

## Exact Distribution of Threshold Crossing Times for Protein Concentrations: Implication for Biological Timekeeping

Krishna Rijal<sup>1,\*</sup>, Ashok Prasad<sup>2</sup>, Abhyudai Singh,<sup>3</sup> and Dibyendu Das<sup>1,†</sup>

<sup>1</sup>Department of Physics, Indian Institute of Technology Bombay, Powai, Mumbai 400076, India

<sup>2</sup>Department of Chemical and Biological Engineering, Colorado State University, Fort Collins, Colorado 80523, USA

<sup>3</sup>Departments of Electrical and Computer Engineering, Biomedical Engineering and Mathematical Sciences, University of Delaware, Newark, Delaware 19716, USA



(Received 11 August 2021; accepted 3 January 2022; published 24 January 2022)

Stochastic protein accumulation up to some concentration threshold sets the timing of many cellular physiological processes. Here we obtain the exact distribution of first threshold crossing times of protein concentration, in either Laplace or time domain, and its associated cumulants: mean, variance, and skewness. The distribution is asymmetric, and its skewness nonmonotonically varies with the threshold. We study lysis times of *E. coli* cells for holin gene mutants of bacteriophage- $\lambda$  and find a good match with theory. Mutants requiring higher holin thresholds show more skewed lysis time distributions as predicted. The theory also predicts a linear relationship between infection delay time and host doubling time for lytic viruses, that has recently been experimentally observed.

DOI: 10.1103/PhysRevLett.128.048101

Molecular biological processes associated with changes in cell state are controlled by changes in gene expression, a complex stochastic process involving transcription of a gene into a ribonucleic acid (RNA) molecule and its subsequent translation into a protein. Intrinsic noise in transcription and translation leads to a stochastically varying abundance of messenger RNA (mRNA) and proteins within cells, even when genetically identical [1–3]. Experimental studies have directly probed fluctuations of protein levels across cells [4,5]. Theoretical models assuming specific promoter configurations have solved for the coefficient of variance  $CV^2$  (variance divided by square of mean) of the mRNA and protein numbers, as well as their full steady-state distributions [6–10]. Stochastic gene expression has been shown to be a fundamental property of living cells, affecting critical physiological processes of biological and biomedical importance [6,11].

The dynamics of downstream processes governed by the synthesis of a protein requires its accumulation to some minimum concentration threshold, e.g., transcription factors with sigmoidal Hill kinetics [12]. In such cases the timing of the downstream process is governed by the time at which the threshold concentration is reached for the first time, i.e., the First Passage Time (FPT) [13,14] of the stochastic gene expression process. The most well-studied example of timing control by stochastic protein accumulation is probably lysis of lambda phage-infected *E. coli*. Here the protein holin self-assembles on the bacterial membrane and punctures it after its concentration crosses a threshold, causing the cell to ultimately lyse or burst and release the newly formed viral particles [15,16]. The viral burst size of lambda phage-infected *E. coli* is known to

have a broad distribution [17]. While previous work has studied the timing of lysis and its relation with viral fitness [16,18,19] and the lysis-lysogeny decision [20], later studies have highlighted the distribution of lysis timing and demonstrated connections with the first passage time [21–24]. Fluctuations in lysis times lead to variations in viral burst sizes affecting both viral population fitness [25] as well as the health of the host. Genetic mutations of holin have been shown to regulate the stochasticity in lysis times in the  $\lambda$  variants [21,26,27].

First passage timing mechanisms based on protein accumulation may be quite common, from eukaryotic organisms to single cells. Recent work suggests that neuroblast migration timing in *C. elegans* development is controlled by the accumulation of *mig1* [28]. Cell division in *E. coli* has similarly been shown to be controlled by FtsZ expression [29]. Other examples include cell survival during prolonged drug exposure [30] and regulation of cell size in yeast cells by *cdr2* proteins [31]. First passage times are also relevant for other phenomena such as RNA polymerase backtracking and cleavage [32], first binding of proteins to sites on DNA [33], capture of kinetochores [34], and more abstractly, estimating characteristics of energy landscapes [35]. Statistics of first passage times have been of great theoretical interest, and obtaining analytical expressions for their distribution (FPTD) is generally quite challenging [36]. In previous work we derived analytical expressions for the FPTD of the absolute number of any molecule generated through geometrically distributed burst kinetics to reach a threshold [37]. However, in living cells the cell volume is never constant, and for many biological processes the more appropriate

variable is not the absolute number but the concentration of proteins. For example, the variation of noise in lysis times of  $\lambda$ -phage mutants could be explained only assuming a concentration threshold of holins [27], and the concentration threshold of *cdr2* proteins plays a role in the cell division of budding yeast [31].

Previously only approximate formulae existed for the moments of the concentration threshold crossing FPTD [23,27], and the distribution itself was unknown. Here we derive exact analytical expressions (as well as systematic approximations) for the FPTD of molecular concentrations and its moments, and apply them to experimental data on the distribution of lysis times. In the latter part of this text, we show that the threshold phenomenon can explain something out of the box: the experimentally seen linear relationship between mean lysis time and the doubling time of the host cell, after an infection.

A theoretical framework for the stochastic kinetics of protein synthesis has already been developed under the assumptions of short-lived mRNA and long-lived proteins. The exact steady-state distributions of the discrete protein number ( $n$ ) has been derived previously [8] and was shown to follow a negative binomial distribution, while the continuous protein concentration  $c = n/V$  (with  $V$  being cellular volume) was shown to follow the Gamma distribution [7,38]. While the *forward* continuous master equation was suitable to study the protein concentration  $c$  [7], the corresponding *backward* master equation [39] is more convenient for calculating the statistics of the FPT to reach the threshold concentration  $X$ . Given an initial  $c = x < X$ , the survival probability  $S(X, x, t)$  that  $c$  survives reaching the threshold  $X$  through time  $t$ , satisfies the backward continuous master equation:

$$\frac{\partial S(X, x, t)}{\partial t} = k \int_x^X dx' [\nu(x' - x) - \delta(x' - x)] S(X, x', t) - \gamma x \frac{\partial S(X, x, t)}{\partial x}. \quad (1)$$

