alpha+\beta]\lambda_r^+ \\
&\mu_T = \left(v_T/\eta_T^2\right) = \left[\alpha^2 + 2\alpha\left(\beta+\lambda_r^+\right) + \beta^2\right]/L[\alpha+\beta]^2 \\
&\alpha_S = \lambda_r^+\beta/2 - \beta - \lambda_r^+ + \sqrt{\left(\beta^2 - 4(\beta+1)\right)\left[\lambda_r^+\right]^2 + 8\beta\lambda_r^+}\Big/2 \\
&\beta_S = \left(-\lambda_r^+\alpha + 2\alpha + \sqrt{\left(\alpha^2+8\alpha\right)\left[\lambda_r^+\right]^2 - 8\alpha\lambda_r^+}\right)\Big/2\left(\lambda_r^+ - 1\right)
\end{aligned} \quad (42)$$

In **Eqs. 42**, (α_S, β_S) are the critical values of the flipping rates (which can be obtained by solving the equation $\kappa_T = 1$ for α by fixing β or for β by fixing α) such that when $\alpha < \alpha_S$ or $\beta < \beta_S$, then the

Fano factor associated with the FPTs becomes as $\kappa_T < 1$ (sub-Poisson type). When $\alpha > \alpha_S$ or $\beta > \beta_S$, then $\kappa_T > 1$ (super-Poisson type). When $\alpha = \alpha_S$ or $\beta = \beta_S$, then $\kappa_T = 1$ (Poisson type). One should also note that α_S can exist only when $\left[\left(\beta^2 - 4(\beta+1)\right)\lambda_r^+ + 8\beta\right] \geq 0$ and β_S can exist only when $\left[(\alpha+8)\lambda_r^+ - 8\right] \geq 0$. Since the overall transcription rate of a given mRNA transcript is the inverse of the respective FPT, one can conclude that the distribution of transcription rates will be the sub or super Poisson type depending on the relative values of the on-off flipping rates (α, β). Remarkably, the Fano factor κ_T associated with the distribution of FPTs is independent on the size of the mRNA transcript L. Further, the coefficient of variation of FPTs μ_T shows a turn over type behavior with respect to changes in α with a definite maximum at $\alpha = \beta$ which can be shown by solving $\partial\mu_T/\partial\alpha = 0$ for α. It is also interesting to note down the following limiting conditions.

$$\begin{aligned} &\lim_{\beta\to\infty} v_T = L\Big/\left(\lambda_r^+\right)^2;\ \lim_{\beta\to\infty} \kappa_T = 1/\lambda_r^+;\ \lim_{\beta\to\infty} \mu_T = 1/L \\ &\lim_{\alpha\to 0} v_T = L\Big/\left(\lambda_r^+\right)^2;\ \lim_{\alpha\to 0} \kappa_T = 1/\lambda_r^+;\ \lim_{\alpha\to 0} \mu_T = 1/L \end{aligned} \tag{43}$$

Especially, when $\alpha = \beta = \chi$ then one can obtain the following simplified expression for T_0 using the set of **Eqs. 38** as follows.

$$T_0 = 2L\Big/\left(\lambda_r^- + \lambda_r^+\right) + \lambda_r^-\left(1 - \left(1 + \chi/\lambda_r^- + \chi/\lambda_r^+\right)^{-L}\right)\left(\lambda_r^- - \lambda_r^+\right)\Big/\chi\left[\lambda_r^- + \lambda_r^+\right]^2 \tag{44}$$

Eqs. 44 clearly suggest that the overall average transcription rate k_r is a function of the flipping rate χ especially when the timescale associated with the flipping dynamics is much slower than the timescale associated with the generation of a complete transcript. It is remarkable to note the limiting values of T_0 as $\lim_{\chi\to 0} T_0 = L/\lambda_r^+$ and, $\lim_{\chi\to\infty} T_0 = 2L\Big/\left(\lambda_r^- + \lambda_r^+\right)$ .When $\lambda_r^- \to 0$ then **Eqs. 44** clearly suggest that T_0 will be independent of the on-off flipping rate χ. However, the MFPT will be doubled under such situations. When $\lambda_r^- \ll \lambda_r^+$ and for sufficiently large values of the transcript length L, one finally obtains the limiting conditions in line with **Eqs. 19** and **30** as $\lim_{\chi\to\infty} k_r = \lambda_r^+/2L$ and, $\lim_{\chi\to 0} k_r = \lambda_r^+/L$ where $k_r^+ = \lambda_r^+/L$. These limits follow directly from the initial conditions given in **Eqs. 35**. For example, when we hypothetically set up the initial conditions as $q_{n,0}^- = \delta(n);\ q_{n,0}^+ = 0$, then we need to swap the on-off state superscript indices (+ and -) in **Eqs. 39** and subsequently one obtains the limiting values $\lim_{\chi\to\infty} k_r \simeq \left(\lambda_r^- + \lambda_r^+\right)\Big/2L$ and, $\lim_{\chi\to 0} k_r = \lambda_r^-/L$ where $k_r^- = \lambda_r^-/L$ even when $\lambda_r^- < \lambda_r^+$.

When $\lambda_r^- \to 0$ then the system of **Eqs. 37** will be uncoupled. In this situation, the overall average transcription rate k_r rescales from k_r^+ to $k_r^+/2$ when χ tends towards infinity i.e. the overall transcription rate k_r will be confined inside the interval $\left(k_r^+/2\right) \leq k_r \leq k_r^+$. When $\lambda_r^- \to 0$ and $\alpha \neq \beta$

then the overall average effective transcription rate can be expressed as the functions of the on-off flipping rates as $k_r \simeq k_r^+\left[\beta/(\alpha+\beta)\right]$. This equation suggests the probable range of k_r as $0 \le k_r \le k_r^+$ depending on the relative values of the on-off flipping rates. In the following sections, we investigate the possibility of various steady state solutions to **Eqs. 1**.

Steady state solutions

Under steady state conditions, **Eqs. 1** reduce to the following form.

$$\begin{aligned} &k_r^+ p_{m-1,\infty}^+ + \gamma_r (m+1) p_{m+1,\infty}^+ - \left(k_r^+ + \gamma_r m\right) p_{m,\infty}^+ - \alpha p_{m,\infty}^+ + \beta p_{m,\infty}^- = 0 \\ &k_r^- p_{m-1,\infty}^+ + \gamma_r (m+1) p_{m+1,\infty}^- - \left(k_r^- + \gamma_m m\right) p_{m,\infty}^- + \alpha p_{m,\infty}^+ - \beta p_{m,\infty}^- = 0 \end{aligned} \tag{45}$$

Here $p_{m,\infty}^{\pm}$ are the stationary state probabilities associated with the finding of m number of mRNAs in the respective on (+) and off (-) channels of transcription. **Eqs. 45** also can be solved using the standard generating function formalism. One can define the steady state generating functions and the corresponding probability densities as follows.

$$G_{s,\infty}^{\pm} = \sum_m \left(s^m p_{m,\infty}^{\pm}\right);\ G_{1,\infty}^- + G_{1,\infty}^+ = \sum_m p_{m,\infty}^+ + \sum_m p_{m,\infty}^- = 1 \tag{46}$$

Using the generating functions defined in **Eqs. 46**, one can recover the respective probability density functions as follows.

$$p_{m,\infty}^{\pm} = \lim_{s\to 0}\left[\partial_s^m G_{s,\infty}^{\pm}/m!\right];\ p_{m,\infty} = \lim_{s\to 0}\left[\partial_s^m \left(G_{s,\infty}^+ + G_{s,\infty}^-\right)/m!\right] \tag{47}$$

Upon applying the transformation given in **Eqs. 46** into **Eqs. 45**, one obtains the following set of coupled ordinary differential equations.

$$\begin{aligned} &k_r^+ (s-1) G_{s,\infty}^+ + \gamma_r (1-s) d_s G_{s,\infty}^+ - \alpha G_{s,\infty}^+ + \beta G_{s,\infty}^- = 0 \\ &k_r^- (s-1) G_{s,\infty}^- + \gamma_r (1-s) d_s G_{s,\infty}^- + \alpha G_{s,\infty}^+ - \beta G_{s,\infty}^- = 0 \end{aligned} \tag{48}$$

In Refs. [15, 17], the solutions to **Eqs. 45** were expressed in terms of confluent hypergeometric functions. Here we derive an alternate solution set that is easy to handle. **Eqs. 48** can be split into the following set of uncoupled ordinary differential equations.

$$\begin{aligned} &\gamma_r^2 (s-1) d_s^2 G_{s,\infty}^- + F_s d_s G_{s,\infty}^- + H_s G_{s,\infty}^- = 0 \\ &G_{s,\infty}^+ = \left[\gamma_r (s-1) d_s G_{s,\infty}^- - \left(k_r^- (s-1) - \beta\right) G_{s,\infty}^-\right]/\alpha \end{aligned} \tag{49}$$

Various functions in **Eqs. 49** are defined as follows.

$$F_s = \left(\gamma_r + \alpha + \beta - (s-1)\left(k_r^+ + k_r^-\right)\right)\gamma_r;\ H_s = \left(k_r^+ (s-1) - \gamma_r - \alpha\right) k_r^- - k_r^+ \beta \tag{50}$$

Using the substitution as $G_{s,\infty}^{-} = \exp\left(k_r^{+}\left(s-1\right)/\gamma_r\right)M_s$, where M_s is an arbitrary function of s, one can reduce **Eqs. 49** into the following form.

$$zd_z^2 M_s + \left(b-z\right)d_z M_s - aM_s = 0;\ a = 1+\alpha/\gamma_r;\ b = 1+\left(\alpha+\beta\right)/\gamma_r;\ z = \left[k_r^{-} - k_r^{+}\right]\left(s-1\right)/\gamma_r \quad (51)$$

Solution to **Eq. 51** can be expressed in terms of Kummer functions as follows [40].

$$M_z = C_1 \mathrm{KummerM}\left(a,b,z\right) + C_2 \mathrm{KummerU}\left(a,b,z\right) \quad (52)$$

Since $\mathrm{KummerU}\left(a,b,0\right) \to \infty$ for any arbitrary values of $\left\{a,b\right\} > 0$ in **Eq. 52**, we need to set $C_2 = 0$ to enforce the normalization condition of the probability density function. We use the following properties of the KummerM functions to further simplify our results [40, 41].

$$\left.\begin{aligned}
&d_z \mathrm{KummerM}\left(a,b,z\right) = \left[a/z\right]\left(\mathrm{KummerM}\left(a+1,b,z\right) - \mathrm{KummerM}\left(a,b,z\right)\right)\\
&\mathrm{KummerM}\left(0,a,z\right) = 1\\
&\mathrm{KummerM}\left(a,b,z\right) = \exp\left(z\right)\mathrm{KummerM}\left(b-a,b,-z\right)\\
&\mathrm{KummerM}\left(a,a,z\right) = \exp\left(z\right)\mathrm{KummerM}\left(0,a,-z\right) = \exp\left(z\right)
\end{aligned}\right\} \quad (53)$$

Using these properties, one finally obtains the solution to the generating functions obeying the set of differential equations **Eqs. 48** as follows.

$$\begin{aligned}
G_{s,\infty}^{-} &= \left\{\begin{aligned}&\left[\alpha/\left(\alpha+\beta\right)\right]\exp\left(k_r^{+}\left(s-1\right)/\gamma_r\right)\times\\ &\times\mathrm{KummerM}\left(1+\alpha/\gamma_r, 1+\left(\alpha+\beta\right)/\gamma_r, \left[k_r^{-}-k_r^{+}\right]\left(s-1\right)/\gamma_r\right)\end{aligned}\right\}\\
G_{s,\infty}^{+} &= \left\{\begin{aligned}&\left[\beta/\left(\alpha+\beta\right)\right]\exp\left(k_r^{+}\left(s-1\right)/\gamma_r\right)\times\\ &\times\mathrm{KummerM}\left(\alpha/\gamma_r, 1+\left(\alpha+\beta\right)/\gamma_r, \left[k_r^{-}-k_r^{+}\right]\left(s-1\right)/\gamma_r\right)\end{aligned}\right\}
\end{aligned} \quad (54)$$

Here the KummerM function can be defined explicitly as follows [40, 41].

$$\mathrm{KummerM}\left(a,b,z\right) = \sum_{n=0}^{\infty}\left[\left(a\right)_n z^n/n!\left(b\right)_n\right];\ \left(w\right)_n = w\left(w+1\right)\ldots\left(w+n-1\right);\ \left(w\right)_0 = 1 \quad (55)$$

Upon expanding the generating functions in terms of Macularin series with respect to s around s = 0 and then substituting s = 1 in the computed series, one finally obtains the expressions for the respective probability density functions as follows.

$$\begin{aligned}
p_{m,\infty}^{+} &= \left\{\begin{aligned}&\left[\beta/m!\left(\alpha+\beta\right)\gamma_r^m\right]\exp\left(-k_r^{+}/\gamma_r\right)\times\\ &\times\sum_{n=0}^{m}\left[C_n^m\left(\alpha/\gamma_r\right)_{m-n}\left(k_r^{+}\right)^n\left(k_r^{-}-k_r^{+}\right)^{m-n}\Psi_{+}\Big/\left(1+\left(\alpha+\beta\right)/\gamma_r\right)_{m-n}\right]\end{aligned}\right\}\\
p_{m,\infty}^{-} &= \left\{\begin{aligned}&\left[\alpha/m!\left(\alpha+\beta\right)\gamma_r^m\right]\exp\left(-k_r^{+}/\gamma_r\right)\times\\ &\times\sum_{n=0}^{m}\left[C_n^m\left(1+\alpha/\gamma_r\right)_{m-n}\left(k_r^{+}\right)^n\left(k_r^{-}-k_r^{+}\right)^{m-n}\Psi_{-}\Big/\left(1+\left(\alpha+\beta\right)/\gamma_r\right)_{m-n}\right]\end{aligned}\right\}
\end{aligned} \quad (56)$$

Various paramaters and terms in **Eqs. 56** are defined as follows.

$$\begin{aligned}
&\Psi_{+} = \text{KummerM}\left(m-n+\alpha/\gamma_r, m-n+1+(\alpha+\beta)/\gamma_r, -\left(k_r^- - k_r^+\right)/\gamma_r\right) \\
&\Psi_{-} = \text{KummerM}\left(m-n+1+\alpha/\gamma_r, m-n+1+(\alpha+\beta)/\gamma_r, -\left(k_r^- - k_r^+\right)/\gamma_r\right) \\
&C_n^m = \left[m!/n!(m-n)!\right];\ (w)_u = \Gamma(w+u)/\Gamma(w)
\end{aligned} \tag{57}$$

From **Eqs. 54**, one can derive the expressions for the steady state mean, variance, coefficient of variation and Fano factor (we denote them as $\eta_{m,\infty}, v_{m,\infty}, \mu_{m,\infty}, \kappa_{m,\infty}$ respectively) as follows.

$$\begin{aligned}
&\eta_{m,\infty} = \langle m \rangle_\infty = \lim_{s\to 1}\left[\partial_s\left(G_{s,\infty}^+ + G_{s,\infty}^-\right)\right] = \left(\alpha k_r^- + \beta k_r^+\right)/(\alpha+\beta)\gamma_r \\
&v_{m,\infty} = \left\{\begin{array}{l} \lim_{s\to 1}\left[\partial_s^2\left(G_{s,\infty}^+ + G_{s,\infty}^-\right)\right] + \eta_{m,\infty} - \eta_{m,\infty}^2 = \\ = \left(\alpha^3 k_r^- + \Xi\alpha^2 + \beta^2 k_r^+(\beta+\gamma_r) + \Omega\beta\alpha\right)/\gamma_r(\alpha+\beta)^2(\alpha+\beta+\gamma_r) \end{array}\right\} \\
&\mu_{m,\infty} = v_{m,\infty}/\eta_{m,\infty}^2 = \left(\alpha^3 k_r^- + \Xi\alpha^2 + \beta^2 k_r^+(\beta+\gamma_r) + \Omega\beta\alpha\right)/\left(\alpha k_r^- + \beta k_r^+\right)^2(\alpha+\beta+\gamma_r) \\
&\kappa_{m,\infty} = \left\{\begin{array}{l} v_{m,\infty}/\eta_{m,\infty} = \\ = \left(\alpha^3 k_r^- + \Xi\alpha^2 + \beta^2 k_r^+(\beta+\gamma_r) + \Omega\beta\alpha\right)/\left(\alpha k_r^- + \beta k_r^+\right)(\alpha+\beta)(\alpha+\beta+\gamma_r) \end{array}\right\}
\end{aligned} \tag{58}$$

Various terms in **Eqs. 58** are defined as follows.

$$\Xi = \left(\beta\left(2k_r^- + k_r^+\right) + k_r^-\gamma_r\right);\ \Omega = \left(\beta\left(k_r^- + 2k_r^+\right) + k_r^-\left(k_r^- + \gamma_r - 2k_r^+\right) + k_r^+\left(\gamma_r + k_r^+\right)\right) \tag{59}$$

Eqs. 58 suggest that, in the presence of flipping across the on-off states, the effective transcription rate transforms as $k_r = \left(\alpha k_r^- + \beta k_r^+\right)/(\alpha+\beta)$ under steady state conditions. In the following sections we will consider various cases of approximations to **Eqs. 54-59**.

