

# First Passage Time Statistics of Stochastic Transcription Process for Time-dependent Reaction Rates

Kuheli Biswas, Mayank Shreshtha,<sup>\*</sup> Anudeep Surendran,<sup>†</sup> and Anandamohan Ghosh<sup>‡</sup>  
*Indian Institute of Science Education and Research Kolkata, Mohanpur, Nadia 741246, India*

Transcription in gene expression is an intrinsically noisy process which involves production and degradation of mRNAs. An important quantity to describe this stochastic process is the first passage time (FPT), i.e., the time taken by mRNAs to reach a particular threshold. The process of transcription can be modelled as a simple birth-death process, assuming that the promoter is always in an active state and to encode the stochastic environment we consider the transcription rate to be time-dependent. This generalization is suitable to capture bursty mRNA dynamics usually modelled as an ON-Off model and simplifies the calculation of FPT statistics for a cell population. We study the role of periodic modulation of the transcription rate on different moments of FPT distribution of a population of cells. Our calculation shows that for sinusoidal modulation there exists an extremal value of mean FPT as a function of time-period and phase of the transcription signal. However, for square-wave modulation of transcription rates simulation results shows that the extremal value of the MFPT behaves monotonically with the variation of the phase.

## I. INTRODUCTION

The process of gene expression is initiated with the copying of a segment of DNA into mRNA and this initial step is known as transcription. Transcription is an intrinsically stochastic process due to the occurrence of many probabilistic events like conformational changes in DNA, random binding of transcription factors and the random biochemical reactions which govern the process of transcription and translation inherently. In a single cell, mRNAs are often produced in small numbers fluctuating over time subsequently resulting in large variability across a population of cells [1–6]. With the development of experimental techniques to detect mRNA levels in single cells, namely bacteria [7, 8], mammalian cells [6], eukaryotic cells [9] a large number of recent experiments have been devoted to measure and visualize the transcriptome kinetics and dynamics [10–14]. Subsequently, the emphasis has been to explain the variability by formulating a probabilistic description of the process [15, 16]. For example, the process of gene transcription, essentially involving two steps: synthesis of mRNA and degradation of mRNA can be modelled as a birth-death (BD) process [17]. Such description indicates mRNA distribution as Poissonian in contrast to the observed *bursty* behavior of mRNA numbers [7, 18]. However, a similar probabilistic approach is still possible where the gene is considered to toggle between two states of activity (ON-OFF model) [7, 19]. An alternative description of ON-OFF model can be given by birth-death mechanism with time-dependent transcription rates. This necessitates solving the probabilistic description of the stochastic process given by a *master equation* with time-dependent coefficients [20–22].

An explicit solution of the master equation gives the probability distribution of the number of mRNA and can be compared with experiments for various parameter regimes. Moreover, a relevant quantity for a stochastic process is to estimate how long does a randomly varying quantity take to reach a particular threshold value? This quantity is called first passage time (FPT) [23, 24]. FPT can be defined as the time taken by mRNA numbers to reach a particular threshold value for the first time and can be a useful quantity in understanding gene regulation. Single cell measurements monitoring a target gene's protein expression show that microRNAs (miRNA) act as control establishing a threshold level of mRNA below which protein production is suppressed [25]. It has been indicated that miRNA acts as a switch below a threshold mRNA level (strong repression) [26] and as a fine-tuner at high mRNA levels (weak repression) [27]. Hence in the context of gene regulation, it is meaningful to investigate FPT quantifying the time-scale of mRNAs to reach such threshold values. Since transcription is an intrinsically noisy process, FPT for mRNA will vary cell to cell and one is interested in obtaining the FPT distribution function and its moments [28].

Bursty transcriptional dynamics has been observed for housekeeping genes and such behavior can be explained by random switching of DNA between active (ON) and inactive (OFF) states or due to global constraints on transcription [29–31]. However, the burstiness and the levels of transcription are not necessarily correlated [29] and a recent study has numerically demonstrated that the mean first passage time (MFPT) is inversely proportional to transcription rates

---

<sup>\*</sup>School of Mathematical Sciences, Queen Mary University of London, London, United Kingdom

<sup>†</sup>School of Mathematical Sciences, Queensland University of Technology, Brisbane, Australia

<sup>‡</sup>anandamohan@iiserkol.ac.in

of mRNA or mean burst sizes [32]. In this paper, we model gene regulation as a simple BD process and to account for the relaxation behavior of transcriptional dynamics we consider time-dependent reaction rates and present an analytical understanding of the effects of higher transcription rates on the MFPT. We obtain an approximate solution of the first-passage time distribution and characterize the dependence of its moments on periodic modulation of transcription rates.

The primary focus of our present work is to study a general BD model which incorporates time-dependent transcription rates encoding fluctuations in the transcriptional process. We are interested to derive a general form of first passage time distribution for a population of cells each subjected to the same time-dependent transcription rates,  $k_t(t)$ . It is possible to obtain a description of FPT distribution under the assumption that  $k_t(t)$  is differentiable at each point of time. We show that the mean mRNA dynamics in ON-OFF model can be obtained as a special case of BD process with time-dependent rates and an approximate expression of MFPT can be obtained for initial normal distribution on mRNA numbers. Periodically driven biological systems are often encountered in various contexts like cAMP signalling [33], cell cycle and circadian rhythms [34–36], etc. In the context of gene expression, modulated reaction rate has recently got some attention [37, 38] and the probability distribution of the mRNA numbers has been obtained [22, 39, 40]. This motivates us to analyze the time-dependent BD process with a sinusoidal transcription rate and to obtain an expression for the FPT distribution. An interesting observation is that there exist extremal values of MFPT as a function of time period and phase of  $k_t$ . The robustness of the extremal values has been ascertained by incorporating fluctuations in frequency and amplitude of periodic transcription rate. The results have been compared with the numerical simulation of the stochastic reactions implemented by a modified Gillespie algorithm considering the time dependence in reaction rates [21, 41, 42].

The paper is organized as follows: in section [II] we present the formalism of FPT distribution function and different moments of FPT for a general case, in section [II A] we discuss time independent BD process which is the limiting case of time-dependent BD process, in section (II B) we have calculated MFPT for a population of ON-OFF cells, in section [III A] we will provide an approximate solution of FPT distribution function for sinusoidal transcription rate of BD process and compare approximately obtained value of moments of FPT distribution with numerical simulation, in section III B we show FPT distribution and variation of MFPT, standard deviation with different time period and phase of  $k_t$  for square wave transcription rate by stochastic simulation, in section [IV] we will discuss the results which we obtain from numerical simulation or analytical calculation for different type of transcription rates.

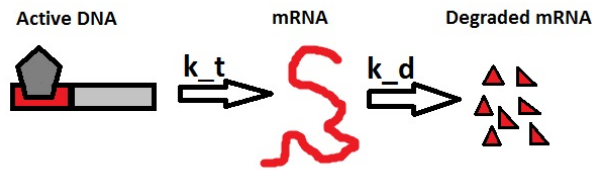


FIG. 1: Schematic diagram of BD model. mRNA is transcribed from active DNA with transcription rate  $k_t$  and mRNAs degrade with degradation rate  $k_d$ .

## II. FIRST PASSAGE TIME DISTRIBUTION

Transcription of mRNA and its subsequent degradation can be modelled as two first order irreversible chemical reactions:



where  $D$  stands for active DNA which transcribes at a rate  $k_t$ , producing mRNAs,  $M$ , and  $k_d$  is the rate of degradation. Often, extracellular signals activate complex biochemical signal transcription pathways that result in a chemical modification of specific transcription factors [43]. In eukaryotic cells the signalling molecules may simply enter the cell and directly bind to the transcription factor inducing a physical change in the shape of the transcription factor protein rendering it into an active molecular state. The transcription factor in its active state binds to the promoter region of a gene and increases the rate of transcription [44]. Moreover, transcription factor can be activator or repressor and they both can be activated or deactivated by signalling molecules and in this work our focus is on the activator activation. Thus global environmental signals that fluctuate in time cause fluctuations in the rate of transcription and consequently the transcription rate, in general, should be time-dependent, i.e.  $k_t(t)$ .