Here the initial condition is  $S(X, x, 0) = 1$  and the boundary condition  $S(X, x = X, t) = 0$ ; the rate of protein production  $k\nu$  is assumed to be proportional to the mRNA production rate  $k$  and the experimentally known protein burst size distribution  $\nu(x - x') = (1/b)e^{-(x-x')/b}$  with mean burst concentration  $b$  [5,38]. The rate of decay of the protein per unit concentration is  $\gamma$  which expresses the joint effect of protein degradation and cell growth. The FPTD for the *first* threshold crossing ( $x \geq X$ ) in time  $t$  is obtained from  $S(X, x, t)$  as  $f(X, x, t) = -\partial S(X, x, t)/\partial t$ . To solve Eq. (1) we convert the integro-differential equation into a partial differential equation and take the Laplace transform,  $\tilde{S}(X, x, s) = \int_0^\infty dt e^{-st} S(X, x, t)$ , leading to a differential equation for  $\tilde{S}$  as a function of the (scaled) initial concentration  $\tilde{x} = x/b$  (details in Sec. S2 of the Supplemental Material [40]):

$$\tilde{x} \frac{\partial^2 \tilde{S}(X, \tilde{x}, s)}{\partial \tilde{x}^2} + \left( \frac{k + \gamma + s}{\gamma} - \tilde{x} \right) \frac{\partial \tilde{S}(X, \tilde{x}, s)}{\partial \tilde{x}} - \frac{s}{\gamma} \tilde{S}(X, \tilde{x}, s) = -\frac{1}{\gamma}. \quad (2)$$

The homogeneous part of the above equation is a confluent hypergeometric equation [43]. Using the boundary condition and the fact that  $\tilde{S}$  is finite as  $x \rightarrow 0$ , the solution in terms of the confluent hypergeometric function  ${}_1F_1(a, c, \tilde{x})$  [43] is as follows (see Sec. S3 of the Supplemental Material [40]):

$$\tilde{S}(X, x, s) = \frac{1}{s} \left( 1 - \frac{{}_1F_1\left[\frac{s}{\gamma}, 1 + \frac{k+s}{\gamma}, \frac{\tilde{x}}{b}\right]}{{}_1F_1\left[\frac{s}{\gamma}, 1 + \frac{k+s}{\gamma}, \frac{X}{b}\right]} \right). \quad (3)$$

Since  $\tilde{f}(X, x, s) = 1 - s\tilde{S}(X, x, s)$ , the desired exact FPTD in Laplace space for any  $\gamma$  and any initial protein concentration  $x > 0$  is

$$\tilde{f}(X, x, s)|_{\gamma \neq 0} = \frac{{}_1F_1\left[\frac{s}{\gamma}, 1 + \frac{k+s}{\gamma}, \frac{x}{b}\right]}{{}_1F_1\left[\frac{s}{\gamma}, 1 + \frac{k+s}{\gamma}, \frac{X}{b}\right]}. \quad (4)$$

The above calculation (for  $x > 0$ ) is applicable to the special case  $x \rightarrow 0$  of interest, as the initial protein level is zero at the beginning of the translation. The case of exactly  $x = 0$  requires a separate treatment, but is numerically identical to  $x \rightarrow 0$  as expected (details in Sec. S3 of the Supplemental Material [40]).

For vanishing decay constant ( $\gamma \rightarrow 0$ ), Eq. (4) simplifies to  $\tilde{f}(X, x, s) = \exp\{-(X-x)s/b(k+s)\}$  which is analytically invertible and gives the exact FPTD in the time domain (see Sec. S4 of the Supplemental Material [40]):

$$f(X, x, t)|_{\gamma=0} = e^{-\frac{(X-x)}{b}t} e^{-kt} \sum_{n=1}^{\infty} \frac{(X-x)^n}{n!(n-1)!} \left(\frac{k}{b}\right)^n t^{n-1}. \quad (5)$$

Note that for  $\gamma = 0$  the result is a function of the difference between the threshold and initial concentrations ( $X - x$ ), as expected from the translational symmetry in Eq. (1).

In general for  $\gamma \neq 0$ , Eq. (4) may be inverted numerically using *Mathematica* [44]. Finally we can also derive the exact expression of the first few moments, which are useful when comparing with empirical distributions. From Eq. (4), the  $n$ th moment  $\langle t^n \rangle = [\partial^n \tilde{f}(X, x, s)/\partial s^n]_{s \rightarrow 0}$ . Defining  $g(s) = {}_1F_1[(s/\gamma), 1 + \{(k+s)/\gamma\}, (x/b)]$ ,  $h(s) = {}_1F_1[(s/\gamma), 1 + \{(k+s)/\gamma\}, (X/b)]$ ,  $g^{(m)}(0) = [\partial^m g(s)/\partial s^m]_{s \rightarrow 0}$  and  $h^{(m)}(0) = [\partial^m h(s)/\partial s^m]_{s \rightarrow 0}$ , the first three moments are derived analytically exactly (see Sec. S5 of the Supplemental Material [40]):

$$\langle t \rangle = h^{(1)}(0) - g^{(1)}(0) \quad (6)$$

$$\langle t^2 \rangle = 2h^{(1)}(0)(h^{(1)}(0) - g^{(1)}(0)) + g^{(2)}(0) - h^{(2)}(0) \quad (7)$$

$$\begin{aligned} \langle t^3 \rangle = & 6(h^{(1)}(0))^3 + g^{(1)}(0)(3h^{(2)}(0) - 6(h^{(1)}(0))^2) \\ & + 3h^{(1)}(0)(g^{(2)}(0) - 2h^{(2)}(0)) + h^{(3)}(0) - g^{(3)}(0). \end{aligned} \quad (8)$$

Note that  $CV^2 = (\langle t^2 \rangle / \langle t \rangle^2) - 1$  and the skewness =  $[(\langle t^3 \rangle - 3\langle t^2 \rangle \langle t \rangle + 2\langle t \rangle^3) / (\langle t^2 \rangle - \langle t \rangle^2)^{3/2}]$  follow from the above expressions.

Just like the distribution in Eq. (4), the quantities  $g^{(1)}(0)$ ,  $g^{(2)}(0)$ ,  $g^{(3)}(0)$ ,  $h^{(1)}(0)$ ,  $h^{(2)}(0)$ , and  $h^{(3)}(0)$  (see Sec. S5 of the Supplemental Material [40]) depend on the four parameters  $(x/b)$ ,  $k$ ,  $\gamma$  and  $(X/b)$ .

