Estimation of Mean First Passage Time for Bursty Gene Expression

Mayank Shreshtha, Anudeep Surendran and Anandamohan Ghosh *

Department of Physical Sciences,
Indian Institute of Science Education and Research (IISER) Kolkata
Mohanpur 741246, India
(Dated: April 30, 2016)

Gene expression is an intrinsically noisy process, typically, producing mRNAs and proteins in bursts. An important description of such stochastic processes can be done in terms of the Mean First Passage Time (MFPT), i.e., the time taken by mRNAs/proteins to reach a particular threshold. We study the role of burstiness on MFPT and obtain an analytical expression for different models of transcriptional and translational bursts. Our analytical results and numerical simulations confirm that MFPT monotonically decreases with burstiness.

I. INTRODUCTION

The phenomenon of gene expression is an inherently stochastic process generating fluctuating number of copies of mRNAs and proteins by the processes of transcription and translation respectively. The number of mRNAs and proteins has a wide range of cell-to-cell variability and these fluctuations are important to understand [1–6]. However, in spite of the random fluctuations the transfer of genetic information over generations or synthesis of proteins to cater to the demand of the cell is done with amazing effectiveness. The cell must adapt to an appropriate strategy to control the noise levels to ensure effective functioning of the cellular mechanisms. An important observation in single cell experiments has been that mRNAs and proteins are produced in bursts [7, 8] as opposed to Poisson statistics expected from a typical birth-death process[9]. Bursty transcriptional dynamics invited new ideas to model the stochastic dynamics, namely, two-state or ON-OFF model[10], two-stage model[5, 11], three-stage model[5], time-dependent rate models[12] with goals to quantify the experimental findings. It is to be noted that transcriptional bursting is mainly seen in eukaryotes and mammalian cells [13] and such bursty mRNA dynamics can be explained by a simple ON-OFF mechanism [10]. On the other hand, in the two-stage model only the protein dynamics can be bursty and such translational bursts are typical of prokaryotes [1, 13, 14]. The three-stage model [15] is more general, capable of exhibiting both transcriptional (mRNA) as well as translational (protein) bursts but transcriptional bursts is believed to be of more biological relevance [13]. While the models had been largely successful in explaining the steady state behaviour it is also important to ask the question whether transcriptional burst is an efficient strategy for a cell? There had been a few attempts to establish that the bursty mRNA dynamics efficiently encodes information transfer [16] and possibly be favoured from a thermodynamical perspective[17]. In this paper we explore the consequences of burstiness on the transient dynamics and theoretically investigate the Mean First Passage Time behaviour.

The first passage time is an useful quantification of stochastic physical processes [18]. The first passage time in the phenomenon of gene expression can be defined as the time taken by mRNA or protein to reach a particular threshold for the first time. Since the processes involved are stochastic, one is rather interested in the probability distribution of these times or the first moment *i.e.,* Mean First Passage Time (MFPT). In the context of gene expression, a typical situation in a cell may demand a critical number of proteins to be synthesized and the average time to reach that number is important. The role of bursty dynamics on MFPT thus becomes a meaningful quantity to analyze and has attracted some attention in recent past. The theoretical description of gene expression has been closely related to queuing theory [19] and related quantities like waiting time distributions has been shown to determine the transcriptional dynamics of ON-OFF model [20]. It has also been shown that lysis time in bacteriophage can be formulated as FPT, a measure that an organism possibly uses for robustness to cellular noise [21]. In a very relevant simulation based study on genetic switches, MFPT to switch between states has been found to be exponentially sensitive to parameters of gene regulation, like burstiness in mRNA and protein [22]. However, an exact calculation of MFPT even in simple scenarios is often difficult and a progress in that direction can be useful to gain insight into the gene regulatory

-----

* Electronic address: anandamohan@iiserkol.ac.in

mechanisms.

Recent studies on λ-Phage [23, 24] and HIV virus [24–26] have established that biological fate selection is stochastic in nature and the crucial decisions like the switching from active to inactive state of the virus are driven by noise in the gene expression. A large number of biologically important systems exhibit similar transition from productive to unproductive states, for example, viral latency of human immunodeficiency virus(HIV)-1[27, 28], latency of herpes simplex virus (HSV, subfamily-α) in infected neurons[29, 30], latency in Kaposi's sarcoma-associated herpesvirus (KSHV)[31, 32], viral latency in cytomegalovirus infection in the lung[33]. Interestingly the decision to switch between the active to inactive states depends on the time for the protein levels to reach a threshold necessitating to formulate the stochastic process as a first passage time problem [21]. Moreover, the first passage time statistics depends on the bursty dynamics of the protein expression levels and has been experimentally observed (matched with stochastic simulations) [27].

In this paper, we investigate MFPT for different models of stochastic gene expression that exhibit burstiness. We analytically solve the Langevin dynamics and obtain the probability distribution function, subsequently, we derive the MFPT as a function of the *degree of burstiness*. We find that with the increase in bursty behaviour the MFPT decreases enabling a cell to reach a target threshold faster. Intuitively this can be understood as mRNA/protein level or fluctuations increase it is more likely that a desired threshold is quickly reached. We compute the variance of the FPT and show that it also decreases monotonically with burstiness. We will provide an exact calculation for different models of gene expression and also verify our findings by numerical simulation of the stochastic dynamics. The paper is organized as follows: in Section II, we study the two-state or the ON-OFF model and outline the method to solve the propagator corresponding to the Langevin dynamics and hence analytically derive the MFPT as a function of mRNA (transcriptional) burstiness; in Section III, we study the two-stage model and obtain MFPT analytically as a function of protein (translational) burstiness; in Section IV, we analyse MFPT of proteins in the three-stage model with respect to transcriptional burstiness followed by a discussion in Section V.

II. ON-OFF MODEL

The experimental observation of transcriptional bursts can be effectively modelled by an ON-OFF mechanism of the promoter activity [7, 10]. Changes in chromatin architecture, *chromatin remodelling* [3], is mainly responsible for the transition of the promoter activity from an active(ON) state to an inactive(OFF) state. The binding sites may be accessible (inaccessible) in the ON (OFF) state to the transcriptional factors leading to transcription (no transcription). A schematic representation of the ON-OFF model is shown in the Fig.(1)(a) involving the following first order reaction mechanism:

$$
\begin{aligned}
\mathrm{D_{OFF}} &\xrightarrow{k_{on}} \mathrm{D_{ON}} \\
\mathrm{D_{ON}} &\xrightarrow{k_{off}} \mathrm{D_{OFF}} \\
\mathrm{D_{ON}} &\xrightarrow{k_1} \mathrm{D_{ON}} + \mathrm{M} \\
\mathrm{M} &\xrightarrow{k_2} \varnothing
\end{aligned}
\tag{1}
$$

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 rate k_1 producing mRNA (M) which degrades at a rate k_2. The stochastic transcriptional dynamics can be cast as a master equation and it is possible to obtain the exact analytical expressions for the mean and variance of mRNAs, the steady state Fano factor is given by [10]:

$$
\frac{(\sigma_m^2)_s}{M_s} = 1 + \frac{k_1 k_{off}}{(k_{off} + k_{on})(k_{off} + k_{on} + k_2)}.
\tag{2}
$$

It is easy to see that $k_{off} > 0$ implies Fano factor > 1 characterizing transcriptional bursts and the steady-state mRNA probability distribution is non-Poissonian. The mRNA burst-size is typically denoted by a parameter $b_m = k_1/k_{off}$ quantifying the average number mRNAs created between the ON and OFF state of the promoter. It is also sometimes referred to as transcriptional efficiency [34]. The variation of the steady state Fano factor with the burst-size
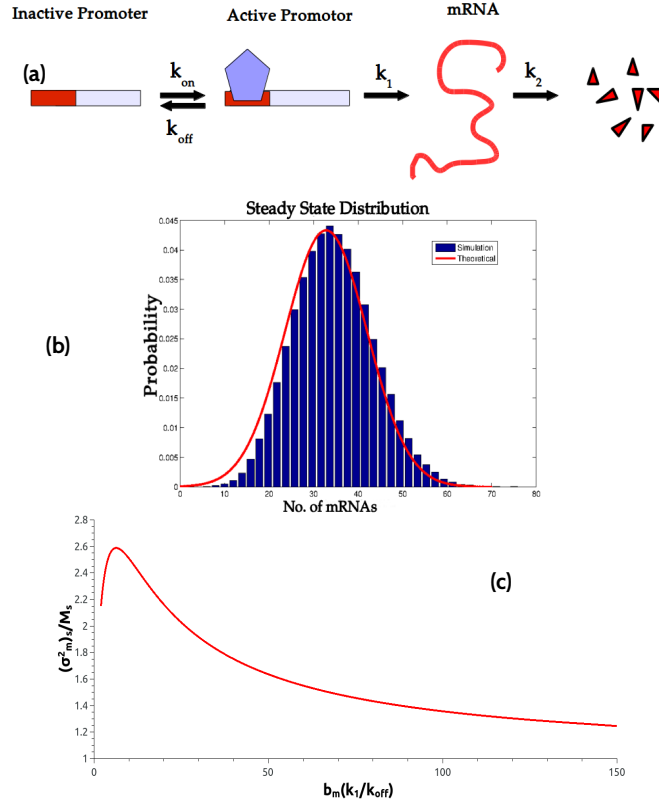
FIG. 1: **(a) A schematic representation of ON-OFF model. (b)Steady state distribution obtained using Gillespie Algorithm averaged over 10^6 realizations for $k_{on} = 3.0$, $k_2 = 0.3$, $k_1 = 20.0$, $k_{off} = 3.1$ and fitted with the steady state PDF $p(M)$ as obtained in Eq.(22). (c) Variation of Fano factor with burst size $b_m = k_1/k_{off}$.**

b_m is shown in Fig.(1)(c). It has been shown that the experimental mRNA expression data shows better agreement with respect to k_{off} modulation [16]. Hence in this study we will vary burst size (b_m) by tuning k_{off} and the mean mRNA production k_1 is held fixed. Burstiness can also be characterized by burst frequency (k_{on}). But it has been experimentally demonstrated that typical expression levels with respect to the variation of the burst frequency is less pronounced [28, 35].