The process of transcription being a stochastic process the number of mRNAs present in a cell is a fluctuating variable and one is interested in finding the probability of finding  $n$  number of mRNAs at time  $t$ , i.e.  $p(n, t)$ , satisfying the master equation:

$$\partial_t p(n, t) = k_t[p(n-1, t) - p(n, t)] + k_d[(n+1)p(n+1, t) - np(n, t)]. \quad (2)$$

Let use the functions  $r_n = k_t(t)$  and  $g_n = nk_d$  and rewrite the above equation as

$$\partial_t p(n, t) = [p_{n-1}r_{n-1} - p_n r_n] + [p_{n+1}g_{n+1} - p_n g_n]. \quad (3)$$

If  $k_t \gg k_d$  the production of mRNA will be higher consequently the number of mRNAs will be larger such that mRNA can be considered to be a continuous variable and our aim is to approximate the dynamics by a Fokker-Planck equation. The two descriptions are in general not exactly equivalent [45, 46] and the approach is to write the master equation as a Kramer's-Moyal expansion and retain terms up to second order from terms involving  $n \pm 1$  to obtain the Fokker-Planck equation [17, 47]. Under this assumption we obtain the Fokker-Planck from Eq.(3)

$$\partial_t p(x, t|x_0, 0) = -\frac{\partial}{\partial x} [(k_t(t) - xk_d)p(x, t|x_0, 0)] + \frac{1}{2} \frac{\partial^2}{\partial x^2} [(k_t(t) + xk_d)p(x, t|x_0, 0)]. \quad (4)$$

where  $p(x, t|x_0, 0)$  is the probability of having  $x$  numbers of mRNA at time  $t$  given that the initial number of mRNA was  $x_0$ .

Now we briefly outline the procedure for computing the FPT distribution [47]. We are interested in the time taken to reach a specific mRNA concentration  $n_c$  starting from an initial number of mRNAs  $x \in (0, n_c)$ . The probability that after time  $t$  the mRNA numbers are still in an interval  $(0, n_c)$  is

$$\int_0^{n_c} p(x', t|x, 0) dx' \equiv G_{<}(x, t). \quad (5)$$

Let  $T$  denote the FPT, within which mRNA leaves the interval  $(0, n_c)$ , such that

$$G_{<}(x, t) = Prob(t \leq T) = \int_0^{n_c} p(x', t|x, 0) dx' \quad (6)$$

and the FPT distribution function  $f_{<}(x, T)$  can be written as

$$f_{<}(x, T) = \frac{\partial}{\partial t} Prob(t \geq T) = -\frac{\partial G_{<}(x, t)}{\partial t} \Big|_{t=T}. \quad (7)$$

Our primary aim is to solve for  $p(x, t|x_0, 0)$  and we start by taking Fourier transform:

$$\mathcal{F}[p] = \rho(\omega, t) = \int_{-\infty}^{\infty} p(x, t|x_0, 0) e^{-i\omega x} dx. \quad (8)$$

It is to be noted that the number of mRNAs can only be positive, even though the limits in the above integral ranges from  $-\infty$  to  $\infty$ . Fourier transformation of Eq. (4) gives

$$\partial_t \rho(\omega, t) = -\left[ i\omega k_t(t) + \frac{k_t(t)\omega^2}{2} \right] \rho(\omega, t) - \omega k_d \partial_\omega \rho(\omega, t). \quad (9)$$

where we have neglected the  $\partial_\omega^2$  term assuming small value of the degradation constant  $k_d$ . For solving Eq. (9) we use the ansatz

$$\rho(\omega, t) = \exp\left(\sum_{n=1}^{\infty} b_n(t)\omega^n\right) \quad (10)$$

and obtain

$$\sum_{n=1}^{\infty} \left[ \frac{db_n}{dt} + nk_d b_n(t) \right] \omega^n + i\omega k_t(t) + \frac{k_t(t)\omega^2}{2} = 0. \quad (11)$$

The above set of odes can be solved term by term

$$\begin{aligned}
b_1(t) &= b_{10}e^{-kat} - iH_1(t)e^{-kat}, & H_1(t) &= \int_0^t k_t(t')e^{kat'} dt' \\
b_2(t) &= b_{20}e^{-2kat} - H_2(t)e^{-2kat}/2, & H_2(t) &= \int_0^t k_t(t')e^{2kat'} dt' \\
b_n(t) &= b_{n0}e^{-nkat} \quad \forall n \geq 3
\end{aligned} \tag{12}$$

resulting in the solution of Eq. (9)

$$\rho(\omega, t) = \exp\left[\omega(b_{10}e^{-kat} - iH_1(t)e^{-kat}) + \omega^2(b_{20}e^{-2kat} - H_2(t)e^{-2kat}/2) + \sum_{n=3}^{\infty} \omega^n b_{n0}e^{-nkat}\right]. \tag{13}$$

From the initial condition  $\rho(\omega, 0) = e^{-i\omega x_0}$  we can fix  $b_{10} = -ix_0$  and  $b_{n0} = 0 \forall n \geq 2$ . Finally the general solution for  $\rho(\omega, t)$  can be given by

$$\rho(\omega, t) = \exp\left[-i\omega(x_0 + H_1(t))e^{-kat} - \omega^2 H_2(t)e^{-2kat}/2\right] \tag{14}$$

Inverse Fourier transform of  $\rho(\omega, t)$  yields the solution of Eq. (4)

$$p(x, t|x_0, 0) = \frac{1}{\sqrt{2\pi H_2(t)e^{-2kat}}} \exp\left[-\frac{(x - (x_0 + H_1(t))e^{-kat})^2}{2H_2(t)e^{-2kat}}\right]. \tag{15}$$

It is now straightforward to obtain  $G_{<}(x_0, t)$  from Eq. (6) and the FPT distribution from Eq. (7)

$$f_{<}(x_0, T) = -\frac{1}{2} \frac{\partial}{\partial t} \left[ \operatorname{erf}\left(\frac{x_0 + H_1(t)}{\sqrt{2H_2(t)}}\right) - \operatorname{erf}\left(\frac{x_0 + H_1(t) - n_c e^{kat}}{\sqrt{2H_2(t)}}\right) \right] \Big|_T \tag{16}$$

The above FPT distribution is an approximate solution for any time-dependent BD process that can be described by an equivalent Fokker-Planck equation with the additional constraint that  $k_t(t)$  should be differentiable at every point of time. It is also useful to estimate the FPT statistics when the initial number of mRNAs are at steady state which may lie below or above threshold. For instances the initial mRNA numbers lying above threshold,  $n_c < x < \infty$ , we will define

$$G_{>}(x, t) = \int_{n_c}^{\infty} p(x', t|x, 0) dx'. \tag{17}$$

and similar to the above calculations we can obtain the FPT distribution

$$f_{>}(x_0, T) = -\frac{1}{2} \frac{\partial}{\partial t} \left[ \operatorname{erf}\left(\frac{x_0 + H_1(t) - n_c e^{kat}}{\sqrt{2H_2(t)}}\right) \right] \Big|_T \tag{18}$$

Once we obtain the FPT distribution we can calculate the mean,  $\mu_T$ , and the standard deviation,  $\sigma_T$ , for any time-dependent transcription process. We will show the dependence of these moments on properties of the time-dependent transcription rate,  $k_t(t)$ , as transcriptional threshold,  $n_c$ , is set at different levels. A similar approach using Laplace transform is able to provide Laplace transformed representation of the FPT distribution [48].

