The Protein Hourglass: Exact First Passage Time Distributions for Protein Thresholds

Krishna Rijal,[1,*] Ashok Prasad,[2] and Dibyendu Das[1,†]

[1] *Physics Department, Indian Institute of Technology Bombay, Mumbai 400076, India,*
[2] *Department of Chemical and Biological Engineering, Colorado State University, Fort Collins, Colorado, USA*

(Dated: August 26, 2020)

Protein thresholds have been shown to act as an ancient timekeeping device, such as in the time to lysis of *E. coli* infected with bacteriophage *lambda*. The time taken for protein levels to reach a particular threshold for the first time is defined as the first passage time of the protein synthesis system, which is a stochastic quantity. The first few moments of the distribution of first passage times were known earlier, but an analytical expression for the full distribution was not available. In this work, we derive an analytical expression for the first passage times for a long-lived protein. This expression allows us to calculate the full distribution not only for cases of no self-regulation, but also for both positive and negative self-regulation of the threshold protein. We show that the shape of the distribution matches previous experimental data on lambda-phage lysis time distributions. We also provide analytical expressions for the FPT distribution with non-zero degradation in Laplace space. Furthermore, we study the noise in the precision of the first passage times described by coefficient of variation (CV) of the distribution as a function of the protein threshold value. We show that under conditions of positive self-regulation, the CV declines monotonically with increasing protein threshold, while under conditions of linear negative self-regulation, there is an optimal protein threshold that minimizes the noise in the first passage times.

PACS number(s): 05.40.-a, 02.50.-r, 87.10.-e, 87.14.Ee

I. INTRODUCTION

Thresholds in protein concentrations are ubiquitous in cell biology and are thought to be one of the main methods by which cells control the timings of events. Imposing a threshold before a biological process can be activated not only protects the process from noise-driven spurious activation, it also allows the cell to maintain a certain time interval between events. Cells do have other timing devices, most significant of which are circadian and ultradian rhythms. But thresholds offer one of the simplest timing devices possible, like the sand-filled hourglass of yesteryears, which is probably why they are ubiquitous in cellular processes. A good example of thresholds determining timing is provided by the passage of cells through the restriction point (R) during the cell cycle. It has been shown that passage through R in fibroblasts is governed by a rather precise threshold of the activity of the kinase, Cdk, and that a live cell Cdk sensor accurately predicts passage through R for 96% of cells [1]. Just as in cell reproduction, in cell death too, the timing and probability of mitochondrial Outer Membrane Permeabilization, the irreversible step committing the cell to apoptosis after an event like DNA damage, depends on pro-apoptotic signals crossing a threshold[2–4]. However, the most well studied question in cell biology where thresholds are believed to play a major role in determining timing is the replication of lambda-phage.

After lambda-phage infects an E. coli bacterium, it can have one of two possible life cycles, either the lytic or the lysogenic path. The molecular basis of this decision making has been critical to the development of our understanding of stochastic gene regulation over the last half-century[5–15]. The lysogenic phage incorporates their DNA into the bacterial chromosome, which sits there silently until a stress signal indicating host distress activates the genes and they enter the lytic cycle. In the lytic pathway, the virus multiplies in its host for a programmed time period, and then lyses the host, which then bursts, releasing a number of virions. The majority of infected bacteria will take the lytic route, and most remarkably, the timing of lysis after initial phage infection is tightly controlled [16] with a variation of only about 5% of the mean, which is a remarkable level of precision for such a complicated system, especially when the burst size is found to vary significantly more [16–18]. The timing system of lambda-phage is controlled by the accumulation of a protein, holin [19, 20], that permeabilizes the membrane after it accumulates to a critical level. The tight regulation of lysis time is believed to have an evolutionary basis and was a prediction from optimal foraging theory[16]. Significantly there does not appear to be any additional genetic regulation that determines the timing of lysis. In two seminal papers [21, 22], Singh, Dennehy and colleagues explored the idea that the timing of lysis is controlled by the first passage time for the holin protein to reach a threshold. For any stochastic system, the first passage time is a random variable that denotes the first time some event occurred [23]. Examples of such events may be like a passive or active particle reaching a boundary [24, 25], a protein binding to a

* 194120004@iitb.ac.in
† dibyendu@phy.iitb.ac.in

specific patch on DNA [26], a moving kinetochore being captured by microtubules [27], or as in this case protein levels reaching a pre-determined threshold. The papers [21, 22], showed that a model based on the statistics of first passage times to reach a critical protein concentration could reproduce the observed data. Furthermore they provided a basis for the observed lack of regulation of the timing by demonstrating that for a long-lived protein, the optimal strategy for precise event timing is no regulation at all, and that both negative and positive feedback actually increase the noise in event timing [22]. The above-mentioned papers were based on calculations of the first few moments of the first passage time distribution, while an analytical expression of the full distribution itself was not available prior to this paper.

There are significant advantages in obtaining the full distribution of a stochastic biochemical process. For example, it has been shown that when inferring parameters from data, it is possible to infer incorrect parameters with only the mean and the variance of the distribution [28]. However, given the complexity in even simple molecular biological processes, there are relatively few exact distributions known. These include Refs. [29–35]. In this paper we present an analytical calculation for the first passage time of a long-lived protein driven by a bursting process to reach a threshold. We show that the first passage distribution applies even when the protein has negative or positive autoregulation, showing either linear kinetics or Hill kinetics. We show that the shape of the distribution matches the experimental data on lysis times. We calculate the coefficient of variation of the distribution and show how it varies with the level of the protein threshold. Finally, for non-zero protein degradation, we calculate the Laplace transform of the FPT distribution exactly.

II. MODEL

Consider a gene that is switched on at time $t = 0$ and begins to express a timekeeper protein, which triggers an intracellular event of interest, eg. lysis, once it reaches a critical level in the cell. A minimal model of gene expression assumes translation in bursts and can incorporate feedback regulation by considering the transcription rate as a function of the protein level (see Fig.(1)). We represent the protein population count or number by states $[0, 1, ..., X]$. Protein numbers go up due to bursts of transcription-translation that add a number of the proteins to the system, while protein numbers are reduced due to degradation of single proteins. The process terminates when the protein level reaches a certain threshold or absorbing-state X for the first time. The corresponding time t defines the "first passage time" (FPT). The statistics of these times determines the timing of the intracellular event.

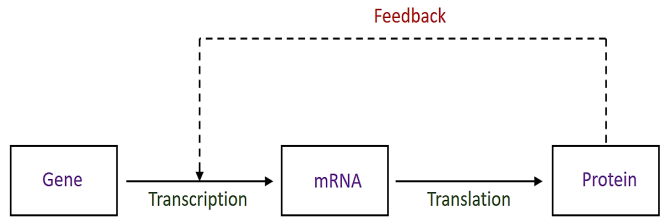


FIG. 1: Model depicting a gene transcribing mRNAs, which are further translated into proteins. The transcription rate is assumed to be regulated by the protein level, creating a feedback loop (denoted by the dashed line).

The bursts of protein numbers are assumed to follow a geometric distribution, as shown previously for protein burst-size distributions [36]. Thus the Poisson rate of making a transition from state i to state $i + n$ is

$$r_{i,i+n} = k_i \frac{b^n}{(b+1)^{n+1}}, \quad \text{where}, n = \{0, 1, 2...\}. \tag{1}$$

Here mean burst size is denoted by $b \in (0, \infty)$, and $n = 0$ denotes no change in protein number with probability $b/(b+1)$. The net Poisson rate for any burst of size $n > 0$ out of state i is thus

$$\sum_{n=1}^{\infty} r_{i,i+n} = k_i b/(b+1). \tag{2}$$

Each protein molecule degrades with a constant rate γ. For most of our calculations, we will assume that the proteins are long-lived and hence we will set the degradation rate γ to 0.

We use the backward Master equation (BME) formalism [37] to calculate the first passage probabilities. The BME describes the time-evolution of the survival probabilities $S(i, t)$, that we define as the probability that the protein count stays below the threshold X at time t, given that it was at i at time $t = 0$, where $i = 0, 1, \cdots, (X-1)$).

If we define a vector $\mathbf{S}(t) = \begin{bmatrix} S(0,t) & S(1,t) & S(2,t) & \dots & S(X-1,t) \end{bmatrix}^T$ whose components are survival probabilities, we may write the backward master equation as

$$\dot{\mathbf{S}}(t) = \mathbf{A}\mathbf{S}(t) \tag{3}$$

with the initial condition $\mathbf{S}(t = 0) = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \end{bmatrix}^T$. Since the process terminates for protein count reaching any value $\geq X$, the boundary condition is $S(i \geq X, t) = 0$. Taking into account degradation, the above matrix of Poisson rates, $\mathbf{A}$ is

$$\begin{bmatrix} -\frac{bk_0}{b+1} & \frac{bk_0}{(b+1)^2} & \cdots & \frac{b^{X-2}k_0}{(b+1)^{X-1}} & \frac{b^{X-1}k_0}{(b+1)^X} \\ \gamma & -\left(\frac{bk_1}{b+1}+\gamma\right) & \cdots & \frac{b^{X-3}k_1}{(b+1)^{X-2}} & \frac{b^{X-2}k_1}{(b+1)^{X-1}} \\ \vdots & 2\gamma & -\left(\frac{bk_2}{b+1}+2\gamma\right) & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & (X-1)\gamma & -\left(\frac{bk_{X-1}}{b+1}+(X-1)\gamma\right) \end{bmatrix} \tag{4}$$

Note that the general element in this rate matrix may be written as

$$A_{ij} = \begin{cases} \frac{k_{i-1}b^{j-i}}{(b+1)^{j-i+1}}\left(1-(b+1)\,\delta_{ij}\right) - (i-1)\gamma\delta_{ij} & \text{for } j \geq i \\ (i-1)\gamma\delta_{i,j+1} & \text{for } j < i \end{cases} \tag{5}$$

Furthermore, the diagonal elements of this matrix are the negative sum of all outgoing rates $-\gamma i - \sum_{n=1}^{\infty} r_{i,i+n} = -\gamma i - k_i b/(b+1)$.

The first passage probability to reach X for the first time between t to $t+dt$ is $f_{0,X}(t)dt$, where the FPT distribution $f_{0,X}(t)$ is related to the survival probability as $f_{0,X}(t) = -\partial S_{0,X}(t)/\partial t$. Thus our objective is to solve for $S_{0,X}(t)$ and obtain the desired expression for $f_{0,X}(t)$.

From here, we can proceed using two different methods. Integrating Eq. (3) gives

$$\mathbf{S}(t) = \exp(\mathbf{A}t)\ \mathbf{S}(0). \tag{6}$$

The FPT distribution follows from Eqs (3) and (6) as

$$\mathbf{f}(t) = -\dot{\mathbf{S}}(t) = -\frac{\partial}{\partial t}[\exp(\mathbf{A}t)\mathbf{S}(0)] \tag{7}$$

Calculation of the full distribution requires the evaluation of $\exp(\mathbf{A}t)$ which may be carried out as shown in Appendix G. An alternative and more elegant method is to use Laplace Transforms on Eq. (3) as outlined below. Both the methods to calculate the full distribution gives identical results as expected. Furthermore, we can also calculate the moments of the FPT through evaluation of the inverse matrix $\mathbf{A^{-1}}$(as was also shown in [22]).