A. The Exact mRNA Propagator from Langevin Dynamics

Let us write the Langevin dynamics corresponding to the ON-OFF model assuming that the number of expressed mRNAs is large. It must be noted that we will be only dealing with intrinsic noise which mainly arises due to fluctuations induced by the stochastic activation and deactivation of the active promoter (D) and production and degradation of the mRNA molecules (M). The Langevin equations are

$$\begin{aligned} \frac{dD}{dt} &= k_{on} - (k_{on} + k_{off})D + \eta_D(t) \\ \frac{dM}{dt} &= k_1 D - k_2 M + \eta_M(t). \end{aligned} \tag{3}$$

The intrinsic fluctuations of this system is modelled by white noise η having the properties:

$$\begin{aligned} \langle \eta_i(t) \rangle &= 0 \\ \langle \eta_i(t)\eta_i(t+\tau) \rangle &= \lambda_i \delta(\tau) \end{aligned} \tag{4}$$

where $i = D$ *or* M. An essential feature of this system that the noise strengths λ_i are independent of the state variables D and M. The expression for noise magnitude of DNA and mRNA are

$\lambda_d = 2k_{on} k_{off}/k'$ and $\lambda_m = 2k_1 k_{on}/k'$ where $k' = k_{on} + k_{off}$. A detailed derivation of the noise strengths is provided in Appendix A and in [36–38]. Note that in steady state the mean number of active promoter is $D_s = k_{on}/k'$ while the mean number of mRNAs is $M_s = k_1 k_{on}/k_2 k'$. Without any loss of generality we can assume that the initial number of active DNA molecules $D_0 = 1.0$. If the promoter stays for longer period in active state (assuming promoter kinetics to be slow), a large number of mRNAs will be synthesized in quick succession i.e., the mRNA dynamics is bursty and can be quantified by the burst parameter(size) $b_m = k_1/k_{off}$.

We are interested in computing the propagator $p(M, t|M_0)$, the probability of having M mRNAs at time t given that the initial number of mRNA was M_0. We start by introducing new variables $q = (D - D_s)/\sqrt{\lambda_D}$ and $v = (M - M_s)/\sqrt{\lambda_M}$ and plugging these in Eq.(3) we obtain the rescaled Langevin equations [39]:

$$
\begin{aligned}
\dot{q}(t) &= -k'q(t) + \xi_D(t) \\
\dot{v}(t) &= -k_2 v(t) + k_q q(t) + \xi_M(t)
\end{aligned}
\tag{5}
$$

where $k_q = k_1 \sqrt{\lambda_D/\lambda_M}$ and $\xi_i (i = D \text{ or } M)$ is scale-free noise. Representing the fluctuations as a vector $\mathbf{u(t)}^T = (q(t), v(t))$ the system of equations Eqs.(5) can be recast in matrix form:

$$
\dot{\mathbf{u}}(t) = -\mathbf{\Upsilon} \cdot \mathbf{u}(t) + \mathbf{\Xi}(t)
\tag{6}
$$

where the drift term has

$$
\mathbf{\Upsilon} = \begin{pmatrix} k' & 0 \\ -k_q & k_2 \end{pmatrix}
\tag{7}
$$

and the noise is represented by $\mathbf{\Xi}^T(\mathrm{t}) = (\xi_D(t), \xi_M(t))$. Let $\rho(q, v, t)$ be the joint probability distribution function which now satisfies the Fokker-Planck Equation (FPE)

$$
\frac{\partial \rho}{\partial t} = \sum_{i,j=1}^{2} \left\{ \Upsilon_{ij} \frac{\partial}{\partial u_i}(u_j \ \rho) + \mathcal{D}_{ij} \frac{\partial^2 \rho}{\partial u_i u_j} \right\}
\tag{8}
$$

where the diffusion matrix $\mathcal{D}$ in our problem is of the form

$$
\mathcal{D} = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.
\tag{9}
$$

We seek the conditional PDF $\rho(q, v, t|q_0, v_0)$ which is the solution of Eq.(8) for the initial condition $\rho(q, v, 0) = \delta(v - v_0)\delta(q - q_0)$ and the natural boundary conditions $\rho(q, v, t) \to 0$ as $|q| \to \infty$ and $|v| \to \infty$. The solution of Eq.(8) will be bivariate Gaussian and can be obtained following [40, 41]. Here we will briefly outline the procedure and introduce the relevant parameters.

Let us write $\mathbf{u} = (q, v)$ in terms of the fluctuations $\delta\mathbf{u} = (\delta q, \delta v)$ and the conditional mean $\bar{\mathbf{u}} = (\bar{q}, \bar{v})$:

$$
\mathbf{u} = \delta\mathbf{u} + \bar{\mathbf{u}}
\tag{10}
$$

and for sharp $\mathbf{u}_0 = (q_0, v_0)$ the mean can be written as

$$
\bar{\mathbf{u}}(t) = \mathsf{G}(t)\mathbf{u}_0
\tag{11}
$$

where we have introduced the Matrix Green function [40, 41]

$$
\mathsf{G}(t) = \exp(-\mathbf{\Upsilon}t) = \begin{pmatrix} e^{-k't} & 0 \\ \frac{k_q(e^{-k_2 t} - e^{-k't})}{k' - k_2} & e^{-k_2 t} \end{pmatrix}.
\tag{12}
$$

The covariance matrix can be defined in terms of the Green function:

$$
\boldsymbol{\sigma}(t) = \begin{bmatrix} \sigma_{11}(t) & \sigma_{12}(t) \\ \sigma_{21}(t) & \sigma_{22}(t) \end{bmatrix} = 2 \int_0^t \mathsf{G}(t')\mathcal{D}\mathsf{G}^{\mathrm{T}}(t')dt'
\tag{13}
$$

and the covariance matrix elements are given by:

$$\sigma_{11}(t) = \frac{1}{2k'}\Big[1 - e^{-2k't}\Big]$$

$$\sigma_{12}(t) = \sigma_{21}(t) = C_1\Big[\frac{1}{k'+k_2}[1 - e^{-(k_2+k')t}] - \frac{1}{2k'}[1 - e^{-2k't}]\Big] \tag{14}$$

$$\sigma_{22}(t) = C_1^2\Big[\frac{1}{2k'}[1 - e^{-2k't}] + \frac{1}{2k_2}[1 - e^{-2k_2t}] - \frac{2}{k'+k_2}[1 - e^{-(k'+k_2)t}]\Big] + \frac{1}{2k_2}[1 - e^{-2k_2t}]$$

where we have defined $C_1 = k_q/(k' - k_2)$. Finally, the formal solution of FPE given in Eq.(8) is a bivariate Gaussian having the following form:

$$\rho(\mathbf{u}, t|\mathbf{u}_0) = \frac{1}{2\pi\sqrt{\Delta(t)}}\exp\Big[-\frac{1}{2}\delta\mathbf{u(t)}^{\mathrm{T}}\,\mathsf{S}\,\delta\mathbf{u(t)}\Big] \tag{15}$$

where $S(t)$ is the inverse of the covariance matrix:

$$\boldsymbol{\sigma}(t)^{-1} = \mathsf{S}(t) = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} = \frac{1}{\Delta(t)}\begin{bmatrix} \sigma_{22} & -\sigma_{12} \\ -\sigma_{12} & \sigma_{11} \end{bmatrix} \tag{16}$$

and the determinant $\Delta(t) = \det(\boldsymbol{\sigma}(t)) = \sigma_{11}\sigma_{22} - \sigma_{12}^2$.

Noting that $\delta\mathbf{u} = \mathbf{u} - \bar{\mathbf{u}}$ the solution in Eq.(15) can be equivalently written as:

$$\rho(q, v, t|q_0, v_0) = \frac{1}{2\pi\sqrt{\Delta(t)}}\exp\Big[-\Big\{\frac{1}{2}S_{11}(t)(\delta q)^2 + S_{12}(t)(\delta q)(\delta v) + \frac{1}{2}S_{22}(t)(\delta v)^2\Big\}\Big]. \tag{17}$$

Integrating Eq.(17) with respect to q will yield the marginal probability distribution $p(v, t)$ which is nothing but conditional PDF of the rescaled mRNA (v), which has the form:

$$p(v, t) = \int_{-\infty}^{\infty} dq\rho(q, v, t) = \frac{1}{\sqrt{2\pi\sigma_{22}}}\exp\big(-\frac{(v - \bar{v})^2}{2\,\sigma_{22}}\big). \tag{18}$$

Introducing the burst size $b_m = k_1/k_{off}$ and $\phi = k_2/k_{off}$ the variance can be written as

$$\sigma_{22}(t) = \frac{b_m\Big[(k' - k_2)^2 - k_2k'(e^{-k't} - e^{-k_2t})^2 - (k_2e^{-k't} - k'e^{-k_2t})^2\Big]}{2k'\phi(1 + k_{on}/k_{off} + \phi)(k' - k_2)^2} + \frac{1 - e^{-2k_2t}}{2k_2}. \tag{19}$$

Now, we can convert this expression into the actual mRNA propagator, $p(M, t|M_0)$, by multiplying it by the Jacobian of the transformation (i.e., $dv/dM = 1/\sqrt{\lambda_M}$) and writing v and $\bar{v}$ in terms of M and M_0. Upon substitution we obtain the $p(M, t|M_0)$ as:

$$p(M, t|M_0) = \frac{1}{\sqrt{2\pi\sigma_m^2(t)}}\exp\Big\{-\frac{(M - \bar{M}(t))^2}{2\sigma_m^2(t)}\Big\} \tag{20}$$

$$\bar{M} = b_m\Big[\frac{(D_0 - D_s)(e^{-k_2t} - e^{-k't})}{(1 + \frac{k_{on}}{k_{off}} - \phi)} - \frac{k_{on}}{\phi\,k'}(e^{-k_2t} - 1)\Big] + M_0e^{-k_2t} \tag{21}$$

and the mRNA variance is $\sigma_m^2(t) = \lambda_M\sigma_{22}(t)$. The detailed procedure to compute the mean and the variance is given in Appendix B.