### A. Time Independent Birth-Death Process

As a simple exercise we consider constant transcription rate,  $k_t(t) = k_1 \forall t$  and calculate the FPT distributions

$$f_{>}(x_0, T) = \frac{k_1 e^{k_d T}}{\sqrt{2\pi H_2}} \exp\left(-\frac{(x_0 + H_1 - n_c e^{k_d T})^2}{2H_2}\right) \left[\frac{x_0 + H_1 - n_c e^{k_d T}}{2H_2} e^{k_d T} + \frac{n_c k_d}{k_1} - 1\right] \tag{19}$$

$$f_{<}(x_0, T) = \frac{k_1 e^{k_d T}}{\sqrt{2\pi H_2}} \exp\left(-\frac{(x_0 + H_1)^2}{2H_2}\right) \left[\frac{x_0 + H_1}{2H_2} e^{k_d T} - 1\right] - f_{>}(x_0, T) \tag{20}$$

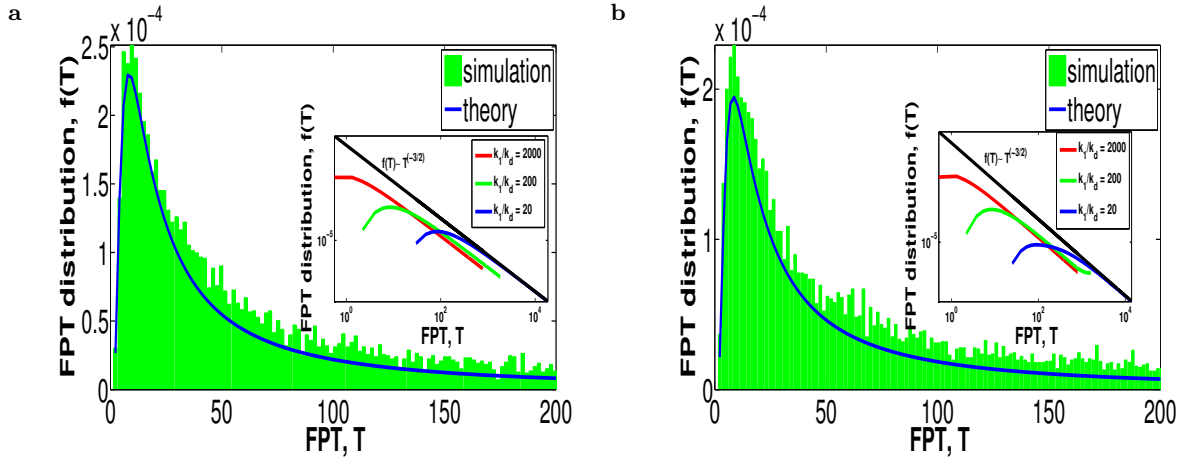


FIG. 2: The FPT distributions for time independent birth death process for initial mRNA numbers drawn from a normal distribution,  $\mathcal{N}(\mu_{x_0}, 1)$  with mean at steady state ( $\mu_{x_0} = x_{ss} = k_1/k_d$ ) where  $k_1 = 20$ ,  $k_d = 0.01$ , (a) above threshold  $n_c = x_{ss} - 10$  (b) below threshold  $n_c = x_{ss} + 10$ . The insets show the power-law tail of the FPT distributions for different  $k_1/k_d$  ratios. The numerical estimates are averaged over  $10^4$  realizations. The black line indicates  $f(T) \approx T^{-3/2}$  as a guide to the eye.

where,  $H_1(T) = (k_1/k_d)[\exp(k_d T) - 1]$  and  $H_2(T) = (k_1/2k_d)[\exp(2k_d T) - 1]$ . Keeping the initial mRNA numbers fixed at steady state value,  $x_{ss} = k_1/k_d$ , above and below the threshold we calculate the FPT and compare with the distributions obtained numerically from Gillespie algorithm [Fig.(2)]. The FPT distributions have power-law tails and are shown in the insets of for different ratios of  $k_1/k_d$  [Fig.(3)]. For our choice of the initial values of mRNA numbers drawn from a normal distribution  $\mathcal{N}(\mu_{x_0}, 1)$  with the mean value set at the steady state  $\mu_{x_0} = x_{ss} = k_1/k_d$  the large time behavior of FPT distribution for small  $k_d$  can be given as

$$f_{\leq}(x, T) \sim T^{(-3/2)} \frac{n_c}{\sqrt{k_1}}. \quad (21)$$

Hence the tail of the FPT distribution  $f(T) \sim T^{-3/2}$  and the agreement is good for a wide range of levels of transcription as observed by varying  $k_1/k_d$ . If the initial number of mRNA,  $x = 0$ , the behavior of the moments of the FPT can be given by simple expressions if we set the threshold value high i.e.,  $n_c > k_1/k_d \gg 1$

- $\mu_T = n_c/k_1$
- $\sigma_T = \sqrt{n_c}/k_1$

These approximate estimates and the results of numerical simulations of the mRNA dynamics implemented by the Gillespie algorithm are shown in Fig.(3). We also show the errorbars in MFPT obtained from the standard deviation of numerically estimated FPT distributions. It is expected that when mRNA levels are higher the Fokker-Planck approximation is appropriate and the agreement with simulation results should be better. We also observe that the error  $\sim |\mu_T - \mu_T^{num}|$  decreases with transcription rates,  $k_1$ . For small,  $k_1$ , the mRNA copy numbers are less, and the errors are higher as seen for different choice of parameters of the model.

## B. ON-OFF Model as a time-dependent Birth-Death process

In the birth-death model, we assume that the promoter is always in an active state and the transcription rate is constant with time. In general, the promoter is not always in an active state but switches between an active state and an inactive state. The ON-OFF model of transcription can explain the experimental observation that the mRNAs are produced in bursts obeying non-Poissonian statistics [7, 19]. The ON-OFF model can be represented as following first order reactions:



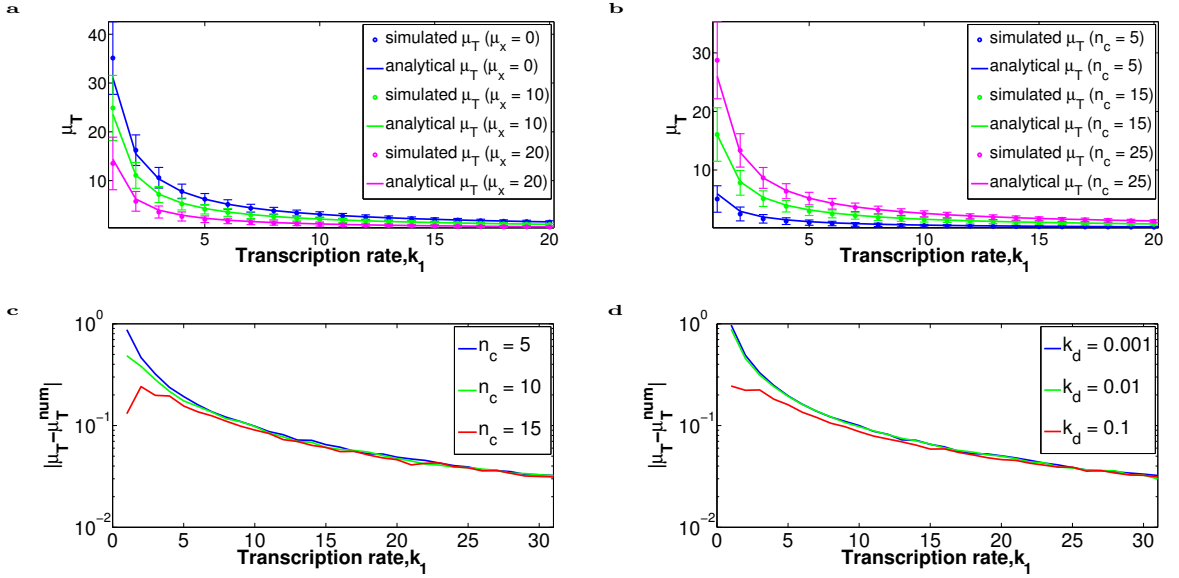


FIG. 3: Variation of MFPT ( $\mu_T$ ) and standard deviation of FPT with transcription rate,  $k_1$ , are shown (a) for different initial mRNA numbers are drawn from a normal distribution,  $\mathcal{N}(\mu_x, 1)$ , with mean  $\mu_x$  and threshold value set at  $n_c = 30$  (b) for different threshold value,  $n_c$ , and with mean initial mRNA number  $\mu_x = 0$ . The degradation rate is fixed at  $k_d = 0.01$ . The deviation of the MFPT simulation results with the approximate estimate is shown in (c) for different cutoff ( $n_c$ ) values and fixed degradation rate,  $k_d = 0.01$  (d) for different  $k_d$  and  $n_c$  fixed at 5. All numerical estimates are averaged over  $10^4$  realizations.