III. RESULTS

A. The exact first passage time distribution

1. *Unequal Transition Rates with zero protein degradation ($\gamma = 0$)*

In the most general case the transition rate constants k_i are all unequal. We outline the solution in the case of $\gamma = 0$ as follows. We can take Laplace transform of Eq. (3), to convert it into an algebraic equation. If we define $\tilde{\mathbf{S}} = \int_0^\infty e^{-st}\mathbf{S}dt$, then Eq (3), after performing Laplace transformation gives,

$$\tilde{\mathbf{S}} = (s\mathbb{I} - \mathbf{A})^{-1}\mathbf{S}(0)$$

The general form of the elements of $(s\mathbb{I} - \mathbf{A})^{-1}$ is

$$(s\mathbb{I} - A)^{-1}_{ij} = \begin{cases} \frac{b^{j-i}k_{i-1}}{(b+1)^{j-i+1}} \frac{\prod_{q=i}^{j-2}(s+k_q)}{\prod_{p=i-1}^{j-1}(s+\frac{bk_p}{b+1})} & \text{if } j > i+1 \\ \frac{k_{i-1}}{(b+1)^2(s+\frac{bk_{i-1}}{b+1})(s+\frac{bk_i}{b+1})} & \text{if } j = i+1 \\ \frac{1}{s+\frac{bk_{i-1}}{b+1}} & \text{if } j = i \\ 0 & \text{if } j < i \end{cases} \tag{8}$$

In order to obtain this result, we first evaluated $(s\mathbb{I} - \mathbf{A})^{-1}$ in the special cases of $X = 1, 2, 3$ and 4, and then generalized the mathematical form to any X. We are interested in the case of initial protein number being 0, and in the corresponding first passage distribution

$$\begin{aligned} f_{0,X}(t) = \mathcal{L}^{-1}[\tilde{f}_{0,X}(s)] =& \mathcal{L}^{-1}[1 - s\tilde{S}_{0,X}(s)] \\ =& \mathcal{L}^{-1}[1 - s\sum_{j=1}^{X}(s\mathbb{I} - A)^{-1}_{1j}] \end{aligned} \tag{9}$$

where, $\mathcal{L}^{-1}$ denotes inverse Laplace transformation. In the last step, we have used the fact that $\mathbf{S}(t=0) = \begin{bmatrix}1 & 1 & 1 & \ldots & 1\end{bmatrix}^T$. Since all the poles of $(s\mathbb{I} - \mathbf{A})^{-1}$are simple, it is easy to find inverse Laplace transform by evaluating Bromwich integral, by finding residues. We get the exact FPT distribution (see Appendix A) as

$$\begin{aligned} f_{0,X}(t) =& \frac{bk_0}{(b+1)^2}\Bigg[\left(\frac{k_1}{k_1-k_0}+b\right)e^{-\frac{bk_0 t}{b+1}} + \left(\frac{k_1}{k_0-k_1}\right)e^{-\frac{bk_1 t}{b+1}} \\ &+ \sum_{j=3}^{X}\sum_{n=0}^{j-1}\frac{k_n}{(b+1)^{j-2}}\frac{\prod_{q=1}^{j-2}((b+1)k_q - bk_n)}{\prod_{p=0,p\neq n}^{j-1}(k_p-k_n)}e^{-\frac{bk_n t}{b+1}}\Bigg] \end{aligned} \tag{10}$$

2. *Equal Transition Rates with zero protein degradation ($\gamma = 0$)*

The case of uniform rates $k_i = k$ (corresponding to no feedback) is special. We see in Eqs. (10) that denominators of various terms go to zero. This may appear to lead to divergent answers, which is not expected physically for any bounded function k_i. From Eq. (8) setting all $k_i = k$, we get

$$(s\mathbb{I} - A)^{-1}_{ij} = \begin{cases} \frac{kb^{j-i}(s+k)^{j-i-1}}{(b+1)^{j-i+1}(s+\frac{bk}{b+1})^{j-i+1}} & \text{if } j > i \\ \frac{1}{s+\frac{bk}{b+1}} & \text{if } j = i \\ 0 & \text{if } j < i \end{cases} \tag{11}$$

One way to verify Eq. (11) directly is to check that $(s\mathbb{I}-\mathbf{A})^{-1}(s\mathbb{I}-\mathbf{A}) = \mathbb{I}$ (see Appendix B). Compared to Eq. (8), in Eq. (11) the poles are not simple but of higher order. Hence, the Laplace inverse transform leading to the FPT distribution will not be a sum of pure exponentials. We use Bromwich integral and find residues (see Appendix C). The final desired answer for the first passage time distribution is:

$$f_{0,X}(t) = \frac{kb^X}{(b+1)^X}\,{}_1F_1[1-X,1,-\frac{kt}{b+1}]e^{-\frac{kbt}{b+1}} \tag{12}$$

Here ${}_1F_1$ is the confluent hypergeometric function. An alternative derivation of Eq. (12), involving interesting mathematical identities, are shown in Appendix H and I.

3. *Non-zero protein degradation ($\gamma \neq 0$)*

The main results derived above are for zero protein degradation rates. However we have also obtained exact results in the Laplace space for non-zero degradation rates. Using Eq. (9), the result for general X may be obtained by studying the pattern of $\tilde{f}_{0,X}(s)$ for small X values and then generalizing to higher values.

Unequal transition rates: For distinct k_i for $X > 2$, we have

$$\tilde{f}_{0,X}(s) = \frac{(-b)^X k_0}{(b+1)^X}\frac{\prod_{l=1}^{X-1}(\gamma l + k_l + s)}{D(X)} \tag{13}$$

where the denominator $D(X)$ of Eq. (13) is not easy to find in a compact form. But we could relate $D(X)$ to $D(X-1)$ and another function $Q(X)$, through a recurrence relation given below:

$$D(X) = -\left[(X-1)\gamma + s + \frac{bk_{X-1}}{b+1}\right]D(X-1) + Q(X)$$
$$Q(X) = -\left[\frac{bk_{X-2}}{(b+1)^2}D(X-2) + \frac{b}{b+1}Q(X-1)\right](X-1)\gamma$$

with $D(1) = -s - \frac{bk_0}{b+1}$; $D(2) = \frac{b^2k_0k_1}{(b+1)^2} + \frac{b\gamma k_0}{b+1} - \frac{b\gamma k_0}{(b+1)^2} + \frac{bk_0 s}{b+1} + \frac{bk_1 s}{b+1} + \gamma s + s^2$ and $Q(2) = -\frac{b\gamma k_0}{(1+b)^2}$. Using the recurrence relations one can iteratively obtain $D(X)$ for any X and hence find the Laplace transform $\tilde{f}_{0,X}(s)$.

Equal transition rates: For $k_i = k$ for all i, the FPT distribution in the Laplace space is simpler. For $X > 1$,

$$\begin{aligned}\tilde{f}_{0,X}(s) = &\left[b^X k\prod_{l=1}^{X-1}(\gamma l + k + s)\right]\left[\prod_{i=0}^{X-1}(\gamma i + s)\right.\\ &+ b^X(k+s)\prod_{j=1}^{X-1}(\gamma j + k + s)\\ &\left.+\sum_{n=1}^{X-1}\left[\binom{X}{n}b^n\prod_{i=0}^{X-1}(\gamma i + k\Theta(i-X+n) + s)\right]\right]^{-1}\end{aligned} \tag{14}$$

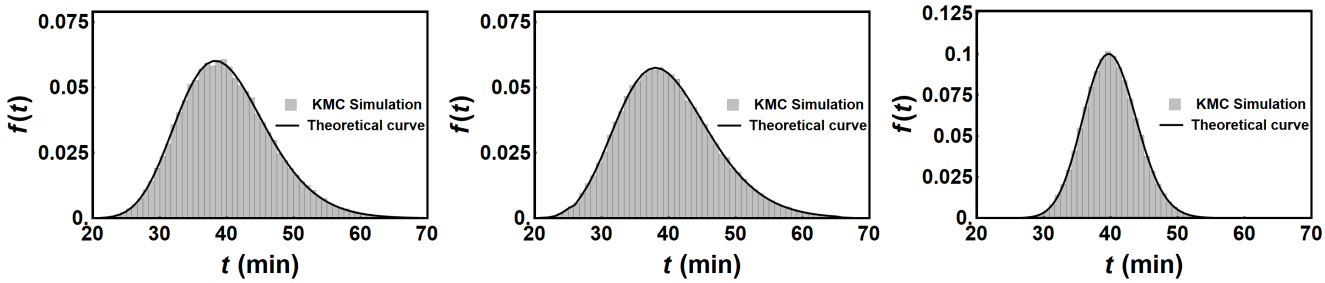
where Θ denotes Heaviside step function.

In both the cases of unequal and equal transition rates, the expressions of $\tilde{f}_{0,X}(s)$ have polynomials of degree equal to X as denominators (see Eq. (13) and (14)). It is well known that finding roots exactly of a polynomial beyond some small values of X is very difficult. Hence, finding inverse Laplace transform for large X remains intractable for $\gamma \neq 0$. To demonstrate the complexity of the poles for even small values of X, we have calculated the poles and hence $f_{0,X}(t)$ for $X = 1, 2, 3$ in Appendix D and E, for both the cases.

B. Kinetic Monte Carlo results compared with the exact formulae

In order to verify Eq. (10) and Eq. (12) we carried out Kinetic Monte Carlo (KMC) simulations for cases of positive, negative and no feedback. In order to compare relative fluctuations, we hold the mean first passage time same in the three cases. The FPT distribution is obtained by measuring the times of first passage to reach the protein number threshold of $X = 500$ starting from 0 number of proteins. At any instant of time, the system is in a state i with i number of proteins. In the next step of KMC, any one of the events by which $i \to i+n$ is permitted, where in principle $n \in [1, \infty)$. But for practicality, to have a finite set of events in simulations, we restrict $n \in [1, X]$. The rates of the transition of $i \to i+n$ are given by Eq. (1). As is usual for KMC, the events are chosen with probability $r_{i,i+n}/Q_i$, and the time increments Δt are drawn randomly from an exponential distribution with decay constant Q_i, where Q_i is given by Eq. (2) . For the simulations, following [22] we chose $X = 500$ protein molecules, and $k_i = c_1 \pm c_2 i$ for positive and negative feedback cases respectively. We have $c_2 = 0.05$ min^{-1} for these two cases. For no feedback $c_2 = 0$. The value of c_1 was chosen to be 0.581 min^{-1}, 25.443 min^{-1} and 6.275 min^{-1} for the positive, negative and zero feedback cases respectively, to ensure that all the three cases have a same mean first passage time of 40 mins (for ease of comparison). We have also chosen $b = 2$ just as in [22].

The plots of our KMC simulations and the curves for $f_{0,X}(t)$ given by the exact theoretical formulae Eqs. (10) and (12) are shown in Fig. (2), and they match perfectly.