B. Steady state PDF

In the steady state $\bar{M}(t) \to \bar{M}_s$, $\sigma_m^2(t) \to (\sigma_m^2)_s$ the probability distribution function (PDF) in Eq.(20) becomes:

$$p(M) = \frac{1}{\sqrt{2\pi(\sigma_m^2)_s}}\exp\Big(-\frac{(M - M_s)^2}{2\,(\sigma_m^2)_s}\Big). \tag{22}$$
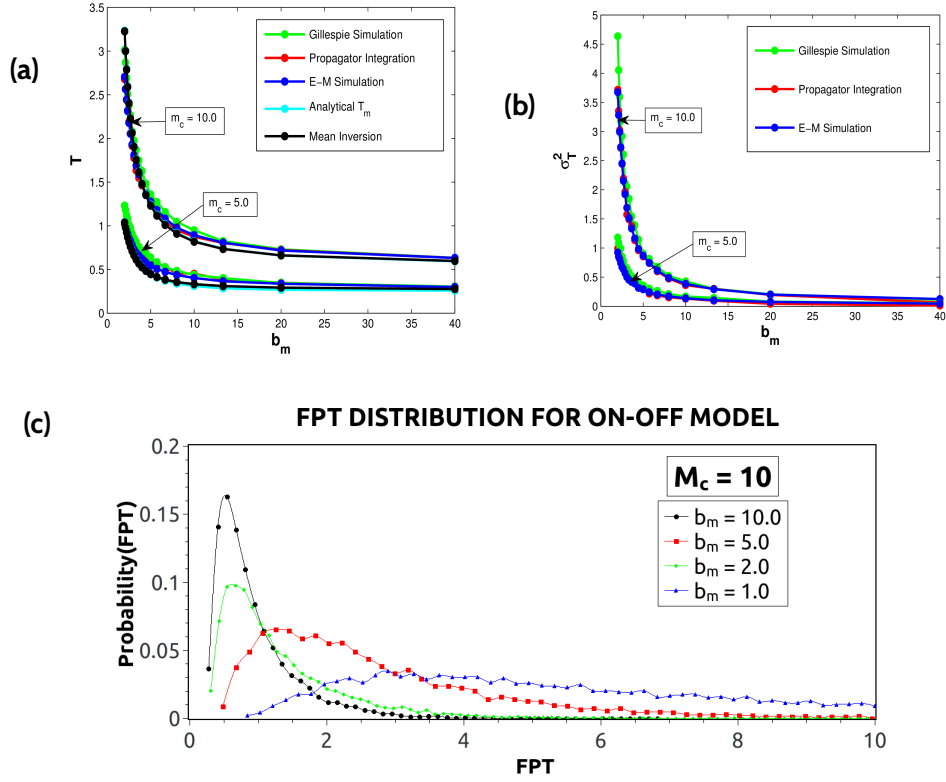
FIG. 2: **(a)MFPT for ON-OFF model plotted as a function of burst size ($b_m = k_1/k_{off}$). Mean Inversion (Black), Propagator Integration (Red), analytical $\tilde{T}_m$ (Cyan) Eq.(26), Euler-Maruyama method (Blue) and Gillespie algorithms (Green) and (b) Variance of FPT for the ON-OFF model have been plotted for two threshold values $M_c = 5$ and 10. Other parameter values are set at $k_{on} = 3.0$, $k_1 = 20.0$, $k_2 = 0.3$. (c) Probability distribution of FPT for different $b_m = 1, 2, 5$ and 10 given $M_c = 10$(other parameters are same as (a) and (b)). The EM simulations are done with time-step of 10^{-3}. In Gillespie simulations we average over 10^4 realizations.**

The steady state Fano factor is now given by:

$$\frac{(\sigma_m^2)_s}{M_s} = 1 + \frac{k_1 \, k_{off}}{k' \, (k' + k_2)} \tag{23}$$

substituting $k' = k_{on} + k_{off}$ we recover the exact expression Eq.(2).

C.   Mean First Passage Time for ON-OFF model

Equipped with the propagator obtained in the previous section we can now estimate the mean first passage time taken by mRNAs to reach a threshold M_c starting from an initial number M_0. The probability that at time t mRNA numbers are such that $0 \leq M \leq M_c$ is $G(M_0, t) = \int_0^{M_c} p(M, t|M_0)dM$. Again, if the duration for which mRNAs are in $[0, M_c]$ is T, $Prob(T \geq t) = \int_0^{M_c} p(M, t|M_0)dM = G(M_0, t)$. Following Gardiner[42], MFPT can be defined as

$$T_m = -\int_0^\infty t \, \partial_t G(M_0, t)dt = \int_0^\infty G(M_0, t)dt. \tag{24}$$

We numerically evaluate the double integral using Mathematica and plot MFPT versus burst parameter b_m as shown in Fig.(2) (*Propagator Integration*). However, it is not possible to evaluate

the integral in closed form, hence, the dependence of MFPT on other parameters can only be numerically established.

Alternatively, having obtained the exact time dependence of the mean mRNA in Eq.(21) and setting $\bar{M} = M_c$ and initial $M_0 = 0$ we can solve T_m algebraically from:

$$M_c = b_m \Big[\frac{(D_0 - D_s)(e^{-k_2 T_m} - e^{-k' T_m})}{(1 + \frac{k_{on}}{k_{off}} - \phi)} - \frac{k_{on}}{\phi k'}(e^{-k_2 T_m} - 1) \Big]. \tag{25}$$

MFPT thus obtained is also shown in Fig.(2) (*Mean Inversion*).

It is now possible to obtain an analytical expression for MFPT with the assumptions $k_{on} \gg k_2$ and $k_{off} \gg k_2$. k_2 being the degradation rate of mRNAs this is a reasonable assumption and allows us to ignore $\phi = k_2/k_{off}$ and $e^{-k' T_m}$. We thus have a simplified expression of MFPT

$$\tilde{T}_m = \frac{-1}{k_2} \log \Big[\frac{M_c/b_m - D_s/\phi}{k_{off}/k' - k_{on}/k_2} \times \frac{k'}{k_{off}} \Big]. \tag{26}$$

Note that large k_{off} implies low burstiness (b_m) and in this regime the agreement of $\tilde{T}_m$ is good with the above methods. We have also estimated MFPT from simulating the stochastic dynamics by Gillespie algorithm [43, 44] and numerically integrating the Langevin equations by Euler-Maruyama algorithm [45, 46]. All the estimates are shown in Fig.(2)(a) for different choices of threshold M_c.

D. Variance of FPT for ON-OFF model

It is also very interesting to explore the dependence of other moments of the FPT distribution on transcriptional burst size (b_m). The other moments[42] of the FPT can be expressed as

$$\langle (T_m)^n \rangle = \int_0^\infty t^{n-1} G(M_0, t) dt. \tag{27}$$

Using this definition, the variance of the FPT can be given as

$$(\sigma_T)^2 = \langle (T_m)^2 \rangle - \langle T_m \rangle^2 \tag{28}$$

and enumeration of this yields us the theoretical curve for $(\sigma_T)^2$ in Fig.(2)(b). We also show the numerical results obtained by Gillespie and E-M algorithm. Moreover, we numerically obtain the FPT distribution for different transcriptional burst sizes($b_m = 1, 2, 5$ and 10) given a fixed threshold value($M_c = 10$) and have been shown in Fig.(2)(c).

III. TWO-STAGE MODEL

Now, we analyze the two-stage model, where, transcription is followed by translation [1, 5, 6, 47]. In this model mRNA statistics is Poissonian while the number of proteins produced from a single mRNA can be bursty. This feature of translational bursting is the major cause of noise in prokaryotes [13, 14]. The schematic given in Fig.(3)(a) shows that the mRNA molecules (M) are first transcribed from a double stranded DNA template at a rate k_1 and proteins (P) are created at a translation rate k_3. mRNAs and proteins degrade with rate constants k_2 and k_4, respectively. The two-stage dynamics can be described by the Langevin equations

$$\begin{aligned} \frac{dM}{dt} &= k_1 - k_2 M + \eta_M(t) \\ \frac{dP}{dt} &= k_3 M - k_4 P + \eta_P(t). \end{aligned} \tag{29}$$

The noise statistics are same as in Eq.(4) except that $i = M \ or \ P$. The noise magnitude for mRNA and protein are $\lambda_M = 2k_1$ and $\lambda_P = 2k_1 k_3/k_2$. A detailed derivation of λ_i is given in Appendix A and in [11]. We introduce the protein burst size $b_p = k_3/k_2$, which essentially measures the average number of protein produced in one transcript of mRNA. This burstiness can be thought of as a measure of enhancement of transcriptional noise by translation. The
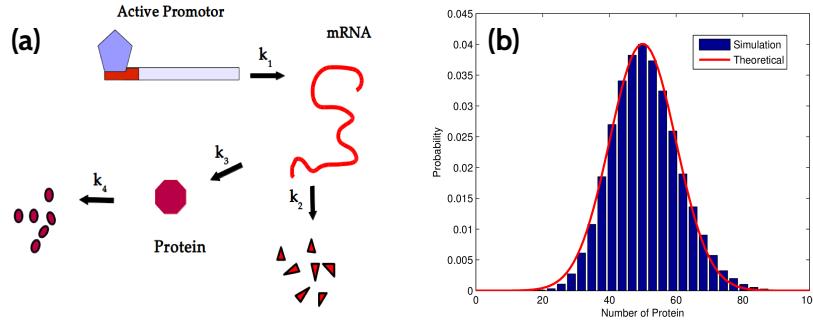
FIG. 3: **(a) Schematic representation of two-stage model (b) Steady state histogram obtained for the parameter values** $k_1 = 0.1$, $k_2 = 0.1$, $k_3 = 0.1$, $k_4 = 0.002$ **using Gillespie Algorithm averaged over** 10^4 **realizations and fitted with the steady state PDF** $p(P)$ **as obtained in Eq.(31)**

rescaled Langevin equations for the two-stage model [39] can be written as:

$$
\begin{aligned}
\dot{q}(t) &= -k_2 q(t) + \xi_M(t) \\
\dot{v}(t) &= -k_4 v(t) + k_q q(t) + \xi_P(t)
\end{aligned}
\tag{30}
$$

where $q = (M - M_s)/\sqrt{\lambda_M}$, $v = (P - P_s)/\sqrt{\lambda_P}$ and $k_q = k_3 \sqrt{\lambda_M/\lambda_P}$. We see that the structure of coupled equations Eqs.(30) is exactly same as Eqs.(5). Hence, following the approach mentioned in the previous section will give the protein propagator $p(P, t|P_0)$ similar to the mRNA propagator in Eq.(20):

$$
p(P, t|P_0) = \frac{1}{\sqrt{2\pi\sigma_p^2(t)}} \exp\left\{ -\frac{(P - \bar{P}(t))^2}{2\sigma_p^2(t)} \right\}
\tag{31}
$$

with mean protein

$$
\bar{P}(t) = b_p \left[\frac{(M_0 - M_s)(e^{-k_4 t} - e^{-k_2 t})}{(1 - \varphi)} + \frac{k_1}{k_4}(1 - e^{-k_4 t}) \right] + P_0 e^{-k_4 t}
\tag{32}
$$

and variance

$$
\sigma_p^2(t) = \lambda_P \left[\frac{b_p \left[(k_2 - k_4)^2 - k_4 k_2 (e^{-k_2 t} - e^{-k_4 t})^2 - (k_4 e^{-k_2 t} - k_2 e^{-k_4 t})^2 \right]}{2k_4(1 + \varphi)(k_2 - k_4)^2} + \frac{1 - e^{-2k_4 t}}{2k_4} \right]
\tag{33}
$$

where $\varphi = k_4/k_2$ is typically a small quantity as mRNA is more unstable than protein in most of the biological systems.