*Statistics of Lysis Times.*—In order to match our results with experimental data, we used the raw data of Ref. [27]. Briefly, site-directed mutagenesis was used to generate a library of mutations in the S105 holin allele, each of which differed from the parent allele by one or two amino acid substitutions. These mutated sequences were then used to generate a library of lysogenic lambda phages, each carrying a slightly different holin gene. These viruses were used to infect *E. coli* cells, and lysis was thermally induced and measured at the single cell level for 91–174 cells per strain.

We estimate some required parameters as follows. Holins degrade slowly; hence the decay of  $x$  is mostly due to cell growth, with doubling time of roughly 40 min. Hence we choose  $\gamma = \ln(2)/40 \text{ min}^{-1}$ . We choose  $x/b = 0.01$  to represent  $x \rightarrow 0$ , the vanishingly small initial protein concentration. Next we numerically eliminate the parameter  $(X/b)$  between the expressions of  $CV^2$  and mean FPT  $\langle t \rangle$  (see discussion in Sec. S7 in the Supplemental Material [40]), such that the theoretical curve of  $CV^2$  versus  $\langle t \rangle$  gets fixed by just one fitting parameter, i.e.,  $k$ . The best fit of the theory to the experimental data for the 20 mutants is shown in Fig. 1, and yields the fitted value  $k = 4.5 \text{ min}^{-1}$ .

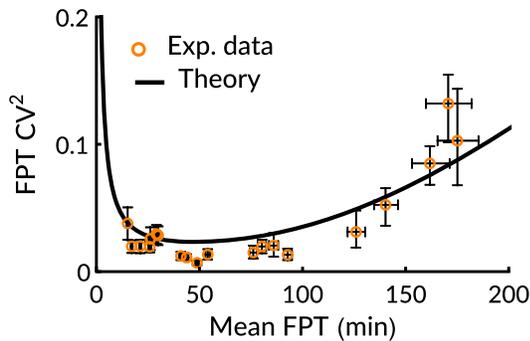


FIG. 1.  $CV^2$  against  $\langle t \rangle$  for 20 mutants (symbols) and exact theory (solid line) with best fit parameter  $k = 4.5 \text{ min}^{-1}$ . Error bars are 90% confidence intervals, obtained after bootstrapping 1000 replicates (see Sec. S6 of the Supplemental Material [40]). Here  $\gamma = \ln(2)/40 \text{ min}^{-1}$  and  $(x/b) = 0.01$ .

The  $CV^2$  curve has a minimum at mean FPT around  $t_m \sim 40 \text{ min}$ .

While we expect the mutants to have roughly the same  $(x/b)$ ,  $k$ , and  $\gamma$  values (as given above), their mean FPT differs as does the threshold  $(X/b)$ . We fix  $(X/b)$  for every mutant by matching the theoretical mean from Eq. (6) with the experimental average from the data, for that mutant. Then we obtain the full theoretical FPTD [by inverting Eq. (4) [44]] and plot against the experimental distribution to check how well they match. This is shown in Fig. 2 for two cases—mutant-1 (JD405) and mutant-2 (JD426) (see Sec. S1 of the Supplemental Material [40] and [27]) with the mean FPT smaller and larger than  $t_m$  respectively. See plots for the remaining 18 mutants in the Supplemental Material [40] (Fig. 3).

The data and theoretical curves (in Fig. 2) both suggest that FPTD is non-Gaussian and is skewed to the right. We explicitly study the variation of the skewness of FPTD for the mutants in Fig. 3—it shows a nonmonotonic behavior just like the  $CV^2$  with a minimum around  $\sim 35 \text{ min}$  close to  $t_m$  mentioned above. Thus mutants with increasingly larger mean FPT have increasingly asymmetric FPTD. Note that the distributions have asymptotic (large  $t$ ) exponential tails  $\sim \exp(-t/\tau_c)$ , with characteristic times  $\tau_c$  being related to the smallest pole  $s_* = -1/\tau_c$  of  $\tilde{f}(X, x, s)$  in Eq. (4). A plot of  $\tau_c / \langle t \rangle$  against the mean FPT shows a similar non-monotonic curve as  $CV^2$  and skewness (see Sec. S8 in the Supplemental Material [40]).

Our results lead to an interesting prediction regarding the dependence of the mean lysis time on the cell doubling time. We showed above that fluctuations are minimal around a mean time  $t_m$ , for a given cellular size doubling time  $\ln(2)/\gamma$ . This value  $t_m$  for any bacterial cell however may vary, and depends upon experimental conditions which may change the cell doubling time. From the theory, we numerically calculate the  $t_m$  at minimum  $CV^2$  and find it to be linearly dependent on the doubling time  $\ln(2)/\gamma$  (see Fig. 4). We also show analytically that  $t_m \approx -(1/\gamma) \ln(1-f)$  where the fraction  $f$  is a ratio between an “optimal threshold”  $X_{\text{opt}}$  and the steady-state concentration  $c_{\text{ss}} = kb/\gamma$  (see Sec. S9 in the Supplemental

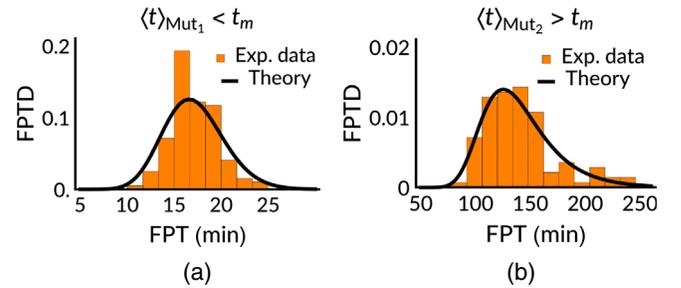


FIG. 2. The FPTD from experiment and theory for (a) mutant-1 with  $\langle t \rangle = 17.1 \text{ min}$  and (b) mutant-2 with  $\langle t \rangle = 140.3 \text{ min}$ . Note  $t_m \sim 40 \text{ min}$ .