(a) (b) (c)

FIG. 2: Theoretical FPT distributions $f(t)$, are compared with KMC simulations. For stable protein ($\gamma = 0$), burst size $b = 2$, and protein threshold level $X = 500$. We show histogram plots of 100000 simulations performed for linear feedbacks of the for $k_i = c_1 + c_2 i$ with (a) $c_2 = 0.05$ min^{-1} for positive feedback, (b) $c_2 = -0.05$ min^{-1} for negative feedback and (c) $c_2 = 0$ for the no feedback cases. Value of c_1 is choosen such that mean FPT distribution is kept constant at 40 min. The black curves on the top of the histograms are the theoretical FPT distributions given by Eq. (10) and Eq. (12), for the chosen parameters.

C. Formulae for moments and cumulants

We can also calculate exact expressions for the moments of the FPT distribution, using Eq. (7) after obtaining the matrix $\mathbf{A}^{-1}$. The general form of elements of the matrix $\mathbf{A}^{-1}$ is

$$A^{-1}_{ij} = \begin{cases} -\frac{(1+b\delta_{i,j})}{bk_{j-1}} & \text{for } j \geq i \\ 0 & \text{for } j < i \end{cases} \tag{15}$$

We obtained this by putting $s = 0$ in Eq. (8). Identical expression was also obtained in [22]. For special case of all identical $k_i = k$, first two moments are given as [22]

$$\langle T_0 \rangle = \frac{b+X}{bk} \text{ and } \langle T_0^2 \rangle = \frac{2b^2 + 4bX + X^2 + X}{(kb)^2} \tag{16}$$

which gives

$$CV^2 = \frac{\sigma^2}{\langle T_0 \rangle^2} = \frac{b^2 + X + 2bX}{(b+X)^2} \tag{17}$$

But the explicit forms of the higher moments and cumulants were not given in [22]. Here, for the ready reference of the reader we are giving the explicit expressions for the next two higher cumulants (see calculations in Appendix F):

$$\begin{aligned} \kappa_3 &= \frac{2}{(bk)^3}[(1-X)b^3 + X(b+1)^3] \\ \kappa_4 &= \frac{6}{(bk)^4}[(1-X)b^4 + X(b+1)^4] \end{aligned} \tag{18}$$

The following definitions of skewness and kurtosis based on κ_3 and κ_4 will be used in the next section:

$$\text{Skewness} = \frac{\kappa_3}{(\kappa_2)^{\frac{3}{2}}} \quad \text{Kurtosis} = 3 + \frac{\kappa_4}{\kappa_2^2} \tag{19}$$

D. Experimental Lysis Time Distribution in Lambda phage compared to the Theoretical Distribution

Lysis of the host E. coli by an infective lambda-phage is controlled by a protein from the holin family, S105 or simply holin, that accumulates in the bacterial cytoplasmic membrane till it reaches a critical value, when it triggers to form micron-scale holes in the membrane, releasing an enzyme that rapidly destroys the bacterial cell wall and frees the viral progeny. The timing of lysis has been shown to be tightly and precisely controlled by the accumulation of the holin protein, i.e. by a first passage time process. Since holins have a half-life larger than the mean lysis time, our results for the FPT of a long-lived protein are directly applicable [17, 20–22, 38, 39]

Holin synthesis does not begin immediately after infection by the lambda phage, but after a delay. In order to compare the theoretical distribution with the experimental data, we first have to account for this delay in the transcription of holin, $T_{pR'}$, which is known from experiments [21] to be ≈ 15 mins. We will also assume, following [21], that the standard deviation of $T_{pR'}$ is small enough to ignore. So lysis time t_{LT} can be related to first passage time t as:

$$t = t_{LT} - T_{pR'}, \tag{20}$$

and it follows that

$$\langle t_{LT} \rangle = \langle t \rangle + 15 \; ; \; \sigma^2_{t_{LT}} = \sigma^2_t. \tag{21}$$

We now make the simplified assumptions, following [21, 22] , that: (i) there is no feedback, i.e. $k_I = k$, and (ii) the holins do not degrade easily, i.e. $\gamma = 0$. Experimentally [21] it is known that $\langle t_{LT} \rangle \approx 65$ min, and $\sigma_{LT} \approx 3.5$ min. Just as in [21], we take $X = 1500$. These numbers may be used in Eqs. (16) and (17) to obtain the the parameters

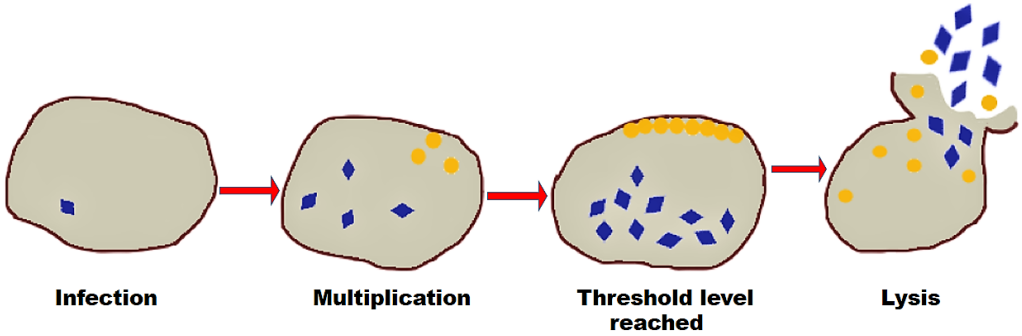


FIG. 3: A schematic figure of steps leading to lysis of an E.Coli cell infected by a λ-phage. The blue rhombi represent viruses (in capsids), and the yellow circles represent the holin protein molecules.

$$b \approx 3\ ;\ k \approx 10\ \text{min}^{-1}. \tag{22}$$

Finally using the above parameters (Eq. (22)), we plot the theoretical curve for $F(t_{LT})$ (Eq. (12)) in Fig. (4) against the available experimental data from [21]. We can see that the theoretical distribution, with no fitting but using the parameters from [22] matches the experimental data quite well.

We tested to see whether we can improve the correspondence by fitting the parameters to the data. We fixed the values of mean first passage time to the experimentally determined 65 mins, and that of the burst size b to be 3. Using minimization of the residual sum of squares [40] we found that the best-fit value of X to be very close to 1500, namely 1463. Next, we kept only the mean fixed at 65 min and performed a 2-parameter minimization of the residual sum of squares. The best-fit value of b and X was found to be 2.6 and 1296 respectively. However the quality of the fits are not very significantly better than the original set of parameter values.

Using $b = 3$; $X = 1500$; $k = 10\ \text{min}^{-1}$ we can also compare the skewness and kurtosis of our predictions with the experimental values. The theoretical expressions are shown in Eq. (19) in the previous section. The correspondence is very close — the theoretically predicted skewness and kurtosis are 0.0025 and 3.015 as compared to the experimental values 0.0012 and 3.540, respectively.

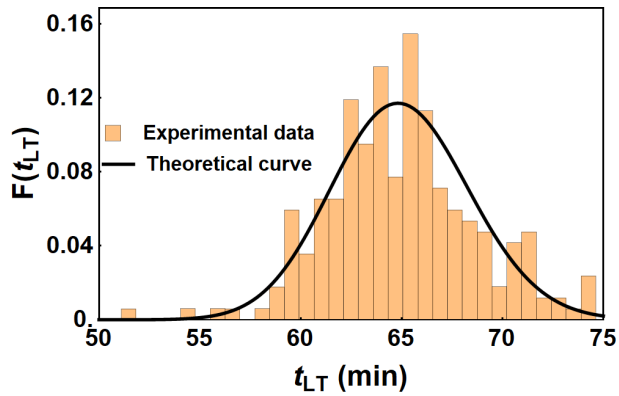


FIG. 4: Comparison of experimental [21] and theoretical lysis time distributions. Here we use $b = 3$, $k = 10\ \text{min}^{-1}$ and $X = 1500$ for the theoretical formula (Eq. (12)).

E. Feedback Effects

Eq. (10) is couched in terms of generic k_i, which is the parameter that controls protein synthesis rates in the model. As can be seen from Eq. (2), k_i controls the Poisson rate out of state i. In general k_i can be a function of i, especially if the protein production is under some form of auto-regulation, as is very common in protein synthesis. Thus k_i could be constant, increase, or decrease as i increases, which corresponds to the cases of no regulation, positive regulation and negative regulation respectively. For both positive and negative self-regulation, we can numerically calculate the full distribution of first passage times by inputting the form of the self-regulation in Eq. (10). One of the important questions we can ask is whether incorporation of feedback has an impact on the noise in first passage times or not, which determines the accuracy of a FPT-based molecular clock. For a given level of threshold, [22] already compared the relative noise for the case of no feedback, positive feedback and negative feedback and found that the no feedback case typically has the lower level of noise. We show in Eq. (17) that for the no-regulation case the CV depends

only on the burst size and the value of the threshold X. With addition of any kind of feedback, negative or positive, we are adding extra noise as a consequence of temporal inhomogeneity of the translation process, on top of the unavoidable Poissonian noise of translational burst.

We can also examine the behavior of the noise in the FPT distribution as a function of the protein threshold, X. We do this below for the cases of positive and negative feedback.

1. *Positive Feedback*

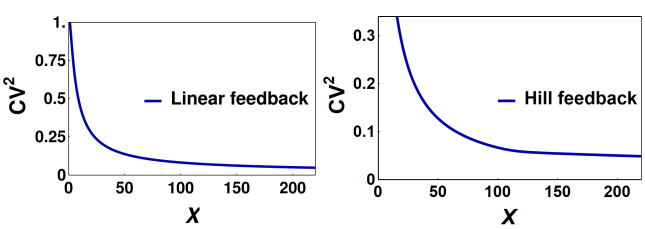
We implemented positive feedback using either a linear function,

$$k_i = \alpha + \beta i, \tag{23}$$

or a Hill function,

$$k_i = \alpha \frac{\gamma + (\beta i)^n}{1 + (\beta i)^n}. \tag{24}$$

For numerical calculation we need to choose explicit numerical forms of these functions. Using arbitrarily chosen numerical parameters we plotted the coefficient of variation of the distribution as calculated using our results above against the protein threshold, X. The results are shown in Fig. 5, and they show that the relative noise in the FPT distribution reduces with increasing X (or equivalently with increasing mean FPT), but then effectively plateaus. Interestingly this behavior can be seen in both forms of the feedback.