A. Transient Dynamics and Steady State

It is interesting to study the transient dynamics of the two-stage model. We plot the Fano factor for different choice of φ. For $\varphi \to 0$, we recover the results as obtained by Thattai *et. al* [1] and in this limit we obtain the transient Fano factor:

$$
\frac{\sigma_p^2(t)}{\bar{P}(t)} = (1 + b_p)\left(\frac{1 - e^{-2k_4 t}}{1 - e^{-k_4 t}} \right)
\tag{34}
$$

For small times and large b the $\sigma_P^2(t)/\bar{P}(t) \cong 2b$ which is almost double of the steady state value.
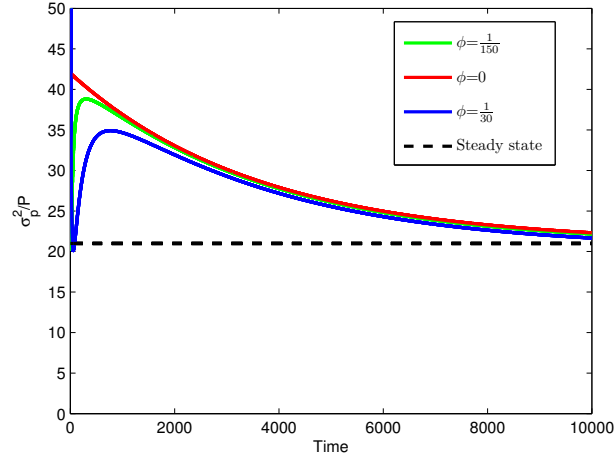
FIG. 4: **Transient Fano factor for two-stage model for constant protein degradation rate $k_4 = 1/3600$. The protein burst parameter $b_p = 20$ and transcription rate $k_1 = 0.01$. For different degradation rate of mRNA, $k_2 = 1/120, 1/24,$ and $10,000$ correspond to $\varphi = 1/30,\ 1/150$ and $2 \times 10^{-8} \approx 0$.**
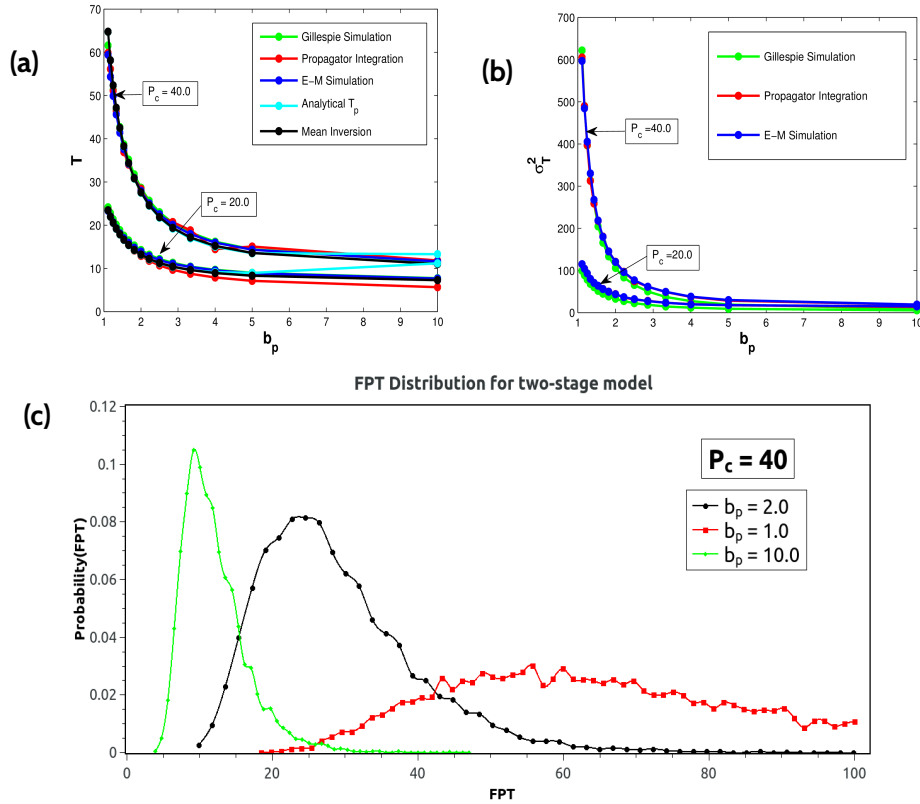

FIG. 5: **(a) MFPT for two-stage model shown as a function of protein burst size ($b_p = k_3/k_2$). Mean Inversion (Black), Propagator Integration (Red), Analytical $\tilde{T}_p$ (Cyan) Eq.(38), Euler-Maruyama method (Blue) and Gillespie algorithm (Green) (b) Variance of FPT for the two-stage model have been plotted for two threshold values $P_c = 20$ and 40. Other parameters were set at $k_1 = 1.0,\ k_3 = 1.0,\ k_4 = 0.02$. (c) Probability distribution of FPT for different $b_p = 1, 2$ and 10 given $P_c = 40$ (other parameters are same as (a) and (b)). The EM simulations are done with time-step of 10^{-3}. In Gillespie simulations we average over 10^4 realizations.**

Again taking the steady state limit first, the Fano factor assumes the value

$$\frac{(\sigma_p^2)_s}{\bar{P}_s} = 1 + \frac{k_3}{(k_2 + k_4)} = 1 + \frac{b_p}{1 + \varphi} \tag{35}$$

which matches exactly with the expression obtained by Thattai $et~al.$ [1] by solving the Master Equation describing the two-stage model. Also , for the case when $k_2 \ll k_4$, i.e. $\varphi \to 0$ the Fano factor reduces to

$$\frac{(\sigma_p^2)_s}{P_s} \approx 1 + b_p. \tag{36}$$

Fig.(4) shows the time-dependent behaviour of Fano factor obtained from the exact analytical expressions of the mean and the variance for different small values of φ. Moreover, Eqs.(34),(35)and (36) confirm that we can easily obtain the previously known results [1] as limiting cases from our analytical expression of the propagator which we will now use to calculate transient quantities like moments of FPT.

B. Moments of First Passage Time for Two-stage model

As shown for ON-OFF model, it is possible to estimate the MFPT for two-stage model from Eq.(31) by $propagator~integration$ and alternatively by $mean~inversion$ by setting $\bar{P} = P_c$ and initial $M_0 = P_0 = 0$ and algebraically solving T_p from:

$$P_c = b_p \left[\frac{(M_0 - M_s)(e^{-k_4 T_p} - e^{-k_2 T_p})}{(1 - \varphi)} + \frac{k_1}{k_4}(1 - e^{-k_4 T_p}) \right] \tag{37}$$

Numerical estimates using Mathematica is shown in Fig.(5)(a) for different values of b_p. It must be noted that b_p values were changed by varying rate constant k_2. One can get a closed form expression for MFPT under the assumption $\varphi \ll 1$:

$$\tilde{T}_p = \frac{-1}{k_4} \log \left[\frac{k_2(k_1 b_p - P_c k_4)}{b_p k_1(k_2 + k_4)} \right] \tag{38}$$

which is in good agreement with numerical estimates for low burst size (b_p). Similar results are also obtained by A.Singh $et.~al.$ [48] for the case of stable proteins ($k_4 = 0$). Our approach is more general and for the case of low burst size and $k_4 \to 0$ we easily obtain

$$\begin{aligned} \tilde{T}_p &= \frac{-1}{k_4} \left[\log(\frac{k_2}{k_2 + k_4}) + \log\left(1 - \frac{P_c \cdot k_4}{k_1 \cdot b_p}\right) \right] \\ &\approx \frac{P_c}{k_1 b_p} \end{aligned}$$

in agreement with Eq. 18 of A.Singh $et.~al.$ [48].

The theoretical curve for the variance of the FPT for this model is obtained by evaluating Eq.(28) but now for proteins and compared with the stochastic simulations as shown in Fig.(5)(b). Also, the FPT distribution for two-stage model for a fixed cut-off($P_c = 40$) is obtained numerically using Euler-Maruyama algorithm for different translational burst sizes($b_p = 1, 2$ and 10) in Fig.(5)(c).
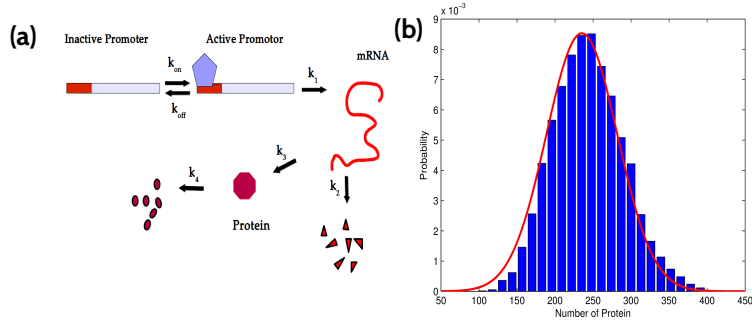
IV. THREE-STAGE MODEL

FIG. 6: **(a)Schematic representation of three-stage model (b) Steady state histogram obtained for the parameter values** $k_{on} = 0.5$, $k_{off} = 0.01$, $k_1 = 1.0$, $k_2 = 0.1$, $k_3 = 1.2$, $k_4 = 0.05$ **using Gillespie Algorithm averaged over** 10^4 **realizations and fitted with the steady state PDF** $p(P)$ **as obtained in Eq.**(C-34)

Three-stage model is the most realistic among all stochastic gene expression models as it takes into account all the features(promoter transition, transcription and translation) and have successfully reproduced the single-gene expression data of eukaryotes and mammalians[2–4, 35, 49]. Three-stage model captures both transcriptional as well as translational bursting but we will focus on quantifying the effect of transcriptional bursting which is of more biological importance [13]. The schematic representation of the model is shown in Fig.(6)(a) and can be modelled by a master equation which has been solved analytically [5]. An equivalent Langevin description of the three-stage model[50] can be given as

$$
\begin{aligned}
\frac{dD}{dt} &= k_{on} - (k_{on} + k_{off})D + \eta_D(t) \\
\frac{dM}{dt} &= k_1 D - k_2 M + \eta_M(t) \\
\frac{dP}{dt} &= k_3 M - k_4 P + \eta_P(t).
\end{aligned}
\tag{39}
$$

The intrinsic fluctuations of this system are modelled by white noise terms(η) having the following properties:

$$
\begin{aligned}
\langle \eta_i(t) \rangle &= 0 \\
\langle \eta_i(t) \eta_i(t+\tau) \rangle &= \lambda_i \delta(\tau)
\end{aligned}
\tag{40}
$$

where λ_i are noise magnitudes for $i = D, M, P$. λ_D and λ_M are of the same form as in ON-OFF model and $\lambda_P = 2k_3 k_1 k_{on}/k_2 k'$ [see Appendix A].