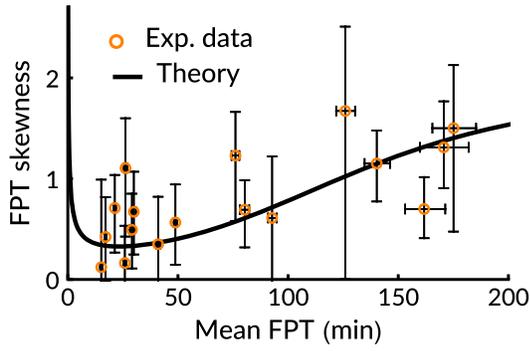


FIG. 3. Skewness against  $\langle t \rangle$  for 20 mutants (symbols) and exact theory (solid line) with  $k = 4.5 \text{ min}^{-1}$ ,  $x/b = 0.01$ , and  $\gamma = \ln(2)/40 \text{ min}^{-1}$ . Error bars are determined in a similar way as in Fig. 1.

Material [40]). Thus, under the plausible assumption that wild-type viruses have optimal noise characteristics, the mean lysis time of lambda phage under different host cell doubling times should be linearly related with cell doubling time. Interestingly, for a wide variety of viruses, a linear correlation between infection delay time and host doubling time has been reported recently [45].

*Approximate FPTD.*—Systematic approximations are useful since often they yield simpler expressions of practical use. Previous approximations for moments of the FPTD [23,27,46] were based on *ad hoc* assumptions. We develop a systematic approximate theory which matches the exact results up to second order in fluctuations as follows. Under the assumption that the burst sizes are small, the bursty term in Eq. (1) may be smoothed through the Kramers-Moyal expansion [39] (see Sec. S10 in the Supplemental Material [40]). Retaining terms up to second order, we obtain the following backward Fokker-Planck equation, valid for  $x < kb/\gamma$ :

$$\frac{\partial S(X, x, t)}{\partial t} = (kb - \gamma x) \frac{\partial S(X, x, t)}{\partial x} + kb^2 \frac{\partial^2 S(X, x, t)}{\partial x^2}. \quad (9)$$

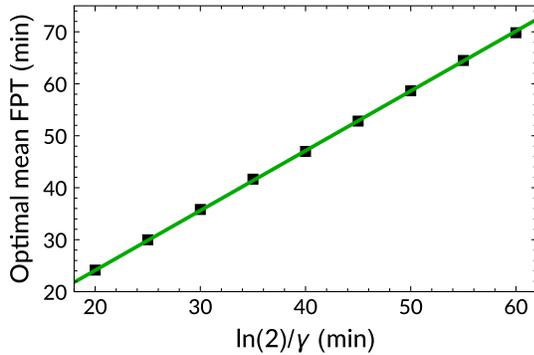


FIG. 4. Optimal mean FPT versus cell doubling time. As discussed in the text, the slope of the green line is equal to  $-\ln(1-f)/\ln(2)$  with  $f = 0.55$ .

The forward Fokker-Planck equation (Sec. S10 in the Supplemental Material [40]) which is the counterpart of the above backward Eq. (9), is the corresponding Kramer-Moyal approximation to the exact theory developed in Ref. [7]. That forward equation, under Ito convention, is related to the following Langevin equation [39] (see Sec. S10 in the Supplemental Material [40] which includes Ref. [41]):

$$\frac{dc}{dt'} = kb - \gamma c + \eta(t') \quad (10)$$

where  $t'$  denotes the forward evolving time, and is to be distinguished from the “backward” time  $t$ . Note  $\eta(t')$  is a Gaussian noise with  $\langle \eta(t') \rangle = 0$  and  $\langle \eta(t'_1) \eta(t'_2) \rangle = 2D\delta(t'_1 - t'_2)$ , and the diffusion constant needs to be identified as  $D = kb^2$ . The simple Langevin equation with production, decay, and noise terms has appeared in earlier theories, e.g., of FPT for mRNA kinetics [47]—yet its exact FPTD was not known. We note that Eq. (9) corresponding to the Langevin Eq. (10), can be exactly solved to obtain the FPTD in Laplace space (see Sec. S11 in the Supplemental Material [40]):

$$\tilde{f}(X, x, s) \Big|_{\gamma \neq 0}^{\text{approx}} = \frac{U\left[\frac{s}{2\gamma}, \frac{1}{2}, \frac{(bk - \gamma x)^2}{2b^2 \gamma k}\right]}{U\left[\frac{s}{2\gamma}, \frac{1}{2}, \frac{(bk - \gamma X)^2}{2b^2 \gamma k}\right]} \quad (11)$$

Here,  $U$  denotes Tricomi’s confluent hypergeometric function related to  ${}_1F_1$  (see Sec. S11 in the Supplemental Material [40]). From Eq. (11) the theoretical  $\text{CV}^2$  and skewness may be obtained just as we did for the exact theory [Eq. (4)]—see Sec. S12 of the Supplemental Material [40] which includes Ref. [42].

The curves of  $\text{CV}^2$  and skewness from the approximate FPTD [Eq. (11)] are plotted in Fig. 5 along with those from the exact theory [using Eqs. (7), (8)], and the ones from the Langevin simulations obtained using Eq. (10). The exact theory matches the approximate theory and simulations for  $\text{CV}^2$  perfectly, but deviates a bit from those in the case of skewness, which is expected since in the Kramers-Moyal approximation we ignored third order moments which contribute to the skewness.

We derived analytically exact results for the FPTD for protein concentrations by solving the backward master equation. We showed that the FPTD and its moments match those observed for lysis times in lambda phage, reinforcing the hypothesis that the lysis time is governed by a protein accumulation-dependent FPTD. We developed a systematic approximation for the FPTD and derived analytical results for the approximate FPTD. The results are general and can be applied to fit lysis time distributions for all lytic phages including lambda phage, as well as other protein threshold crossing processes. The distributions themselves are non-Gaussian with exponential tails and may have a high

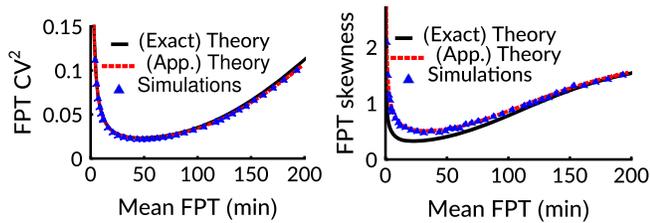


FIG. 5. The (a)  $CV^2$  and (b) skewness, obtained from the exact continuous theory, the approximate theory, and the Langevin simulations are compared. The parameters are same as in Figs. (1) and (3).

skewness, indicating asymmetry that may be significant for some lytic phages. Our predictions may be checked using Gillespie simulations of protein numbers divided by time varying cellular volume [7] or using more exact time-dependent rates [48,49].