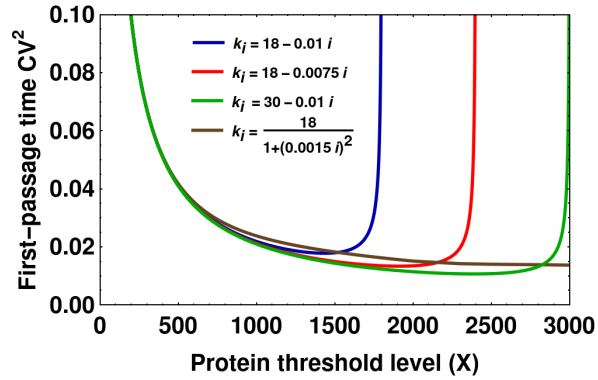
(a) (b)

FIG. 5: For stable protein ($\gamma = 0$) and for burst size, $b = 3$, noise in FPT (CV^2) is plotted as a function of protein threshold value (X) for different kinds of positive feedback: (a) Linear feedback, $k_i = 1.5 + 0.0725i$ and (b) hill feedback $k_i = \frac{15(0.1+(0.008i)^{15})}{1+(0.008i)^{15}}$

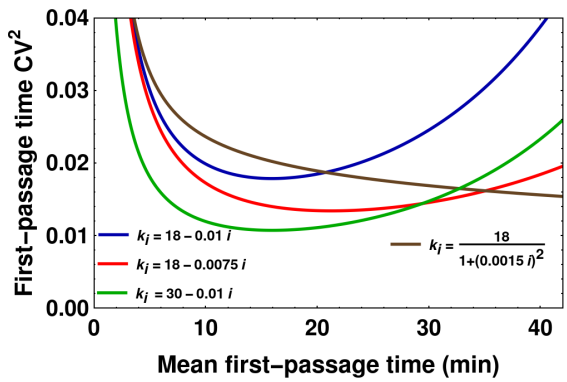
2. *Negative Feedback*

The case of negative feedback turned out to be more interesting. We used a linear form of negative feedback, $k_i = \alpha - \beta i$, as well as a Hill function form, given by

$$k_i = \frac{\alpha}{1 + (\beta i)^n} \tag{25}$$



(a)



(b)

FIG. 6: For stable protein ($\gamma = 0$) and for burst size, $b = 10$, noise in FPT (CV^2) is plotted as a function of (a) protein threshold value (X) and (b) mean FPT, for linear and hill negative feedbacks are plotted. For the linear feeback, the transription rate amplitude and feedback strengths are varied to see how the optimal mean and X depends on them.

Our results are shown in Fig. 6. It may be seen that for the Hill function there is a monotonic decrease, while for the linear case, the noise in the FPT times first decreases and then increases, yielding a U-shaped curve. It was recently shown in [41] that the CV of lysis timing for different bacteriophage mutants as a function of the effective protein threshold is also U-shaped. In that paper they attributed this effect to the effect of dilution due to bacterial growth. We show here that a general negative feedback that may arise due to dilution, or possibly could be caused by some as yet undiscovered mechanism, reproduces this relationship.

IV. DISCUSSION

In biology, it is often very difficult to find exact distributions of even simple stochastic processes. In those cases, people resort to expressions of mean and variance. But, the mean and variance contain partial information

about a process, and in particular do not capture large deviations from central values. Analytic expression for full distributions are thus desirable when possible.

Many a time, the differences in time scales of various sub-processes involved are exploited to simplify the dynamics of the stochastic process, in order to make the problem analytically solvable or simpler. For long time, people relied on mean and variance of protein concentration in the gene expression of 2-state promoter system. Only recently, for some limiting timescale of protein lifetime, the steady state protein distribution could be calculated for the same problem [29]. In another example,[35], time-dependent distributions of mRNA and protein numbers are obtained, for detailed model of gene expression, by using biologically relevant timescale approximations.

In this paper, we used a similar strategy, where in the bulk of the paper we ignored protein degradation, and calculated exact FPT distribution for protein thresholds. This assumption is valid for stable proteins whose mean lifetime is much longer than the event timescale. Otherwise, our results are very general. An additional strength of these results lie in being able to predict the exact distribution with arbitrary types of feedback. Many proteins exhibit auto-regulation, thus acting as self-activators or self-repressors, which could in principle manifest as arbitrary functional forms of feedback. Indirect feedback of protein levels upon their own transcription and translation is also common, and could again be approximated as self-regulation. In all such cases where we can impute a functional form for the parameter $k(i)$, the exact FPT distribution can be calculated by inputting that form into the general formulae obtained above. In principle, we can also solve the inverse problem and get insight about the dynamics of the process by measuring the experimental FPT distribution and using it to deduce the functional form of the feedback. Thus high precision experimental data on the FPT of lysis by lambda phage or any other similar process would be very valuable to help us discriminate between these different possibilities using the analytical expressions derived here.

While the bulk of the paper operates under the assumption of long-lived proteins, we also calculate the exact FPT distribution in Laplace case for non-zero degradation. We show that for the case with protein degradation, the main analytical difficulty is the complexity of the poles of the Laplace Transform, which involve solutions of very high order polynomial equations. We have also studied the noise in first passage times as described by CV and shown that for linear negative feedback, it has a minimum at an optimal protein threshold as well as at a value of the mean first passage time.

Given their simplicity, protein hourglasses, where the cells time a process using the first passage time of a protein level reaching a threshold, are likely to be widespread in cell biology. One can exploit the analytic expression of the distribution we obtained, to find the probability distribution of other cellular processes that depend on the FPT. One such example is the stochastic number of viral progeny on lysis, or the viral burst-size distribution, after a host-cell infection. We hope to extend our work towards these applications in future.

ACKNOWLEDGMENTS

D.D. would like to acknowledge SERB India (grant no. MTR/2019/000341) for financial support. A.P. would like to acknowledge hospitality and support from the IIT Bombay Physics department, where he was a visiting scholar for a month when some of this research was carried out. K.R. thanks IIT Bombay for Institute Ph.D. fellowship.

Appendix A: FPT distribution calculation for unequal transition rates and $\gamma = 0$

Bromwich Integral: For a Laplace transform $f(s)$, the inverse transform is given by:

$$\mathcal{L}^{-1}[f(s)] = \frac{1}{2\pi i}\lim_{T\to\infty}\int_{\gamma-iT}^{\gamma+iT} e^{st} f(s)ds \tag{A1}$$

where the integration is done along the vertical line $Re(s) = \gamma$ in the complex plane such that γ is greater than the real part of all singularities of $f(s)$ and $f(s)$ is bounded on the line. As we will see later, $f(s)$ in our case is always of the form $P(s)/Q(s)$, where P and Q are polynomials of s and degree of Q is greater than that of P. Such functions are always bounded over the entire complex plane except at the finite number of poles. So, we can always find the vertical line $Re(s) = \gamma$ such that the boundedness condition is satisfied. Also,$P(s)/Q(s) \to 0$ as $s \to \infty$. So, we can do integration in an infinite D-shaped contour. Then, in our case, using residue theorem, Eq. (A1) becomes [42]

$$\mathcal{L}^{-1}[f(s)] = \frac{1}{2\pi i}\sum_{p_i}\mathrm{Res}(e^{st}f(s))$$

where, the summation is over all the poles, p_i, of $f(s)$. So,

$$\mathcal{L}^{-1}[f(s)] = \sum_{p_i} \frac{1}{(n_i-1)!} \frac{d^{n_i-1}}{ds^{n_i-1}} \left((s-p_i)^{n_i} e^{st} f(s)\right)\Big|_{s=p_i} \tag{A2}$$

where, n_i is the order of ith pole, p_i.
For this general case of unequal transition rates with no protein degradation, recall from the main text

$$(s\mathbb{I}-A)^{-1}_{ij} = \begin{cases} \frac{b^{j-i}k_{i-1}}{(b+1)^{j-i+1}} \frac{\prod_{q=i}^{j-2}(s+k_q)}{\prod_{p=i-1}^{j-1}(s+\frac{bk_p}{b+1})} & \text{if } j > i+1 \\ \frac{k_{i-1}}{(b+1)^2(s+\frac{bk_{i-1}}{b+1})(s+\frac{bk_i}{b+1})} & \text{if } j = i+1 \\ \frac{1}{s+\frac{bk_{i-1}}{b+1}} & \text{if } j = i \\ 0 & \text{if } j < i \end{cases} \tag{A3}$$

and

$$f_{0,X}(t) = \mathcal{L}^{-1}[1 - s\sum_{j=1}^{X}(s\mathbb{I}-A)^{-1}_{1j}] \tag{A4}$$

For $X = 1$

$$\begin{aligned} f_{0,1}(t) =& \mathcal{L}^{-1}[1 - s(s\mathbb{I}-A)^{-1}_{11}] \\ =& \mathcal{L}^{-1}\left[1 - \frac{s}{s+\frac{bk_0}{b+1}}\right] \\ =& \frac{bk_0}{b+1}\mathcal{L}^{-1}\left[\frac{1}{s+\frac{bk_0}{b+1}}\right] \\ =& \frac{bk_0}{b+1}e^{\frac{-bk_0 t}{b+1}} \end{aligned}$$

For $X = 2$

$$\begin{aligned} f_{0,2}(t) =& \mathcal{L}^{-1}[1 - s(s\mathbb{I}-A)^{-1}_{11} - s(s\mathbb{I}-A)^{-1}_{12}] \\ =& \frac{bk_0}{b+1}e^{\frac{-bk_0 t}{b+1}} + \frac{k_0}{(b+1)^2}\mathcal{L}^{-1}\left[\frac{s}{(s+\frac{bk_0}{b+1})(s+\frac{bk_1}{b+1})}\right] \\ =& \frac{bk_0}{b+1}e^{\frac{-bk_0 t}{b+1}} - \frac{k_0}{(b+1)^2}\left[-\frac{bk_0}{b+1}\frac{e^{\frac{-bk_0 t}{b+1}}}{\frac{-bk_0}{b+1}+\frac{bk_1}{b+1}} - \frac{bk_1}{b+1}\frac{e^{\frac{-bk_1 t}{b+1}}}{\frac{-bk_1}{b+1}+\frac{bk_0}{b+1}}\right] \text{... using Eq. (A2)} \\ =& \frac{bk_0}{(b+1)^2}\left[\left(\frac{k_1}{k_1-k_0}+b\right)e^{-\frac{bk_0 t}{b+1}} + \left(\frac{k_1}{k_0-k_1}\right)e^{-\frac{bk_1 t}{b+1}}\right] \end{aligned}$$

For $X > 2$

$$\begin{aligned} f_{0,X}(t) =& \mathcal{L}^{-1}[1 - s(s\mathbb{I}-A)^{-1}_{11} - s(s\mathbb{I}-A)^{-1}_{12}] - \mathcal{L}^{-1}[s\sum_{j=3}^{X}(s\mathbb{I}-A)^{-1}_{1j}] \\ =& \frac{bk_0}{(b+1)^2}\left[\left(\frac{k_1}{k_1-k_0}+b\right)e^{-\frac{bk_0 t}{b+1}} + \left(\frac{k_1}{k_0-k_1}\right)e^{-\frac{bk_1 t}{b+1}}\right] - \sum_{j=3}^{X}\frac{b^{j-1}k_0}{(b+1)^j}\mathcal{L}^{-1}\left[\frac{s\prod_{q=1}^{j-2}(s+k_q)}{\prod_{p=0}^{j-1}(s+\frac{bk_p}{b+1})}\right] \end{aligned}$$

for each j, there are j number of simple poles at $bk_n/(b+1)$ with n going from 0 to $j-1$. So using Eq. (A2), we have

$$=\frac{bk_0}{(b+1)^2}\left[\left(\frac{k_1}{k_1-k_0}+b\right)e^{-\frac{bk_0t}{b+1}}+\left(\frac{k_1}{k_0-k_1}\right)e^{-\frac{bk_1t}{b+1}}\right]-\sum_{j=3}^{X}\frac{b^{j-1}k_0}{(b+1)^j}\sum_{n=0}^{j-1}\left[-\frac{bk_n}{b+1}\frac{\prod_{q=1}^{j-2}(-\frac{bk_n}{b+1}+k_q)}{\prod_{p=0,p\neq n}^{j-1}(-\frac{bk_n}{b+1}+\frac{bk_p}{b+1})}e^{-\frac{bk_nt}{b+1}}\right]$$

$$=\frac{bk_0}{(b+1)^2}\left[\left(\frac{k_1}{k_1-k_0}+b\right)e^{-\frac{bk_0t}{b+1}}+\left(\frac{k_1}{k_0-k_1}\right)e^{-\frac{bk_1t}{b+1}}+\sum_{j=3}^{X}\sum_{n=0}^{j-1}\frac{k_n}{(b+1)^{j-2}}\frac{\prod_{q=1}^{j-2}((b+1)k_q-bk_n)}{\prod_{p=0,p\neq n}^{j-1}(k_p-k_n)}e^{-\frac{bk_nt}{b+1}}\right]$$