Case I. $\alpha \neq \beta;\ k_r^- = 0$

In this situation, **Eqs. 54** reduce to the following form.

$$\begin{aligned}
&G_{s,\infty}^- = \left[\alpha/(\alpha+\beta)\right]\exp\left(k_r^+(s-1)/\gamma_r\right)\text{KummerM}\left(1+\alpha/\gamma_r, 1+(\alpha+\beta)/\gamma_r, -k_r^+(s-1)/\gamma_r\right) \\
&G_{s,\infty}^+ = \left[\beta/(\alpha+\beta)\right]\exp\left(k_r^+(s-1)/\gamma_r\right)\text{KummerM}\left(\alpha/\gamma_r, 1+(\alpha+\beta)/\gamma_r, -k_r^+(s-1)/\gamma_r\right)
\end{aligned} \tag{60}$$

Upon expanding these generating functions in terms of Macularin series around the point $s = 0$ and then substituting $s = 1$ in the computed series, one finally obtains the following probability density functions associated with the steady state mRNA populations.

$$p_{m,\infty}^{+}=\left\{\begin{array}{l}\left[\beta/m!(\alpha+\beta)\right]\left(k_r^{+}/\gamma_r\right)^{m}\exp\left(-k_r^{+}/\gamma_r\right)\times\\ \times\sum_{n=0}^{m}\left[(-1)^{m-n}\,\Phi_{+}C_n^m\left(\alpha/\gamma_r\right)_{m-n}\Big/\left(1+(\alpha+\beta)/\gamma_r\right)_{m-n}\right]\end{array}\right\}$$
$$p_{m,\infty}^{-}=\left\{\begin{array}{l}\left[\alpha/m!(\alpha+\beta)\right]\left(k_r^{+}/\gamma_r\right)^{m}\exp\left(-k_r^{+}/\gamma_r\right)\times\\ \times\sum_{n=0}^{m}\left[(-1)^{m-n}\,\Phi_{-}C_n^m\left(1+\alpha/\gamma_r\right)_{m-n}\Big/\left(1+(\alpha+\beta)/\gamma_r\right)_{m-n}\right]\end{array}\right\} \tag{61}$$

Various terms in **Eqs. 61** are defined as follows.

$$\Phi_{+}=\text{KummerM}\left(m-n+\alpha/\gamma_r,m-n+1+(\alpha+\beta)/\gamma_r,k_r^{+}/\gamma_r\right)$$
$$\Phi_{-}=\text{KummerM}\left(m-n+1+\alpha/\gamma_r,m-n+1+(\alpha+\beta)/\gamma_r,k_r^{+}/\gamma_r\right) \tag{62}$$

From **Eqs. 60**, one can derive the expressions for the steady state mean, variance, coefficient of variation and Fano factor as follows.

$$\eta_{m,\infty}=\beta k_r^{+}\big/(\alpha+\beta)\gamma_r$$
$$v_{m,\infty}=\beta k_r^{+}\left(\alpha^2+\alpha\left(2\beta+\gamma_r+k_r^{+}\right)+\beta(\beta+\gamma_r)\right)\Big/\gamma_r(\alpha+\beta)^2(\alpha+\beta+\gamma_r)$$
$$\mu_{m,\infty}=\left(\alpha^2+\alpha\left(2\beta+\gamma_r+k_r^{+}\right)+\beta(\beta+\gamma_r)\right)\Big/\beta k_r^{+}(\alpha+\beta+\gamma_r) \tag{63}$$
$$\kappa_{m,\infty}=\left(\alpha^2+\alpha\left(2\beta+\gamma_r+k_r^{+}\right)+\beta(\beta+\gamma_r)\right)\Big/(\alpha+\beta)(\alpha+\beta+\gamma_r)$$
$$\max\left(\kappa_{m,\infty}\right)=\left(\alpha_C^2+\alpha_C\left(2\beta+\gamma_r+k_r^{+}\right)+\beta(\beta+\gamma_r)\right)\Big/(\alpha_C+\beta)(\alpha_C+\beta+\gamma_r)$$

Clearly, when $k_r^{-}=0$, then the overall transcription rate rescales as $k_r=\beta k_r^{+}\big/(\alpha+\beta)$ and using the respective limiting condition from **Eqs. 63** one can derive the optimum α at which the stationary state Fano factor $\kappa_{m,\infty}$ attains a maximum as $\alpha_C=\sqrt{\beta(\beta+\gamma_r)}$. This can be obtained by solving $\left[\partial\kappa_{m,\infty}/\partial\alpha\right]=0$ for α. Using the definition of the burst size or the transcription efficiency $\sigma=k_r/\alpha$, one can also show that the steady state Fano factor $\kappa_{m,\infty}$ will attain a maximum value at the optimum transcription efficiency $\sigma_C=k_r/\alpha_C=k_r\big/\sqrt{\beta(\beta+\gamma_r)}$.

Case II. $\alpha=\beta=\chi;\ k_r^{-}\neq 0$

In this case, the generating functions reduce to the following form.

$$G_{s,\infty}^{-}=\left[\exp\left(k_r^{+}(s-1)/\gamma_r\right)/2\right]\text{KummerM}\left(1+\chi/\gamma_r,1+2\chi/\gamma_r,\left[k_r^{-}-k_r^{+}\right](s-1)/\gamma_r\right)$$
$$G_{s,\infty}^{+}=\left[\exp\left(k_r^{+}(s-1)/\gamma_r\right)/2\right]\text{KummerM}\left(\chi/\gamma_r,1+2\chi/\gamma_r,\left[k_r^{-}-k_r^{+}\right](s-1)/\gamma_r\right) \tag{64}$$

Upon expanding these generating functions in terms of Macularin series around the point $s=0$ and then substituting $s=1$ in the computed series, one finally obtains the following probability density functions associated with the steady state mRNA populations.

$$p_{m,\infty}^{+}=\left\{\begin{array}{l}\left[\exp\left(-k_r^{+}/\gamma_r\right)/2m!\gamma_r^{m}\right]\times\\ \times\sum_{n=0}^{m}\left[\left(k_r^{+}\right)^{n}\left(k_r^{-}-k_r^{+}\right)^{m-n}\Delta_{+}C_n^{m}\left(\chi/\gamma_r\right)_{m-n}/\left(1+2\chi/\gamma_r\right)_{m-n}\right]\end{array}\right\}$$
$$p_{m,\infty}^{-}=\left\{\begin{array}{l}\left[\exp\left(-k_r^{+}/\gamma_r\right)/2m!\gamma_r^{m}\right]\times\\ \times\sum_{n=0}^{m}\left[\left(k_r^{+}\right)^{n}\left(k_r^{-}-k_r^{+}\right)^{m-n}\Delta_{-}C_n^{m}\left(1+\chi/\gamma_r\right)_{m-n}/\left(1+2\chi/\gamma_r\right)_{m-n}\right]\end{array}\right\} \quad (65)$$

Various terms and functions in **Eqs. 65** are defined as follows.

$$\Delta_{+}=\mathrm{KummerM}\left(m-n+\chi/\gamma_r,m-n+1+2\chi/\gamma_r,-\left(k_r^{-}-k_r^{+}\right)/\gamma_r\right)$$
$$\Delta_{-}=\mathrm{KummerM}\left(m-n+1+\chi/\gamma_r,m-n+1+2\chi/\gamma_r,-\left(k_r^{-}-k_r^{+}\right)/\gamma_r\right) \quad (66)$$

Using the generating functions given in **Eqs. 64** and noting the definition of the overall generating function $G_{s,\infty}=G_{s,\infty}^{-}+G_{s,\infty}^{+}$, one can derive the following properties of the stationary state mRNA populations and their various limiting properties.

$$\eta_{m,\infty}=\left(k_r^{-}+k_r^{+}\right)/2\gamma_r \quad (67)$$

$$v_{m,\infty}=\left[\left(k_r^{-2}+k_r^{-}\left(4\chi+2\gamma_r-2k_r^{+}\right)\right)\gamma_r+4k_r^{+}\left(\chi+\gamma_r/2+k_r^{+}/4\right)\right]/\left(4\gamma_r^{2}+8\gamma_r\chi\right)$$
$$\lim_{\chi\to 0}v_{m,\infty}=\left[\left(k_r^{-2}+2k_r^{-}\left(\gamma_r-k_r^{+}\right)+k_r^{+}\left(2\gamma_r+k_r^{+}\right)\right)\right]/4\gamma_r^{2};\ \lim_{\chi\to\infty}v_{m,\infty}=\left(k_r^{-}+k_r^{+}\right)/2\gamma_r \quad (68)$$

$$\mu_{m,\infty}=4\gamma_r\left[k_r^{+}\left(\chi+\gamma_r/2+k_r^{+}/4\right)+k_r^{-}\left(\chi+\gamma_r/2-k_r^{+}/2+k_r^{-}/4\right)\right]/\left(k_r^{-}+k_r^{+}\right)^{2}\left(2\chi+\gamma_r\right)$$
$$\lim_{\chi\to\infty}\mu_{m,\infty}=2\gamma_r/\left(k_r^{+}+k_r^{-}\right) \quad (69)$$
$$\lim_{\chi\to 0}\mu_{m,\infty}=\left[\left(\left(k_r^{-}\right)^{2}+2k_r^{-}\left(\gamma_r-k_r^{+}\right)+k_r^{+}\left(2\gamma_r+k_r^{+}\right)\right)\right]/\left(k_r^{+}+k_r^{-}\right)^{2}$$

$$\kappa_{m,\infty}=\left[\left(\left[k_r^{-}\right]^{2}+k_r^{-}\left(4\chi+2\gamma_r-2k_r^{+}\right)\right)\gamma_r+4k_r^{+}\left(\chi+\gamma_r/2+k_r^{+}/4\right)\right]/2\left(k_r^{-}+k_r^{+}\right)\left(2\chi+\gamma_r\right)$$
$$\lim_{\chi\to 0}\kappa_{m,\infty}=\left[\left(\left[k_r^{-}\right]^{2}+2k_r^{-}\left(\gamma_r-k_r^{+}\right)+k_r^{+}\left(2\gamma_r+k_r^{+}\right)\right)\right]/2\gamma_r\left(k_r^{+}+k_r^{-}\right);\ \lim_{\chi\to\infty}\kappa_{m,\infty}=1 \quad (70)$$

Case III. $\alpha=\beta=\chi;\ k_r^{-}=0$

In this situation, various statistical properties of the stationary state mRNA populations and their various limiting values can be derived as follows.

$$\left.\begin{aligned}
&\eta_{m,\infty} = k_r^+ / 2\gamma_r \\
&v_{m,\infty} = k_r^+ \left(4\chi + k_r^+ + 2\gamma_r\right) / 4\gamma_r \left(2\chi + \gamma_r\right) \\
&\lim_{\chi\to 0} v_{m,\infty} = k_r^+ \left(k_r^+ + 2\gamma_r\right) / 4\gamma_r^2 ;\ \lim_{\chi\to\infty} v_{m,\infty} = k_r^+ / 2\gamma_r \\
&\mu_{m,\infty} = \gamma_r \left(k_r^+ + 2\gamma_r + 4\chi\right) / k_r^+ \left(2\chi + \gamma_r\right) \\
&\lim_{\chi\to\infty} \mu_{m,\infty} = 2\gamma_r / k_r^+ ;\ \lim_{\chi\to 0} \mu_{m,\infty} = \left(2\gamma_r + k_r^+\right) / k_r^+ \\
&\kappa_{m,\infty} = \left(k_r^+ + 2\gamma_r + 4\chi\right) / \left(4\chi + 2\gamma_r\right) \\
&\lim_{\chi\to 0} \kappa_{m,\infty} = 1 + k_r^+ / 2\gamma_r ;\ \lim_{\chi\to\infty} \kappa_{m,\infty} = 1
\end{aligned}\right\} \tag{71}$$

Case IV. $\alpha \neq \beta;\ \alpha \ll \gamma_r;\ \beta \ll \gamma_r;\ k_r^- \neq 0$

In this case, the generating functions reduce to the following simple form.

$$G_{s,\infty}^- \simeq \left[\alpha / (\alpha+\beta)\right] \exp\left(k_r^- (s-1) / \gamma_r\right);\ G_{s,\infty}^+ \simeq \left[\beta / (\alpha+\beta)\right] \exp\left(k_r^+ (s-1) / \gamma_r\right) \tag{72}$$

The corresponding probability density functions can be derived via expanding these generating functions in terms of Macularin series around $s = 0$ and then substituting $s = 1$ in the computed series expansion.

$$\begin{aligned}
&p_{m,\infty}^- \simeq \left[\alpha / (\alpha+\beta)\right] \left(\exp\left(-k_r^- / \gamma_r\right) \left(k_r^- / \gamma_r\right)^m / m!\right) \\
&p_{m,\infty}^+ = \left[\beta / (\alpha+\beta)\right] \left(\exp\left(-k_r^+ / \gamma_r\right) \left(k_r^+ / \gamma_r\right)^m / m!\right)
\end{aligned} \tag{73}$$

Therefore, one finally obtains the following bimodal Poisson type expression for the probability density function associated with the stationary state mRNA populations.

$$p_{m,\infty} = \left(\alpha \exp\left(-k_r^- / \gamma_r\right) \left(k_r^- / \gamma_r\right)^m + \beta \exp\left(-k_r^+ / \gamma_r\right) \left(k_r^+ / \gamma_r\right)^m\right) / m!(\alpha+\beta) \tag{74}$$

These types of bimodal probability density of the mRNA populations seem to play critical role in the cell to cell variability within an organism [14]. From **Eqs. 60**, one can derive the expressions for the steady state statistical properties of the mRNA population as follows.

$$\begin{aligned}
&\eta_{m,\infty} = \left(\beta k_r^+ + \alpha k_r^-\right) / (\alpha+\beta)\gamma_r \\
&v_{m,\infty} = \left(\alpha^2 k_r^- \gamma_r + \beta\alpha\left(\left(k_r^- + k_r^+\right)\gamma_r + \left(k_r^+ - k_r^-\right)^2\right) + \beta^2 k_r^+ \gamma_r\right) / \gamma_r^2 (\alpha+\beta)^2 \\
&\mu_{m,\infty} = \left(\alpha^2 k_r^- \gamma_r + \beta\alpha\left(k_r^-\left(k_r^- + \gamma_r - 2k_r^+\right)\gamma_r + k_r^+\left(\gamma_r + k_r^+\right)\right) + \beta^2 k_r^+ \gamma_r\right) / \left(\alpha k_r^- + \beta k_r^+\right)^2 \\
&\kappa_{m,\infty} = \left(\alpha^2 k_r^- \gamma_r + \beta\alpha\left(k_r^-\left(k_r^- + \gamma_r - 2k_r^+\right)\gamma_r + k_r^+\left(\gamma_r + k_r^+\right)\right) + \beta^2 k_r^+ \gamma_r\right) / \gamma_r \left(\alpha k_r^- + \beta k_r^+\right)(\alpha+\beta)
\end{aligned} \tag{75}$$

From **Eqs. 75** one can conclude that the steady state Fano factor will attain a maximum when the at the flipping rates α_C, β_C which are connected via $\alpha_C = \beta_C \sqrt{k_r^+ / k_r^-}$.