where  $D_{on}$  and  $D_{off}$  are the active and the inactive states of the promoter and  $k_{on}$  and  $k_{off}$  are activation and inactivation rate constants, respectively. The transcription takes place with a constant rate  $k_1$  producing mRNA ( $M$ ) which degrades at a rate  $k_d$ . Alternatively, the ON-OFF model can be treated as a BD model where DNA is always in an active state and transcribing with a time-dependent transcription rate,  $k_t(t)$ , given by

$$k_t(t) = D_0 k_1 e^{-(k_{on} + k_{off})t} + \frac{k_1 k_{on}}{k_{on} + k_{off}} \left[ 1 - e^{-(k_{on} + k_{off})t} \right], \quad (23)$$

where  $D_0$  is initial value of promoters which are in active state. From the expression of the effective transcription rate  $k_t$  we can calculate  $H_1$  and  $H_2$  as follow

$$\begin{aligned} H_1(t) &= c_1(1 - e^{-a_1 t}) + b_1(e^{k_d t} - 1) \\ H_2(t) &= c_2(1 - e^{-a_2 t}) + b_2(e^{k_d t} - 1). \end{aligned} \quad (24)$$

involving the constants  $c_1 = \frac{k_1}{k_{on} + k_{off} - k_d} (D_0 - \frac{k_{on}}{k_{on} + k_{off}})$ ,  $b_1 = \frac{k_1 k_{on}}{k_d(k_{on} + k_{off})}$ ,  $a_1 = k_{on} + k_{off} - k_d$ ,  $c_2 = \frac{k_1}{k_{on} + k_{off} - 2k_d} (D_0 - \frac{k_{on}}{k_{on} + k_{off}})$ ,  $b_2 = \frac{k_1 k_{on}}{2k_d(k_{on} + k_{off})}$ ,  $a_2 = k_{on} + k_{off} - 2k_d$ . The detailed calculation of  $k_t(t)$  is shown in Appendix (A).

The FPT distributions  $f_{<}(x_0, T)$  for ON-OFF model is obtained from Eq.(16) with the time-dependent functions defined as follows:

$$\begin{aligned} f_{<}(x_0, T) &= \frac{1}{\sqrt{2\pi H_2(T)}} \left[ e^{-\frac{(H_1(T) + x_0)^2}{2H_2(T)}} \left( a_1 c_1 e^{-a_1 T} + b_1 k_d e^{k_d T} - \frac{(x_0 + H_1(T))(a_2 c_2 e^{-a_2 T} + 2b_2 k_d e^{2k_d T})}{2H_2(T)} \right) \right. \\ &\quad \left. + e^{-\frac{(H_1(T) + x_0 - n_c e^{k_d T})^2}{2H_2(T)}} \left( a_1 c_1 e^{-a_1 T} + b_1 k_d e^{k_d T} - n_c k_d e^{k_d T} - \frac{(x_0 + H_1(T) - n_c e^{k_d T})(a_2 c_2 e^{-a_2 T} + 2b_2 k_d e^{2k_d T})}{2H_2(T)} \right) \right] \end{aligned} \quad (25)$$

Now under the condition  $(k_{on} + k_{off}) \gg 1$  the effective transcription rate will be  $\frac{k_1 k_{on}}{k_{on} + k_{off}}$  and if  $k_d \ll 1$  the mean first passage time is given by

$$\mu_T = \frac{n_c - \mu_x}{\beta} \left[ \frac{1}{k_{on}} + \frac{1}{k_{off}} \right] \quad (26)$$

$$\sigma_T = \frac{\sqrt{n_c - \mu_x}}{\beta} \left[ \frac{1}{k_{on}} + \frac{1}{k_{off}} \right]. \quad (27)$$

where the mean burst size is denoted by  $\beta = k_1/k_{off}$  and  $\mu_x$  denotes the average initial values of mRNA ( $x_0$ ) of a population of cells. The typical values of model parameters are taken from single cell studies [18] and the dependence of the MFPT on  $\beta$  is shown in Fig.4. We also show the dependence of MFPT on the mean burst frequency, defined as  $\beta_f = k_{on}$  [49, 50], for varying initial conditions and different thresholds in Fig.4.

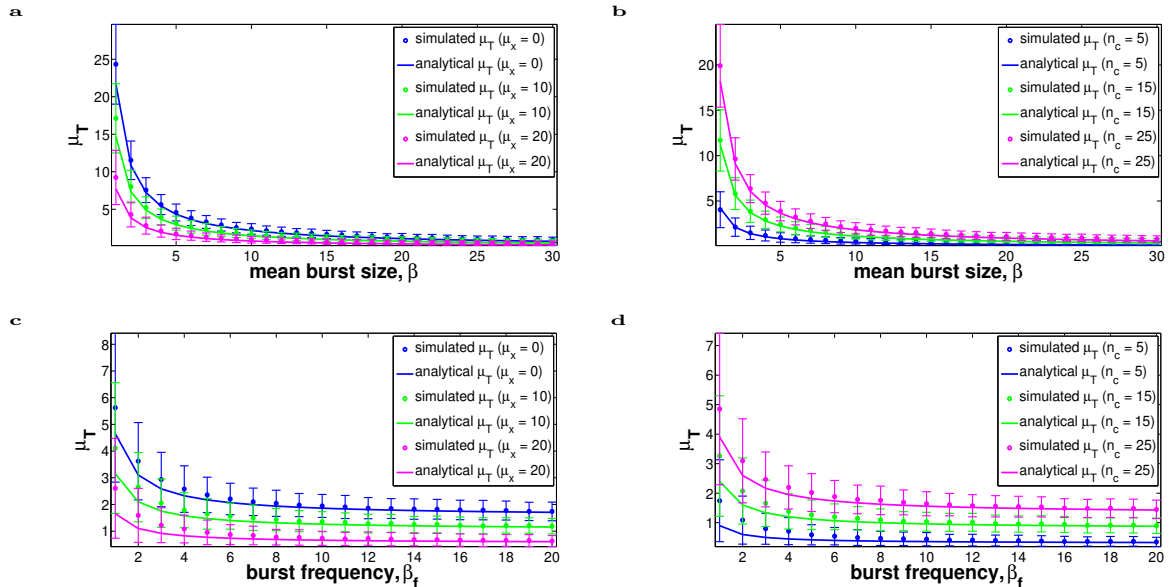


FIG. 4: Variation of MFPT and standard deviation of FPT are shown for varying mean burst size,  $\beta (= k_1/k_{off})$  (a) for different values of mean initial mRNA numbers drawn from normal distribution and fixed threshold  $n_c = 30$ . (b) for different threshold values and initial mRNA  $x = 0$ . The parameter values are  $k_d = 0.01$ ,  $k_{on} = 5$ ,  $k_{off} = 2$ . Variation of MFPT and standard deviation of FPT are shown for varying mean burst frequency,  $\beta_f (= k_{on})$  (c) for different values of mean initial mRNAs drawn from normal distribution and fixed threshold  $n_c = 30$  and  $k_1 = 20$  (d) for different threshold values and initial mRNA number  $x = 0$ . Numerical averages are computed over  $10^4$  realizations.