We showed that the FPT theory predicts that some mutants will have optimal noise characteristics, i.e., minimal CV and skewness, which is borne out by the data. The nonmonotonic relation between lysis time and noise suggests that lysis time may be selected during evolution to minimize noise, as suggested by the observation that the wild-type lambda phage has the lowest level of noise in lysis times [24].

We also predicted an interesting linear relationship between the optimal mean lysis time and the host cell doubling time. This result may be quite general for lytic viruses, since viral escape from an infected cell is typically characterized by a delay that could be a sign of a threshold phenomena. It is thus intriguing that for a wide variety of viruses, the initial burst timing is linearly correlated with the cell doubling time [45] across many different types of hosts and lysis times ranging from minutes to a week. Lytic phages are also important in phage therapy, where bacteriophages are used to kill pathogenic bacteria [50], and these results on lysis timing may have useful applications for that novel area of medicine.

D. D. acknowledges SERB India (Grant No. MTR/2019/000341) for financial support. K. R. thanks IIT Bombay for Institute Ph.D. fellowship.

\*krishnarijal@iitb.ac.in

†dibyendu@phy.iitb.ac.in

- [1] M. Thattai and A. Van Oudenaarden, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 8614 (2001).
- [2] M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain, *Science* **297**, 1183 (2002).
- [3] J. M. Raser and E. K. O'shea, *Science* **309**, 2010 (2005).
- [4] J. M. Raser and E. K. O'Shea, *Science* **304**, 1811 (2004).
- [5] J. Yu, J. Xiao, X. Ren, K. Lao, and X. S. Xie, *Science* **311**, 1600 (2006).
- [6] A. Sanchez, S. Choubey, and J. Kondev, *Annu. Rev. Biophys.* **42**, 469 (2013).

- [7] N. Friedman, L. Cai, and X. S. Xie, *Phys. Rev. Lett.* **97**, 168302 (2006).
- [8] V. Shahrezaei and P. S. Swain, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 17256 (2008).
- [9] A. Singh, C. A. Vargas, and R. Karmakar, in *Proceedings of the 52nd IEEE Conference on Decision and Control* (IEEE, New York, 2013), pp. 7217–7222.
- [10] N. Kumar, A. Singh, and R. V. Kulkarni, *PLoS Comput. Biol.* **11**, e1004292 (2015).
- [11] D. Fraser and M. Kaern, *Mol. Microbiol.* **71**, 1333 (2009).
- [12] U. Alon, *An Introduction to Systems Biology* (CRC Press, Bristol, 2019).
- [13] S. Redner, *A Guide to First-Passage Processes* (Cambridge University Press, Cambridge, England, 2001).
- [14] G. Oshanin, R. Metzler, and S. Redner, *First-Passage Phenomena and Their Applications* (World Scientific Publishing Company, Singapore, 2014).
- [15] U. Bläsi, P. Fraisl, C.-Y. Chang, N. Zhang, and R. Young, *J. Bacteriol.* **181**, 2922 (1999).
- [16] R. White, S. Chiba, T. Pang, J. S. Dewey, C. G. Savva, A. Holzenburg, K. Pogliano, and R. Young, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 798 (2011).
- [17] M. Delbrück, *J. Bacteriol.* **50**, 131 (1945).
- [18] I. N. Wang, D. L. Smith, and R. Young, *Annu. Rev. Microbiol.* **54**, 799 (2000).
- [19] G. L. Ryan and A. D. Rutenberg, *J. Bacteriol.* **189**, 4749 (2007).
- [20] M. G. Cortes, J. T. Trinh, L. Zeng, and G. Balázsi, *Biophys. J.* **113**, 2110 (2017).
- [21] J. J. Dennehy and N. Wang, *BMC Microbiol.* **11**, 174 (2011).
- [22] A. Singh and J. J. Dennehy, *J. R. Soc. Interface* **11**, 20140140 (2014).
- [23] A. D. Co, M. C. Lagomarsino, M. Caselle, and M. Osella, *Nucleic Acids Res.* **45**, 1069 (2017).
- [24] K. R. Ghusinga, J. J. Dennehy, and A. Singh, *Proc. Natl. Acad. Sci. U.S.A.* **114**, 693 (2017).
- [25] I. N. Wang, *Genetics* **172**, 17 (2006).
- [26] Y. Shao and I. N. Wang, *Genetics* **181**, 1467 (2009).
- [27] S. Kannoly, T. Gao, S. Dey, N. Wang, A. Singh, and J. J. Dennehy, *iScience* **23**, 101186 (2020).
- [28] S. Gupta, J. Varennes, H. C. Korswagen, and A. Mugler, *PLoS Comput. Biol.* **14**, e1006201 (2018).
- [29] K. Sekar, R. Rusconi, J. T. Sauls, T. Fuhrer, E. Noor, J. Nguyen, V. I. Fernandez, M. F. Buffing, M. Berney, S. Jun *et al.*, *Mol. Syst. Biol.* **14**, e8623 (2018).
- [30] D. A. Charlebois, N. Abdennur, and M. Kaern, *Phys. Rev. Lett.* **107**, 218101 (2011).
- [31] K. Z. Pan, T. E. Saunders, I. Flor-Parra, M. Howard, and F. Chang, *eLife* **3**, e02040 (2014).
- [32] É. Roldán, A. Lisica, D. Sánchez-Taltavull, and S. W. Grill, *Phys. Rev. E* **93**, 062411 (2016).
- [33] J. J. Parmar, D. Das, and R. Padinhateeri, *Nucleic Acids Res.* **44**, 1630 (2016).
- [34] I. Nayak, D. Das, and A. Nandi, *Phys. Rev. Research* **2**, 013114 (2020).
- [35] A. L. Thorneywork, J. Gladrow, Y. Qing, M. Rico-Pasto, F. Ritort, H. Bayley, A. B. Kolomeisky, and U. F. Keyser, *Sci. Adv.* **6**, eaaz4642 (2020).