Applying the formalism developed for the previous models we can write the rescaled Langevin equations as

$$
\begin{aligned}
\dot{q}(t) &= -k' q(t) + \xi_D(t) \\
\dot{v}(t) &= -k_2 v(t) + k_q q(t) + \xi_M(t) \\
\dot{y}(t) &= -k_4 y(t) + k_v v(t) + \xi_P(t)
\end{aligned}
\tag{41}
$$

where

$$
\begin{aligned}
q &= \frac{(D - D_s)}{\sqrt{\lambda_D}} = (D - k_{on}/k')/\sqrt{\lambda_D} \\
v &= (M - M_s)/\sqrt{\lambda_M} = (M - k_1 k_{on}/k_2 k')/\sqrt{\lambda_M} \\
y &= (P - P_s)/\sqrt{\lambda_P} = (P - k_1 k_3 k_{on}/k_2 k_4 k')/\sqrt{\lambda_P}
\end{aligned}
\tag{42}
$$

are the rescaled parameters for DNA, mRNA and protein respectively. Also, $k_q = k_1\sqrt{\lambda_D/\lambda_M}$ and $k_v = k_3\sqrt{\lambda_M/\lambda_P}$. Representing the fluctuations as a vector $\mathbf{u(t)}^T = (q(t), v(t), y(t))$ the system of equations Eqs.(41) can be recast in matrix form:

$$
\dot{\mathbf{u}}(t) = -\mathbf{\Upsilon} \cdot \mathbf{u}(t) + \mathbf{\Xi}(t)
\tag{43}
$$

where the drift term is $\boldsymbol{\Upsilon} = \begin{pmatrix} k' & 0 & 0 \\ -k_q & k_2 & 0 \\ 0 & -k_v & k_4 \end{pmatrix}$ and the noise is represented by $\boldsymbol{\Xi}(t) = \begin{pmatrix} \xi_D(t) \\ \xi_M(t) \\ \xi_P(t) \end{pmatrix}$. Let $\rho(q, v, y, t | q_0, v_0, y_0)$ be the joint probability distribution function which now satisfies the Fokker-Planck Equation (FPE) having the same form as Eq.(8) except for the fact that index i and j run till 3. Also, diffusion matrix $\mathcal{D} = \mathbb{I}/2$, where $\mathbb{I}$ is a 3×3 identity matrix. Similarly the matrix Green function can be written as

$$\mathsf{G}(t) = \exp(-\boldsymbol{\Upsilon}t) = \begin{bmatrix} e^{-k't} & 0 & 0 \\ k_q \frac{(e^{-k_2 t} - e^{-k't})}{k' - k_2} & e^{-k_2 t} & 0 \\ G_{31} & k_v \frac{(e^{-k_4 t} - e^{-k_2 t})}{k_2 - k_4} & e^{-k_4 t} \end{bmatrix} \tag{44}$$

where

$$G_{31} = k_q k_v \frac{\{(k_2 - k_4)e^{-k't} + (k_4 - k')e^{-k_2 t} + (k' - k_2)e^{-k_4 t}\}}{(k_2 - k_4)(k_2 - k')(k_4 - k')}.$$

A. Moments of FPT for Three-stage Model

We write the fluctuations in the rescaled variables as in Eq.(10) for the three-stage model where $\bar{\mathbf{u}}(t) = \mathsf{G}(t)\mathbf{u}_0$, $\delta\mathbf{u}^T = (\delta q, \delta v, \delta y)$ and $\mathbf{u}_0^T = (q_0, v_0, y_0)$. The covariance matrix, $\boldsymbol{\sigma}(t)$, is given by Eq.(13) and its elements are given explicitly in Appendix [Eq.(C-33)]. The formal solution of FPE is a multivariate Gaussian distribution having the following form:

$$\rho(\mathbf{u}, t | \mathbf{u}_0) = \frac{1}{\sqrt{(2\pi)^3 \Delta(t)}} \exp\left[-\frac{1}{2} \delta\mathbf{u(t)}^T \mathsf{S} \, \delta\mathbf{u(t)} \right] \tag{45}$$

We used Mathematica to calculate the determinant, Δ, and the inverse, $\mathsf{S}(t) = \boldsymbol{\sigma}^{-1}(t)$. We integrate out q and v from $\rho(q, v, y, t | q_0, v_0, y_0)$ and converted $\rho(y, t | y_0)$ to $p(P, t | P_0))$ and finally got the exact propagator $p(P, t | P_0)$. The final form of the propagator is not shown here for brevity and is given in Appendix[(Eq.(C-34))]. In the steady state limit $p(P, t | P_0)$ is in agreement with the steady state histogram obtained using Gillespie algorithm and have been shown in Fig.(6)(b). MFPT is calculated using Eq.(24) and the theoretical curve(propagator integration) is obtained. Similarly, variance, σ_T^2, is computed using Eq.(28) and both MFPT and variance (of FPT) behaviour with transcriptional burst size(b_m) for two protein cutoffs ($P_c = 75$ and 50) are shown in Fig.(7)(a) and (b) respectively along with the stochastic simulation results. The ratio of standard deviation of FPT, σ_T, to the MFPT (T), i.e., coefficient of variation(CV) is calculated and plotted with b_m in Fig.(7)(c). Lastly, we show the FPT distribution for different transcriptional burst size ($b_m = 2.0, 2.5, 5, 10$) for a fixed protein threshold($P_c = 75$) in Fig.(7)(d).

V. CONCLUSION

In this paper we obtain a complete statistical description of stochastic bursty gene expression models starting from a Langevin description. The Langevin approach is strictly valid when the number of relevant molecules (mRNA or protein) per cell are present in large numbers. Applying Langevin approach for ON-OFF model is an idealisation as the number of mRNAs in a cell are typically small in number. However, we have used the ON-OFF model as a test problem to demonstrate an exact formalism to derive the propagator with which MFPT can be calculated. Since the number of copies of proteins are typically present in large numbers it is convenient to work with Langevin equations and obtaining the exact probability distributions is straightforward for two-stage and three-stage models. Our analytical results match the steady state behaviour obtained from an equivalent master equation approach for ON-OFF and two-stage models as well as three-stage model. The ON-OFF model exhibits bursty dynamics such that the mRNA Fano factor varies non-monotonically with respect to burstiness. Since mRNA burstiness $b_m \propto 1/k_{off}$, for small b_m gene is mostly in OFF state therefore transcribing less with less fluctuations i.e., Fano factor ≈ 1. Again, in the limit of high burstiness transcription increases manifold such that Fano factor $\to 1$. The MFPT, on the other hand, always decreases monotonically with burstiness. In the regime of low burstiness, where fluctuations and mean are of same order, the yield reaches a threshold at larger times as compared to that when fluctuations are appreciable. Again, in the limit of high burstiness, frequent transcriptions raise the basal level lowering the MFPT even further. We also observe similar behaviour for the variance computed from the first passage time distribution, as burst parameter is increased the spread in the FPT fluctuations decreases. We can conclude that burstiness not only enables mRNAs/proteins to reach a threshold faster but also with more reliability. This feature
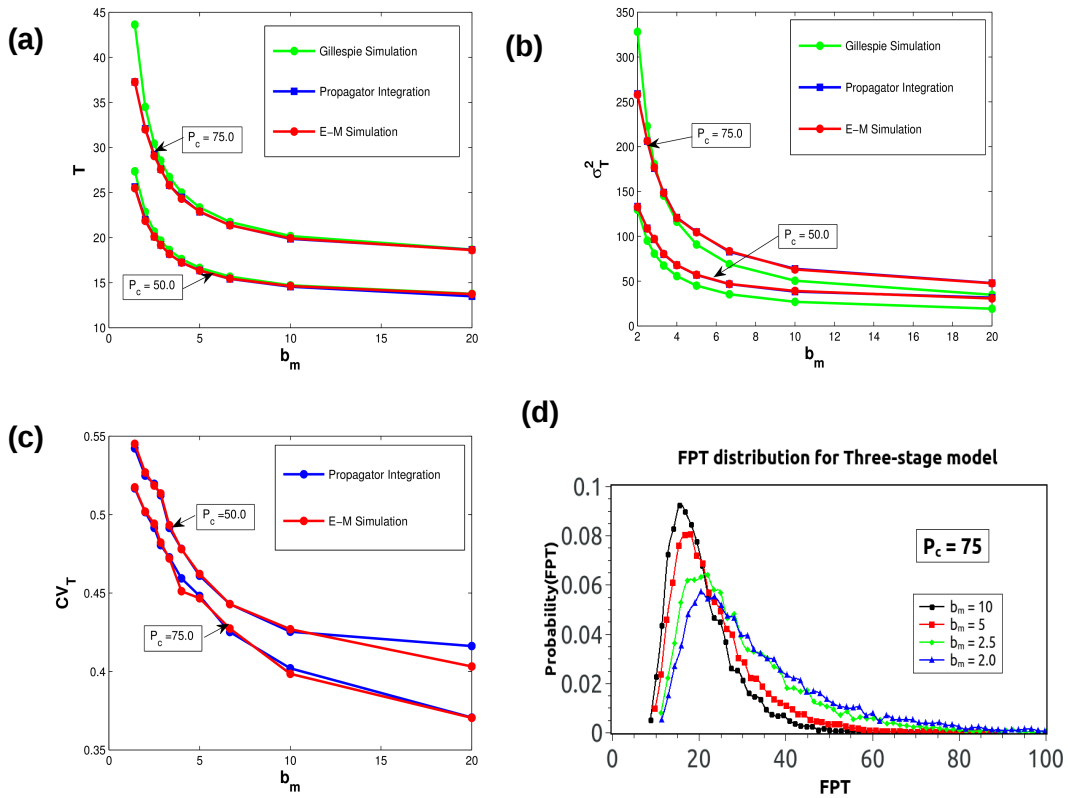
FIG. 7: **MFPT for three-stage model shown as a function of transcriptional burst size ($b_m = k_1/k_{off}$). Propagator Integration (blue), EM-method (red) and Gillespie algorithm(Green). (b) Variance of FPT as function of b_m for the three-stage model have been plotted for two threshold values $P_c = 50$ and 75. (c) Coefficient of variation (CV) is plotted as function of b_m. The EM simulations are done with time-step of 10^{-3}. In Gillespie simulations we average over 10^4 realizations. (d) Probability distribution of FPT for different $b_m = 2, 2.5, 5$ and 10 given $P_c = 75$. Other parameters were set at $k_{on} = 0.5$, $k_1 = 1.0$, $k_2 = 0.2$, $k_3 = 1.2$, $k_4 = 0.05$.**

is common across all the three models we have studied showing either transcriptional or translational bursts. Our exact calculation confirms this conjecture and is supported by extensive numerical estimation of the moments of the FPT from the time-dependent PDF.