Case V. $\alpha \neq \beta;\ \alpha \ll \gamma_r;\ \beta \ll \gamma_r;\ k_r^- = 0$

This is exactly the scenario addressed in Ref. [14]. In this case, the generating functions defined in **Eqs. 54** takes the following simple form.

$$G_{s,\infty}^- \simeq \left[\alpha/(\alpha+\beta)\right];\ G_{s,\infty}^+ \simeq \left[\beta/(\alpha+\beta)\right]\exp\left(k_r^+(s-1)/\gamma_r\right) \tag{76}$$

We have used the properties of the KummerM functions defined in **Eqs. 53** to derive **Eqs. 76**. The corresponding probability density functions can be derived as follows.

$$p_{m,\infty}^- \simeq \left[\alpha/(\alpha+\beta)\right];\ p_{m,\infty}^+ = \left[\beta/(\alpha+\beta)\right]\exp\left(-k_r^+/\gamma_r\right)\left[\left(k_r^+/\gamma_r\right)^m/m!\right] \tag{77}$$

Using **Eqs. 77**, one finally obtains the Poisson density function with zero spike [14].

$$p_{m,\infty} \simeq \left[\alpha/(\alpha+\beta)\right] + \left[\beta/(\alpha+\beta)\right]\exp\left(-k_r^+/\gamma_r\right)\left[\left(k_r^+/\gamma_r\right)^m/m!\right] \tag{78}$$

In **Eq. 78**, the number of mRNA molecules m takes only the integer values and **Eq. 78** is valid for the entire range of m i.e. $m = 0$ to infinity. Noting that $G_{s,\infty} = G_{s,\infty}^- + G_{s,\infty}^+$, from **Eqs. 76**, one can derive various statistical properties such as mean, variance, coefficient of variation and Fano factor associated with the stationary state mRNA populations as follows.

$$\left.\begin{aligned}
\eta_{m,\infty} &= \beta k_r^+/(\alpha+\beta)\gamma_r \\
\nu_{m,\infty} &= \beta k_r^+\left(\alpha\left(\gamma_r + k_r^+\right) + \beta\gamma_r\right)/\gamma_r^2(\alpha+\beta)^2 \\
\mu_{m,\infty} &= \left(\alpha\left(\gamma_r + k_r^+\right) + \beta\gamma_r\right)/\beta k_r^+ \\
\kappa_{m,\infty} &= 1 + k_r^+\alpha/\gamma_r(\alpha+\beta)
\end{aligned}\right\} \tag{79}$$

STOCHASTIC SIMULATION METHODS

To check the validity **Eqs. 29** and **32** under various conditions, we performed detailed stochastic simulations on the complete system of master equations **Eqs. 1**. Here there are two different timescales viz. the timescale associated with the generation of a complete mRNA transcript and the timescale associated with the flipping dynamics across on-off state channels of transcription. We used the Gillespie algorithm [42, 43] to simulate the system of **Eqs. 1**. Let us denote the number of mRNA molecules at time t as m. Initially at $t = 0$, $m = 0$ and the system was in the on-state so that $k_r \to k_r^+$. Clearly, there are four different reaction transitions viz. $(m-1) \to m$ which represents the zeroth order transcription with a rate k_r, $(m+1) \to m$ which represents the first order recycling of mRNA molecules with a rate $\gamma_r m$, $[+] \to [-]$ represents the flipping from the on-state to the off-state with a rate α and $[-] \to [+]$ represents the flipping from the off-state to the on-state with a rate β. The total reaction flux here is $f_T = k_r + \gamma_r m + \alpha + \beta$. We generated two different random numbers which are equally distributed inside $r_1, r_2 \in (0,1)$. The reaction times were sampled from the exponential type distribution $p(\tau) \propto \exp(-f_T\tau)$. This can be achieved via

transforming r_1 using the rule $\tau = -(\ln r_1)/f_T$. We used r_2 to decide on which reaction takes place at this time point. In these iteration steps we set $k_r \to k_r^{\pm}$ depending on the current state of the transcription channel.

a) When $r_2 \le (k_r/f_T)$ then the transcription reaction will take place.
b) When $(k_r/f_T) < r_2 \le ((k_r + \gamma_r m)/f_T)$, then the recycling of mRNA will take place.
c) When $((k_r + \gamma_r m)/f_T) < r_2 \le ((k_r + \gamma_r m + \alpha)/f_T)$, then the transition from the on-state to the off-state will take place. Upon such transition we set the transcription rate as $k_r \to k_r^-$.
d) When $((k_r + \gamma_r m + \alpha)/f_T) < r_2$, then the transition from the off-state to the on-state will take place. Upon such transition we set the transcription rate as $k_r \to k_r^+$.

Several trajectories were generated each with a total time t_T and various statistical properties of mRNA numbers such as mean, variance, coefficient of variation and Fano factor were computed across the time axis. To compute the optimum flipping rate, same set of simulations were performed at different values of the flipping rate. Using this dataset, the optimum flipping rates at which the variance and the Fano factor attained the maximum were numerically computed.

To compute the mean first passage time (MFPT) associated with the generation of a complete mRNA transcript we used the following algorithm. We set up the initial number of mRNA bases as $n = 0$ and the system starts from the on-state or off-state channel with the respective microscopic elongation rates $\lambda_r^{\pm}$. The completely transcribed mRNA transcript will have L number of bases. Clearly, there are three different transitions viz. $(n-1) \to n$ which represents the transition with rate $\lambda_r^{\pm}$ depending on the transcription channel at that time point, $[+] \to [-]$ represents the flipping from the on-state to the off-state with a rate α and $[-] \to [+]$ represents the flipping from the off-state to the on-state channel with a rate β. The total reaction flux here is $f_Q = \lambda_r + \alpha + \beta$ where $\lambda_r = \lambda_r^{\pm}$ depending on the current state of the transcription channel. We generated two different random numbers which are equally distributed inside $r_1, r_2 \in (0,1)$. The reaction times were sampled from the exponential distribution $p(\tau) \propto \exp(-f_Q \tau)$. This can be achieved via transforming r_1 using the rule $\tau = -(\ln r_1)/f_Q$. We used r_2 to decide on which reaction transition takes place at this time point. In these iteration steps, we set $\lambda_r \to \lambda_r^{\pm}$ depending on the current state of the transcription channel.

a) When $r_2 \le (\lambda_r/f_Q)$ then the microscopic elongation transition $n \to (n+1)$ will take place.

b) When $\left(\lambda_r / f_Q\right) < r_2 \leq \left(\left(\lambda_r + \alpha\right) / f_Q\right)$, then the transition from the on-state to the off-state will take place. Upon such transition we set the microscopic elongation transition rate as $\lambda_r \rightarrow \lambda_r^-$.

c) When $\left(\left(\lambda_r + \alpha\right) / f_Q\right) < r_2$, then the transition from the off-state to the on-state will take place. Upon such transition we set the microscopic elongation transition rate as $\lambda_r \rightarrow \lambda_r^+$.

When $n = L$, then the iteration was stopped and the first passage time (FPT) was noted. Several such trajectories were generated and the obtained FPTs were used to compute the MFPT and other statistical properties of FPTs such as variance, coefficient of variation and the Fano factor. This analysis was repeated at various of values of the flipping rates.

RESULTS AND DISCUSSION

Transcription bursting seems to emerge as a consequence of the interplay between the on-off flipping rates (α, β), microscopic transcription elongation rates (λ_r^+, λ_r^-), resultant transcription rates (k_r^+, k_r^-), and the recycling rate of mRNAs γ_r. When the on-off flipping rates and the transcription elongation rates are comparable to each other, then a continuous type transcription with monomodal type distribution of the mRNA populations will be observed as in **Fig. 3A**. In this case, the transcription will be always in the on-state by definition. Transcription bursting emerges when one sets $\alpha > \beta$ and the condition $k_r^- \ll k_r^+$ is satisfied which is clearly demonstrated in **Figs. 3B-D**. Here β is the rate of flipping from the off-state to the on-state of transcription and α is the rate of flipping from the on-state to the off-state. In most of the natural scenarios one finds that $k_r^- \simeq 0$. Therefore, the inequality condition $\alpha > \beta$ along with the decay rate constant γ_r decides the emergence of the bursting in transcription and also the burst size that is defined as $\sigma = k_r / \alpha$. The mRNA number fluctuations in the two-stage or continuous transcription and decay follows a typical Poisson density function with Fano factor = 1. When the transcription undergoes on-off flipping dynamics, then the mRNA number fluctuations follow a typical super Poisson type density function with a Fano factor more than one. Earlier studies suggested that such systems are closely follow a negative binomial distribution function [15]. The entire transcription scheme can be fragmented into several sub-processes viz.

(a) The flipping dynamics across the on-off channels of the transcription. This is characterized by the flipping rates (α, β). Here the on-state describes the fully functional transcription machinery and the off-state describes a scenario where the RNAP is stalled due to the presence huge positive supercoil barrier ahead or other chromosomal barriers.

(b) Microscopic transcription elongation events representing the growth of individual mRNAs starting from $n = 0$ number of bases toward $n = L$ bp. This is characterized by the microscopic transcription elongation rates (λ_r^+, λ_r^-) corresponding to the on and off channels respectively.

(c) Mesoscopic transcription dynamics. This is characterized by the mesoscopic transcription rates (k_r^+, k_r^-) corresponding to the on and off state channels which are the cumulative effects of (**a**) and (**b**). When the transcription follows pure on or off state channels so that there is no flipping across them i.e., $\alpha = \beta = 0$, then one finds that $k_r^\pm \simeq \lambda_r^\pm / L$. When the transcription follows a

random trajectory via flipping across on and off states, then the resultant transcription rate ξ will be a random variable. Clearly, ξ varies across different mRNA transcripts. The overall average transcription rate can be defined as $k_r = \langle \xi \rangle$. When $\alpha \neq \beta$, then one finds the probable range of the resultant transcription rate $k_r^- \leq \xi \leq k_r^+$ and $k_r^- \leq k_r \leq k_r^+$.

(d) Recycling dynamics of mRNAs which is characterized by the decay rate γ_r. Independently it is a first order decay which follows a Poisson type density function.

Individually, the population of states in the on-off channels associated with the uncoupled sub-process (**a**) follows a binomial density function for which the Fano factor is less than one. This is similar to the tossing of a coin. On the other hand, the population of various states in the sub-processes (**b**), (**c**) and (**d**) individually follow typical Poisson density function with a Fano factor of one (see **Appendix B** for a simplified derivation). When the sub-process (**a**) is dynamically coupled with (**b**) and (**c**), then the resultant or effective transcription rate ξ randomly fluctuates among the population of transcription events across (k_r^-, k_r^+). The probability density function associated with the fluctuations in the resultant transcription rate ξ will be dependent on the on-off flipping rates, microscopic elongation rates (λ_r^-, λ_r^+) and the length of the mRNA transcript L.

Figs. 4 and **5** demonstrate how the statistical properties of the mRNA number fluctuations varies with time and flipping rate parameters. When $\alpha = \beta = \chi$ and $\gamma_r = 0$, then one finds from **Figs. 4B** and **D** that the variance and the Fano factor associated with the mRNA numbers attain maxima at the optimum flipping rates $\chi_{C,v}$ and $\chi_{C,\kappa}$ respectively. These optimum flipping rates seems to be decrease with time towards zero in line with the theoretical predictions of **Eqs. 29** and **32**. These results are shown in **Figs. 5**. Clearly, the system will follow a Poisson type density function (Fano factor of mRNA number fluctuations equals to one) when the flipping rates become $\chi \to 0$ as well as $\chi \to \infty$. This reasonable since when $\chi \to 0$ then the on-off channels will be uncoupled and all the transcription events will be via the on-state channel. When $\chi \to \infty$ then the system of strongly coupled and both the transcription channels will be equally utilized. When $\gamma_r \neq 0$, then the optimum flipping rates $\chi_{C,v}$ and $\chi_{C,\kappa}$ first converge to the steady state limits and then move slowly towards zero which is evident from **Figs. 4E**, **G**, **Figs. 5B** and **E**. **Figs. 6** demonstrate how the statistical properties of the mRNA number fluctuations vary with time when $\alpha \neq \beta$. Remarkably, **Eqs. 63** suggested that there exists steady state optimum points α_C and β_C at which the Fano factor attains the maximum. These optimum points are connected via $\alpha_C = \sqrt{\beta_C(\beta_C + \gamma_r)}$ as shown in **Eqs. 63** and **Figs. 7**. When we fix α and iterate β, then one finds the optimum value of α at which the Fano factor attains maximum as $\alpha_C = \sqrt{\beta(\beta + \gamma_r)}$. Particularly for a fixed $\beta = 1$, one finds for $\gamma_r = 1$ that $\alpha_C = \sqrt{2}$ as shown in **Fig. 6D**. In the same way, when we fix β and then iterate α, then one finds the optimum β at which the Fano factor attains a maximum as $\beta_C = \sqrt{\alpha^2 + \gamma_r^2/4} - \gamma_r/2$ and so on. Particularly for a fixed $\alpha = 1$, one finds for $\gamma_r = 1$ that $\beta_C = 0.618$ as shown in **Fig. 6H**.

Generation of an individual mRNA transcript can be characterized by sequential microscopic transcription elongation rates (λ_r^+, λ_r^-) as shown in **Figs. 2**. We define the time that is required to generate a complete mRNA transcript of size L bp starting from zero number of bases as the first passage time (FPT). Inverse of this FPT is the transcription rate ξ associated with the respective mRNA transcript. Since FPT is a random variable across the population of transcription events, the resultant transcription rate will be a time dependent random variable in view of the mesoscopic transcription dynamics [44]. The average of FPTs is defined as the mean first passage time (MFPT). When there is a flipping across the on-off states, then the overall average transcription rate k_r will be an inverse of the mean first passage time (T_0, MFPT) as $k_r = [1/T_0]$. When there is no flipping across the on-off states then one finds that $k_r = k_r^{\pm}$ where $k_r^{\pm} = \lambda_r^{\pm}/L$ depending upon the transcription channel used to generate the mRNA transcript.

Figs. 8 demonstrate how the statistical properties of FPTs such as mean, variance, Fano factor and coefficient of variation associated with the generation of a complete transcript varies with the on-off flipping rates. When we set $\alpha = \beta = \chi$ and $\lambda_r^- \neq 0$, then there exists an optimum flipping rate $\chi_{C,T}$ at which the variance and the Fano factor associated with the distribution of FPTs attain the maxima. This result is line with the prediction of **Eqs. A7** of **Appendix A**. Further, the Fano factor associated with the FPTs seems to be less than one throughout the iteration range of χ. This indicates the sub-Poisson type density function of FPTs. These results are demonstrated in **Figs. 8A-D**. When $\alpha \neq \beta$ and $\lambda_r^- \to 0$, then the distribution of FPTs varies depending on the ratio (α / β) of the flipping rates.

When (α / β) = 1, then the distribution of FPTs approximately follow a Poisson type density function with a Fano factor = 1. When (α / β) > 1, then the distribution of FPTs follow a super Poisson (over dispersed with a Fano factor > 1). When (α / β) < 1, then the distribution of FPTs follow a sub Poisson (under dispersed with a Fano factor < 1). These results are demonstrated in **Figs. 8E** and **F**. It is also interesting to note the limit $\lim_{\lambda_r^- \to 0} T_0 \simeq (\alpha + \beta) L / \beta \lambda_r^+$. These results are consistent with the theoretical prediction by **Eqs. A3** of **Appendix A**. as we have shown in **Appendix A**. Various other limiting values of the MFPT (corresponding to the case $\lambda_r^- \to 0$) associated with the generation of a complete mRNA transcript are shown in **Figs. 9** (for the situation where $\alpha = \beta = \chi$) and **Figs. 10** (for the situation where $\alpha \neq \beta$). **Figs. 10** provides the computational proof for the various limiting values of the MFPT required to generate a full mRNA transcript as described in the **Appendix A**. Variation of the critical flipping rates (α_S, β_S) associated with distribution of FPTs with respect to (α, β) as described in **Eqs. 42** is demonstrated in **Figs. 10E** and **F**. Emergence of steady state bimodal type distribution of mRNA populations is demonstrated in **Figs. 11** along with a fitting to the experimental data digitized from Ref. [14].

One can broadly classify the models on transcription bursting into two state and multi state models [45]. In two state models, the promoter is assumed to flip across on and off states. Two state models generally produce a bimodal type density of the mRNA numbers. In the multistate models [19], there are several states across which the transcription process fluctuates. Multistate models generally result in the multimodal type density functions associated with the mRNA number

fluctuations [46]. In all these models, the transcription rate k_r will be assumed to be a constant. This assumption will work only when the timescale associated with the flipping across various states is much higher than the timescale associated with the generation of a complete mRNA transcript. It is generally assumed that the transcription machinery generates several mRNA transcripts (following the Poisson density function) in the on-state before flipping to the off-state e.g. the interrupted Poisson model developed in Ref. [47]. However close look at the underlying transcription mechanism reveals that the RNAP or RNA pol II performs several rounds of stall-continue type dynamics before generating a complete transcript. Therefore, the time that is required to generate a given transcript will be a random variable. This means that the resultant transcription rate ξ (which is the inverse of the first passage time required to generate a complete mRNA transcript) will be a random variable.

The overall average transcription rate will be defined as $k_r = \int \xi p(\xi) d\xi = \langle \xi \rangle$ where $p(\xi)$ is the probability density function associated with the distribution of transcription rates. Depending on the timescale of on-off state flipping one can consider two different scenarios viz. (**a**) when the timescale of on-off flipping is much longer than the timescale associated with the generation of a complete mRNA transcript, then the transcription rate will be homogenous within a given on-state period but varies from burst to burst. This can be approximately described as a static disorder in the transcription rates. (**b**) When the timescale of flipping is much shorter than the timescale associated with the mRNA synthesis, then the transcription rate varies from mRNA to mRNA. This can be described as a dynamical disorder in the transcription rates [44, 48, 49]. Our mean first passage time calculations suggested that when the rate of off-state transcription channel is zero i.e. $\lambda_r^- \to 0$, then the overall average transcription rate will be transformed as $k_r \simeq h_\infty^+ k_r^+$ where $k_r^+ \simeq \lambda_r^+ / L$ and $h_\infty^+ = [\beta/(\alpha+\beta)]$ is the stationary state probability of finding the transcription machinery in the on-state (**Appendix A** and **Eqs. B8** of **Appendix B** or sufficiently large mRNA length L) in the presence of flipping across the on-off states. This means that the transcription "rate constant" will be a function the on-off flipping rates. Comparison of this quantity with the steady state and time dependent solutions of **Eqs. 1** clearly suggested that **Eqs. 1** is accurate enough (this can be inferred from the expression of $\eta_{m,t}$ in **Eqs. 11** and $\eta_{m,\infty}$ of **Eqs. 58**) to capture the inhomogeneity of the resultant transcription rate ξ at sufficiently large mRNA lengths. However, **Eqs. 1** cannot explain the origin of over-dispersion in the mRNA numbers. Our detailed study reveals that such over-dispersion mainly originates from the non-Poisson type distribution of FPTs associated with the elongation of individual mRNA transcripts especially when the on-off flipping rates are such that $(\alpha / \beta) > 1$.