Appendix B: Proof of $(s\mathbb{I}-\mathbf{A})^{-1}(s\mathbb{I}-\mathbf{A})=\mathbb{I}$ for special case of equal transition rates and $\gamma=0$

$$(s\mathbb{I}-A)^{-1}_{ij}=\begin{cases}\frac{kb^{j-i}(s+k)^{j-i-1}}{(b+1)^{j-i+1}(s+\frac{bk}{b+1})^{j-i+1}} & \text{if } j>i\\ \frac{1}{s+\frac{bk}{b+1}} & \text{if } j=i\\ 0 & \text{if } j<i\end{cases}\tag{B1}$$

From Eq. (5) in main text,

$$(s\mathbb{I}-A)_{ij}=\begin{cases}-\frac{kb^{j-i}}{(b+1)^{j-i+1}} & \text{if } j>i\\ s+\frac{kb}{b+1} & \text{if } j=i\\ 0 & \text{if } j<i\end{cases}\tag{B2}$$

Product of two upper triangular matrices is also upper triangular. So,
$[(s\mathbb{I}-A)^{-1}(s\mathbb{I}-A)]_{ij}=0$ for $j<i$.
Also, $[(s\mathbb{I}-A)^{-1}(s\mathbb{I}-A)]_{ii}=(s\mathbb{I}-A)^{-1}_{ii}(s\mathbb{I}-A)_{ii}=\frac{1}{s+\frac{bk}{b+1}}\left(s+\frac{kb}{b+1}\right)=1$
For $j>i$,

$$\begin{aligned}[(s\mathbb{I}-A)^{-1}(s\mathbb{I}-A)]_{ij}=&\sum_{m=i}^{j}(s\mathbb{I}-A)^{-1}_{im}(s\mathbb{I}-A)_{mj}\\ =&(s\mathbb{I}-A)^{-1}_{ii}(s\mathbb{I}-A)_{ij}+(s\mathbb{I}-A)^{-1}_{ij}(s\mathbb{I}-A)_{jj}+\sum_{m=i+1}^{j-1}(s\mathbb{I}-A)^{-1}_{im}(s\mathbb{I}-A)_{mj}\\ =&-\frac{1}{s+\frac{kb}{b+1}}\frac{kb^{j-i}}{(b+1)^{j-i+1}}-\frac{kb^{j-i}(s+k)^{j-i-1}}{(b+1)^{j-i+1}(s+\frac{bk}{b+1})^{j-i}}+\sum_{m=i+1}^{j-1}(s\mathbb{I}-A)^{-1}_{im}(s\mathbb{I}-A)_{mj}\\ =&-\frac{1}{s+\frac{kb}{b+1}}\frac{kb^{j-i}}{(b+1)^{j-i+1}}\left[1-\left(\frac{s+k}{s+\frac{bk}{b+1}}\right)^{j-i-1}\right]+\sum_{m=i+1}^{j-1}(s\mathbb{I}-A)^{-1}_{im}(s\mathbb{I}-A)_{mj}\end{aligned}\tag{B3}$$

$$\begin{aligned}\sum_{m=i+1}^{j-1}(s\mathbb{I}-A)^{-1}_{im}(s\mathbb{I}-A)_{mj}=&\frac{k^2b^{j-i}(s+k)^{-i-1}}{(b+1)^{j-i+2}(s+\frac{bk}{b+1})^{-i+1}}\sum_{m=i+1}^{j-1}\left(\frac{s+k}{s+\frac{bk}{b+1}}\right)^m\\ =&\frac{k^2b^{j-i}(s+k)^{-i-1}}{(b+1)^{j-i+2}(s+\frac{bk}{b+1})^{-i+1}}\frac{\left[1-\left(\frac{s+k}{s+\frac{bk}{b+1}}\right)^{j-i-1}\right]}{1-\frac{s+k}{s+\frac{bk}{b+1}}}\\ =&\frac{1}{s+\frac{kb}{b+1}}\frac{kb^{j-i}}{(b+1)^{j-i+1}}\left[1-\left(\frac{s+k}{s+\frac{bk}{b+1}}\right)^{j-i-1}\right]\end{aligned}\tag{B4}$$

Putting Eq. (B4) into Eq. (B3), we get

$$[(s\mathbb{I}-A)^{-1}(s\mathbb{I}-A)]_{ij}=0$$

Hence, $(s\mathbb{I}-\mathbf{A})^{-1}(s\mathbb{I}-\mathbf{A})=\mathbb{I}$

Appendix C: FPT distribution calculation for special case of equal transition rates and $\gamma=0$

$$(s\mathbb{I}-A)^{-1}_{ij}=\begin{cases}\frac{kb^{j-i}(s+k)^{j-i-1}}{(b+1)^{j-i+1}(s+\frac{bk}{b+1})^{j-i+1}} & \text{if } j>i\\ \frac{1}{s+\frac{bk}{b+1}} & \text{if } j=i\\ 0 & \text{if } j<i\end{cases} \tag{C1}$$

We know that $\tilde{f}_{0,X}(s)=1-s\tilde{S}_{0,X}(s)$. So,

$$\begin{aligned}\tilde{f}_{0,X}(s)=&1-s\sum_{j=1}^{X}(s\mathbb{I}-A)^{-1}_{1j}\\
=&1-\frac{s}{s+\frac{bk}{b+1}}-s\sum_{j=2}^{X}\frac{kb^{j-1}(s+k)^{j-2}}{(b+1)^j(s+\frac{bk}{b+1})^j}\\
=&\frac{\frac{bk}{b+1}}{s+\frac{bk}{b+1}}-\frac{sk}{b(s+k)^2}\sum_{j=2}^{X}\frac{b^j(s+k)^j}{(b+1)^j(s+\frac{bk}{b+1})^j}\\
=&\frac{\frac{bk}{b+1}}{s+\frac{bk}{b+1}}-\frac{sk}{b(s+k)^2}\frac{b^2(s+k)^2}{(b+1)^2(s+\frac{bk}{b+1})^2}\frac{1-\frac{b^{X-1}(s+k)^{X-1}}{(b+1)^{X-1}(s+\frac{bk}{b+1})^{X-1}}}{1-\frac{b(s+k)}{(b+1)(s+\frac{bk}{b+1})}}\\
=&\frac{\frac{bk}{b+1}}{s+\frac{bk}{b+1}}-\frac{sk}{b(s+k)^2}\frac{b^2(s+k)^2}{(b+1)(s+\frac{bk}{b+1})}\frac{1-\frac{b^{X-1}(s+k)^{X-1}}{(b+1)^{X-1}(s+\frac{bk}{b+1})^{X-1}}}{(b+1)(s+\frac{bk}{b+1})-b(s+k)}\\
=&\frac{\frac{bk}{b+1}}{s+\frac{bk}{b+1}}-\frac{\frac{bk}{b+1}}{s+\frac{bk}{b+1}}\left[1-\frac{b^{X-1}(s+k)^{X-1}}{(b+1)^{X-1}(s+\frac{bk}{b+1})^{X-1}}\right]\\
=&\frac{kb^X(s+k)^{X-1}}{(b+1)^X(s+\frac{bk}{b+1})^X}\end{aligned}$$

Using Eq. (A2),

$$\begin{aligned}f_{0,X}(t)=&\mathcal{L}^{-1}[\tilde{f}_{0,X}(s)]\\
=&\frac{kb^X}{(b+1)^X}\mathcal{L}^{-1}\left[\frac{(s+k)^{X-1}}{(s+\frac{bk}{b+1})^X}\right]\\
=&\frac{kb^X}{(b+1)^X}\frac{1}{(X-1)!}\frac{d^{X-1}}{ds^{X-1}}\left[e^{st}(s+k)^{X-1}\right]\Big|_{s=-\frac{bk}{b+1}}\end{aligned}$$

We can use Leibniz rule to calculate $(X-1)^{th}$ derivative

$$\begin{aligned}=&\frac{kb^X}{(b+1)^X}\frac{1}{(X-1)!}\sum_{n=0}^{X-1}\binom{X-1}{n}t^ne^{st}\frac{(X-1)!}{n!}(s+k)^n\Big|_{s=-\frac{bk}{b+1}}\\
=&\frac{kb^X}{(b+1)^X}\left[\sum_{n=0}^{X-1}\frac{1}{n!}\binom{X-1}{n}\left(\frac{kt}{b+1}\right)^n\right]e^{-\frac{kbt}{b+1}}\end{aligned}$$

The term inside the square bracket is a confluent hypergeometric function ${}_1F_1$, and hence

$$f_{0,X}(t) = \frac{kb^X}{(b+1)^X} {}_1F_1[1-X, 1, -\frac{kt}{b+1}]e^{-\frac{kbt}{b+1}} \tag{C2}$$

Appendix D: For small values of X, FPT distributions for unequal transition rates and $\gamma \neq 0$

For $X = 1$, protein degradation is not relevant. So,

$$f_{0,1}(t) = \frac{bk_0}{b+1} e^{\frac{-bk_0 t}{b+1}}$$

For $X = 2$,

$$\tilde{f}_{0,2}(s) = \frac{bk_0(s+k_1+\gamma)}{(b+1)(s+\alpha_+)(s+\alpha_-)}$$

where

$$\alpha\pm = \frac{-\sqrt{(b+1)^2\left(b^2k_0^2 - 2bk_0\left((b-1)\gamma + bk_1\right) + (b\gamma + bk_1 + \gamma)^2\right)} + (b+1)^2\gamma \pm b(b+1)\left(k_0+k_1\right)}{2(b+1)^2}$$

So, using Eq. (A2),

$$f_{0,2}(t) = \frac{b^2k_0}{(b+1)^2}\left[\frac{k_1+\gamma-\alpha_+}{\alpha_- - \alpha_+}e^{-\alpha_+ t} + \frac{k_1+\gamma-\alpha_-}{\alpha_+ - \alpha_-}e^{-\alpha_- t}\right]$$