In real biological systems quantities like MFPT can throw light into the underlying cellular mechanisms. The transcriptional efficiency and timing mechanism of a cell can be related as had been studied in lysis time variation of $\lambda-$Phage [21]. Lysis time has been framed as first passage time problem and MFPT was defined as the time taken by holin (lysis protein) levels to reach a critical cut-off accordingly deciding the active-inactive state of the virus. Moreover, in the case of active viral replication of HIV protein *Tat* it has been observed that large (small) burst size results in productive (unproductive) viral replication [28, 35]. The source of noise and the effects of transcriptional burstiness on viral latency can be understood as a first passage time problem in the associated protein expression levels. The analytical and numerical method presented in this work can be extended to similar small gene regulatory motifs and help us to precisely understand how burstiness can control the timing mechanism of a cell.

APPENDIX

A. DERIVATION OF NOISE MAGNITUDES

ON-OFF Model

The system of Langevin equations for ON-OFF model can be written as

$$
\begin{aligned}
\dot{D} &= k_{on}D_0 - k_{off}D + \eta_D(t) \\
\dot{M} &= k_1 D - k_2 M + \eta_M(t)
\end{aligned}
\tag{A-1}
$$

where $D + D_0 = 1$ and stochastic variables η_D and η_M, satisfy the following conditions

$$
\begin{aligned}
\langle \eta_i(t) \rangle &= 0 \\
\langle \eta_i(t)\eta_i(t+\tau) \rangle &= \lambda_i \delta(\tau) \quad (i = D, M).
\end{aligned}
\tag{A-2}
$$

We will derive the exact expressions for λ_d and λ_m. Let us consider our system to be very close to the steady state, $\langle D \rangle = k_{on}/k'$ and $\langle M \rangle = k_1 k_{on}/k_2 k'$, and we assume a time interval δt which is so small that at most only a single reaction can happen, then $\eta_i \delta t = +1, 0, -1$. Let us define the probability that DNA count changes by an amount j and mRNA number changes by an amount k as $P(j,k) = \mathcal{P}(\eta_D \delta t = j, \eta_M \delta t = k)$. From (A-1), one can have

$$
\begin{aligned}
P(+1,0) &= k_{on}\langle D_0 \rangle \delta t \\
P(-1,0) &= k_{off}\langle D \rangle \delta t \\
P(0,+1) &= k_1 \langle D \rangle \delta t \\
P(0,-1) &= k_2 \langle M \rangle \delta t \\
P(0,0) &= 1 - (k_{on}\langle D_0 \rangle \delta t + k_{off}\langle D \rangle \delta t + k_1 \langle D \rangle \delta t + k_2 \langle M \rangle \delta t)
\end{aligned}
\tag{A-3}
$$

and all other $P(j,k) = 0$. These probabilities can be used to calculate the first moment

$$
\begin{aligned}
\langle \eta_D \delta t \rangle &= (+1) \times k_{on}\langle D_0 \rangle \delta t + (-1) \times k_{off}\langle D \rangle \delta t \\
&\quad + (0) \times (1 - (k_{on}\langle D_0 \rangle \delta t + k_{off}\langle D \rangle \delta t + k_1 \langle D \rangle \delta t + k_2 \langle M \rangle \delta t)) \\
&= 0 \\
\langle \eta_M \delta t \rangle &= 0
\end{aligned}
$$

The second moment can be calculated as

$$
\begin{aligned}
\langle \eta_D^2 \delta t^2 \rangle &= (+1)^2 \times k_{on}\langle D_0 \rangle \delta t + (-1)^2 \times k_{off}\langle D \rangle \delta t = 2\, k_{off}\langle D \rangle \delta t \\
\langle \eta_M^2 \delta t^2 \rangle &= (+1)^2 \times k_1 \langle D \rangle \delta t + (-1)^2 \times k_2 \langle M \rangle \delta t = 2k_2 \langle M \rangle \delta t \\
\langle \eta_D \eta_M \rangle &= 0.
\end{aligned}
$$

Assuming that the steady state values are large enough, then the fluctuations away from steady state values will also be very small in comparison for all time. In this scenario, the probability conditions will be approximately obeyed for all δt time intervals and as a result, $\eta_i(t_1)$ will be uncorrelated with $\eta_i(t_2)$ where $|t_2 - t_1| > \delta t$. So, we use delta function to describe the lack of correlation for different time $(t_1 \neq t_2)$. Finally, we obtain the following properties:

$$
\begin{aligned}
\langle \eta_D \rangle &= \langle \eta_M \rangle = 0 \tag{A-4} \\
\langle \eta_D(t_1)\eta_D(t_2) \rangle &= 2\, k_{off}\langle D \rangle \delta(t_1 - t_2) = \lambda_D \delta(t_1 - t_2) \tag{A-5} \\
\langle \eta_M(t_1)\eta_M(t_2) \rangle &= 2\, k_2 \langle M \rangle \delta(t_1 - t_2) = \lambda_M \delta(t_1 - t_2) \tag{A-6} \\
\langle \eta_D(t_1)\eta_M(t_2) \rangle &= 0. \tag{A-7}
\end{aligned}
$$

Thus, $\lambda_d = 2k_{off}k_{on}/k'$ and $\lambda_m = 2k_1 k_{on}/k'$.

Two-stage Model

We derive the noise magnitude for Two-stage model using Ozbudak *et.al*[11] approach for which system of Langevin equations is written as

$$\frac{dM}{dt} = k_1 - k_2 M + \eta_M(t)$$
$$\frac{dP}{dt} = k_3 M - k_4 P + \eta_P(t). \tag{A-8}$$

Noise statistics for this model is

$$\langle \eta_i(t) \rangle = 0$$
$$\langle \eta_i(t)\eta_i(t + \tau) \rangle = \lambda_i \delta(\tau) \tag{A-9}$$

where $i = M$ or P. λ_i are the noise magnitudes and will be chosen such that they give results consistent with the steady state Poisson distribution. In the steady state mRNA number is $M_s = k_1/k_2$. If we set $M = M_s + \delta M$ and expand near steady state, we will get

$$\frac{d}{dt}(\delta M) + k_2 \delta M = \eta_M. \tag{A-10}$$

Applying Fourier transform $(f(t) = \int \frac{1}{2\pi} e^{-i\omega t} f(\omega) d\omega)$ we obtain

$$\delta M(\omega)\Big(- i\omega + k_2 \Big) = \eta_M(\omega). \tag{A-11}$$

Taking complex conjugate and ensemble-average we get

$$\langle |\delta M(\omega)|^2 \rangle = \frac{\lambda_M}{\omega^2 + k_2^2}. \tag{A-12}$$

Here the correlation function is $C_M(\tau) = \langle \delta M(t)\delta M(t + \tau) \rangle_t$ and power spectrum is given as $S_M(\omega) = \langle |\delta M(\omega)|^2 \rangle$. Therefore, applying Wiener-Khinchin theorem (correlation function is itself the Fourier transform of the power spectrum), yields

$$C_M(\tau) = \int \frac{d\omega}{2\pi} e^{-i\omega\tau} S_M(\omega). \tag{A-13}$$

Autocorrelation function at steady state will be a special case of the Eq.(A-13) and is given by

$$C_M(\tau = 0) = \langle \delta M^2 \rangle = \int \frac{d\omega}{2\pi} S_M(\omega). \tag{A-14}$$

Therefore, the steady state fluctuations can be expressed as

$$\langle \delta M^2 \rangle = \frac{1}{2\pi} \int d\omega \frac{\lambda_M}{\omega^2 + k_2^2} = \frac{\lambda_M}{2k_2}. \tag{A-15}$$

Since mRNA production is a one-step process having Poisson distribution, mean and variance must be equal which implies $\langle \delta M^2 \rangle = M_s$ resulting in $\lambda_M = 2k_1$. As protein production can also be thought as Poisson distribution for fixed number of mRNAs, so applying the above procedure we will get $\lambda_P = 2k_1 k_3/k_2$.

Three-stage Model

The system of Langevin equations for the three-stage model is given as

$$\frac{dD}{dt} = k_{on} - (k_{on} + k_{off})D + \eta_D(t)$$
$$\frac{dM}{dt} = k_1 D - k_2 M + \eta_M(t)$$
$$\frac{dP}{dt} = k_3 M - k_4 P + \eta_P(t) \tag{A-16}$$

where noise properties for η_i is same as the Eq.(A-2) except that now $i = D, M$ and P. $\lambda_D = 2k_{off}k_{on}/k'$ and $\lambda_M = 2k_1 k_{on}/k'$ has been already derived before. Using the same approach as mentioned for two-state model, one can define probability $P(j, k, l) = \mathcal{P}(\eta_D \delta t = j, \eta_M \delta t = k, \eta_P \delta t = l)$ where j, k and l are the amount of change in DNA, mRNA and protein count respectively. Following the similar procedure

we will get $\lambda_P = 2k_3 k_1 k_{on}/k_2 k'$.