CONCLUSION

Transcription bursting is essential to generate variation among the individuals of a given population. The mechanism of bursting comprises of at least three sub-processes with different timescale regimes viz. flipping dynamics across the on-off state transcription elongation channels, microscopic transcription elongation events and the mesoscopic transcription dynamics along with the mRNA recycling. Flipping dynamics across the on-off states is similar to the tossing of a coin. When the flipping dynamics is combined with the microscopic elongation events, then the distribution of resultant transcription rates will be over-dispersed. This in turn reflects as the over-dispersed non-Poisson type distribution of mRNA numbers.

Our detailed calculations show that there exist optimum flipping rates (α_C, β_C) at which the stationary state Fano factor and variance associated with the mRNA numbers attain maxima. These optimum points are connected via $\alpha_C = \sqrt{\beta_C\left(\beta_C + \gamma_r\right)}$. Here α is the rate of flipping from the on-state to the off-state and β is the rate of flipping from the off-state to the on-state of transcription and γ_r is the decay rate of mRNA. When $\alpha = \beta = \chi$, then there exist optimum flipping rates at which the non-stationary Fano factor and variance attain maxima. Here $\chi_{C,v} \simeq 3k_r^+ \big/ 2\left(1 + k_r^+ t\right)$ (where k_r^+ is the rate of transcription through the on-state channel) is the optimum flipping rate at which the variance of mRNA numbers attains a maximum and $\chi_{C,\kappa} \simeq 1.72/t$ is the optimum flipping rate at which the Fano factor attains a maximum. These optimum points reduce to zero when $t \to \infty$.

Close look at the transcription mechanism reveals that the RNA polymerase enzyme complex performs several rounds of stall-continue type dynamics before generating a complete mRNA transcript. Based on this observation we model the transcription event as the stochastic trajectory taken by the transcription machinery across these on-off state elongation channels. Each transcript follows different trajectory. The total time taken by a given trajectory is the first passage time (FPT). Inverse of this FPT is the resultant transcription rate ξ associated with the particular mRNA. Therefore, the time that is required to generate a given mRNA transcript will be a random variable. This means that the resultant transcription rate ξ will be a random variable. The overall average transcription rate will be the ensemble average $k_r = \langle \xi \rangle$ which is equal to the inverse of the mean first passage time required to generate a complete mRNA. For a stall-continue type dynamics of RNA polymerase, the overall average transcription rate can be expressed as $k_r \simeq h_\infty^+ k_r^+$ where $k_r^+ \simeq \lambda_r^+ \big/ L$, λ_r^+ is the microscopic transcription elongation rate on the on-state channel and L is the length of complete mRNA and $h_\infty^+ = \left[\beta \big/ \left(\alpha + \beta\right)\right]$ is the stationary state probability of finding the transcription machinery in the on-state.

APPENDIX A

Consider the following backward type coupled master equations (**Eqs. 40** of the main text) corresponding to the second moments of the first passage times (FPTs) required to generate a complete mRNA transcript starting from an arbitrary *n* size of mRNA [37, 39].

$$\lambda_r^+ R_{n+1}^+ - \lambda_r^+ R_n^+ - \alpha R_n^+ + \beta R_n^- = -2T_n^+;\ \lambda_r^- R_{n+1}^- - \lambda_r^- R_n^- + \alpha R_n^+ - \beta R_n^- = -2T_n^-;\ R_L^\pm = 0 \tag{A1}$$

In this equation, the first moments of the FPTs are defined as follows (as described in **Eqs. 42** of the main text).

$$\begin{aligned}
&T_n^+ = \left(\alpha\lambda_r^-\left[\beta(L-n)+\lambda_r^-\left(1-W^{-L+n}\right)\right]+\beta^2(L-n)\lambda_r^+\right)\Big/y^2\\
&T_n^- = \left(\left[L-n\right]y-\lambda_r^-\lambda_r^+\alpha\left(1-W^{-L+n}\right)\right)\Big/y^2\\
&T_n = \left(\left[L-n\right]y-\lambda_r^-\lambda_r^+\alpha\left(1-W^{-L+n}\right)+\alpha\lambda_r^-\left[\beta(L-n)+\lambda_r^-\left(1-W^{-L+n}\right)\right]+\beta^2(L-n)\lambda_r^+\right)\Big/y^2\\
&T_n = T_n^+ + T_n^-;\ y = \left[\alpha\lambda_r^- + \beta\lambda_r^+\right];\ W = \left(1+\beta/\lambda_r^- + \alpha/\lambda_r^+\right)
\end{aligned} \tag{A2}$$

The set of difference **Eqs. A1** is exactly solvable using standard summing technique [37] and the final expression corresponding to the variance (v_T), Fano factor (κ_T) and coefficient of variation (μ_T) of FPTs associated with the generation of a complete mRNA transcript of length *L* starting from *n* = 0 can be written as follows.

$$\begin{aligned}
&R_n = R_n^+ + R_n^-;\ R_0 = R_0^+ + R_0^-\\
&\eta_T = T_0 = \left(Ly-\lambda_r^-\lambda_r^+\alpha\left(1-W^{-L}\right)+\alpha\lambda_r^-\left[\beta L+\lambda_r^-\left(1-W^{-L}\right)\right]+\beta^2 L\lambda_r^+\right)\Big/y^2\\
&v_T = R_0 - T_0^2 = \frac{W^{-2L}\begin{pmatrix}\alpha^2\lambda_r^+A_1+\alpha^3\lambda_r^-A_2+\\ \alpha^4 L\left[\lambda_r^-\right]^2 W^L\left(2\beta\lambda_r^- W^L+3\beta\lambda_r^+W^L+\lambda_r^-\lambda_r^+\left(W^L-2\right)+2\left[\lambda_r^-\right]^2\right)+\\ \alpha^5 L\left[\lambda_r^-\right]^3 W^{2L}+\alpha\beta\left[\lambda_r^+\right]^2 W^L A_3+\beta^4 L\left[\lambda_r^+\right]^3(\beta+\lambda_r^-)W^{2L}\end{pmatrix}}{(\alpha\lambda_r^-+\beta\lambda_r^+)^4(\alpha\lambda_r^-+\lambda_r^+(\beta+\lambda_r^-))}\\
&\kappa_T = \left(v_T/\eta_T\right);\ \mu_T = \left(v_T/\eta_T^2\right)
\end{aligned} \tag{A3}$$

Here *y* and *W* are defined as in **Eqs. A2** and various other terms A_1, A_2 and A_3 in **Eqs. A3** are defined as follows.

$$A_1 = \begin{pmatrix}\beta^2 L\lambda_r^- W^L\left(\left[\lambda_r^-\right]^2\left(5W^L+4\right)-2\lambda_r^-\lambda_r^+\left(2W^L+1\right)+\left[\lambda_r^+\right]^2\left(5W^L-2\right)\right)+\\ \beta^3 L W^{2L}\left(3\left[\lambda_r^-\right]^2+6\lambda_r^-\lambda_r^++\left[\lambda_r^+\right]^2\right)+\\ \beta\left[\lambda_r^-\right]^2\left(4W^L+(2L-3)W^{2L}-1\right)(\lambda_r^--\lambda_r^+)^2+\left[\lambda_r^-\right]^3\left(W^{2L}-1\right)(\lambda_r^--\lambda_r^+)^2\end{pmatrix} \tag{A4}$$

$$A_2 = \begin{pmatrix} \left[\lambda_r^-\right]^2 \left(\beta^2 L W^{2L} - 2\beta L \lambda_r^+ \left(W^L - 1\right) W^L + \left[\lambda_r^+\right]^2 \left(-2LW^L + W^{2L} - 1\right)\right) + \\ 3\beta^2 L \left[\lambda_r^+\right]^2 W^{2L} + 2\left[\lambda_r^-\right]^3 \left(\beta L \left(W^L + 1\right) W^L + 2L\lambda_r^+ W^L - \lambda_r^+ W^{2L} + \lambda_r^+\right) + \\ 2\beta L \lambda_r^- \lambda_r^+ W^L \left(3\beta W^L + 2\lambda_r^+ \left(W^L - 1\right)\right) + \left[\lambda_r^-\right]^4 \left(-2LW^L + W^{2L} - 1\right) \end{pmatrix} \tag{A5}$$

$$A_3 = \begin{pmatrix} 2\beta^2 L \left(\left[\lambda_r^-\right]^2 \left(2W^L + 1\right) - \lambda_r^- \lambda_r^+ \left(W^L + 1\right) + \left[\lambda_r^+\right]^2 W^L\right) \\ + \beta^3 L W^L (3\lambda_r^- + 2\lambda_r^+) + 2\beta\lambda_r^- \left(LW^L - 2W^L + L + 2\right)(\lambda_r^- - \lambda_r^+)^2 - \\ 4\left[\lambda_r^-\right]^2 \left(W^L - 1\right)(\lambda_r^- - \lambda_r^+)^2 \end{pmatrix} \tag{A6}$$

In **Eqs. A4-A6**, W is defined as in **Eqs. A2**. When $\alpha = \beta = \chi$, then the expression corresponding to the variance of FPTs given in **Eqs. A3** simplifies to the following form.

$$\eta_T = \left(L\left[\lambda_r^- + \lambda_r^+\right] - \lambda_r^- \lambda_r^+ \left(1 - Y^{-L}\right) + \lambda_r^- \left[\chi L + \lambda_r^- \left(1 - Y^{-L}\right)\right] + \chi L \lambda_r^+\right) \Big/ \left[\lambda_r^- + \lambda_r^+\right]^2$$

$$v_T = \frac{Y^{-2L} \begin{pmatrix} 2\chi^2 L Y^L (\lambda_r^- + \lambda_r^+)^2 B_1 + 4\chi^3 L Y^{2L} (\lambda_r^- + \lambda_r^+)^3 + \\ \chi\lambda_r^- (\lambda_r^- - \lambda_r^+)^2 (\lambda_r^- + \lambda_r^+) B_2 + \left[\lambda_r^-\right]^2 \lambda_r^+ \left(Y^L - 1\right)(\lambda_r^- - \lambda_r^+)^2 B_3 \end{pmatrix}}{\chi^2 (\lambda_r^- + \lambda_r^+)^4 (\chi(\lambda_r^- + \lambda_r^+) + \lambda_r^- \lambda_r^+)} \tag{A7}$$

$$\kappa_T = \left(v_T / \eta_T\right);\ \mu_T = \left(v_T / \eta_T^2\right)$$

The terms Y, B_1, B_2 and B_3 in **Eqs. A7** are defined as follows.

$$Y = \left(1 + \chi/\lambda_r^- + \chi/\lambda_r^+\right);\ B_1 = \left(\left[\lambda_r^-\right]^2 \left(Y^L + 2\right) + \left[\lambda_r^+\right]^2 Y^L - 2\lambda_r^- \lambda_r^+\right)$$

$$B_2 = \left(\lambda_r^- \left(-2LY^L + Y^{2L} - 1\right) + 2\lambda_r^+ \left(LY^L - 2Y^L + L + 2\right) Y^L\right) \tag{A8}$$

$$B_3 = \left(\lambda_r^- Y^L - 4\lambda_r^+ Y^L + \lambda_r^-\right)$$

Numerical analysis of **Eqs. A7** reveals the existence of maximum v_T, κ_T and μ_T at the optimum flipping rate $\chi_{C,T}$ which is demonstrated in **Figs. 8B** and **D**. **Eqs. A2** clearly suggest that when the microscopic transcription elongation rate associated with the off-state channel is close to zero i.e. $\lambda_r^- \to 0$ then the overall average transcription rate scales with the on-off flipping rates as $k_r = \left[1/T_0\right] \simeq k_r^+ \left[\beta/(\alpha + \beta)\right]$ where $k_r^+ = \lambda_r^+ / L$ is the overall transcription rate of the pure on-state channel and λ_r^+ is the elongation rate of the on-state channel.

APPENDIX B

Let us consider the uncoupled on-off state transcription elongation channels. When there is no flipping across the on and off state channels, then **Eqs. 1** will be uncoupled as follows.

$$\partial_t u_{m,t}^{\pm} = k_r^{\pm} u_{m-1,t}^{\pm} + \gamma_r (m+1) u_{m+1,t}^{\pm} - \left(k_r^{\pm} + \gamma_r m\right) u_{m,t}^{\pm};\ u_{m,0}^{\pm} = \delta(m) \tag{B1}$$

Here $u_{m,t}^{\pm}$ is the probability of finding *m* number of mRNAs at time *t* in the respective independent transcription channels. Upon defining the generating functions as $G_{s,t}^{\pm} = \sum_{m=0}^{\infty} s^m u_{m,t}^{\pm}$ one can transform **Eqs. B1** into the following partial differential equations.

$$\partial_t G_{s,t}^{\pm} = k_r^{\pm} (s-1) G_{s,t}^{\pm} + \gamma_r (1-s) \partial_s G_{s,t}^{\pm};\ \ G_{s,0}^{\pm} = 1 \tag{B2}$$

Upon solving **Eqs. B2** for the for the appropriate boundary conditions one finds the expression for the generating functions as follows.

$$G_{s,t}^{\pm} = \exp\left(k_r^{\pm} (s-1)\left(1-\exp(-\gamma_r t)\right)/\gamma_r\right) \tag{B3}$$

Using this generating function one can derive various statistical properties of the mRNA number fluctuations corresponding to on-off channels as follows.

$$\begin{aligned} &\eta_{m,t}^{\pm} = \lim_{s\to 1} \partial_s G_{s,t}^{\pm} = k_r^{\pm}\left(1-\exp(-\gamma_r t)\right)/\gamma_r \\ &v_{m,t}^{\pm} = \lim_{s\to 1} \partial_s^2 G_{s,t}^{\pm} + \eta_{m,t} - \eta_{m,t}^2 = k_r^{\pm}\left(1-\exp(-\gamma_r t)\right)/\gamma_r \\ &\kappa_{m,t}^{\pm} = 1;\ \mu_{m,t}^{\pm} = \left(k_r^{\pm}\left(1-\exp(-\gamma_r t)\right)/\gamma_r\right)^{-1} \end{aligned} \tag{B4}$$

Upon expanding the generating function in to a Macularin series with respect to variable *s* and then setting *s* = 1, one finally obtains the probability density functions as follows.

$$u_{m,t}^{\pm} = \left(\left(k_r^{\pm}/\gamma_r\right)\left[1-\exp(-\gamma_r t)\right]\right)^m \exp\left(-\left(k_r^{\pm}/\gamma_r\right)\left[1-\exp(-\gamma_r t)\right]\right)/m! \tag{B5}$$

Using **Eqs. B5** one can directly obtain the steady state probability density functions as follows.

$$u_{m,\infty}^{\pm} = \left(k_r^{\pm}/\gamma_r\right)^m \exp\left(-k_r^{\pm}/\gamma_r\right)/m! \tag{B6}$$

When the on-off state flipping is uncoupled from the transcription elongation, then the flipping dynamics can be described by the following set of coupled differential equations.

$$\partial_t \begin{bmatrix} h_t^+ \\ h_t^- \end{bmatrix} = \begin{bmatrix} -\alpha & \beta \\ \alpha & -\beta \end{bmatrix} \begin{bmatrix} h_t^+ \\ h_t^- \end{bmatrix};\ \begin{bmatrix} h_0^+ \\ h_0^- \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{B7}$$

Here we have the normalization condition $h_t^+ + h_t^- = 1$. Upon solving **Eqs. B7** with the appropriate initial conditions, one obtains the expression for the time dependent probability $h_t^{\pm}$ of finding the transcription state in the respective on-off channels as follows.

$$\begin{bmatrix} h_t^+ \\ h_t^- \end{bmatrix} = \begin{bmatrix} \left[\beta/(\alpha+\beta)\right] - \left[\alpha/(\alpha+\beta)\right]\exp\left(-(\alpha+\beta)t\right) \\ \left[\alpha/(\alpha+\beta)\right]\left(1-\exp\left(-(\alpha+\beta)t\right)\right) \end{bmatrix}; \begin{bmatrix} h_\infty^+ \\ h_\infty^- \end{bmatrix} = \begin{bmatrix} \beta/(\alpha+\beta) \\ \alpha/(\alpha+\beta) \end{bmatrix} \tag{B8}$$

Upon combining **Eqs. B5** with **Eqs. B8**, one can derive $p_{m,t}^{\pm} \simeq u_{m,t}^{\pm} h_t^{\pm}$, which is the probability of finding *m* number of mRNAs in the respective on-off state channels. Explicitly one can write the following expression.

$$p_{m,t} \simeq u_{m,t}^{+} h_t^{+} + u_{m,t}^{-} h_t^{-} = h_t^{+} \left(\varphi_t^{+}\right)^m \exp\left(-\varphi_t^{+}\right)/m! + h_t^{-} \left(\varphi_t^{-}\right)^m \exp\left(-\varphi_t^{+}\right)/m! \tag{B9}$$