For $X = 3$,

$$\tilde{f}_{0,3}(s) = \frac{b^3k_0(s+k_1+\gamma)(s+k_2+2\gamma)}{(b+1)^3(s+\beta)(s+\delta)(s+\delta^*)}$$

where δ and δ^* are complex conjugates. The values of β, δ and δ^* are too big to be put here. Using Eq. (A2), we get

$$\begin{aligned} f_{0,3}(t) =& \frac{b^3k_0}{(b+1)^3}\left[\frac{(k_1+\gamma-\beta)(k_2+2\gamma-\beta)}{(\delta-\beta)(\delta^*-\beta)}e^{-\beta t} + \frac{(k_1+\gamma-\delta)(k_2+2\gamma-\delta)}{(\beta-\delta)(\delta^*-\delta)}e^{-\delta t}\right. \\ &\left. + \frac{(k_1+\gamma-\delta^*)(k_2+2\gamma-\delta^*)}{(\beta-\delta^*)(\delta-\delta^*)}e^{-\delta^* t}\right] \end{aligned}$$

Appendix E: For small values of X, FPT distribution for equal transition rate and $\gamma \neq 0$

$$\tilde{f}_{0,X}(s) = \frac{b^X k \prod_{l=1}^{X-1}(\gamma l + k + s)}{\prod_{i=0}^{X-1}(\gamma i + s) + b^X(k+s)\prod_{j=1}^{X-1}(\gamma j + k + s) + \sum_{n=1}^{X-1}\left[\binom{X}{n} b^n \prod_{i=0}^{X-1}(\gamma i + k\Theta(i-X+n) + s)\right]} \tag{E1}$$

For $X = 1$, protein degradation is not relevant. So,

$$f_{0,1}(t) = \frac{bk}{b+1} e^{\frac{-bkt}{b+1}}$$

For $X = 2$,

$$\tilde{f}_{0,2}(s) = \frac{b^2 k(s+k+\gamma)}{(b+1)^2(s+\alpha_+)(s+\alpha_-)}$$

where $\alpha_\pm = \frac{bk}{b+1} + \frac{\gamma}{2} \pm \sqrt{\frac{\gamma^2}{4} + \frac{bk\gamma}{(b+1)^2}}$
So, using Eq. (A2),

$$f_{0,2}(t) = \frac{b^2 k}{(b+1)^2} \left[\frac{k+\gamma-\alpha_+}{\alpha_- - \alpha_+} e^{-\alpha_+ t} + \frac{k+\gamma-\alpha_-}{\alpha_+ - \alpha_-} e^{-\alpha_- t} \right]$$

For $X = 3$,

$$\tilde{f}_{0,3}(s) = \frac{b^3 k(s+k+\gamma)(s+k+2\gamma)}{(b+1)^3(s+\beta)(s+\delta)(s+\delta^*)}$$

where δ and δ^* are complex conjugates and

$$\begin{aligned}\beta = &- \frac{\sqrt[3]{\frac{2}{3}}(b+1)\gamma\left((b+1)^2\gamma + 3bk\right)}{\sqrt[3]{\sqrt{3}\sqrt{(b+1)^{12}\gamma^3\left(-108b^3k^3 - 27b^2(b+3)(3b+1)\gamma k^2 - 4(b+1)^6\gamma^3 - 36b(b+1)^4\gamma^2 k\right)} + 9(b-1)b(b+1)^6\gamma^2 k}} \\ &- \frac{\sqrt[3]{\sqrt{3}\sqrt{(b+1)^{12}\gamma^3\left(-108b^3k^3 - 27b^2(b+3)(3b+1)\gamma k^2 - 4(b+1)^6\gamma^3 - 36b(b+1)^4\gamma^2 k\right)} + 9(b-1)b(b+1)^6\gamma^2 k}}{\sqrt[3]{2}2^{2/3}(b+1)^3} \\ &+ \frac{bk}{b+1} + \gamma\end{aligned}$$

$$\begin{aligned}\delta = &\frac{i\left(\sqrt{3}-i\right)(b+1)\gamma\left((b+1)^2\gamma + 3bk\right)}{2^{2/3}\sqrt[3]{3}\sqrt[3]{\sqrt{3}\sqrt{(b+1)^{12}\gamma^3\left(-108b^3k^3 - 27b^2(b+3)(3b+1)\gamma k^2 - 4(b+1)^6\gamma^3 - 36b(b+1)^4\gamma^2 k\right)} + 9(b-1)b(b+1)^6\gamma^2 k}} \\ &- \frac{i\left(\sqrt{3}+i\right)\sqrt[3]{\sqrt{3}\sqrt{(b+1)^{12}\gamma^3\left(-108b^3k^3 - 27b^2(b+3)(3b+1)\gamma k^2 - 4(b+1)^6\gamma^3 - 36b(b+1)^4\gamma^2 k\right)} + 9(b-1)b(b+1)^6\gamma^2 k}}{2\sqrt[3]{2}2^{2/3}(b+1)^3} \\ &+ \frac{bk}{b+1} + \gamma\end{aligned}$$

Using Eq. (A2), we get

$$\begin{aligned}f_{0,3}(t) = &\frac{b^3 k}{(b+1)^3} \left[\frac{(k+\gamma-\beta)(k+2\gamma-\beta)}{(\delta-\beta)(\delta^*-\beta)} e^{-\beta t} + \frac{(k+\gamma-\delta)(k+2\gamma-\delta)}{(\beta-\delta)(\delta^*-\delta)} e^{-\delta t} \right. \\ &\left. + \frac{(k+\gamma-\delta^*)(k+2\gamma-\delta^*)}{(\beta-\delta^*)(\delta-\delta^*)} e^{-\delta^* t} \right]\end{aligned}$$

Appendix F: Moments calculation

We define a vector $\langle \mathbf{T}^3 \rangle$, whose components are third moments of the FPTs starting from various protein counts i. Doing integration by parts successively, we get:

$$\langle \mathbf{T}^3 \rangle = -\mathbf{A} \left[\int_0^\infty t^3 exp(\mathbf{A}t) dt \right] \mathbf{S}(0) = -6\mathbf{A}^{-3}\mathbf{S}(0) \tag{F1}$$

Noting the fact that

$$A_{ij}^{-2} = \sum_{m=i}^{j} A_{im}^{-1} A_{mj}^{-1} = \sum_{m=i}^{j} \left(\frac{1+b\delta_{im}}{bk_{m-1}}\right)\left(\frac{1+b\delta_{mj}}{bk_{j-1}}\right) = \sum_{m=i}^{j} \frac{1}{b^2 k_{m-1}k_{j-1}} \left[1 + b\delta_{mj} + b\delta_{im} + b^2\delta_{im}\delta_{mj}\right]$$
$$= \frac{1}{bk_{j-1}} \left[\frac{1}{k_{j-1}} + \frac{1}{k_{i-1}} + \frac{b\delta_{ij}}{k_{i-1}} + \frac{1}{b}\sum_{m=i}^{j} \frac{1}{k_{m-1}}\right],$$

we finally have using Eq. (F1)

$$\langle T_0^3 \rangle = -6\sum_{j=1}^{X} A_{1j}^{-3} = -6\sum_{j=1}^{X}\sum_{p=1}^{X} A_{1p}^{-2} A_{pj}^{-1} = -6\sum_{j=1}^{X}\sum_{p=1}^{j} A_{1p}^{-2} A_{pj}^{-1}$$
$$= \frac{6}{b^2}\sum_{j=1}^{X}\sum_{p=1}^{j} \frac{1+b\delta_{pj}}{k_{p-1}k_{j-1}} \left[\frac{1}{k_{p-1}} + \frac{1}{k_0} + \frac{b\delta_{1p}}{k_0} + \frac{1}{b}\sum_{m=1}^{p} \frac{1}{k_{m-1}}\right] \tag{F2}$$

Next we define a vector $\langle \mathbf{T}^4 \rangle$, whose components are fourth moments of the FPTs starting from various protein counts i. Again doing integration by parts successively, we get:

$$\langle \mathbf{T}^4 \rangle = -\mathbf{A}\left[\int_0^\infty t^4 exp(\mathbf{A}t)dt\right]\mathbf{S}(0) = 24\mathbf{A}^{-4}\mathbf{S}(0) \tag{F3}$$

So,

$$\langle T_0^4 \rangle = 24\sum_{j=1}^{X} A_{1j}^{-4} = 24\sum_{j=1}^{X}\sum_{p=1}^{j} A_{1p}^{-2} A_{pj}^{-2}$$
$$= 24\sum_{j=1}^{X}\sum_{p=1}^{j} \frac{1}{b^2 k_{p-1}k_{j-1}} \left[\frac{1}{k_{p-1}} + \frac{1}{k_0} + \frac{b\delta_{1p}}{k_0} + \frac{1}{b}\sum_{m=1}^{p} \frac{1}{k_{m-1}}\right]\left[\frac{1}{k_{j-1}} + \frac{1}{k_{p-1}} + \frac{b\delta_{pj}}{k_{p-1}} + \frac{1}{b}\sum_{n=p}^{j} \frac{1}{k_{n-1}}\right] \tag{F4}$$

Appendix G: Alternative formalism to calculate FPT distribution for general case of unequal transition rates with no degradation

Here we begin with Eq. (7) in the main text and evaluate $exp(\mathbf{A}t)$.

Firstly we note from Eq. (5) in main text that matrix $\mathbf{A}$ is upper triangular, for $\gamma = 0$ and hence the eigenvalues are given by its diagonal elements:

$$\lambda_j = -\frac{bk_{j-1}}{(b+1)}. \tag{G1}$$

Next we show that the matrix which diagonalises $\mathbf{A}$ such that $\mathbf{D^{-1}AD} = \mathbf{A_D}$ (with $(\mathbf{A_D})_{ij} = \lambda_j\delta_{ij}$) has its elements given by:

$$D_{ij} = \begin{cases} \frac{k_{i-1}}{(b+1)^{j-i}} \frac{\prod_{p=i}^{j-2}((b+1)k_p - bk_{j-1})}{\prod_{p=i-1}^{j-2}(k_p - k_{j-1})} & \text{if } j > i+1 \\ \frac{k_{i-1}}{(b+1)(k_{i-1}-k_i)} & \text{if } j = i+1 \\ 1 & \text{if } j = i \\ 0 & \text{if } j < i \end{cases} \tag{G2}$$

With the above form of $\mathbf{D}$ it may be shown directly that

$$\mathbf{AD}_j = \lambda_j \mathbf{D}_j \tag{G3}$$

such that the vector $\mathbf{D}_j$ is the j-th column of $\mathbf{D}$ and an eigenvector of $\mathbf{A}$.