- [36] A. J. Bray, S. N. Majumdar, and G. Schehr, *Adv. Phys.* **62**, 225 (2013).
- [37] K. Rijal, A. Prasad, and D. Das, *Phys. Rev. E* **102**, 052413 (2020).
- [38] L. Cai, N. Friedman, and X. S. Xie, *Nature (London)* **440**, 358 (2006).
- [39] C. W. Gardiner, *Handbook of Stochastic Methods* (Springer, Berlin, Berlin, 1985), Vol. 3.
- [40] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevLett.128.048101> for detailed calculations and experimental details, which includes Refs. [41] and [42].
- [41] N. G. Van Kampen, *J. Stat. Phys.* **24**, 175 (1981).
- [42] M. Abramowitz, I. A. Stegun, and R. H. Romer, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (Wolfram Library Archive, 1988).
- [43] G. F. Simmons, *Differential Equations with Applications and Historical Notes* (CRC Press, Bristol, 2016).
- [44] FixedTalbotNumericalLaplaceInversion.m, *Mathematica* online, version 12.2 (2003).
- [45] T. Jin and J. Yin, *Integr. Biol.* **13**, 44 (2021).
- [46] K. Biswas and A. Ghosh, *Eur. Phys. J. E* **44**, 16 (2021).
- [47] M. Shreshtha, A. Surendran, and A. Ghosh, *Phys. Biol.* **13**, 036004 (2016).
- [48] T. Lu, D. Volfson, L. Tsimring, and J. Hasty, *Syst. Biol.* **1**, 121 (2004).
- [49] T. Carletti and A. Filisetti, *Comput. Math. Methods Med.* **2012**, 423627 (2012).
- [50] S. T. Abedon, S. J. Kuhl, B. G. Blasdel, and E. M. Kutter, *Bacteriophage* **1**, 66 (2011).

# Supplemental Material for “Exact distribution of threshold-crossing times of cellular protein concentration: comparison with experimental lysis time distributions of phage— $\lambda$ mutants ”

Krishna Rijal<sup>1</sup>, Ashok Prasad<sup>2</sup>, Abhyudai Singh<sup>3</sup>, Dibyendu Das<sup>1\*</sup>

1. Department of Physics, Indian Institute of Technology Bombay, Powai, Mumbai 400076, India

2. Department of Chemical and Biological Engineering, Colorado State University, Fort Collins, Colorado, USA

3. Departments of Electrical and Computer Engineering, Biomedical Engineering and Mathematical Sciences, University of Delaware, Newark, DE 19716, USA

\* Corresponding author, email: dibyendu@phy.iitb.ac.in

December 21, 2021

## Notations used for parameters in the main text:

Notation	Description
$k$	The rate of transcription of a gene into mRNA
$b$	Mean size of protein burst concentration
$\gamma$	The rate of decay of protein concentration
$X$	Protein concentration threshold

## S1: Experimental details:

Full details of the experimental methods can be found in Ref. [1]. Here we include a brief description in non-specialist language for completeness. Firstly, plasmids were constructed with a mutated version of the holin gene using site directed mutagenesis, by standard molecular biology techniques. These plasmids were then used to transform E. coli lysogens, i.e. cells already harboring a lambda-phage (MC4100 ( $\lambda$  cI857 S::Cam) strain) genome. Following induction of the lytic cycle in these lysogens, the mutated holin gene carried by the plasmid is incorporated (via recombination) into the genomes of some of the phages that emerge after cell lysis . The mutated phages were then used to lysogenize wild-type E. coli cells (MC4100 strain), resulting in a collection of lysogenic strains harboring phage genomes with mutations in the holin gene. The phage genome in these lysogens (prophage) expresses a temperature-sensitive cI857 repressor,

which is required to maintain the lysogenic state. The lysogen enters the lytic cycle when the cI857 repressor is inactivated following a heat shock. Lysis time was measured as the time taken for a cell to lyse under a microscope starting from the time of application of a heat shock of 42°C. Measurements were carried out with the help of a microscope-mounted, temperature controlled perfusion chamber. Please refer to Ref. [1] for a complete list of the mutant viruses and details of the mutations introduced.

## S2: The continuous Backward Master equation for Survival probability:

Let  $S(X, x, t)$  denote the Survival probability of the gene expression process such that the protein concentration stays below threshold concentration  $X$  at time  $t$ , starting with concentration  $x$  at  $t = 0$ . Since the terminal state (on first passage) of such a process is known (i.e.  $X$ ), the equation governing  $S$  is developed backward in time  $t$  and initial position  $x$ , and the general procedure to construct it is well-known [2]. In this specific problem, between a “past time  $t + \delta t$ ” to a “future time  $t$ ” the concentration could have jumped from  $x$  to  $x'$  with probability  $k\nu(x' - x)\delta t$  due to bursty protein production, where  $\nu(x' - x) = (1/b)e^{-(x' - x)/b}$  is the probability distribution of the burst size  $x' - x$ . Similarly  $x$  could have decayed to  $x - \delta x$  with probability  $x\gamma\delta t$ . There is a remaining likelihood that the concentration does not change and stays  $x$ . Putting these information together one obtains the stochastic Master equation. Note that while the protein production is a jump process, the decay is a smooth process [2], and we are dealing with a continuous variable  $x$ . Consequently the continuous integro-differential backward Master equation (in the limit  $\delta t, \delta x = 0$ ) for the Survival probability is:

$$\frac{\partial S(X, x, t)}{\partial t} = k \int_x^X dx' [\nu(x' - x) - \delta(x' - x)] S(X, x', t) - \gamma x \frac{\partial S(X, x, t)}{\partial x} \quad (1)$$

Note that this is the exact counterpart of the forward Master equation of the process developed in [3] for the probability density  $p(c, t')$  of the protein concentration  $c(t')$  at time  $t'$ :

$$\frac{\partial p(c, t')}{\partial t'} = k \int_0^c dc' [\nu(c - c') - \delta(c - c')] p(c', t') + \gamma \frac{\partial [cp(c, t')]}{\partial c} \quad (2)$$

The equation (1) is to be solved for the initial condition  $S(X, x, 0) = 1$  and the absorbing boundary condition  $S(X, x = X, t) = 0$ . We first convert the integro-differential Eq. (1) to a pure differential equation by doing the following mathematical manipulations. Substituting  $\nu$  we have:

$$\frac{\partial S(X, x, t)}{\partial t} = k \int_x^X dx' \frac{1}{b} e^{-\frac{x' - x}{b}} S(X, x', t) - kS(X, x, t) - \gamma x \frac{\partial S(X, x, t)}{\partial x}$$

which on integration by parts gives

$$\frac{\partial S(X, x, t)}{\partial t} = -kS(X, x', t) e^{-\frac{x' - x}{b}} \Big|_{x' = x}^{x' = X} + k \int_x^X dx' \frac{\partial S(X, x', t)}{\partial x'} e^{-\frac{x' - x}{b}} - kS(X, x, t) - \gamma x \frac{\partial S(X, x, t)}{\partial x}.$$