In this equation, we have defined the function $\varphi_t^{\pm} = \left(k_r^{\pm}/\gamma_r\right)\left[1-\exp\left(-\gamma_r t\right)\right]$. When the timescale associated with the on-off state flipping dynamics is much lower than the timescale associated with the mRNA synthesis and decay dynamics so that $(\alpha+\beta) \gg \gamma_r$, then **Eq. B9** reduces to $p_{m,t} \simeq u_{m,t}^{+} h_\infty^{+} + u_{m,t}^{-} h_\infty^{-}$. Under complete steady state conditions $p_{m,\infty} \simeq u_{m,\infty}^{+} h_\infty^{+} + u_{m,\infty}^{-} h_\infty^{-}$ and one obtains the following bimodal Poisson type expression.

$$p_{m,\infty} \simeq \left[\left(k_r^{+}/\gamma_r\right)^m \beta \exp\left(-k_r^{+}/\gamma_r\right) + \left(k_r^{-}/\gamma_r\right)^m \alpha \exp\left(-k_r^{-}/\gamma_r\right)\right] \Big/ (\alpha+\beta) m! \tag{B10}$$

When the transcription rate associated with the off-state channel tends toward zero, then one recovers the Poisson density function with zero spike [14] as follows.

$$\lim_{k_r^- \to 0} p_{m,\infty} \simeq \left[\alpha/(\alpha+\beta)\right] + \left(k_r^{+}/\gamma_r\right)^m \beta \exp\left(-k_r^{+}/\gamma_r\right) \Big/ (\alpha+\beta) m! \tag{B11}$$

REFERENCES

[1] Alberts B 2002 *Molecular biology of the cell* (New York: Garland Science)

[2] Lewin B, Krebs J E, Kilpatrick S T, Goldstein E S and Lewin B 2011 *Lewin's genes X* (Sudbury, Mass.: Jones and Bartlett)

[3] Blake W J, M K A, Cantor C R and Collins J J 2003 Noise in eukaryotic gene expression *Nature* **422** 633-7

[4] Ozbudak E M, Thattai M, Kurtser I, Grossman A D and van Oudenaarden A 2002 Regulation of noise in the expression of a single gene *Nat Genet* **31** 69-73

[5] Fraser H B, Hirsh A E, Giaever G, Kumm J and Eisen M B 2004 Noise minimization in eukaryotic gene expression *PLoS Biol* **2** e137

[6] Becskei A, Kaufmann B B and van Oudenaarden A 2005 Contributions of low molecule number and chromosomal positioning to stochastic gene expression *Nat Genet* **37** 937-44

[7] Raj A and van Oudenaarden A 2008 Nature, nurture, or chance: stochastic gene expression and its consequences *Cell* **135** 216-26

[8] Elowitz M B, Levine A J, Siggia E D and Swain P S 2002 Stochastic gene expression in a single cell *Science* **297** 1183-6

[9] Mitarai N, Dodd I B, Crooks M T and Sneppen K 2008 The generation of promoter-mediated transcriptional noise in bacteria *PLoS Comput Biol* **4** e1000109

[10] McAdams H H and Arkin A 1997 Stochastic mechanisms in gene expression *Proc Natl Acad Sci U S A* **94** 814-9

[11] Paulsson J 2004 Summing up the noise in gene networks *Nature* **427** 415-8

[12] Golding I, Paulsson J, Zawilski S M and Cox E C 2005 Real-time kinetics of gene activity in individual bacteria *Cell* **123** 1025-36

[13] Corrigan A M, Tunnacliffe E, Cannon D and Chubb J R 2016 A continuum model of transcriptional bursting *Elife* **5**

[14] Chong S, Chen C, Ge H and Xie X S 2014 Mechanism of transcriptional bursting in bacteria *Cell* **158** 314-26

[15] Shahrezaei V and Swain P S 2008 Analytical distributions for stochastic gene expression *Proc Natl Acad Sci U S A* **105** 17256-61

[16] Kumar N, Singh A and Kulkarni R V 2015 Transcriptional Bursting in Gene Expression: Analytical Results for General Stochastic Models *PLoS Comput Biol* **11** e1004292

[17] Shahrezaei V, Ollivier J F and Swain P S 2008 Colored extrinsic fluctuations and stochastic gene expression *Mol Syst Biol* **4** 196

[18] Raj A, Peskin C S, Tranchina D, Vargas D Y and Tyagi S 2006 Stochastic mRNA Synthesis in Mammalian Cells *PLOS Biology* **4** e309

[19] Klindziuk A and Kolomeisky A B 2018 Theoretical Investigation of Transcriptional Bursting: A Multistate Approach *The Journal of Physical Chemistry B* **122** 11969-77

[20] So L H, Ghosh A, Zong C, Sepulveda L A, Segev R and Golding I 2011 General properties of transcriptional time series in Escherichia coli *Nat Genet* **43** 554-60

[21] Fujita K, Iwaki M and Yanagida T 2016 Transcriptional bursting is intrinsically caused by interplay between RNA polymerases on DNA *Nat Commun* **7** 13788

[22] Pulkkinen O and Metzler R 2013 Distance matters: the impact of gene proximity in bacterial gene regulation *Phys Rev Lett* **110** 198101

[23] Levens D and Larson D R 2014 A new twist on transcriptional bursting *Cell* **158** 241-2

[24] Fukaya T, Lim B and Levine M 2016 Enhancer Control of Transcriptional Bursting *Cell* **166** 358-68

[25] Tunnacliffe E, Corrigan A M and Chubb J R 2018 Promoter-mediated diversification of transcriptional bursting dynamics following gene duplication *Proc Natl Acad Sci U S A* **115** 8364-9

[26] Ma J and Wang M D 2016 DNA supercoiling during transcription *Biophys Rev* **8** 75-87

[27] Brackley C A, Johnson J, Bentivoglio A, Corless S, Gilbert N, Gonnella G and Marenduzzo D 2016 Stochastic Model of Supercoiling-Dependent Transcription *Phys Rev Lett* **117** 018101
[28] Gerasimova N S, Pestov N A, Kulaeva O I, Clark D J and Studitsky V M 2016 Transcription-induced DNA supercoiling: New roles of intranucleosomal DNA loops in DNA repair and transcription *Transcription* **7** 91-5
[29] Kim S, Beltran B, Irnov I and Jacobs-Wagner C 2019 Long-Distance Cooperative and Antagonistic RNA Polymerase Dynamics via DNA Supercoiling *Cell* **179** 106-19 e16
[30] Tantale K, Mueller F, Kozulic-Pirher A, Lesne A, Victor J M, Robert M C, Capozi S, Chouaib R, Backer V, Mateos-Langerak J, Darzacq X, Zimmer C, Basyuk E and Bertrand E 2016 A single-molecule view of transcription reveals convoys of RNA polymerases and multi-scale bursting *Nat Commun* **7** 12248
[31] Heberling T, Davis L, Gedeon J, Morgan C and Gedeon T 2016 A Mechanistic Model for Cooperative Behavior of Co-transcribing RNA Polymerases *PLoS Comput Biol* **12** e1005069
[32] Jonkers I and Lis J T 2015 Getting up to speed with transcription elongation by RNA polymerase II *Nat Rev Mol Cell Biol* **16** 167-77
[33] Maiuri P, Knezevich A, De Marco A, Mazza D, Kula A, McNally J G and Marcello A 2011 Fast transcription rates of RNA polymerase II in human cells *EMBO Rep* **12** 1280-5
[34] Smith S and Grima R 2018 Single-cell variability in multicellular organisms *Nat Commun* **9** 345
[35] Peccoud J and Ycart B 1995 Markovian Modeling of Gene-Product Synthesis *Theoretical Population Biology* **48** 222-34
[36] Iyer-Biswas S, Hayot F and Jayaprakash C 2009 Stochasticity of gene products from transcriptional pulsing *Phys Rev E Stat Nonlin Soft Matter Phys* **79** 031911
[37] Gardiner C W 1985 *Handbook of stochastic methods for physics, chemistry, and the natural sciences* (Berlin; New York: Springer-Verlag)
[38] Kampen N G v 1981 *Stochastic processes in physics and chemistry* (Amsterdam; New York; New York: North-Holland ; Sole distributors for the USA and Canada, Elsevier North-Holland)
[39] Risken H 1989 *The Fokker-Planck equation : methods of solution and applications* (Berlin; New York: Springer-Verlag)
[40] Whittaker E T and Watson G N 1969 *A course of modern analysis* (Cambridge: University Press)
[41] Abramowitz M and Stegun I A 1965 *Handbook of mathematical functions, with formulas, graphs, and mathematical tables* (New York: Dover Publications)
[42] Gillespie D T 2007 Stochastic simulation of chemical kinetics *Annu Rev Phys Chem* **58** 35-55
[43] Gillespie D T 1977 Exact stochastic simulation of coupled chemical reactions *The Journal of Physical Chemistry* **81** 2340-61
[44] Murugan R 2006 Stochastic transcription initiation: Time dependent transcription rates *Biophys Chem* **121** 51-6
[45] Wang Y, Ni T, Wang W and Liu F 2018 Gene transcription in bursting: a unified mode for realizing accuracy and stochasticity *Biol Rev Camb Philos Soc*
[46] Munsky B, Neuert G and van Oudenaarden A 2012 Using gene expression noise to understand gene regulation *Science* **336** 183-7
[47] Dobrzynski M and Bruggeman F J 2009 Elongation dynamics shape bursty transcription and translation *Proc Natl Acad Sci U S A* **106** 2583-8
[48] Zwanzig R 1990 Rate processes with dynamical disorder *Accounts of Chemical Research* **23** 148-52
[49] Zwanzig R 1992 Dynamical disorder: Passage through a fluctuating bottleneck *The Journal of Chemical Physics* **97** 3587-9

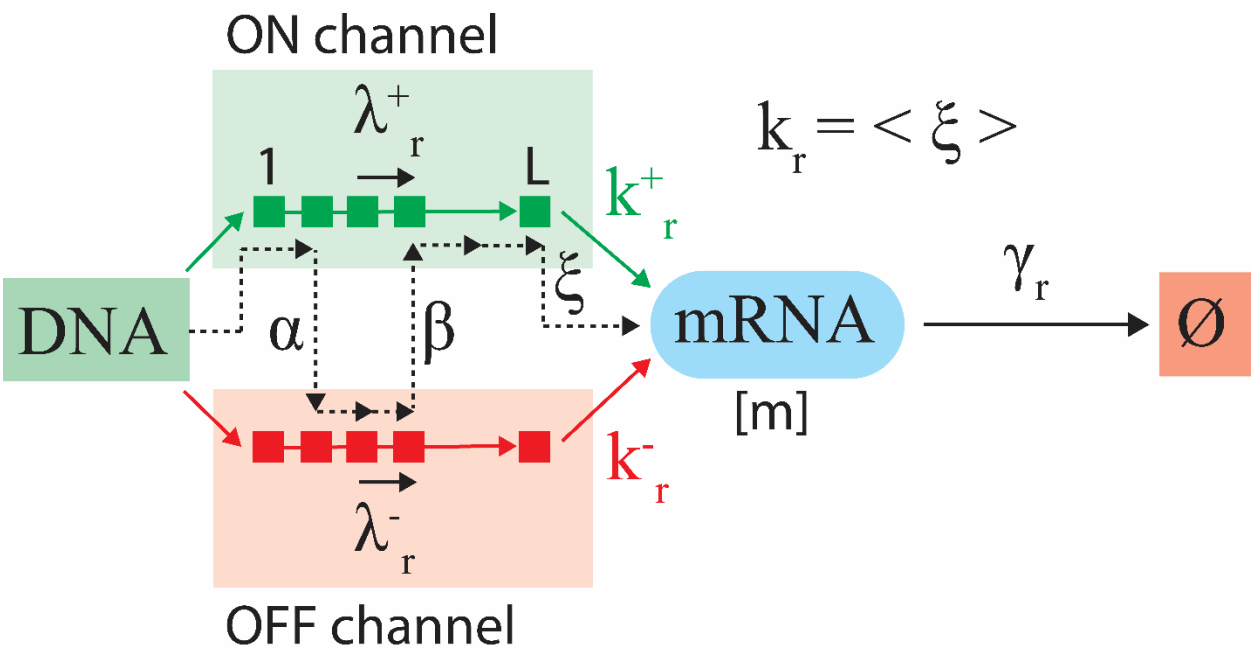

FIGURE 1. Stochastic transcription rate and transcription bursting. In this generalized model, the mRNA transcript of size L bp can be generated via either pure on or off state channels with rates k_r^+ and k_r^- respectively. The microscopic transition rates $\lambda_r^\pm$ characterize the addition of individual nucleotides at the end of growing mRNA in the process of transcription elongation. The resultant transcription rate will be the inverse of the first passage time required to generate a complete mRNA. When there is a flipping across these on-off states then the resultant transcription rate ξ fluctuate across k_r^+ and k_r^- in a random manner where $k_r^- \ll k_r^+$ in general. The ensemble average of ξ across several trajectories of mRNA synthesis will be the average transcription rate $k_r = \langle \xi \rangle$. Here (+) denote the on-state and (-) denotes the off-state of transcription and γ_r is the decay rate constant of mRNAs. The rate of flipping from the on-state to the off state is α and the rate of flipping from the off-state to the on-state in β. The mesoscopic transcription rates $k_r^\pm$ are connected to the microscopic ones $\lambda_r^\pm$ via the relationship $k_r^\pm = \lambda_r^\pm / L$. As a result, flipping across on-off states, the overall effective transcription rate k_r will be somewhere in between $\left(k_r^-, k_r^+\right)$. Here the dotted line is a stochastic transcription trajectory which varies from transcript to transcript. This means that the effective transcription rate ξ (it is not a constant anymore) varies from transcript to transcript and it is a stochastic quantity.

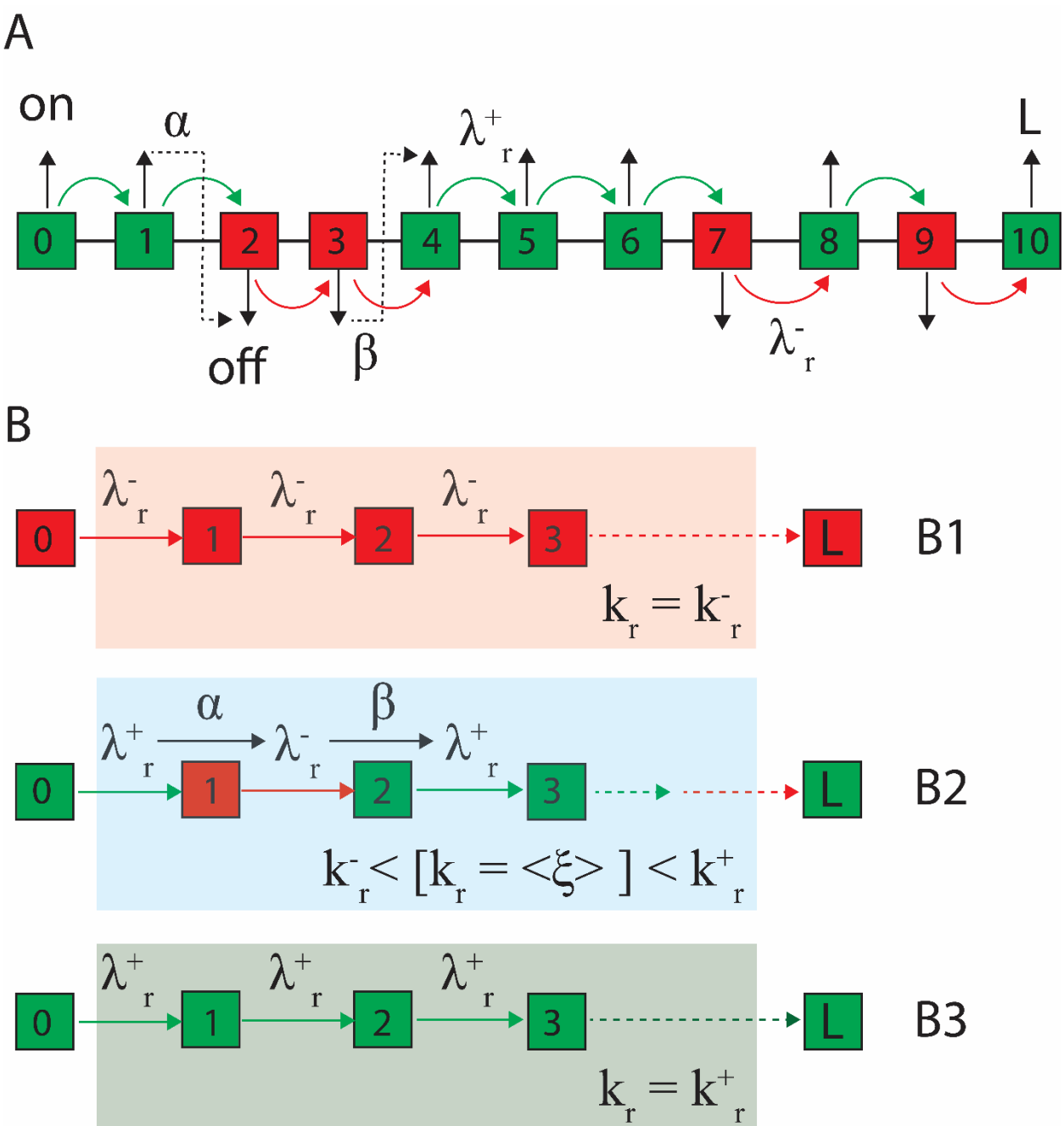

FIGURE 2. Stochasticity in the transcription rate. **A**. In this model the total length of the full transcript is L = 10 bp. Green color denotes the on-state and red color denotes the off-state channels. At time t = 0, the system was in the on-state. Green to green transition is characterized by the microscopic transcription elongation rate λ_r^+ and red to red transition is characterized by the elongation rate λ_r^-. Here $\lambda_r^\pm$ are measured in bp/s. **B**. When the entire transcription process follows the pure off-state channel (**B1**), then the overall transcription rate scales as $k_r = k_r^- = \lambda_r^- / L$ that is

measured in s^{-1}. When there is a flipping across the on- and off-state channels (**B2**), then the transcription rate of a given mRNA trajectory will be somewhere inside as $k_r^- < \xi < k_r^+$. For a stall-continue type elongation of RNA polymerase, it approaches the limit $\lim_{\lambda_r^- \to 0} k_r \simeq (\alpha+\beta)\lambda_r^+ / \beta L$ as shown in **Appendix A** where $k_r = \langle \xi \rangle$. When the entire transcription process follows the pure on-state channel (**B3**), then the overall transcription rate can be expressed as $k_r = k_r^+ = \lambda_r^+ / L$.