Proof of $\mathbf{AD}_j = \lambda_j \mathbf{D}_j$:

Using Eqs (5) (in main text), (G1) and (G2), we get for $j > (i+1)$,

$$\begin{aligned}(\mathbf{AD}_j)_i = (\mathbf{AD})_{ij} = \sum_{m=i}^{j} A_{im} D_{mj} =& k_{i-1} \sum_{m=i}^{j-2} \frac{b^{m-i}(1-(b+1)\delta_{im})}{(b+1)^{m-i+1}} \frac{k_{m-1}}{(b+1)^{j-m}} \frac{\prod_{p=m}^{j-2}((b+1)k_p - bk_{j-1})}{\prod_{p=m-1}^{j-2}(k_p - k_{j-1})} \\ &+ \frac{k_{i-1} b^{j-1-i}}{(b+1)^{j-i+1}} \frac{k_{j-2}}{(b+1)(k_{j-2}-k_{j-1})} + \frac{k_{i-1} b^{j-i}}{(b+1)^{j-i+1}} \\ =& \frac{k_{i-1} b\, Y}{(b+1)^{j-i+1} \prod_{p=i-1}^{j-2}(k_p - k_{j-1})} \\ =& \frac{-k_{j-1} b k_{i-1} \prod_{p=i}^{j-2}((b+1)k_p - bk_{j-1})}{(b+1)^{j-i+1} \prod_{p=i-1}^{j-2}(k_p - k_{j-1})} = \lambda_j D_{ij} = \lambda_j (\mathbf{D}_j)_i \end{aligned} \tag{G4}$$

The quantity Y in the intermediate state above is a sum of several terms as follows, which simplify on repeated factorisation and cancellation to give the short expression in the last step above:

$$\begin{aligned} Y =& -k_{i-1} \prod_{p=i}^{j-2}((b+1)k_p - bk_{j-1}) + \sum_{m=i}^{j-2} b^{m-i} \left(\prod_{p=i-1}^{m-1} (k_p - k_{j-1})\; k_m \prod_{q=m+1}^{j-2} ((b+1)k_q - bk_{j-1}) \right) \\ &+ b^{j-i-1} \prod_{p=i-1}^{j-2} (k_p - k_{j-1}) \end{aligned} \tag{G5}$$

In the special case of $j = i+1$,

$$(\mathbf{AD}_{i+1})_i = (\mathbf{AD})_{i,i+1} = \frac{-bk_{i-1}^2}{(b+1)^2(k_{i-1}-k_i)} + \frac{bk_{i-1}}{(b+1)^2} = \left(\frac{-bk_i}{(b+1)}\right)\left(\frac{k_{i-1}}{(b+1)(k_{i-1}-k_i)}\right) = \lambda_{i+1} D_{i,i+1}, \tag{G6}$$

and for $j = i$

$$(\mathbf{AD}_i)_i = (\mathbf{AD})_{ii} = \frac{-bk_{i-1}}{(b+1)} = \lambda_i D_{ii}. \tag{G7}$$

Hence, proved.

The elements of the inverse matrix $\mathbf{D^{-1}}$ is given by

$$D_{ij}^{-1} = \begin{cases} \frac{k_{i-1}}{(b+1)^{j-i}} \frac{\prod_{p=i}^{j-2}((b+1)k_p - bk_{i-1})}{\prod_{p=i}^{j-1}(k_p - k_{i-1})} & \text{if } j > i+1 \\ \frac{k_{i-1}}{(b+1)(k_i - k_{i-1})} & \text{if } j = i+1 \\ 1 & \text{if } j = i \\ 0 & \text{if } j < i \end{cases} \tag{G8}$$

Using the above diagonal form, the vector of different first passage probability distributions in Eq. (7) may be written as

$$\mathbf{f}(t) = -\frac{\partial}{\partial t}[\mathbf{DPD^{-1}S}(0)], \tag{G9}$$

where $\mathbf{P} = \exp(\mathbf{A_D} t) = \mathbf{D^{-1}} \exp(\mathbf{A}t)\mathbf{D}$. Thus $(\mathbf{P})_{mn} = \exp(\lambda_m t)\delta_{mn}$. Moreover we are interested in the case of initial protein number being 0, and in the corresponding first passage distribution

$$f_{0,X}(t) = (\mathbf{f}(t))_1 = -\frac{\partial}{\partial t} \sum_{j=1}^{X} (\mathbf{DPD^{-1}})_{1j} (\mathbf{S}(0))_j. \tag{G10}$$

From Eq. (G8), $[PD^{-1}]_{ij} = \exp(-\frac{k_{i-1}bt}{b+1}) D_{ij}^{-1}$, and then it follows from Eq.(G2) that

$$[DPD^{-1}]_{ij} = \begin{cases} \frac{k_{i-1}}{(b+1)^{j-i}} \sum_{n=i-1}^{j-1} e^{-\frac{k_n bt}{b+1}} \frac{\prod_{p=i}^{j-2}((b+1)k_p - bk_n)}{\prod_{p=i-1,p\neq n}^{j-1}(k_p - k_n)} & \text{if } j > i+1 \\ \frac{k_{i-1}}{(b+1)(k_i - k_{i-1})} \left(e^{-\frac{k_{i-1}bt}{b+1}} - e^{-\frac{k_i bt}{b+1}} \right) & \text{if } j = i+1 \\ e^{-\frac{k_{i-1}bt}{b+1}} & \text{if } j = i \\ 0 & \text{if } j < i \end{cases} \tag{G11}$$

Recalling that all the elements of $\mathbf{S}(0)$ are 1, we have from Eq. (G10),

$$\begin{aligned} f_{0,X}(t) =& -\frac{\partial}{\partial t} \sum_{j=1}^{X} (DPD^{-1})_{1j} \\ =& -\frac{\partial}{\partial t} \left[e^{-\frac{k_0 bt}{b+1}} + \frac{k_0}{(b+1)(k_1 - k_0)} \left(e^{-\frac{k_0 bt}{b+1}} - e^{-\frac{k_1 bt}{b+1}} \right) + \sum_{j=3}^{X} \frac{k_0}{(b+1)^{j-1}} \sum_{n=0}^{j-1} e^{-\frac{k_n bt}{b+1}} \frac{\prod_{p=1}^{j-2}((b+1)k_p - bk_n)}{\prod_{p=0,p\neq n}^{j-1}(k_p - k_n)} \right] \\ =& \frac{k_0 b}{(b+1)^2} \left[\left(\frac{k_1}{k_1 - k_0} + b\right) e^{-\frac{bk_0 t}{b+1}} + \left(\frac{k_1}{k_0 - k_1}\right) e^{-\frac{bk_1 t}{b+1}} + \sum_{j=3}^{X} \sum_{n=0}^{j-1} \frac{k_n}{(b+1)^{j-2}} \frac{\prod_{p=1}^{j-2}((b+1)k_p - bk_n)}{\prod_{p=0,p\neq n}^{j-1}(k_p - k_n)} e^{-\frac{bk_n t}{b+1}} \right] \end{aligned} \tag{G12}$$

The above expression is valid for $X > 2$, and for the special cases of $X = 1$,

$$f_{0,1}(t) = \frac{k_0 b}{b+1} e^{-\frac{bk_0 t}{b+1}} \tag{G13}$$

and $X = 2$,

$$f_{0,2}(t) = \frac{k_0 b}{(b+1)^2} \left(\left(b + \frac{k_1}{k_1 - k_0} \right) e^{-\frac{bk_0 t}{b+1}} + \frac{k_1}{k_0 - k_1} e^{-\frac{bk_1 t}{b+1}} \right). \tag{G14}$$

Appendix H: Alternative formalism to calculate FPT distribution for special case of equal transition rates with no degradation

As mentioned in the main text, we can take proper limit of Eq. (10) to get non divergent FPT expression for this special case. We first write the first passage distributions for $X = 1, 2, 3, 4$ below in a particular way to observe a pattern:

$$\begin{aligned} X = 1: \quad & f_{0,1}(t) = \frac{k_0 b}{1+b} e^{-ak_0} \\ X = 2: \quad & f_{0,2}(t) = \frac{k_0 b}{(b+1)^2} \left[b e^{-ak_0} + k_1 \left(\frac{e^{-ak_0}}{\epsilon_{10}} + \frac{e^{-ak_1}}{\epsilon_{01}} \right) \right] \\ X = 3: \quad & f_{0,3}(t) = \frac{k_0 b}{(b+1)^3} \left[b^2 e^{-ak_0} + bk_1 \left(\frac{e^{-ak_0}}{\epsilon_{10}} + \frac{e^{-ak_1}}{\epsilon_{01}} \right) + bk_2 \left(\frac{e^{-ak_0}}{\epsilon_{20}} + \frac{e^{-ak_2}}{\epsilon_{02}} \right) \right. \\ & \qquad \left. + k_1 k_2 \left(\frac{e^{-ak_0}}{\epsilon_{10}\epsilon_{20}} + \frac{e^{-ak_1}}{\epsilon_{21}\epsilon_{01}} + \frac{e^{-ak_2}}{\epsilon_{12}\epsilon_{02}} \right) \right] \\ X = 4: \quad & f_{0,4}(t) = \frac{k_0 b}{(b+1)^4} \left[b^3 e^{-ak_0} + b^2 k_1 \left(\frac{e^{-ak_0}}{\epsilon_{10}} + \frac{e^{-ak_1}}{\epsilon_{01}} \right) + b^2 k_2 \left(\frac{e^{-ak_0}}{\epsilon_{20}} + \frac{e^{-ak_2}}{\epsilon_{02}} \right) \right. \\ & + b^2 k_3 \left(\frac{e^{-ak_0}}{\epsilon_{30}} + \frac{e^{-ak_3}}{\epsilon_{03}} \right) + bk_1 k_2 \left(\frac{e^{-ak_0}}{\epsilon_{10}\epsilon_{20}} + \frac{e^{-ak_1}}{\epsilon_{21}\epsilon_{01}} + \frac{e^{-ak_2}}{\epsilon_{12}\epsilon_{02}} \right) + bk_2 k_3 \left(\frac{e^{-ak_0}}{\epsilon_{30}\epsilon_{20}} + \frac{e^{-ak_2}}{\epsilon_{02}\epsilon_{32}} + \frac{e^{-ak_3}}{\epsilon_{23}\epsilon_{03}} \right) \\ & \left. + bk_1 k_3 \left(\frac{e^{-ak_0}}{\epsilon_{10}\epsilon_{30}} + \frac{e^{-ak_1}}{\epsilon_{31}\epsilon_{01}} + \frac{e^{-ak_3}}{\epsilon_{03}\epsilon_{13}} \right) + k_1 k_2 k_3 \left(\frac{e^{-ak_0}}{\epsilon_{10}\epsilon_{20}\epsilon_{30}} + \frac{e^{-ak_1}}{\epsilon_{01}\epsilon_{21}\epsilon_{31}} + \frac{e^{-ak_2}}{\epsilon_{02}\epsilon_{12}\epsilon_{32}} + \frac{e^{-ak_3}}{\epsilon_{03}\epsilon_{13}\epsilon_{23}} \right) \right] \end{aligned} \tag{H1}$$

In the above $a = bt/(b+1)$, and $\epsilon_{ij} = (k_i - k_j)$. Note that in Eqs. (H1), every coefficient within parentheses $\Big(...\Big)$ may be represented by a symbol $C_I^{(m)}$ involving a certain set of integers $I = \{0, i_1, i_2, \cdots, i_m\}$. The general form of the coefficient is as follows:

$$C_I^{(m)} = \sum_{j=0,i_1,i_2,\cdots}^{i_m} \frac{e^{-ak_j}}{\prod_{p\in I, p\neq j} \epsilon_{pj}} \tag{H2}$$