Using the boundary condition,  $S(X, X, t) = 0$ , in the first term,

$$\begin{aligned}\frac{\partial S(X, x, t)}{\partial t} &= kS(X, x, t) + ke^{\frac{x}{b}} \int_x^X dx' \frac{\partial S(X, x', t)}{\partial x'} e^{-\frac{x'}{b}} - kS(X, x, t) - \gamma x \frac{\partial S(X, x, t)}{\partial x} \\ &= ke^{\frac{x}{b}} \int_x^X dx' \frac{\partial S(X, x', t)}{\partial x'} e^{-\frac{x'}{b}} - \gamma x \frac{\partial S(X, x, t)}{\partial x}\end{aligned}$$

Multiplying by  $e^{-x/b}$ , on both sides,

$$e^{-x/b} \frac{\partial S(X, x, t)}{\partial t} = k \int_x^X dx' \frac{\partial S(X, x', t)}{\partial x'} e^{-\frac{x'}{b}} - e^{-x/b} \gamma x \frac{\partial S(X, x, t)}{\partial x}$$

Then we take derivative with respect to  $x$ , on both sides, and get

$$\begin{aligned}-\frac{1}{b} e^{-x/b} \frac{\partial S(X, x, t)}{\partial t} + e^{-x/b} \frac{\partial^2 S(X, x, t)}{\partial x \partial t} &= -k e^{-x/b} \frac{\partial S(X, x, t)}{\partial x} - \gamma \left[ e^{-x/b} \frac{\partial S(X, x, t)}{\partial x} \right. \\ &\quad \left. + x \left( -\frac{1}{b} e^{-x/b} \frac{\partial S(X, x, t)}{\partial x} + e^{-x/b} \frac{\partial^2 S(X, x, t)}{\partial x^2} \right) \right]\end{aligned}$$

which simplifies to give

$$-\frac{1}{b} \frac{\partial S(X, x, t)}{\partial t} + \frac{\partial^2 S(X, x, t)}{\partial x \partial t} = \left( -k - \gamma + \frac{\gamma x}{b} \right) \frac{\partial S(X, x, t)}{\partial x} - \gamma x \frac{\partial^2 S(X, x, t)}{\partial x^2} \quad (3)$$

We define the Laplace transform  $\tilde{S}(X, x, s) = \int_0^\infty dt e^{-st} S(X, x, t)$ , and take Laplace transform of the above equation in time  $t$  to get,

$$-\frac{1}{b} (s\tilde{S}(X, x, s) - 1) + s \frac{\partial \tilde{S}(X, x, s)}{\partial x} = \left( -k - \gamma + \frac{\gamma x}{b} \right) \frac{\partial \tilde{S}(X, x, s)}{\partial x} - \gamma x \frac{\partial^2 \tilde{S}(X, x, s)}{\partial x^2}$$

which simplifies to

$$x \frac{\partial^2 \tilde{S}(X, x, s)}{\partial x^2} + \left( \frac{k + \gamma + s}{\gamma} - \frac{x}{b} \right) \frac{\partial \tilde{S}(X, x, s)}{\partial x} - \frac{s}{b\gamma} \tilde{S}(X, x, s) = -\frac{1}{b\gamma}$$

On doing the scale transformation  $x = \tilde{x}b$ , we get an ordinary differential equation for  $\tilde{S}$ :

$$\tilde{x} \frac{\partial^2 \tilde{S}(X, \tilde{x}, s)}{\partial \tilde{x}^2} + \left( \frac{k + \gamma + s}{\gamma} - \tilde{x} \right) \frac{\partial \tilde{S}(X, \tilde{x}, s)}{\partial \tilde{x}} - \frac{s}{\gamma} \tilde{S}(X, \tilde{x}, s) = -\frac{1}{\gamma} \quad (4)$$

### S3: The exact first passage time distribution (FPTD) in the Laplace domain:

The main object of our interest is the first passage time distribution  $f(X, x, t)$  to cross the threshold  $X$  at time  $t$ , particularly starting with zero protein concentration (i.e.  $x = 0$ ). This is related to the Survival probability as  $f(X, x, t) = -\frac{\partial}{\partial t} S(X, x, t)$ . In Laplace space this relationship is  $\tilde{f}(X, x, s) = 1 - s\tilde{S}(X, x, s)$ . Thus we need to solve for  $\tilde{S}$  from Eq. (4) to proceed further.

The homogeneous part of Eq. (4) is a Confluent Hypergeometric equation, and a particular solution of the inhomogeneous equation is  $1/s$  (4). The general solution is thus:

$$\tilde{S}(X, x, s) = \frac{1}{s} + A {}_1F_1\left[\frac{s}{\gamma}, 1 + \frac{k+s}{\gamma}, \frac{x}{b}\right] + B \left(\frac{x}{b}\right)^{-\frac{k+s}{\gamma}} {}_1F_1\left[-\frac{k}{\gamma}, 1 - \frac{k+s}{\gamma}, \frac{x}{b}\right] \quad (5)$$

where  $A$  and  $B$  are constants to be determined and  ${}_1F_1$  denotes the Confluent Hypergeometric function of the first kind,  ${}_1F_1[a, b, z] = \sum_{n=0}^{\infty} \frac{a_n z^n}{b_n n!}$ , with  $a_n$  and  $b_n$  being the Pochhammer symbols:  $a_n = a(a+1)(a+2)\dots(a+n-1)$  and  $b_n = b(b+1)(b+2)\dots(b+n-1)$ .