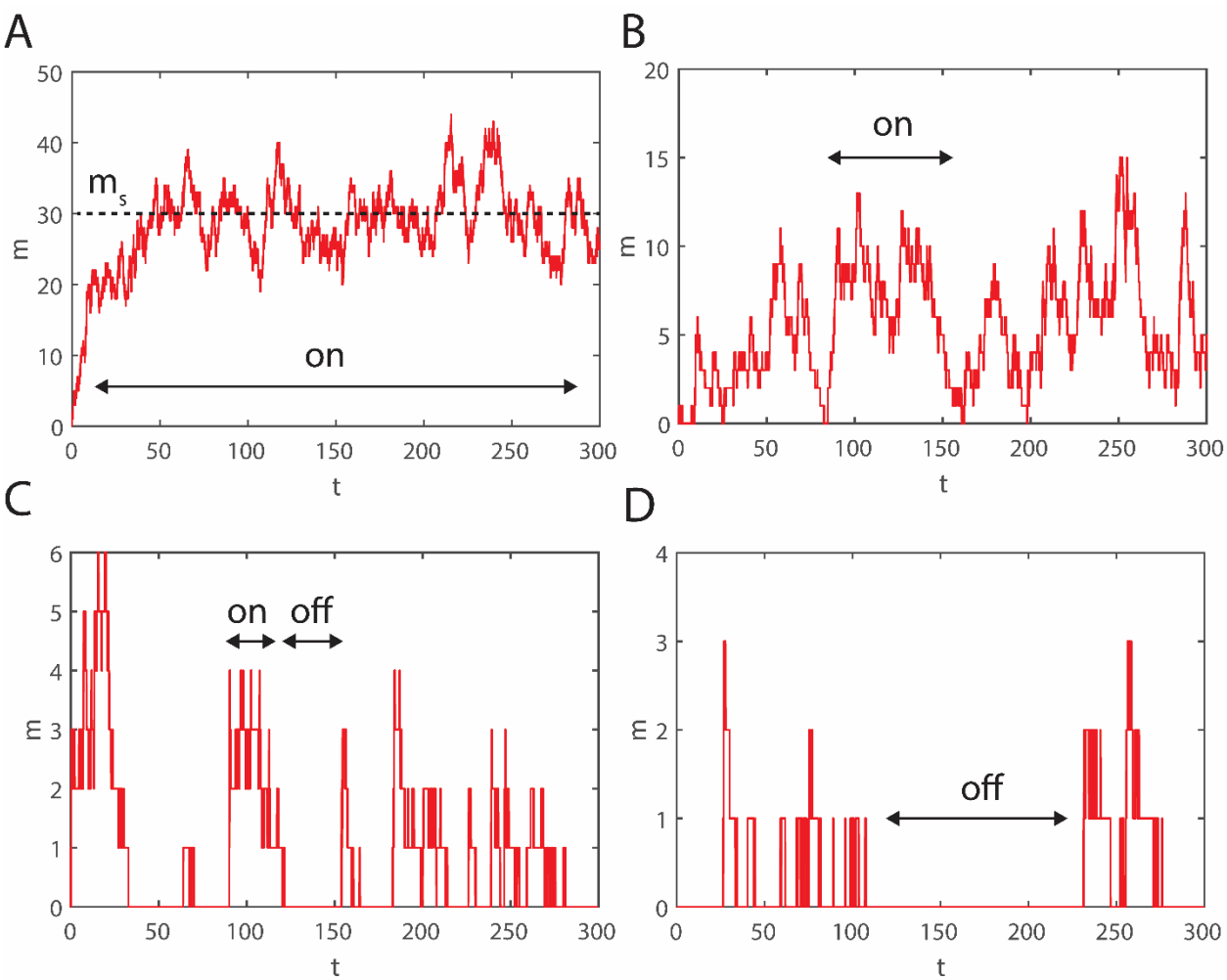

FIGURE 3. Emergence of the transcription bursting phenomenon. These are all single stochastic trajectories. **A**. Continuous transcription process. Here the parameters settings are $\alpha = \beta = 10^{-6}$ s^{-1}, $k_r^+ = 3$ s^{-1}, $\gamma_r = 0.1$ s^{-1}, and $k_r^- = 10^{-5}$ s^{-1}. In these settings, one finds that $k_r \simeq k_r^+$ and therefore the steady state mRNA numbers will be $m_s \simeq k_r^+ / \gamma_r \sim 30$. Initially the system was set into the on-state of transcription. **B**. Transcription bursting emerges when one sets high value for α and low value for β apart from the mandatory condition that $k_r^- \ll k_r^+$. Here $\alpha = 3$, $\beta = 1$ and rest of the parameters are set as in in the panel **A**. In **C**. $\beta = 0.1$ and in **D**. $\beta = 0.05$ and rest of the other parameters are set as in the panel **A**.

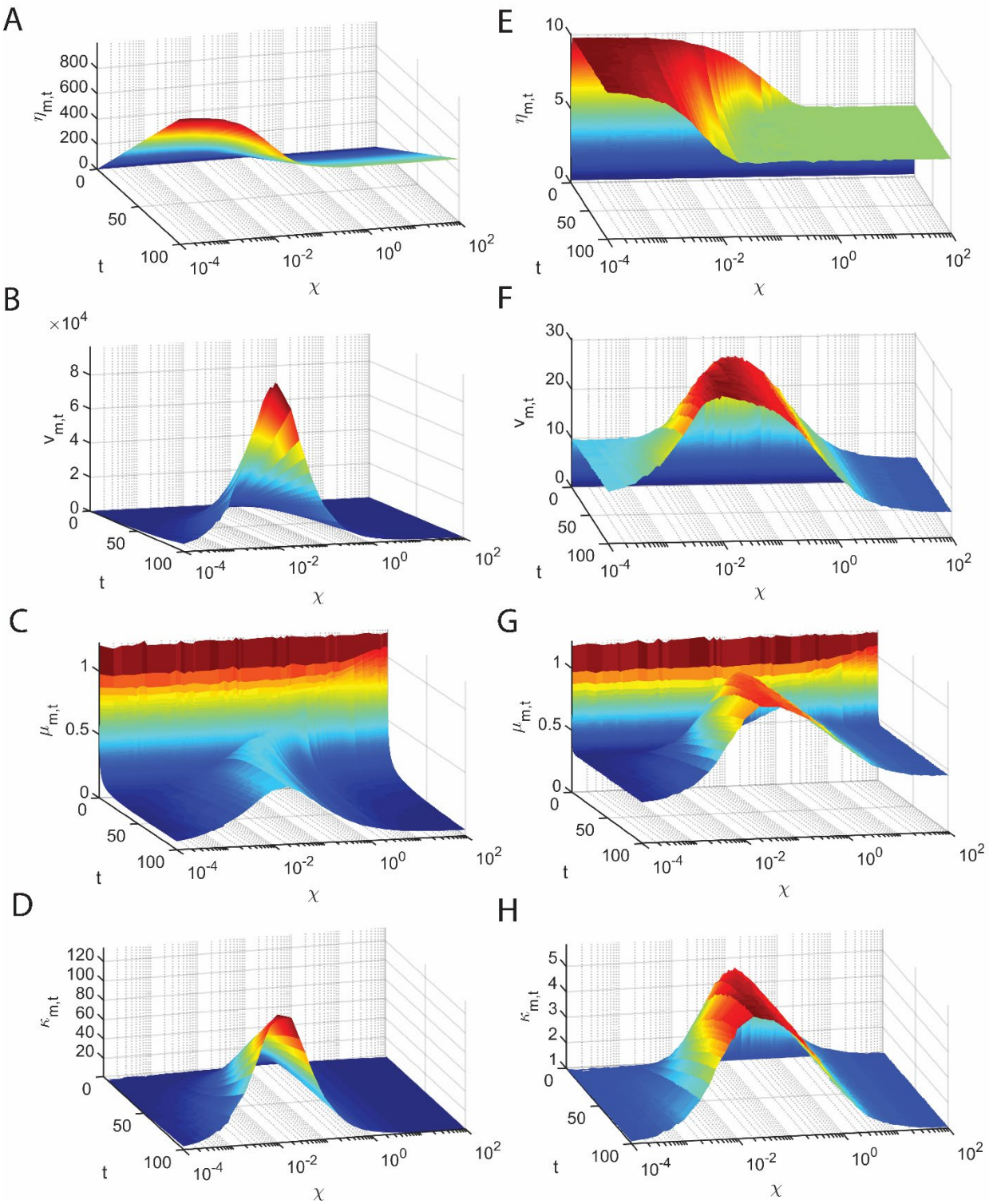

FIGURE 4. Simulation results on various time dependent statistical properties of mRNA number fluctuations in the pre-steady state regime of the transcription event. Here we have set $\alpha = \beta = \chi$ for simplification, $\eta_{m,t}$ denotes the mean, $\kappa_{m,t}$ denotes the Fano factor, $\mu_{m,t}$ denotes the coefficient

of variation, and $v_{m,t}$ denotes the variance. In the panels **A-D** the parameter settings are $k_r^+ = 10$, $k_r^- = 0$ and $\gamma_r = 0$. In line with the predictions of **Eqs. 29** and **32** one observes maxima in the Fano factor and the variance with respect to the flipping rate χ which is also a time dependent quantity. As predicted by these equations, the optimum flipping rate shifts towards zero as time tends towards infinity. In the panels **E-H** the parameter settings are $k_r^+ = 10$, $k_r^- = 0$ and $\gamma_r = 1$. Statistical properties were computed over 10^6 number of trajectories.

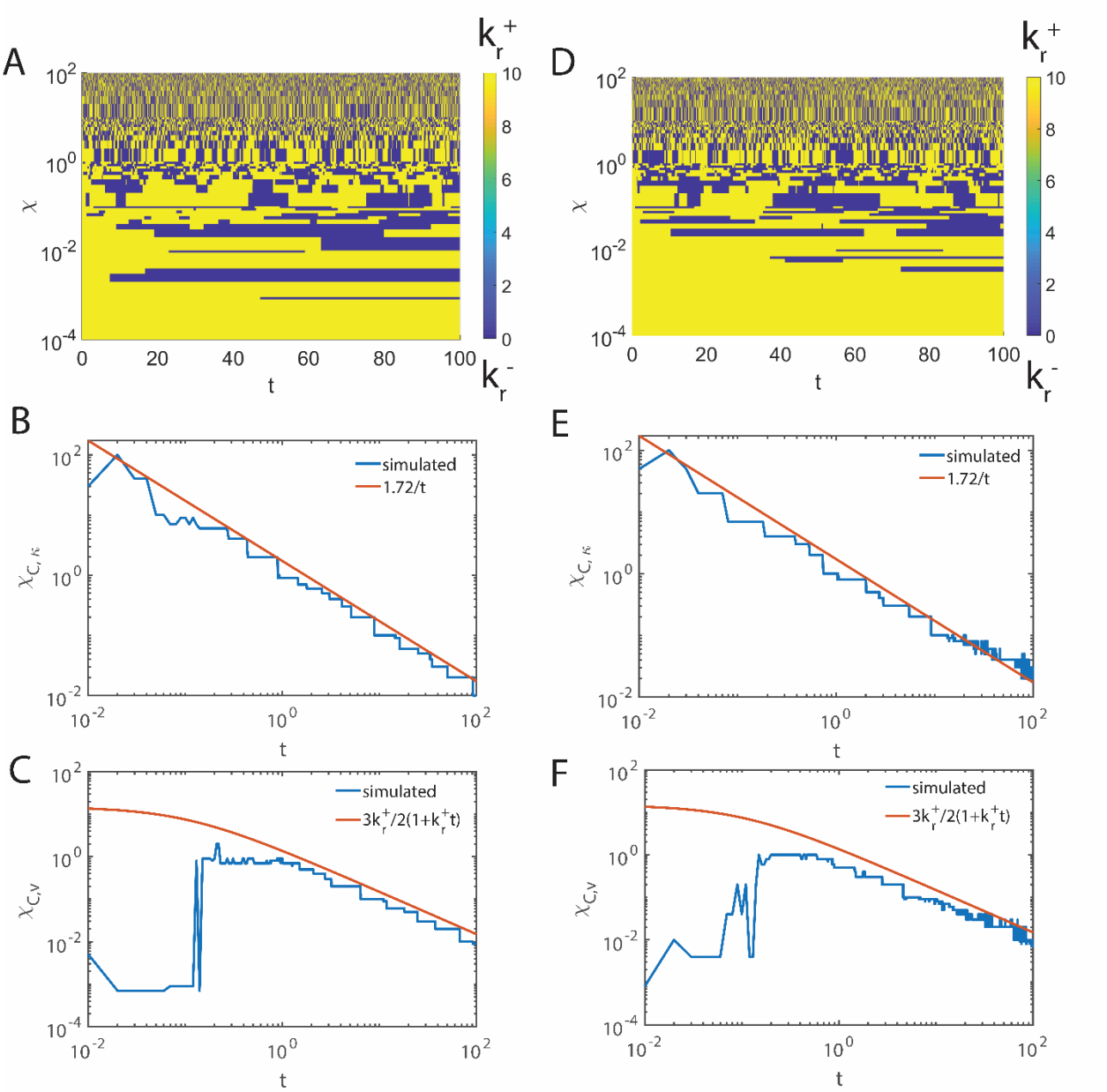

FIGURE 5. Variation of the optimum flipping rate χ that maximizes the variance and the Fano factor of mRNA numbers. In panels **A-C** the settings are $k_r^+ = 10$, $k_r^- = 0$ and $\gamma_r = 0$. In panels **D-F** the settings are $k_r^+ = 10$, $k_r^- = 0$ and $\gamma_r = 1$. Panels **A** and **D** represent the flipping dynamics across the on and off states. Panels **B** and **E** represent the variation of the optimum flipping rate with respect to time which maximize the Fano factor associated with the mRNA number fluctuations. Panels **C** and **F** represents the variation of the optimum flipping rate with respect to time which maximizes the variance associated with the mRNA number fluctuations. Here the solid red lines are the predictions by **Eqs. 29** and **32** which are valid when $\alpha = \beta = \chi$, $k_r^- = 0$ and $\gamma_r = 0$. In panels **C** and **F**, the precision in numerically computing the maximum points is lost when $t < 0.1$ since the surface is almost flat. Statistical properties were computed over 10^6 number of trajectories.