B.   DERIVATION OF CONDITIONAL MEAN AND COVARIANCE MATRIX

We derive the conditional mean and the covariance matrix in terms of the Green's function using the approach shown in Risken [41]. Matrix equation for the rescaled parameters of ON-OFF model is:

$$\dot{\mathbf{u}}(t) = -\Upsilon \cdot \mathbf{u}(t) + \Xi(t) \tag{B-17}$$

where $\mathbf{u(t)}^T = (q(t), v(t))$, $\Xi^T(\mathbf{t}) = (\xi_D(t), \xi_M(t))$ and

$$\Upsilon = \begin{pmatrix} k' & 0 \\ -k_q & k_2 \end{pmatrix}.$$

We can define $\mathcal{L} = (d/dt + \Upsilon)$ as the linear operator which by the definition of Green's function$(G_0(t, t'))$ must satisfy

$$\mathcal{L}G_0(t, t') = \delta(t - t'). \tag{B-18}$$

Matrix equation can be written in terms of linear operator $\mathcal{L}$ as

$$\mathcal{L}\mathbf{u}(t) = \Xi(t).$$

For $t \neq t'$, we will have $\mathcal{L}G_0(t, t') = 0$. Defining $\tau = t - t'$ we can write

$$\mathcal{L}G_{0ij}(\tau, 0) = 0 \Rightarrow \mathcal{L}G_{0ij}(\tau) = 0 \tag{B-19}$$

as $\mathcal{L}$ is time-translation invariant.
Let $u_i^h(t)$ be the homogeneous solution of Eq.(B-17), which implies $\mathcal{L}u_i^h(t) = 0$. From Eq.(B-19) we have,

$$\mathcal{L}G_{0ij}(\tau) = 0 \Rightarrow \mathcal{L}G_{0ij}(t) = 0 \quad \left[\because \frac{d}{dt} = \frac{d}{d\tau} \times \frac{d\tau}{dt} = \frac{d}{d\tau} \right]. \tag{B-20}$$

Therefore, solution for homogeneous part can be written as $u_i^h(t) = G_{0ij}(t)$. Sharp initial conditions for this problem are given as $(q_0, v_0) = \mathbf{u}_0$ *i.e.*, $u_i^h(0) = u_0 i$. This implies

$$u_i^h(t) = G_{ij}(t)u_{0j} \quad \text{where} \quad G_{ij}(0) = \delta_{ij}. \tag{B-21}$$

This Green's function must satisfy the matrix equation

$$\dot{G}_{ij} + \gamma_{ik}G_{kj} = 0$$
$$\Rightarrow \mathsf{G(t)} = \exp(-\Upsilon t) \tag{B-22}$$

where G_{ij} and γ_{ij} are the elements of the matrix $\mathsf{G}(t)$ and Υ.
For the inhomogeneous solution, we will use the method of variation of the parameters. We make an ansatz

$$u_i^{inh}(t) = G_{ij}(t)c_j(t). \tag{B-23}$$

Differentiating Eq.(B-23) with respect to time, and plugging it in Eq.(B-17), we get

$$\left(\dot{G}_{ij}(t) + \gamma_{ik}G_{kj}(t)\right)c_j(t) + G_{ij}(t)\dot{c}_j(t) = \xi_i(t).$$

Using Eq.(B-22), we get

$$G_{ij}(t)\dot{c}_j(t) = \xi_i(t). \tag{B-24}$$

Eq.(B-22) also implies that $\mathsf{G}^{-1}(t) = \mathsf{G}(-t)$ and $\mathsf{G}(t)\mathsf{G}^{-1}(t') = \mathsf{G}(t)\mathsf{G}(-t') = \mathsf{G}(t - t')$.
Finally, the inhomogeneous solution can be obtained as

$$\mathbf{u}^{inh}(t) = \mathsf{G}(t)\int_0^t \mathsf{G}^{-1}(t')\Xi(t')dt' = \int_0^t \mathsf{G}(t)\mathsf{G}^{-1}(t')\Xi(t')dt'$$
$$= \int_0^t \mathsf{G}(t - t')\Xi(t)dt'. \tag{B-25}$$

In terms of matrix elements, we can write

$$
\begin{aligned}
u_i^{inh}(t) &= \int_0^t G_{ij}(t-t')\xi_j(t')dt' \\
&= -\int_t^0 G_{ij}(p)\xi_j(t-p)dp \\
&= \int_0^t G_{ij}(p)\xi_j(t-p)dp \\
&= \int_0^t G_{ij}(t')\xi_j(t-t')dt'.
\end{aligned}
\tag{B-26}
$$

where $p = t - t', dp = -dt'$. Thus, the general solution with the initial condition is given by

$$
\begin{aligned}
u_i(t) &= u_i^h(t) + u_i^{inh}(t) \\
&= G_{ij}(t)u_{0j} + \int_0^t G_{ij}(t')\xi_j(t-t')dt'.
\end{aligned}
\tag{B-27}
$$

Calculation of Moments

Using the scale properties of the noise $\langle \xi_i(t) \rangle = 0$ and $\langle \xi_i(t)\xi_j(t') \rangle = \delta(t-t')$ we obtain the first two moments as follows:

1. The conditional mean (first moment) is given by

$$
\begin{aligned}
\bar{\mathbf{u}}(t) &= \langle u_i(t) \rangle = G_{ij}(t)u_0 j = \mathsf{G}(t)\mathbf{u}_0 \\
\Rightarrow \bar{\mathbf{u}}(t) &= \mathsf{G}(t)\mathbf{u}_0.
\end{aligned}
\tag{B-28}
$$

2. The covariance matrix elements (second moment) are

$$
\begin{aligned}
\sigma_{ij}(t) &= \sigma_{ji}(t) = \langle [u_i(t) - \langle u_i(t) \rangle][u_j(t) - \langle u_j(t) \rangle] \rangle \\
&= \int_0^t \int_0^t G_{ik}(t'_1)G_{js}(t'_2)\langle \xi_k(t-t'_1)\xi_s(t-t'_2) \rangle dt'_1 dt'_2 \\
&= \int_0^t \int_0^t G_{ik}(t'_1)G_{js}(t'_2)\delta(t'_2 - t'_1)dt'_1 dt'_2 \\
&= \int_0^t G_{ik}(t')G_{js}(t')dt'.
\end{aligned}
\tag{B-29}
$$

Thus, the covariance matrix can be written as

$$
\boldsymbol{\sigma}(t) = \int_0^t \mathsf{G}(t')\mathsf{G}^T(t')dt'.
\tag{B-30}
$$

If we define

$$
\mathcal{D} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}
$$

as the diffusion matrix, then, in terms of diffusion matrix, the covariance matrix will be formulated as

$$
\boldsymbol{\sigma}(t) = 2\int_0^t \mathsf{G}(t') \cdot \mathcal{D} \cdot \mathsf{G}^T(t')dt'.
\tag{B-31}
$$

C. EXACT PROPAGATOR FOR THREE-STAGE MODEL

The covariance matrix elements for three-stage model is defined by the formula:

$$
\boldsymbol{\sigma}(t) = \int_0^t \mathsf{G}(t')\mathsf{G}^T(t')dt' = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix}
\tag{C-32}
$$

The elements of block matrix, $\begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$ are same as the covariance matrix elements Eq.(14) derived for ON-OFF model. Other non-trivial elements are given below

$$\sigma_{31}(t) = \sigma_{13}(t) = \frac{k_q k_v}{(k_2 - k_4)(k_2 - k')(k' - k_4)} \left[\left(\frac{k_4 - k_2}{2k'}\right)(1 - e^{-2k't}) + \left(\frac{k_2 - k'}{k_4 + k'}\right)(1 - e^{-(k_4 + k')t}) \right.$$
$$\left. + \left(\frac{k' - k_4}{k' + k_2}\right)(1 - e^{-(k' + k_2)t}) \right]$$

$$\sigma_{32}(t) = \sigma_{23}(t) = \frac{k_q^2 k_v}{(k_2 - k')^2(k_2 - k_4)(k_4 - k')} \left[(k_2 - k_4)\{\frac{1}{2k'}(1 - e^{-2k't}) - \frac{1}{k_2 + k'}(1 - e^{-(k_2 + k')t})\} \right.$$
$$+ (k_4 - k')\{\frac{1}{k' + k_2}(1 - e^{-(k' + k_2)t}) - \frac{1}{2k_2}(1 - e^{-2k_2 t})\}$$
$$\left. + (k' - k_2)\{\frac{1}{(k' + k_4)t}(1 - e^{-(k' + k_4)t}) - \frac{1}{(k_2 + k_4)}(1 - e^{-(k_2 + k_4)t})\} \right]$$
$$+ \frac{k_v}{k_2 - k_4} \left[\frac{1}{(k_2 + k_4)}(1 - e^{-(k_2 + k_4)t}) - \frac{1}{2k_2}(1 - e^{-2k_2 t}) \right]$$

$$\sigma_{33}(t) = \frac{k_q^2 \cdot k_v^2}{(k_2 - k_4)^2 \cdot (k' - k_2)^2 \cdot (k_4 - k')^2} \left[\frac{(k_2 - k_4)^2}{2k'}(1 - e^{-2k't}) + \frac{(k' - k_2)^2}{2k_4}(1 - e^{-2k_4 t}) \right.$$
$$+ \frac{(k_4 - k')^2}{2k_2}(1 - e^{-2k_2 t}) + 2\{\frac{(k_2 - k_4)(k' - k_2)}{k_4 + k'}(1 - e^{-(k_4 + k')t}) + \frac{(k' - k_2)(k_4 - k')}{k_2 + k_4}(1 - e^{-(k_2 + k_4)t})$$
$$\left. + \frac{(k_4 - k')(k_2 - k_4)}{k' + k_2}(1 - e^{-(k' + k_2)t})\} \right]$$
$$+ \frac{k_v^2}{(k_2 - k_4)^2} \left[\frac{1}{2k_4}(1 - e^{-2k_4 t}) + \frac{1}{2k_2}(1 - e^{-2k_2 t}) - \frac{2}{k_2 + k_4}(1 - e^{-(k_2 + k_4)t}) \right] + \frac{1}{2k_4}(1 - e^{-2k_4 t}). \tag{C-33}$$

Using Mathematica, $\boldsymbol{\sigma}^{-1}(t) = \mathsf{S}(t)$ was calculated and substituted in the final form of propagator given below

$$p(P, t|P_0) = \frac{1}{\sqrt{8 \pi \Delta A (-A') \lambda_P}} \exp\left(\frac{-B'^2}{4A'} + C'\right) \tag{C-34}$$

where

$$\Delta = \det(\boldsymbol{\sigma}(t))$$
$$A = \frac{S_{11}}{2}$$
$$A' = \frac{1}{2}\left(\frac{S_{12}^2}{S_{11}} - S_{22}\right)$$
$$B' = \left(\frac{S_{12}S_{13}}{S_{11}} - S_{23}\right)\delta y$$
$$C' = \left(\frac{S_{13}^2}{S_{11}} - S_{33}\right)\delta y^2$$
$$\delta y = y(t) - \bar{y}(t)$$
$$\bar{y}(t) = G_{31}q_0 + k_v \frac{(e^{-k_4 t} - e^{-k_2 t})}{k_2 - k_4}v_0 + y_0 e^{-k_4 t}$$
$$q_0 = \frac{D_0 - D_s}{\sqrt{\lambda_D}}$$
$$v_0 = \frac{M_0 - M_s}{\sqrt{\lambda_M}}$$
$$y_0 = \frac{P_0 - P_s}{\sqrt{\lambda_P}}.$$

For this calculation, we have chosen $D_0 = 1.0$ and $M_0 = P_0 = 0$.