### III. PERIODICALLY MODULATED TRANSCRIPTION RATE

In this section we consider periodically driven transcription rate and calculate the dependence of the MFPT on the phase and the time period of the transcription rate.

#### A. Sinusoidal transcription rate

First we consider the case where transcription rate is modulated as

$$k_t(t) = k_1 + k_2 \sin\left(\frac{2\pi t}{\tau} + \phi\right), \quad (28)$$

where  $k_1$ ,  $k_2$  are constants,  $\tau$  is the time period,  $\phi$  is the phase of transcription rate. All cells in the population experience identical extrinsic modulation  $k_t(t)$  and the FPT distribution for the sinusoidal transcription rate is given by:

$$f(x_0, T) = \frac{1}{\sqrt{2\pi H_2(T)}} \left[ e^{-\frac{(H_1(T)+x_0-n_c e^{k_d T})^2}{2H_2(T)}} \left( \frac{d}{dT} (H_1(T) + x_0 - n_c e^{k_d T}) \right) - \frac{(H_1(T) + x_0 - n_c e^{k_d T})}{2H_2(T)} \frac{dH_2}{dT} \right) - e^{-\frac{(H_1(T)+x_0)^2}{2H_2(T)}} \left( \frac{d}{dT} (H_1(T) + x_0) - \frac{(H_1(T) + x_0)}{2H_2(T)} \frac{dH_2}{dT} \right) \right] \quad (29)$$

where the time-dependent functions depend on transcription rate Eq. (28)

$$H_1(t) = \int_0^t k_t(t') e^{k_d t'} dt' \quad (30)$$

$$H_2(t) = \int_0^t k_t(t') e^{2k_d t'} dt'. \quad (31)$$

Now if we consider initial value of mRNA,  $x_0$  to be 0 and  $k_d \ll 1$ , then we will get

$$H_1(t) \simeq H_2(t) = k_1 t - \frac{k_2 \tau}{2\pi} \left[ \cos\left(\frac{2\pi t}{\tau} + \phi\right) - \cos(\phi) \right] \equiv H(t). \quad (32)$$

Now if we define  $k(t)$  as

$$k(t) = \frac{2\pi}{\tau} H(t) = \frac{2k_1 \pi t}{\tau} - k_2 \left[ \cos\left(\frac{2\pi t}{\tau} + \phi\right) - \cos(\phi) \right], \quad (33)$$

then the form of FPT distribution will be given as follow

$$f(T) = \frac{k_t(T)}{2} \sqrt{\frac{1}{\tau k(T)}} e^{-\frac{(k(T) - 2\pi n_c / \tau)^2}{4\pi k(T) / \tau}}. \quad (34)$$

The approximate FPT distribution,  $f(T)$ , and the FPT distribution obtained from Gillespie simulation is shown in Fig.5(a) for fixed values of the parameters  $\phi$ ,  $\tau$ ,  $n_c$ ,  $k_1$ ,  $k_2$ . We numerically estimate the MFPT and the standard deviation from the FPT distribution given in Eq.(29) and compare with stochastic simulation results. The variation of the MFPT with the time period,  $\tau$ , is shown in Fig.5(b) for two different values of threshold,  $n_c$  while the modulation phase,  $\phi = 0$ . The dependence on the threshold is obvious as setting a higher cutoff needs more mRNAs to be produced requiring more time but an important observation is that the MFPT show a minimum for  $\phi = 0$ . If the phase of the transcription rate is varied the behavior of MFPT with time period changes and is shown in Fig.5(c). The non-trivial dependence of MFPT on the phase and the time period and occurrence of extremal values prompts us to represent MFPT as 3D surface as in Fig.5(d) indicating the existence of a global extremal value. Consideration of sinusoidal transcription rate is an idealization and to establish the robustness of the above mentioned feature in a slightly more realistic scenario we introduce noise in the time-dependent transcription rate as

$$k_t(t) = k_1 + k_2 \sin\left(\frac{2\pi t}{\tau} + \phi\right) + \xi \quad (35)$$

where  $\xi$  is some random noise drawn from an uniform distribution. We show  $\mu_T(\tau, \phi)$  in Appendix Fig.A for different values of noise strength  $\xi$ . The results indicate that the qualitative behavior of  $\mu_T(\tau, \phi)$  remains unchanged with the existence of global minimum robust to fluctuations.

## B. Square wave transcription rate

Now we consider a periodic signal in the form of a square wave

$$k_t(t) = \begin{cases} a & (0 - \phi) \leq t \leq \frac{\tau}{2}(1 - \phi) \\ b & \frac{\tau}{2}(1 - \phi) \leq t \leq \frac{\tau}{2}(2 - \phi). \end{cases} \quad (36)$$

Schematic diagram of  $k_t(t)$  is shown in Fig.6(a) illustrating the phase  $\phi$  and the time-period  $\tau$  and here also  $k_t(t)$  is an extrinsic modulation affecting all cells in the population. The minimum value  $a$  shows there can be some basal transcription and  $b$  is the maximum transcription rate. For square wave transcription rate  $k_t(t)$  is not differentiable at all time and we can not explicitly calculate FPT distribution and show the numerical data obtained from stochastic simulation in Fig.6(b). Simulation results show minimum with respect to time period for phase of the signal,  $\phi < 1$  Fig. 6(c),(d). However, unlike sinusoidal modulation there is no maximum in MFPT for phase  $\phi \sim 1$ .