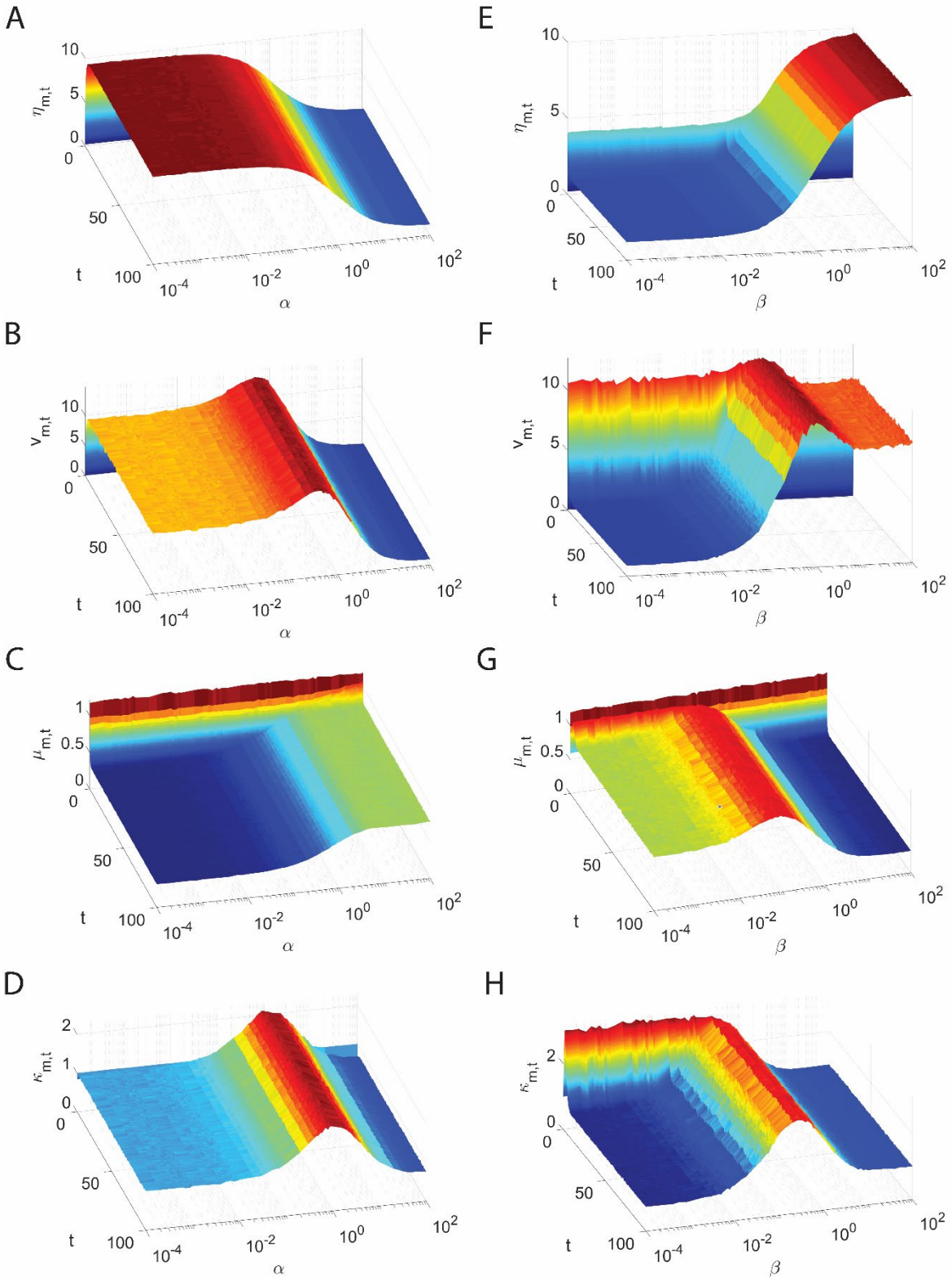

FIGURE 6. Various time dependent statistical properties of mRNA fluctuations in the pre-steady state regime of the transcription event. Here we have set $\alpha \neq \beta$, $\eta_{m,t}$ denotes the mean, $\kappa_{m,t}$ denotes the Fano factor, $\mu_{m,t}$ denotes the coefficient of variation, and $v_{m,t}$ denotes the variance of the mRNA

number fluctuations. In the panels **A-D** the settings are $k_r^+ = 10$, $k_r^- = 1$, $\gamma_r = 1$ and $\beta = 1$. In the panels **E-H** the settings are $k_r^+ = 10$, $k_r^- = 1$, $\gamma_r = 1$ and $\alpha = 1$. In line with the predictions of **Eqs. 29** and **32** one observes maxima in the Fano factor and the variance even when $\alpha \neq \beta$. Statistical properties were computed over 10^6 number of trajectories.

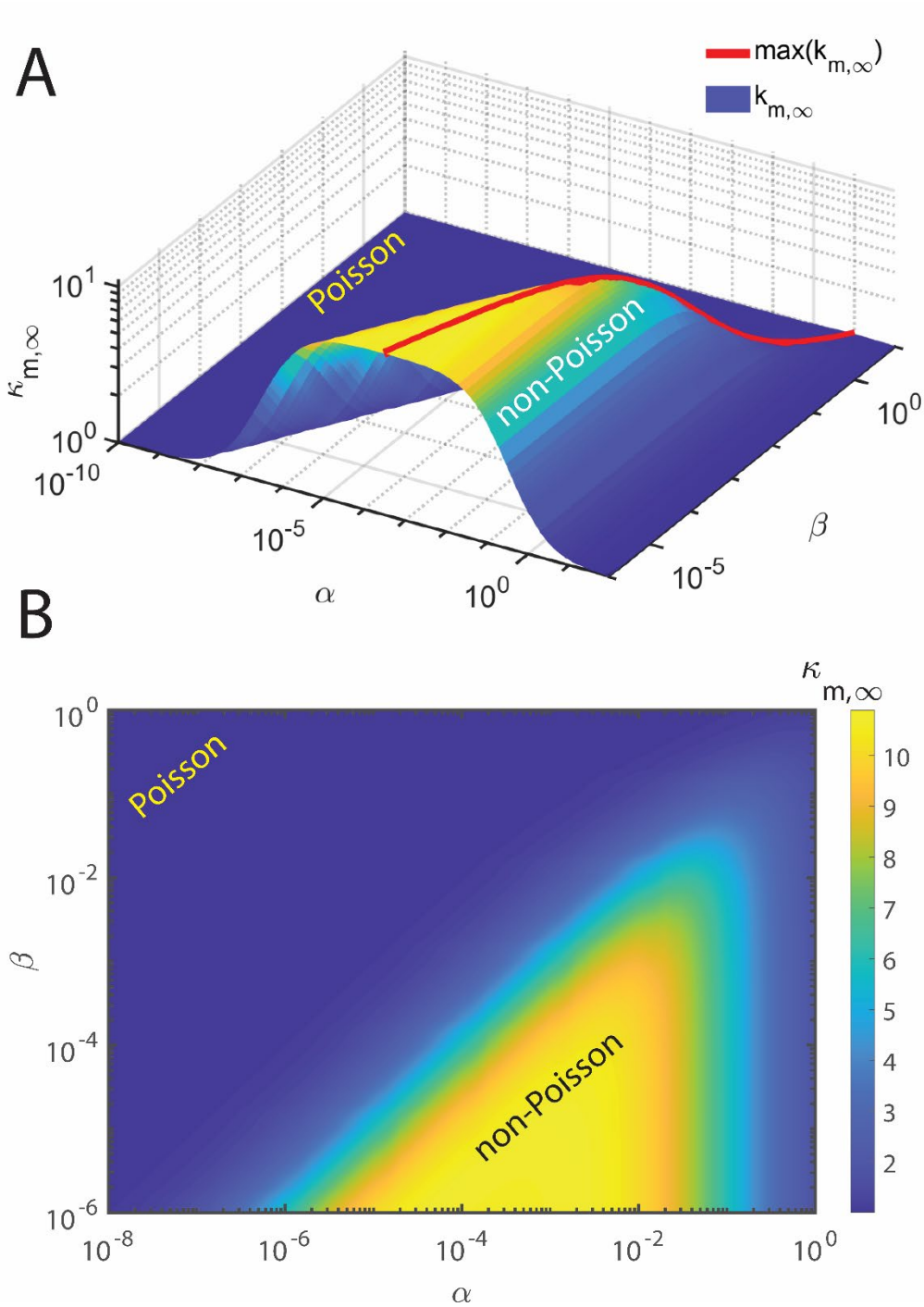

FIGURE 7. Variation of the steady state Fano factor $\kappa_{m,\infty}$ associated with the mRNA number fluctuations with respect to changes in the flipping rates α and β in the limit $\lim_{k_r^- \to 0} \kappa_{m,\infty}$ as given in **Eqs. 63**. Here the settings in panels **A** and **B** are $k_r^+ = 1$, $k_r^- = 0$ and $\gamma_r = 0.1$. Panel **B** is the contour plot of panel **A**. The optimum values of α at which $\kappa_{m,\infty}$ attains a maximum can be expressed as $\alpha_C = \sqrt{\beta(\beta + \gamma_r)}$. This can be obtained by solving $\left[\partial \kappa_{m,\infty} / \partial \alpha\right] = 0$ for α. Red solid

line in panel **A** is the maximum value of the steady state Fano factor which can be obtained by substituting $\alpha = \alpha_C$ in the expression of $\kappa_{m,\infty}$ as in **Eqs. 63**.

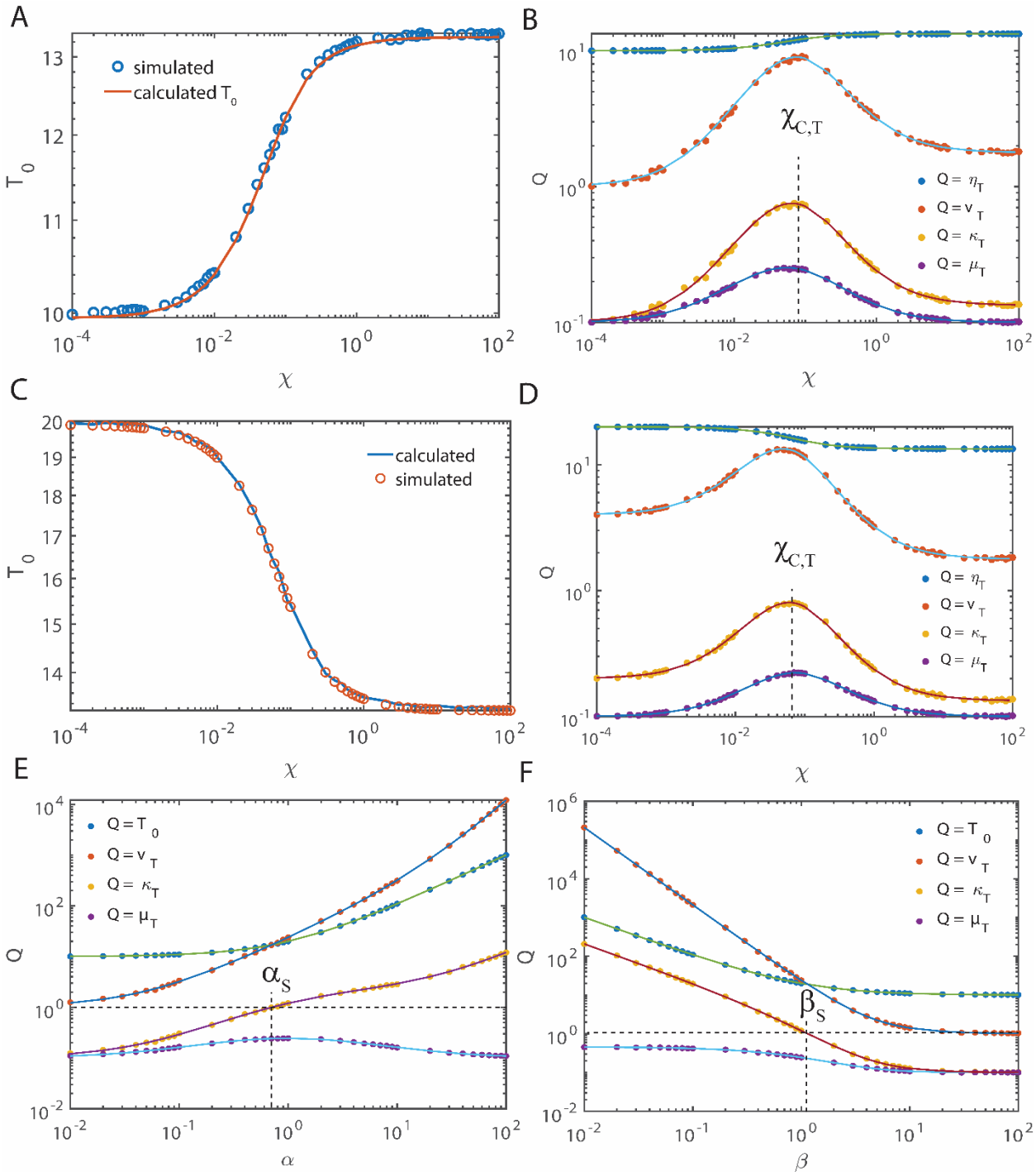

FIGURE 8. Dependence of the overall average transcription time on the on-off flipping rates. Here T_0 is the mean first passage time associated with the formation of a complete mRNA transcript with size $n = L$ bp starting from $n = 0$ bp where n is denotes dynamic number of transcribed bases. Filled or hollow circles are the simulated results and solid lines are the

theoretical predictions computed using **Eqs. A3** (panels **E**, **F**) and **A7** (panels **A-D**) of **Appendix A**. By definition $k_r = [1/T_0]$ is the transcription rate. T_0 was calculated using **Eqs. 39**. Hollow blue and red circles are the stochastic simulation results. In panels **A-D** the settings are $\lambda_r^+ = 10$ bp/s, $\lambda_r^- = 5$ bp/s, $\alpha = \beta = \chi$ and L = 100 bp with $q_{n,0}^+ = \delta(n)$ and $q_{n,0}^- = 0$ in **A** and **B**. In **C** and **D** $q_{n,0}^- = \delta(n)$; $q_{n,0}^+ = 0$. Here $q_{n,t}^\pm$ is the probability of finding n number transcribed bases in the respective on (+) and off (-) state channels. When the flipping rate becomes sufficiently large, then one finds the limiting value as $\lim_{\chi\to\infty} T_0 \simeq 2L/(\lambda_r^+ + \lambda_r^-)$. Panels **B** and **D** demonstrate the variation of the mean (η_T), variance (ν_T), coefficient of variation (μ_T) and Fano factor (κ_T) of the first passage times associated with the generation of a complete mRNA transcript with respect to the flipping rate parameter χ. Clearly, there exists an optimum flipping rate $\chi_{C,T}$ at which the variance, Fano factor and coefficient of variation of the first passage times attain a maximum. Panels **E** and **F** demonstrate behavior of the overall MFPT when $\alpha \neq \beta$ and $\lambda_r^- \to 0$. In this situation, we find $\lim_{\lambda_r^-\to 0} T_0 \simeq (\alpha+\beta)L/\beta\lambda_r^+$. Settings in **E** and **F** are $\lambda_r^+ = 10$, $\lambda_r^- = 10^{-4}$ and L = 100. In panel **E**, $\beta = 1$ and in panel **F**, $\alpha = 1$ with initial conditions $q_{n,0}^+ = \delta(n)$; $q_{n,0}^- = 0$. Statistical properties of FPTs were computed over 10^6 number of trajectories. With these settings one finds that $\beta_S \sim 1.15$ and $\alpha_S \sim 0.7$ in line with the prediction by **Eqs. 42**.

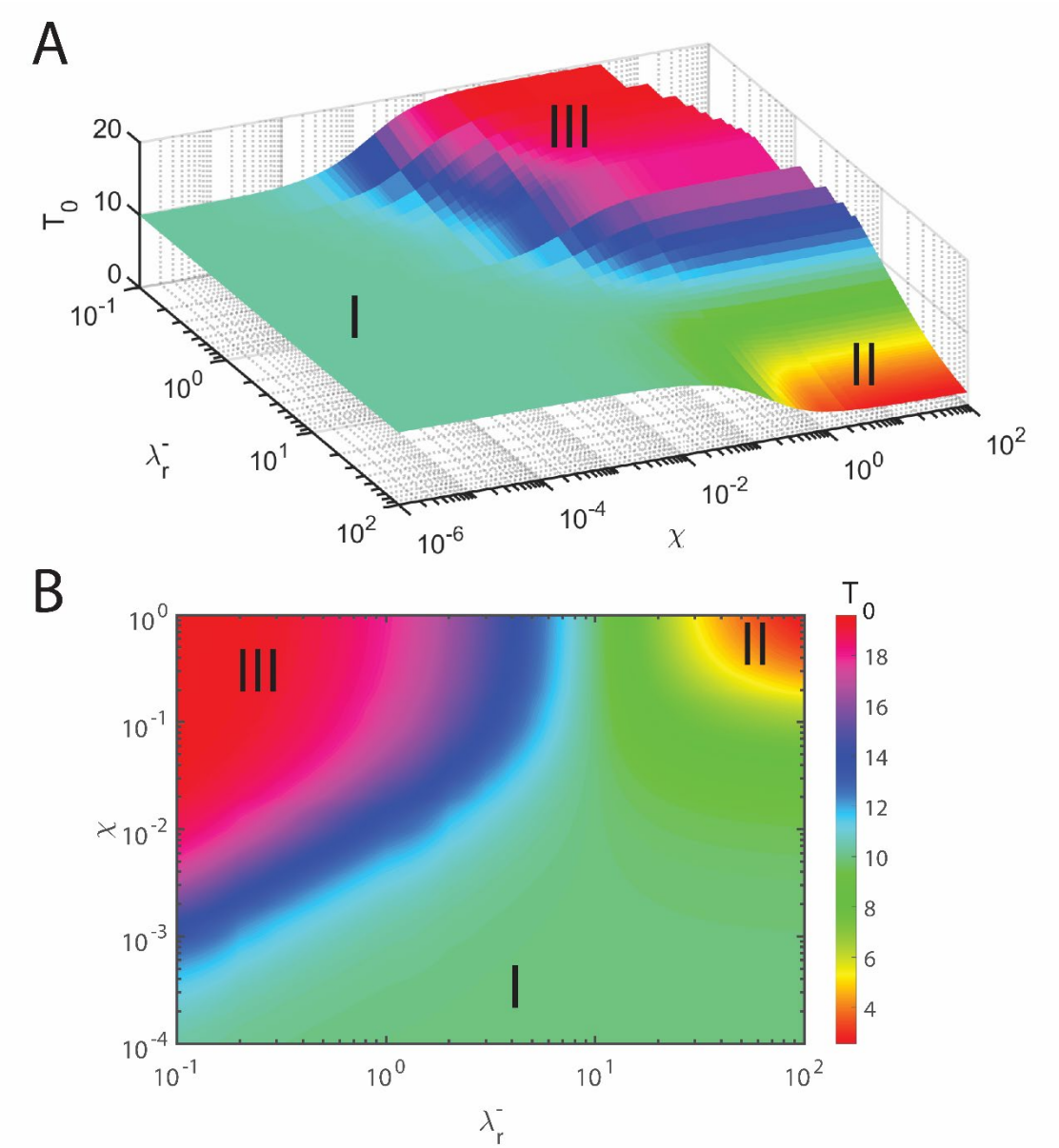

FIGURE 9. Various limiting values of the mean first passage time (MFPT, T_0) associated with the generation of a complete mRNA transcript of size L bp. MFPT was computed using **Eqs. 44**. Here the settings in panel **A** are $\lambda_r^+ = 10$ bp/s and L = 100 bp. When $\lambda_r^- \to 0$ and $\chi \to 0$, then one finds that $T_0 \to L/\lambda_r^+$ = 10 s (**I**). When $\lambda_r^- \to 0$ with finite χ, then one finds that $T_0 \to 2L/\lambda_r^+$ = 20 s (**III**). Hypothetically, when $\lambda_r^- \geq \lambda_r^+$ and $\chi \to \infty$, then one finds that $T_0 \to 2L/\left(\lambda_r^+ + \lambda_r^-\right)$ (**II**). Panel **B** is the contour plot of panel **A**.