---

[1] Mukund Thattai and Alexander van Oudenaarden. Intrinsic noise in gene regulatory networks. *Proceedings of the National Academy of Sciences*, 98(15):8614–8619, 2001.

[2] W Ras, Cell Biol, William J Blake, Mads Kærn, Charles R Cantor, and J J Collins. Noise in

eukaryotic gene expression. *Nature*, 249(2002):247–249, 2003.

[3] Jonathan M Raser and Erin K O'Shea. Control of stochasticity in eukaryotic gene expression. *Science*, 304(5678):1811–4, 2004.

[4] Arjun Raj, Charles S Peskin, Daniel Tranchina, Diana Y Vargas, and Sanjay Tyagi. Stochastic mrna synthesis in mammalian cells. *PLoS Biol*, 4(10):e309, 09 2006.

[5] Vahid Shahrezaei and Peter S Swain. Analytical distributions for stochastic gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, 105(45):17256–17261, 2008.

[6] Johan Paulsson. Summing up the noise in gene networks. *Nature*, 427(6973):415–8, 2004.

[7] Ido Golding, Johan Paulsson, Scott M. Zawilski, and Edward C. Cox. Real-time kinetics of gene activity in individual bacteria. *Cell*, 123(6):1025–1036, 2005.

[8] Jonathan R. Chubb, Tatjana Trcek, Shailesh M. Shenoy, and Robert H. Singer. Transcriptional Pulsing of a Developmental Gene. *Current Biology*, 16(10):1018–1025, 2006.

[9] N.G. Van Kampen. *Stochastic Processes in Physics and Chemistry*. North-Holland Personal Library. Elsevier Science, 1992.

[10] J. Peccoud and B. Ycart. Markovian modeling of gene-product synthesis. *Theoretical Population Biology*, 48(2):222 – 234, 1995.

[11] Ertugrul M. Ozbudak, Mukund Thattai, Iren Kurtser, Alan D. Grossman, and Alexander van Oudenaarden. Regulation of noise in the expression of a single gene. *Nat Genet*, 31(1):69–73, May 2002.

[12] Shubhendu Nandi and Anandamohan Ghosh. Transcriptional dynamics with time-dependent reaction rates. *Physical Biology*, 12(1):016015, 2015.

[13] Arjun Raj and Alexander van Oudenaarden. Nature, nurture, or chance: Stochastic gene expression and its consequences. *Cell*, 135(2):216 – 226, 2008.

[14] HarleyH. McAdams and Adam Arkin. Stochastic mechanisms in geneexpression. *Proceedings of the National Academy of Sciences*, 94(3):814–819, 1997.

[15] Peter S. Swain, Michael B. Elowitz, and Eric D. Siggia. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proceedings of the National Academy of Sciences*, 99(20):12795–12800, 2002.

[16] Lok-hang So, Anandamohan Ghosh, Chenghang Zong, Leonardo A Sepúlveda, Ronen Segev, and Ido Golding. General properties of transcriptional time series in Escherichia coli. *Nature Genetics*, 43(6):554–560, 2011.

[17] Anandamohan Ghosh. Non-equilibrium dynamics of stochastic gene regulation. *Journal of Biological Physics*, 41(1):49–58, 2015.

[18] Peter Hänggi, Peter Talkner, and Michal Borkovec. Reaction-rate theory: fifty years after kramers. *Rev. Mod. Phys.*, 62:251–341, Apr 1990.

[19] Tao Jia and Rahul V Kulkarni. Intrinsic noise in stochastic models of gene expression with molecular memory and bursting. *Physical review letters*, 106(5):058102, 2011.

[20] Maciej Dobrzynski and Frank J Bruggeman. Elongation dynamics shape bursty transcription and translation. *Proceedings of the National Academy of Sciences of the United States of America*, 106(8):2583–2588, 2009.

[21] Abhyudai Singh and John J. Dennehy. Stochastic holin expression can account for lysis time variation in the bacteriophage . *Journal of The Royal Society Interface*, 11(95), 2014.

[22] Pankaj Mehta, Ranjan Mukhopadhyay, and Ned S Wingreen. Exponential sensitivity of noise-driven switching in genetic networks. *Physical Biology*, 5(2):026005, 2008.

[23] Adam Arkin, John Ross, and Harley H. McAdams. Stochastic kinetic analysis of developmental pathway bifurcation in phage -infected escherichia coli cells. *Genetics*, 149(4):1633–1648, 1998.

[24] Abhyudai Singh and Leor S Weinberger. Stochastic gene expression as a molecular switch for viral latency. *Current Opinion in Microbiology*, 12(4):460 – 466, 2009. HostMicrobe Interactions: Fungi/Parasites/Viruses.

[25] Ya-Chi Ho, Liang Shan, Nina N. Hosmane, Jeffrey Wang, Sarah B. Laskey, Daniel I.S. Rosenbloom, Jun Lai, Joel N. Blankson, Janet D. Siliciano, and Robert F. Siliciano. Replication-competent noninduced proviruses in the latent reservoir increase barrier to hiv-1 cure. *Cell*, 155(3):540 – 551, 2013.

[26] Ariel D. Weinberger and Leor S. Weinberger. Stochastic fate selection in HIV-infected patients. *Cell*, 155(3):497–499, 2013.

[27] Arvind K. Chavali, Victor C. Wong, and Kathryn Miller-Jensen. Distinct promoter activation mechanisms modulate noise-driven HIV gene expression. *Scientific reports*, 5:17661, 2015.

[28] Kathryn Miller-Jensen, Siddharth S. Dey, David V. Schaffer, and Adam P. Arkin. Varying virulence: Epigenetic control of expression noise and disease processes. *Trends in Biotechnology*, 29(10):517–525, 2011.

[29] Richard L. Thompson, Chris M. Preston, and Nancy M. Sawtell. De novo synthesis of vp16 coordinates the exit from hsv latency in vivo. *PLoS Pathog*, 5(3):e1000352, 03 2009.

[30] David M Knipe and Anna Cliffe. Chromatin control of herpes simplex virus lytic and latent infection. *Nature reviews. Microbiology*, 6(3):211–221, 2008.

[31] Thomas Gnther and Adam Grundhoff. The epigenetic landscape of latent kaposi sarcoma-associated

herpesvirus genomes. *PLoS Pathog*, 6(6):e1000935, 06 2010.

[32] Zsolt Toth, Dennis T. Maglinte, Sun Hwa Lee, Hye-Ra Lee, Lai-Yee Wong, Kevin F. Brulois, Stacy Lee, Jonathan D. Buckley, Peter W. Laird, Victor E. Marquez, and Jae U. Jung. Epigenetic analysis of kshv latent and lytic genomes. *PLoS Pathog*, 6(7):e1001013, 07 2010.

[33] Natascha K A Grzimek, Doris Dreis, and Susanne Schmalz. Random , Asynchronous , and Asymmetric Transcriptional Activity of during Murine Cytomegalovirus Latency in the Lungs. *Journal of Virology*, 75(6):2692–2705, 2001.

[34] Mads Kaern, Timothy C Elston, William J Blake, and James J Collins. Stochasticity in gene expression: from theories to phenotypes. *Nature reviews. Genetics*, 6(6):451–464, 2005.

[35] Ron Skupsky, John C. Burnett, Jonathan E. Foley, David V. Schaffer, and Adam P. Arkin. Hiv promoter integration site primarily modulates transcriptional burst size rather than frequency. *PLoS Comput Biol*, 6(9):e1000952, 09 2010.

[36] Peter S. Swain. Efficient attenuation of stochasticity in gene expression through post-transcriptional control. *Journal of Molecular Biology*, 344(4):965–976, 2004.

[37] William Bialek and Sima Setayeshgar. Physical limits to biochemical signaling. *Proceedings of the National Academy of Sciences of the United States of America*, 102(29):10040–10045, 2005.

[38] Pankaj Mehta, Sidhartha Goyal, and Ned S Wingreen. A quantitative comparison of srna-based and protein-based gene regulation. *Molecular Systems Biology*, 4(1), 2008.

[39] Rati Sharma and Binny J Cherayil. Reaction dynamics under confinement: an exact path integral treatment of a two-stage model of stochastic gene expression. *Journal of Statistical Mechanics: Theory and Experiment*, 2013(10):P10029, 2013.

[40] V. Balakrishnan. *Elements of Nonequilibrium Statistical Mechanics*. Delhi : Ane Books India, 2008.

[41] H. Risken and T. Frank. *The Fokker-Planck Equation: Methods of Solution and Applications*. Springer Series in Synergetics. Springer Berlin Heidelberg, 2012.

[42] C.W. Gardiner. *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*. Springer complexity. Springer, 2004.

[43] Daniel T Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, 22(4):403 – 434, 1976.

[44] Daniel T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361, 1977.

[45] D. J. Higham. An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM Review*, 43(3):525–546, 2001.

[46] P.E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg, 2013.

[47] Michael B. Elowitz, Arnold J. Levine, Eric D. Siggia, and Peter S. Swain. Stochastic gene expression in a single cell. *Science*, 297(5584):1183–1186, 2002.

[48] K. R. Ghusinga and A. Singh. First-passage time calculations for a gene expression model. In *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*, pages 3047–3052, Dec 2014.

[49] Daniel Zenklusen, Daniel R Larson, and Robert H Singer. Single-RNA counting reveals alternative modes of gene expression in yeast. TL - 15. *Nature structural & molecular biology*, 15 VN - r(12):1263–1271, 2008.

[50] Vahid Shahrezaei, Julien F Ollivier, and Peter S Swain. Colored extrinsic fluctuations and stochastic gene expression. *Molecular Systems Biology*, 4(1), 2008.