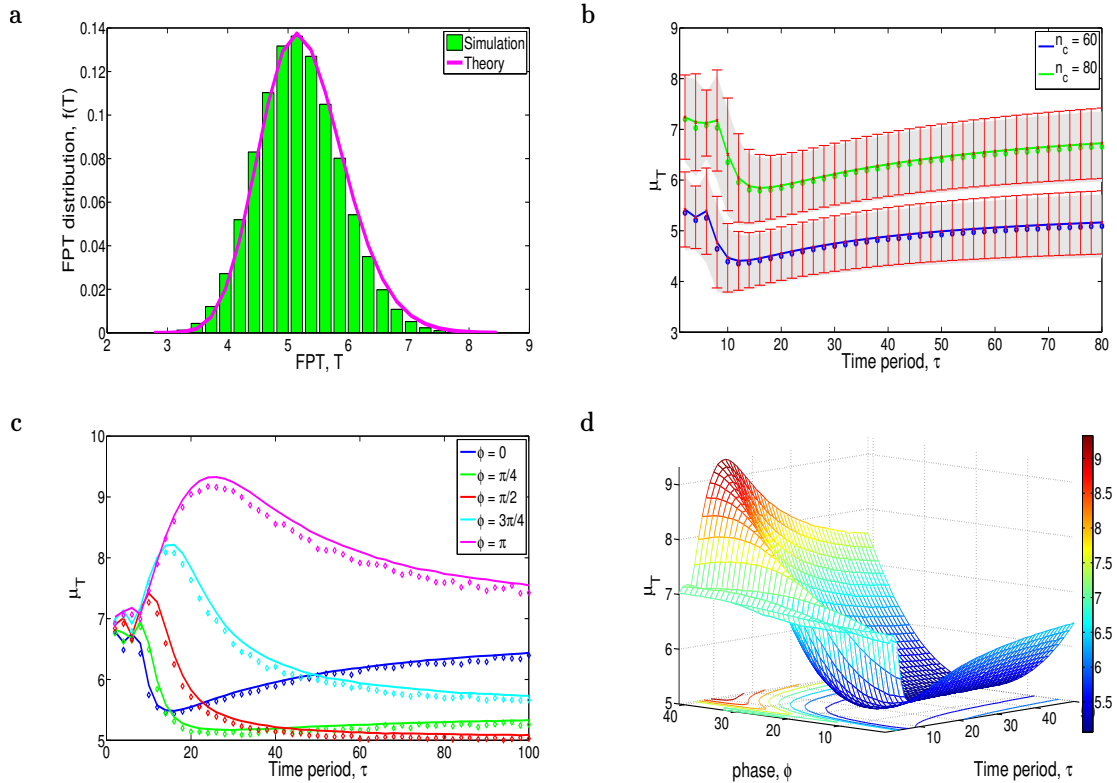


FIG. 5: FPT for sinusoidal transcription rate. (a) Results of the stochastic simulation algorithm (SSA) are shown with FPT in Eq. 29. The histogram was obtained from  $10^5$  realizations. The parameter values are  $\phi = 0, n_c = 75, k_1 = 11, k_2 = 4, k_d = 0.005, \tau = 50$ . (b) Variation of MFPT with  $\tau$  obtained from numerical simulation (dotted plot) and enumeration from Eq. (29) (solid line) are shown for two cutoff values, (blue:  $n_c = 60$ ) and (green:  $n_c = 80$ ). The parameter values are  $\phi = 0, k_1 = 11, k_2 = 4, k_d = 0.005$ . Numerical averages are computed over  $10^4$  realizations. (c) MFPT vs.  $\tau$  for different  $\phi$ . (d) Numerically estimated  $\mu_T$  with respect to  $\phi$  and  $\tau$ .

#### IV. DISCUSSION

We have calculated the first passage time distribution for stochastic gene regulation modeled as birth-death process. An approximate expression of the FPT distribution of a population of cells and its moments are obtained for BD process with the generalization that the reaction rates are time-dependent. The MFPT obtained for the usual BD process (time-independent rates) behaves as  $\mu_T = n_c/k_1$  and is easily understandable as follows: higher transcription rate  $k_1$  implies increased production of mRNAs and higher probability of reaching a fixed threshold,  $n_c$ . Again, the time required for the mRNAs to attain a higher threshold will be proportionately more. In the context of recent studies of gene-regulation in single cells, it is meaningful to investigate FPT statistics for such systems showing bursty mRNA dynamics [7]. The ON-OFF model is a simple model capturing the mRNA burstiness and we demonstrate that the mean evolution of a population of ON-OFF cells is equivalent to a BD process with effective time-dependent rates. We show that our method is easily applicable for ON-OFF model and confirm that the MFPT is inversely proportional to mean mRNA burst size,  $\beta = k_1/k_{off}$  [32]. Approximate FPT distributions are calculated for ON-OFF models where initial distribution of mRNA numbers are Gaussian with arbitrary mean levels and threshold levels and the results are compared with numerical simulations performed by stochastic Gillespie algorithm.

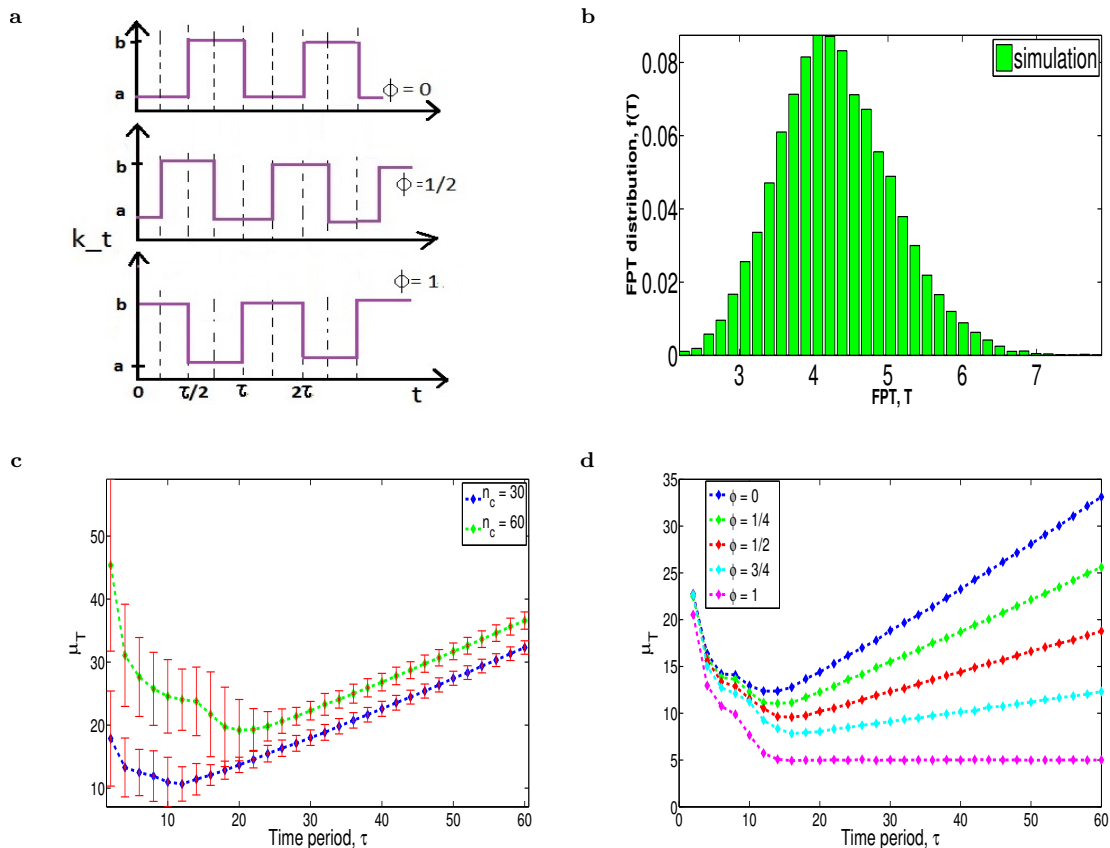


FIG. 6: (a) Schematic diagram of square wave transcription rate,  $k_t$  whose time period is  $\tau$  and amplitude is  $(b - a)$  shown for different phases,  $\phi$ . (b) FPT distribution from SSA averaged over  $10^5$  realizations. Parameter values  $a = 1, b = 7, n_c = 30, \tau = 40, \phi = 1, k_d = 0.05$ . (c) Variation of MFPT with  $\tau$  obtained from SSA shown for two cutoff values, (blue:  $n_c = 30$ ) and (green:  $n_c = 50$ ) and  $\phi = 0$ . (d) Variation of MFPT with  $\tau$  for different  $\phi$  for fixed  $n_c = 30$ .

The analysis of stochastic gene regulation with time-dependent transcription rates is important to understand and model the effects of the noisy cellular environment [22]. In earlier studies, it has been demonstrated that in the presence of periodically modulated environmental fluctuations the process of transcription acts as a low-pass filter with definite fitness advantage [51, 52]. We find that the MFPT has non-trivial dependence on the phase and the period of the modulated transcription rates. For small phase,  $\phi \approx 0$ , as time period,  $\tau$ , of periodic modulation is increased the MFPT shows a non-monotonic dependence having a distinct minimum. For small  $\tau$  within one cycle of variation of transcription rates, the system does not yield enough mRNAs to reach the threshold, hence MFPT is high. In the large  $\tau$  limit the transcriptional process behaves as a RC filter [21, 51], once mRNA level is low (discharged) the system has to wait longer to reach the desired mRNA level (charging). However, the behavior is opposite in the case of  $\phi = \pi$  showing a distinct maximum at an intermediate value of  $\tau$ . But for square wave modulation varying phase shifts introduce only a minimum in MFPT with respect to time period. Unlike the case of periodic modulation in the case of square wave for  $\phi = 1$  the transcription rate is constantly in upstate disallowing any relaxation and MFPT attains a low value for large  $\tau$ . Systematically varying the phase and the time period we obtain the landscape of MFPT indicating that suitable modulation of upstream/downstream pathways can precisely regulate gene transcription. The recent advancements in single-molecule observations of RNA in living cells [37] and the ability to optically control genetic circuits [38] make our results relevant in designing optimal performance of genetic regulatory circuits. Recent experiments analyze the causal link between the stability of mRNAs and mRNA concentration as the post-transcriptional mechanism of controlling gene expression [53]. Studying first passage time to reach a particular mRNA concentration helps us to understand the optimum time for having stable mRNAs and potentially opens a way to understand the regulation of gene expression.