The coefficients $C_I^{(m)}$ have prefactors $b^{X-1-m} k_{i_1} k_{i_2} \cdots k_{i_m}$, and terms corresponding to all possible $\binom{X-1}{m}$ combinations (*comb*) of m integers out of the set $\{1, 2, \cdots, X-1\}$ are added. Thus the general form of $f_{0,X}(t)$, with $C_I^{(m)}$ from Eq. (H2), is finally:

$$f_{0,X}(t) = \frac{k_0 b}{(b+1)^X} \sum_{m=0}^{X-1} b^{X-1-m} \sum_{comb} k_{i_1} k_{i_2} \cdots k_{i_m} C_I^{(m)}. \tag{H3}$$

We then take the limit that k_j for all j approach the same value k by first writing $k_j = k + \Delta j$ and taking $lim \Delta \to 0$. By writing the exponential in the numerator of Eq. (H2) in series form, and noting all terms for $l > m$ vanish with $\Delta \to 0$, we have $lim_{\Delta\to 0} C_I^{(m)}$ given by:

$$\sum_{j=0,i_1,i_2,\cdots}^{i_m} \frac{e^{-ak}[1 + \sum_{l=1}^{m} \frac{(-a\Delta j)^l}{l!}]}{(-\Delta)^m \prod_{p\in I, p\neq j}(j-p)} = \frac{e^{-ak}}{(-\Delta)^m} \Bigg[\sum_{j=0,i_1,i_2,\cdots}^{i_m} \frac{1}{\prod_{p\in I, p\neq j}(j-p)} + \sum_{l=1}^{m} \frac{(-a\Delta)^l}{l!} \sum_{j=0,i_1,i_2,\cdots}^{i_m} \frac{j^l}{\prod_{p\in I, p\neq j}(j-p)} \Bigg]$$

At this point we assume the following two identities (see the discussion on proofs of these in Appendix I below):

$$\sum_{j=0,i_1,i_2,\cdots}^{i_m} \frac{1}{\prod_{p\in I, p\neq j}(j-p)} = 0, \quad \text{and} \quad \sum_{j=0,i_1,i_2,\cdots}^{i_m} \frac{j^l}{\prod_{p\in I, p\neq j}(j-p)} = \delta_{lm}, \tag{H4}$$

Using Eqs. (H4) we have

$$lim_{\Delta\to 0} C_I^{(m)} = \frac{e^{-ak} a^m}{m!}, \tag{H5}$$

and as $k_{i_1} k_{i_2} \cdots k_{i_m} \to k^m$ in the same limit, we have from Eq. (H3),

$$f_{0,X}(t) = \frac{kb^X}{(b+1)^X} \left[\sum_{m=0}^{X-1} \frac{1}{m!} \binom{X-1}{m} \left(\frac{kt}{b+1}\right)^m \right] e^{-\frac{kbt}{b+1}} \tag{H6}$$

The term inside the square bracket is a confluent hypergeometric function ${}_1F_1$, and hence

$$f_{0,X}(t) = \frac{kb^X}{(b+1)^X} {}_1F_1[1-X, 1, -\frac{kt}{b+1}] e^{-\frac{kbt}{b+1}} \tag{H7}$$

Appendix I: Discussion on proof of the identities in Eqs. (H4)

A proof can be give for a special case first. Let the set I be a simple set, $I = \{0, 1, 2, ..., m\}$ and $1 \leq j \leq m$
It is easy to see that

$$\frac{1}{\prod_{p\in I, p\neq j}(j-p)} = \frac{(-1)^{m-j}}{j!(m-j)!} \tag{I1}$$

Using Eq. (I1),

$$\begin{aligned}\sum_{j=0,i_1,i_2,\cdots}^{i_m}\frac{1}{\prod_{p\in I,p\neq j}(j-p)} &= \frac{1}{\prod_{p\in I,p\neq 0}(-p)}+\sum_{j=1}^{m}\frac{1}{\prod_{p\in I}(j-p)}\\ &= \frac{(-1)^m}{m!}+\sum_{j=1}^{m}\frac{(-1)^{m-j}}{j!(m-j)!}\\ &= \frac{(-1)^m}{m!}\sum_{j=0}^{m}(-1)^j\binom{m}{j}\\ &= \frac{(-1)^m}{m!}(1-1)^m\\ &= 0\end{aligned}$$

Hence, the first part of Eq. (H4) is proved, for the simple case. The second part is

$$\begin{aligned}\sum_{j=0,i_1,i_2,\cdots}^{i_m}\frac{j^l}{\prod_{p\in I,p\neq j}(j-p)} &= \sum_{j=1}^{m}\frac{j^l}{\prod_{p\in I,p\neq j}(j-p)}\\ &= \frac{1}{m!}\sum_{j=1}^{m}j^l(-1)^{m-j}\binom{m}{j}\\ &= \frac{1}{m!}\sum_{i=0}^{m-1}(m-i)^l(-1)^i\binom{m}{i}\\ &= \frac{1}{m!}S(l,m)\end{aligned}$$

$S(l,m)$ is known to be the number of surjective maps from a set with l elements to another set with m elements. Only when l is equal to m, surjective mapping is possible. The number of surjective maps is then equal to the number of ways of putting m distinct balls into m distinct boxes (no box being empty) $=m!\delta_{lm}$. Hence, $\frac{1}{m!}S(l,m)=\delta_{lm}$ and the identity for the special case is proved.

For the general case, we conjecture that the identity is valid for any random set of m positive integers. Although we cannot give a proof, we computationally checked it using Mathematica. The set of integers $I=\{0,i_1,i_2,,i_m\}$ were chosen randomly 1000 times. For every such set, the identity was verified. We varied m such that $1\leq m\leq 100$. In that way, the identity was checked $1000\times 100=10^5$ times, for the general case. A formal proof would be nice to find in future.

[1] C. Schwarz, A. Johnson, M. Koivomagi, E. Zatulovskiy, C. J. Kravitz, A. Doncic, and J. M. Skotheim, Mol Cell **69**, 253 (2018).
[2] M. Bentele, I. Lavrik, M. Ulrich, S. Stosser, D. W. Heermann, H. Kalthoff, P. H. Krammer, and R. Eils, J Cell Biol **166**, 839 (2004).
[3] J. E. Chipuk and D. R. Green, Trends Cell Biol **18**, 157 (2008).
[4] S. L. Spencer and P. K. Sorger, Cell **144**, 926 (2011).
[5] A. Arkin, J. Ross, and H. H. McAdams, Genetics **149**, 1633 (1998).
[6] M. Avlund, S. Krishna, S. Semsey, I. B. Dodd, and K. Sneppen, PLoS One **5**, e15037 (2010).
[7] M. Avlund, I. B. Dodd, S. Semsey, K. Sneppen, and S. Krishna, Journal of Virology **83**, 11416 (2009).
[8] M. Bednarz, J. A. Halliday, C. Herman, and I. Golding, PLoS One **9**, e100876 (2014).
[9] I. Golding, Annu Rev Virol **3**, 453 (2016).
[10] I. Golding, Curr Opin Microbiol **43**, 9 (2018).
[11] I. Herskowitz and D. Hagen, Annu Rev Genet **14**, 399 (1980).
[12] A. D. Johnson, B. J. Meyer, and M. Ptashne, Proc Natl Acad Sci U S A **76**, 5061 (1979).
[13] M. Ptashne, *A Genetic Swtch: Gene Control and Phage Lambda* (Blackwell Scientific, Palo Alto, USA, 1986).
[14] Q. Shao, J. T. Trinh, and L. Zeng, J Biol Chem **294**, 3343 (2019).
[15] L. Zeng, S. O. Skinner, C. Zong, J. Sippy, M. Feiss, and

I. Golding, Cell **141**, 682 (2010).
[16] I. N. Wang, Genetics **172**, 17 (2006).
[17] Y. Shao and I. N. Wang, Genetics **181**, 1467 (2009).
[18] A. Singh and J. J. Dennehy, J R Soc Interface **11**, 20140140 (2014).
[19] M. Krupovic and D. H. Bamford, Mol Microbiol **69**, 781 (2008).
[20] I. N. Wang, D. L. Smith, and R. Young, Annu Rev Microbiol **54**, 799 (2000).
[21] A. Singh and J. J. Dennehy, Journal of The Royal Society Interface **11**, 20140140 (2014).
[22] K. R. Ghusinga, J. J. Dennehy, and A. Singh, Proceedings of the National Academy of Sciences **114**, 693 (2017).
[23] A. J. Bray, S. N. Majumdar, and G. Schehr, Advances in Physics **62**, 225 (2013).
[24] S. Redner, *A guide to first-passage processes* (Cambridge University Press, Cambridge, UK ; New York, 2001) pp. ix, 312 p.
[25] A. Biswas, J. Cruz, P. Parmananda, and D. Das, Soft Matter (2020).
[26] J. J. Parmar, D. Das, and R. Padinhateeri, Nucleic acids research **44**, 1630 (2016).
[27] I. Nayak, D. Das, and A. Nandi, Physical Review Research **2**, 013114 (2020).
[28] B. Munsky, G. Li, Z. R. Fox, D. P. Shepherd, and G. Neuert, Proceedings of the National Academy of Sciences **115**, 75337538 (2018).
[29] V. Shahrezaei and P. S. Swain, Proceedings of the National Academy of Sciences **105**, 17256 (2008).
[30] P. Ranjith, P. S. Kumar, and G. I. Menon, Physical review letters **94**, 138102 (2005).
[31] C. Shi, Y. Jiang, and T. Zhou, bioRxiv (2020).
[32] N. Kumar, K. Zarringhalam, and R. V. Kulkarni, Biophysical journal **114**, 2530 (2018).
[33] S. H. Tindemans and B. M. Mulder, Physical Review E **81**, 031910 (2010).
[34] R. Padinhateeri, A. B. Kolomeisky, and D. Lacoste, Biophysical journal **102**, 1274 (2012).
[35] Z. Cao and R. Grima, Proceedings of the National Academy of Sciences **117**, 4682 (2020).
[36] P. J. Ingram, M. P. H. Stumpf, and J. Stark, PLOS Computational Biology **4**, e1000192 (2008).
[37] C. W. Gardiner *et al.*, *Handbook of stochastic methods*, Vol. 3 (springer Berlin, 1985).
[38] A. Gründling, M. D. Manson, and R. Young, Proceedings of the National Academy of Sciences **98**, 9348 (2001).
[39] J. S. Dewey, C. G. Savva, R. L. White, S. Vitha, A. Holzenburg, and R. Young, Proceedings of the National Academy of Sciences **107**, 2219 (2010).
[40] N. R. Draper and H. Smith, *Applied regression analysis*, Vol. 326 (John Wiley & Sons, 1998).
[41] S. Kannoly, T. Gao, S. Dey, N. Wang, A. Singh, and J. J. Dennehy, Iscience , 101186 (2020).
[42] G. B. Arfken and H. J. Weber, *Mathematical methods for physicists* (American Association of Physics Teachers, 1999).