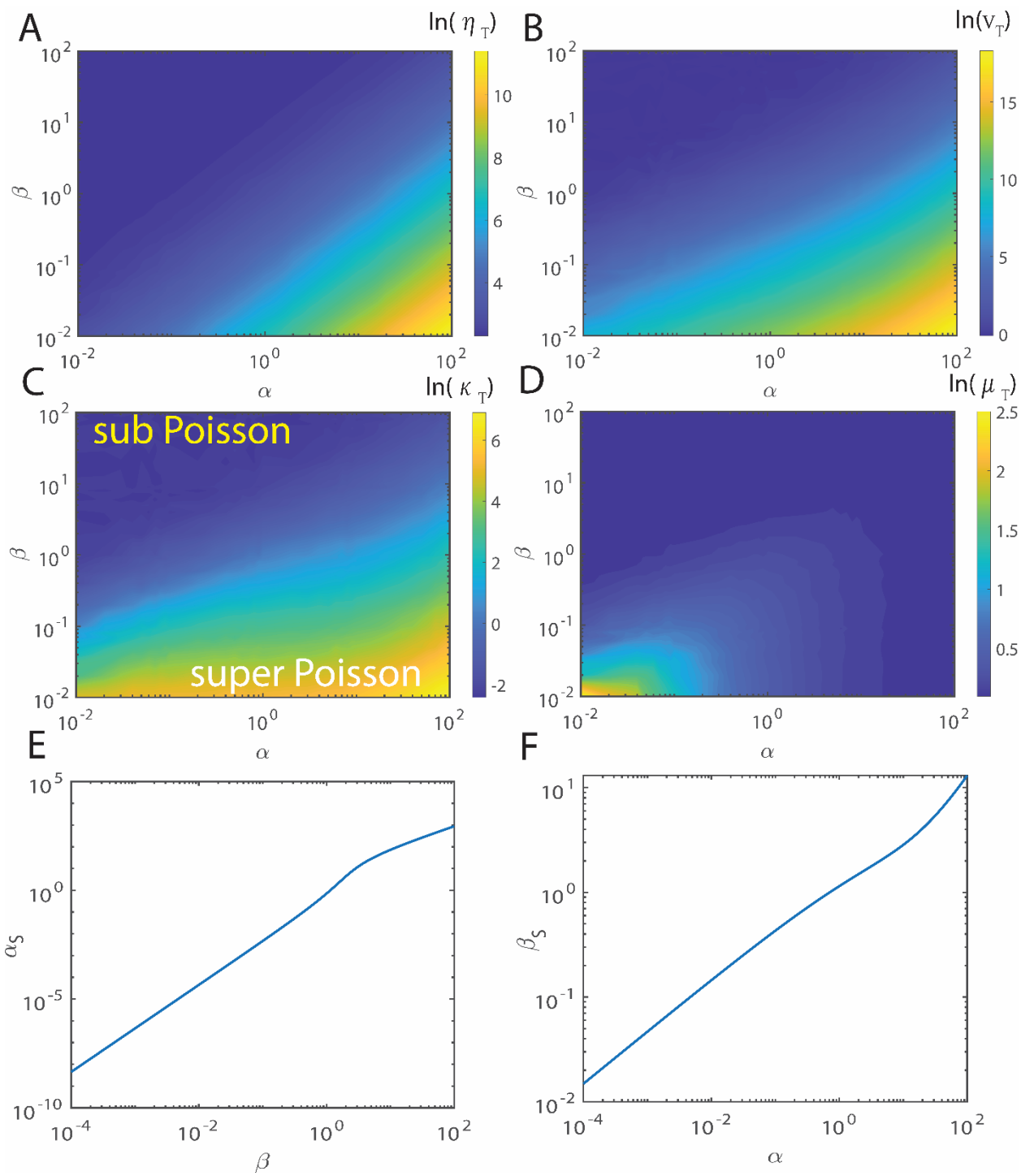

FIGURE 10. Statistical properties of the first passage times (FPTs) associated with the generation of a complete mRNA transcript of size $n = L$ bp starting from $n = 0$. Settings are, $\lambda_r^+ = 10$ bp/s is the elongation rate in the on channel and $L = 100$ bp and $\lambda_r^- = 10^{-4}$ is the elongation rate of the off channel. As shown in **Appendix A**, one finds the limiting value of the mean of FPTs (MFPT, T_0) as $\lim_{\lambda_r^- \to 0} T_0 \simeq (\beta + \alpha) L / \beta \lambda_r^+$. When $\alpha = 10^2$ and $\beta = 10^{-2}$, then one finds the limiting value as

$\lim_{\lambda_r^- \to 0} \ln T_0 \simeq 11.5$. For $\alpha = 10^{-2}$ and $\beta = 10^2$, one finds that $\lim_{\lambda_r^- \to 0} \ln T_0 \simeq 2.3$ and so on. **A**. mean of FPTs (η_T). **B**. Variance of FPTs (v_T). **C**. Fano factor of FPTs (κ_T). **D**. coefficient of variation of FPTs (μ_T). Statistical properties of FPTs were computed over 10^6 number of trajectories. In **E** and **F** (α_S, β_S) are the critical values of the flipping rates such that when $\alpha < \alpha_S$ or $\beta < \beta_S$, then the distribution of the FPTs becomes sub-Poisson type. When $\alpha > \alpha_S$ or $\beta > \beta_S$, then the distribution of FPTs becomes super-Poisson as described in **Eqs. 42**. Here $\lambda_r^+ = 10$ bp/s. **E**. Variation of α_S with respect to β in the limit $\lambda_r^- \to 0$. **F**. Variation of β_S with respect to α in the limit $\lambda_r^- \to 0$.

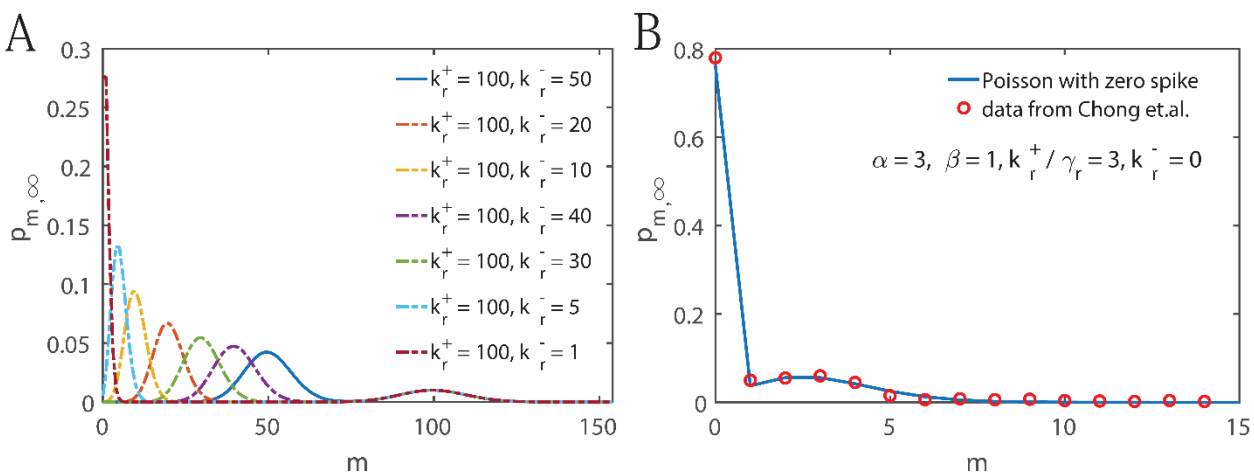

FIGURE 11. Emergence of the bimodal distribution functions associated with the mRNA number fluctuations. In panel **A** the settings are $\alpha = 3$, $\beta = 1$ and $\gamma_r = 1$. In panel **B** the settings are $\alpha = 3$, and $\beta = 1$. This represents the Poisson density function with zero spike as derived in **Eq. 78**. Hollow red circles are the digitized data points from Ref. [14].

Table 1. List of symbols and variables used in the main text

Parameter	Definition	Remarks
MFPT	Mean first passage time. It is the average time required to generate a full mRNA transcript with a given elongation speed of RNA polymerase.	s
FPT	First passage time. It is the time required to generate a given mRNA transcript. Clearly this quantity will vary from transcript to transcript.	s
m	Concentration of the mRNA molecules or the number of mRNA molecules.	M, or number
RNAP	RNA polymerase complex	
Fano Factor	= variance / mean.	In the units of mean.
CV	Coefficient of variation = standard deviation / mean.	dimensionless
k_r	Overall average or effective transcription rate.	M/s or 1/s
k_r^+	Transcription rate in the on-state channel of transcription.	M/s or 1/s
k_r^-	Transcription rate in the off-state channel of transcription.	M/s or 1/s
L	Length of the complete mRNA transcript.	bp
λ_r^+	Microscopic transcription elongation rate constant associated with the addition of single bp with already emerging mRNA in the on-state channel. $k_r^+ = \lambda_r^+ / L$.	bp/s
λ_r^-	Microscopic transcription elongation rate constant associated with the addition of single bp with already emerging mRNA in the off-state channel. $k_r^- = \lambda_r^- / L$.	bp/s
γ_r	Decay rate constant associated with the mRNA molecules. By definition $1/\gamma_r$ is the lifetime of mRNA.	1/s
α	Rate of flipping from the on-state to off-state channel of the transcription.	1/s
β	Rate of flipping from the off-state to the on-state channel of the transcription.	1/s
χ	When $\alpha = \beta$ then $\alpha = \beta = \chi$ is the rate of flipping across on-off states of transcription.	1/s
$\chi_{C,v}$	Optimum on-off flipping rate at which the variance associated with the mRNA number fluctuations attains a maximum. When $k_r^- = 0$ and $\gamma_r = 0$, then one finds that $\chi_{C,v} \simeq 3k_r^+ / 2\left(1 + k_r^+ t\right)$.	1/s
$\chi_{C,\kappa}$	Optimum on-off flipping rate at which the Fano factor associated with the mRNA number fluctuations attains a maximum. When $k_r^- = 0$ and $\gamma_r = 0$, then one finds that $\chi_{C,\kappa} \simeq 1.72/t$.	1/s
σ	$= k_r / \alpha$, is the transcription efficiency or the burst size.	Molecules

σ_C	$= k_r / \alpha_C$, is the optimum transcription efficiency or the burst size at which the steady state Fano factor attains a maximum.	Molecules or dimensionless number
$\eta_{m,\,t}$	Mean value associated with the mRNA number fluctuations at time *t*.	M or number
$\mu_{m,\,t}$	Coefficient of variation (= variance / square of mean) associated with the mRNA number fluctuations at time *t*. $\mu_{m,t} = v_{m,t} / \eta_{m,t}^2$.	dimensionless
$\kappa_{m,\,t}$	Fano factor (= variance / mean) associated with the mRNA number fluctuations at time *t*. $\kappa_{m,t} = v_{m,t} / \eta_{m,t}$.	M or number
$v_{m,\,t}$	Variance associated with the mRNA fluctuations at time *t*.	M^2 or number
$\eta_{m,\,\infty}$	Mean associated with the mRNA number fluctuations at the steady state.	M or number
$\mu_{m,\,\infty}$	Coefficient of variation associated with the mRNA number fluctuations at the steady state. $\mu_{m,\infty} = v_{m,\infty} / \eta_{m,\infty}^2$.	dimensionless
$\kappa_{m,\,\infty}$	Fano factor associated with the mRNA number fluctuations at the steady state. $\kappa_{m,\infty} = v_{m,\infty} / \eta_{m,\infty}$.	M or number
$v_{m,\,\infty}$	Steady state variance associated with the mRNA numbers.	M^2 or number
α_C	When β is fixed, then the optimum value of α at which the stationary state Fano factor attains a maximum. This can be obtained by solving $\left[\partial \kappa_{m,\infty} / \partial \alpha\right] = 0$ for α. Explicitly, when $k_r^- = 0$ then one finds that $\alpha_C = \sqrt{\beta(\beta + \gamma_r)}$.	1/s
β_C	When α is fixed, then iteration over β shows an optimum point β_C at which the steady state Fano factor associated with the mRNA number fluctuations attain a maximum. Explicitly one finds that $\beta_C = \sqrt{\alpha^2 + \gamma_r^2/4} - \gamma_r/2$. Both these optimum flipping rates α_C, β_C are connected via $\alpha_C = \sqrt{\beta_C(\beta_C + \gamma_r)}$.	1/s
$p_{m,t}^{\pm}$	Probability of finding *m* number of mRNAs at time *t* in the respective on (+) and off (-) channels of transcription.	dimensionless
$p_{m,\infty}^{\pm}$	Probability of finding m number of mRNAs in the respective on (+) and off (-) channels of transcription at the steady state.	dimensionless
$p_{m,\,t}$	$= p_{m,t}^{+} + p_{m,t}^{-}$, Total probability of finding *m* number of mRNA molecules at time *t*.	dimensionless
$p_{m,\,\infty}$	$= p_{m,\infty}^{+} + p_{m,\infty}^{-}$, total stationary state probability of finding *m* number of mRNA molecules at time *t*.	dimensionless
T_0	Mean first passage time associated with the generation of a complete mRNA transcript starting from $n = 0$ number of bases to $n = L$ bp. By definition the overall transcription rate is $k_r = 1 / T_0$.	s

ξ	Transcription rate associated with an individual mRNA synthesis. It is the inverse of the FPT required to generate a complete mRNA transcript for a given elongation speed of RNA polymerase. Its ensemble average $k_r = \langle \xi \rangle$ is the overall average transcription rate.	M/s or 1/s
$\chi_{C,T}$	On-off flipping rate at which the Fano factor of FPTs associated with the generation of a complete mRNA transcript attains the maximum point.	1/s
$q^{\pm}_{n,t}$	Probability of finding *n* number of bases in the elongating mRNA at time *t* in the respective on (+) or off (-) channel of transcription.	dimensionless
$G^{\pm}_{s,t}$	$= \sum_{m=0}^{\infty} s^m p^{\pm}_{m,t}$, is the generating function associated with the probability density $p^{\pm}_{m,t}$.	dimensionless
R_0	Second moment of the distribution of FPTs required to generate a complete mRNA transcript starting from $n = 0$.	s^2
v_T	$= R_0 - T_0^2$, variance associated with the distribution of FPTs required to generate a complete mRNA transcript starting from $n = 0$.	s^2
κ_T	$= v_T / \eta_T$, Fano factor associated with the distribution of FPTs required to generate a complete mRNA transcript starting from $n = 0$.	s
μ_T	$= v_T / \eta_T^2$, coefficient of variation associated with the distribution of FPTs required to generate a complete mRNA transcript starting from $n = 0$.	dimensionless
η_T	$= T_0$, mean associated with the distribution of FPTs required to generate a complete mRNA transcript starting from $n = 0$.	s
α_S, β_S	Critical values of the flipping rates (α, β) such that when $\alpha < \alpha_S$ or $\beta < \beta_S$, then the Fano factor κ_T associated with the distribution of FPTs becomes as $\kappa_T < 1$ (sub-Poisson type). When $\alpha > \alpha_S$ or $\beta > \beta_S$, then $\kappa_T > 1$ (super-Poisson type distribution). When $\alpha = \alpha_S$ or $\beta = \beta_S$, then $\kappa_T = 1$ as described in **Eqs. 42**.	1/s