**Author Contributions:** KB,MS,AS,AG designed the project. KB, AG performed the majority of the analysis and wrote the paper.

## APPENDIX

### A. ON-OFF MODEL : EFFECTIVE TRANSCRIPTION RATE

Reaction mechanism for birth-death model is



Rate equation of mRNA for this model is

$$\frac{dM}{dt} = k_t(t) - k_d M \quad (\text{A-3})$$

We assume that DNA is always in active state and concentration of DNA is one. Reaction mechanism for ON-OFF model is



Rate equation for this model will be

$$\frac{dD_{on}}{dt} = k_{on} - (k_{on} + k_{off})D_{on} \quad (\text{A-8})$$

$$\frac{dM}{dt} = k_1 D_{on} - k_d M \quad (\text{A-9})$$

where  $(D_{on} + D_{off}) = 1$  and if we solve for  $D_{on}$  then we will get

$$D_{on}(t) = D_0 e^{-(k_{on} + k_{off})t} + \frac{k_{on}}{k_{on} + k_{off}} \left[ 1 - e^{-(k_{on} + k_{off})t} \right]. \quad (\text{A-10})$$

So effectively rate equation for mRNA for ON-OFF model is

$$\frac{dM}{dt} = D_0 k_1 e^{-(k_{on} + k_{off})t} + \frac{k_1 k_{on}}{k_{on} + k_{off}} \left[ 1 - e^{-(k_{on} + k_{off})t} \right] - k_d M. \quad (\text{A-11})$$

If we compare Eq. (A-3) with Eq. (A-11) we obtain

$$k_t(t) = D_0 k_1 e^{-(k_{on} + k_{off})t} + \frac{k_1 k_{on}}{k_{on} + k_{off}} \left[ 1 - e^{-(k_{on} + k_{off})t} \right]. \quad (\text{A-12})$$

### B. ROBUSTNESS OF MFPT FOR SINUSOIDAL MODULATION AND NOISE

To see whether the qualitative behavior of  $\mu_T(\tau, \phi)$  remains unchanged in presence of any fluctuating environment, we add amplitude fluctuation into the sinusoidal transcription rate:

$$k_t(t) = k_1 + k_2 \sin\left(\frac{2\pi t}{\tau} + \phi\right) + \xi \quad (\text{B-13})$$

where  $\xi \in \mathcal{U}(0, \sigma)$  is noise drawn from an uniform distribution. For this type of fluctuating signal  $\mu_T(\tau, \phi)$  is shown in Fig.A for different additive noise strengths  $\sigma$ . Since the intrinsic transcriptional parameters are kept constant, the noise of the mRNAs are constant at  $\sigma_M = 100$ . Higher noise reduces the MFPT but retains the generic feature of MFPT dependence on time-period and phase.

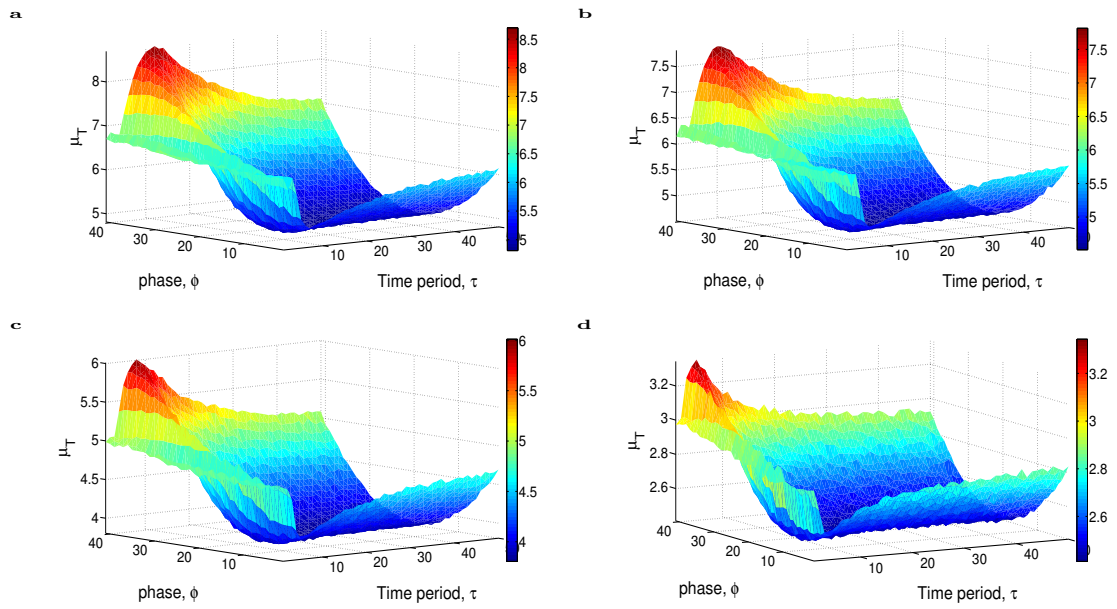


FIG. A: Noisy sinusoidal transcription rate: variation of  $\mu_T$  with respect to  $\tau$  and  $\phi$ . (a)  $\sigma = 1$  (b)  $\sigma = 3$  (c)  $\sigma = 10$  (d)  $\sigma = 40$ . Where  $\sigma$  is the noise strength of  $\xi$  and the standard deviation of mRNA is  $\sigma_M = 100$ . This SSA is done for  $10^4$  realizations with the parameter values  $n_c = 75, k_1 = 11, k_2 = 4, k_d = 0.005$ .

- 
- [1] Thattai M and Van Oudenaarden A. Intrinsic noise in gene regulatory networks. *Proceedings of the National Academy of Sciences*, 98(15):8614–8619, 2001.
- [2] Elowitz M B, Levine A J, Siggia E D, and Swain P S. Stochastic gene expression in a single cell. *Science*, 297(5584):1183–1186, 2002.
- [3] Blake W J, Kaern M, Cantor C R, and Collins J J. Noise in eukaryotic gene expression. *Nature*, 422(6932):633, 2003.
- [4] Raser J M and O’Shea E K. Control of stochasticity in eukaryotic gene expression. *Science*, 304(5678):1811–1814, 2004.
- [5] Paulsson J. Summing up the noise in gene networks. *Nature*, 427(6973):415–8, 2004.
- [6] Raj A, Peskin C S, Tranchina D, Vargas D Y, and Tyagi S. Stochastic mRNA synthesis in mammalian cells. *PLoS Biology*, 4(10):e309, 2006.
- [7] Golding I, Paulsson J, Zawilski S M, and Cox E C. Real-time kinetics of gene activity in individual bacteria. *Cell*, 123(6):1025–1036, 2005.
- [8] Larson D R, Singer R H, and Zenklusen D. A single molecule view of gene expression. *Trends in Cell Biology*, 19(11):630–637, 2009.
- [9] Chubb J R, Trcek T, Shenoy S M, and Singer R H. Transcriptional pulsing of a developmental gene. *Current Biology*, 16(10):1018–1025, 2006.
- [10] Ko M SH. A stochastic model for gene induction. *Journal of Theoretical Biology*, 153(2):181–194, 1991.
- [11] Young J W, Locke J CW, Altinok A, Rosenfeld N, Bacarian T, Swain P S, Mjolsness E, and Elowitz M B. Measuring single-cell gene expression dynamics in bacteria using fluorescence time-lapse microscopy. *Nature Protocols*, 7(1):80, 2012.
- [12] Park H Y, Lim H, Yoon Y J, Follenzi A, Nwokafor C, Lopez-Jones M, Meng X, and Singer R H. Visualization of dynamics of single endogenous mRNA labeled in live mouse. *Science*, 343(6169):422–424, 2014.
- [13] Crosetto N, Bienko M, and Van Oudenaarden A. Spatially resolved transcriptomics and beyond. *Nature Reviews Genetics*, 16(1):57, 2015.
- [14] Owens N DL, Blitz I L, Lane M A, Patrushev I, Overton J D, Gilchrist M J, Cho K WY, and Khokha M K. Measuring absolute rna copy numbers at high temporal resolution reveals transcriptome kinetics in development. *Cell Reports*, 14(3):632–647, 2016.
- [15] Swain P S, Elowitz M B, and Siggia E D. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proceedings of the National Academy of Sciences*, 99(20):12795–12800, 2002.
- [16] Shahrezaei V and Swain P S. Analytical distributions for stochastic gene expression. *Proceedings of the National Academy of Sciences*, 105(45):17256–17261, Nov. 11, 2008.
- [17] Van Kampen N G. *Stochastic Processes in Physics and Chemistry*. Elsevier Science B.V., Amsterdam, Netherlands, 1992.
- [18] So L, Ghosh A, Zong C, Sepúlveda L A, Segev R, and Golding I. General properties of transcriptional time series in *Escherichia coli*. *Nature Genetics*, 43(6):554–560, 2011.
- [19] Peccoud J and Ycart B. Markovian modeling of gene-product synthesis. *Theoretical Population Biology*, 48(2):222 – 234, 1995.
- [20] Murugan R. Stochastic transcription initiation: Time dependent transcription rates. *Biophysical Chemistry*, 121(1):51–56, 2006.
- [21] Nandi S and Ghosh A. Transcriptional dynamics with time-dependent reaction rates. *Physical Biology*, 12(1):016015, 2015.
- [22] Dattani J and Barahona M. Stochastic models of gene transcription with upstream drives: exact solution and sample path characterization. *Journal of the Royal Society Interface*, 4(126):20160833, 2017.
- [23] Hänggi P, Talkner P, and Borkovec M. Reaction-rate theory: fifty years after Kramers. *Reviews of Modern Physics*, 62(2):251, 1990.
- [24] Redner S. *A Guide to First-Passage Processes*. A Guide to First-passage Processes. Cambridge University Press, 2001.
- [25] Mukherji S, Ebert M S, Zheng G X Y, Tsang J S, Sharp P A, and Van Oudenaarden A. MicroRNAs can generate thresholds in target gene expression. *Nature Genetics*, 43(9):854, 2011.
- [26] Bagga S, Bracht J, Hunter S, Massirer K, Holtz J, Eachus R, and Pasquinelli A E. Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation. *Cell*, 122(4):553–563, 2005.
- [27] Raj A and Van Oudenaarden A. Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell*, 135(2):216–226, 2008.
- [28] Singh A and Dennehy J J. Stochastic holin expression can account for lysis time variation in the bacteriophage  $\lambda$ . *Journal of The Royal Society Interface*, 11(95), 2014.
- [29] Sanchez A and Golding I. Genetic determinants and cellular constraints in noisy gene expression. *Science*, 342(6163):1188–1193, 2013.
- [30] Lionnet T and Singer R H. Transcription goes digital. *EMBO Report*, 13(4):313–21, 2012.
- [31] Suter D M, Molina N, Gatfield D, Schneider K, Schibler U, and Naef F. Mammalian genes are transcribed with widely different bursting kinetics. *Science*, 332(6028):472–474, 2011.
- [32] Shreshtha M, Surendran A, and Ghosh A. Estimation of mean first passage time for bursty gene expression. *Physical Biology*, 13(3):036004, 2016.
- [33] Corrigan A M and Chubb J R. Regulation of transcriptional bursting by a naturally oscillating signal. *Current Biology*, 24(2):205–211, 2014.
- [34] Mihalcescu I, Hsing W, and Leibler S. Resilient circadian oscillator revealed in individual cyanobacteria. *Nature*, 430(6995):81, 2004.

- [35] Bieler J, Cannavo R, Gustafson K, Gobet C, Gatfield D, and Naef F. Robust synchronization of coupled circadian and cell cycle oscillators in single mammalian cells. *Molecular Systems Biology*, 10(7):739, 2014.
- [36] Lück S and Westermark P. Circadian mRNA expression: insights from modeling and transcriptomics. *Cellular and Molecular Life Sciences*, 73(3):497–521, 2016.
- [37] Larson D R, Fritzsche C, Sun L, Meng X, Lawrence D S, and Singer R H. Direct observation of frequency modulated transcription in single cells using light activation. *Elife*, 2, 2013.
- [38] Olson E J, Hartsough L A, Landry B P, Shroff R, and Tabor J J. Characterizing bacterial gene circuit dynamics with optically programmed gene expression signals. *Nature Methods*, 11(4):449, 2014.
- [39] Mugler A, Walczak A M, and Wiggins C H. Information-optimal transcriptional response to oscillatory driving. *Physical Review Letters*, 105(5):058101, 2010.
- [40] Jedrak J and Ochab-Marcinek A. Time-dependent solutions for a stochastic model of gene expression with molecule production in the form of a compound poisson process. *Physical Review E*, 94(3):032401, 2016.
- [41] Gillespie D T. Stochastic simulation of chemical kinetics. *Annual Review of Physical Chemistry*, 58:3555, 2007.
- [42] Gillespie D T. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, 22:403–34, 1976.
- [43] Gomperts B D, Kramer I M, and Tatham P E R. Signal transduction. (*New York: Academic*), 18, 2002.
- [44] Hobert O. Gene regulation by transcription factors and microRNAs. *Science*, 319(5871):1785–1786, 2008.
- [45] Hanggi P, Grabert H, Talkner P, and Thomas H. Bistable systems: Master equation versus Fokker-Planck modeling. *Physical Review A*, 29(1):371, 1984.
- [46] Grima R, Thomas P, and Straube A V. How accurate are the nonlinear chemical Fokker-Planck and chemical Langevin equations? *The Journal of Chemical Physics*, 135(8):084103, 2011.
- [47] Gardiner C W. *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*. Springer complexity. Springer, 2004.
- [48] Dev S and Chatterjee S. Run-and-tumble motion with step-like response to a stochastic input. *Physical Review E*, 99:012402, 2019.
- [49] Dar R D, Razoooky B S, Singh A, Trimeloni T V, McCollum J M, Cox C D, Simpson M L, and Weinberger L S. Transcriptional burst frequency and burst size are equally modulated across the human genome. *Proceedings of the National Academy of Sciences*, 2012.
- [50] Singh A, Razoooky B, Cox C D, Simpson M L, and Weinberger L S. Transcriptional bursting from the HIV-1 promoter is a significant source of stochastic noise in HIV-1 gene expression. *Biophysical journal*, 98(8):L32–L34, 2010.
- [51] Hooshangi S, Thiberge S, and Weiss R. Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proceedings of the National Academy of Sciences*, 102(10):3581–6, 2005.
- [52] Bennett M R, Pang W L, Ostroff N A, Baumgartner B L, Nayak S, Tsimring L S, and Hasty J. Metabolic gene regulation in a dynamically changing environment. *Nature*, 454(7208):1119, 2008.
- [53] Nouaille S, Mondeil S, Finoux AL, Moulis C, Girbal L, and Coccagn-Bousquet M. The stability of an mRNA is influenced by its concentration: a potential physical mechanism to regulate gene expression. *Nucleic Acids Research*, 45(20):11711–11724, 2017.