The solution is supposed to be valid for  $x = 0$ . In that limit, the third term in Eq. (5) diverges, because of the factor  $x^{-\frac{k+s}{\gamma}}$ . So, we set  $B = 0$ .

$$\tilde{S}(X, x, s) = \frac{1}{s} + A {}_1F_1\left[\frac{s}{\gamma}, 1 + \frac{k+s}{\gamma}, \frac{x}{b}\right]$$

Furthermore using the boundary condition,  $\tilde{S}(X, x = X, s) = 0$ , we get the value of  $A$  to be

$$A = -\frac{1}{s {}_1F_1\left[\frac{s}{\gamma}, 1 + \frac{k+s}{\gamma}, \frac{X}{b}\right]}$$

Thus survival probability in Laplace space is

$$\tilde{S}(X, x, s) \Big|_{x \neq 0} = \frac{1}{s} \left( 1 - \frac{{}_1F_1\left[\frac{s}{\gamma}, 1 + \frac{k+s}{\gamma}, \frac{x}{b}\right]}{{}_1F_1\left[\frac{s}{\gamma}, 1 + \frac{k+s}{\gamma}, \frac{X}{b}\right]} \right) \quad (6)$$

and using the relation  $\tilde{f}(X, x, s) = 1 - s\tilde{S}(X, x, s)$ , we finally get the central result of this work, namely the FPTD in Laplace space:

$$\tilde{f}(X, x, s) \Big|_{x \neq 0} = \frac{{}_1F_1\left[\frac{s}{\gamma}, 1 + \frac{k+s}{\gamma}, \frac{x}{b}\right]}{{}_1F_1\left[\frac{s}{\gamma}, 1 + \frac{k+s}{\gamma}, \frac{X}{b}\right]} \quad (7)$$

Expressions (6) and (7) are valid for any initial protein concentration  $x > 0$ . The case of  $x = 0$  requires a separate treatment.

### Solution for the case of $x = 0$ :

A separate solution for this case is necessary due to the following reason. Consider the case of  $x = 0$  and  $X$  tending to zero. While  $S(X = 0^+, 0, t)$  is something finite,  $S(X = 0^+, x = X, t) = 0$  (boundary condition). Thus the function  $S$  abruptly changes like a step function and hence the derivative of  $S(X, x, t)$  with respect to  $x$  will be a Dirac delta function. That would mean that various quantities in the differential Eq. (3) will become singular objects. So, we cannot proceed through this differential equation method for the case  $x = 0$ . But instead we may now go back to the integral Eq. (1) and study this case separately. In particular we know that, if the boundary  $X$  goes very close to 0, we expect the survival probability  $S(X = 0, 0, t) = e^{-kt}$  and the FPTD to be  $ke^{-kt}$ . That is because a single jump (with rate  $k$ ), will lead to crossing

of the infinitesimal threshold. This means that in such a limit the Laplace transforms of the Survival probability and the FPTD are  $1/(k+s)$  and  $k/(k+s)$  respectively. We would see below that the solution of the integro-differential equation for  $x=0$  is perfectly consistent with this expectation.  $\tilde{f}(X, x=0, s)$  Putting  $x=0$  in Eq. (1) and then taking Laplace transform corresponding to  $t$ , we get:

$$(s+k)\tilde{S}(X, 0, s) = 1 + \frac{k}{b} \int_0^X dx' e^{-\frac{x'}{b}} \tilde{S}(X, x', s)$$

Note that  $\tilde{S}(X, x' > 0, s)$  is given by the solution Eq. (6), and hence

$$\begin{aligned} (s+k)\tilde{S}(X, 0, s) &= 1 + \frac{k}{b} \int_0^X dx' e^{-\frac{x'}{b}} \left[ \frac{1}{s} \left( 1 - \frac{{}_1F_1\left[\frac{s}{\gamma}, 1 + \frac{k+s}{\gamma}, \frac{x'}{b}\right]}{{}_1F_1\left[\frac{s}{\gamma}, 1 + \frac{k+s}{\gamma}, \frac{X}{b}\right]} \right) \right] \\ &= 1 + \frac{k}{s} (1 - e^{-\frac{X}{b}}) - \frac{k}{{}_1F_1\left[\frac{s}{\gamma}, 1 + \frac{k+s}{\gamma}, \frac{X}{b}\right] b s} \sum_{n=0}^{\infty} \frac{\left(\frac{s}{\gamma}\right)_n}{\left(1 + \frac{s+k}{\gamma}\right)_n} \frac{1}{n!} \int_0^X dx' \frac{e^{-\frac{x'}{b}} x^n}{b^n} \\ &= 1 + \frac{k}{s} (1 - e^{-\frac{X}{b}}) - \frac{k}{{}_1F_1\left[\frac{s}{\gamma}, 1 + \frac{k+s}{\gamma}, \frac{X}{b}\right] b s} \sum_{n=0}^{\infty} \frac{\left(\frac{s}{\gamma}\right)_n}{\left(1 + \frac{s+k}{\gamma}\right)_n} \frac{\Gamma(n+1, \frac{X}{b})}{n!} \end{aligned}$$

Note the use of the Pochhammer symbols. Thus Survival probability

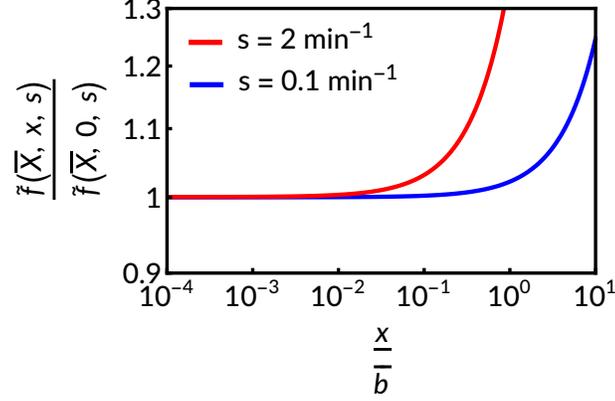
$$\tilde{S}(X, 0, s) = \frac{1}{s} \left( 1 - \frac{k}{k+s} e^{-\frac{X}{b}} - \frac{k}{{}_1F_1\left[\frac{s}{\gamma}, 1 + \frac{k+s}{\gamma}, \frac{X}{b}\right] (s+k)} \sum_{n=0}^{\infty} \frac{\left(\frac{s}{\gamma}\right)_n}{\left(1 + \frac{s+k}{\gamma}\right)_n} \frac{\Gamma(n+1, \frac{X}{b})}{n!} \right)$$

and using the relation  $\tilde{f}(X, 0, s) = 1 - s\tilde{S}(X, 0, s)$ , we get,

$$\boxed{\tilde{f}(X, 0, s) = \frac{k}{k+s} \left( e^{-\frac{X}{b}} + \frac{1}{{}_1F_1\left[\frac{s}{\gamma}, 1 + \frac{k+s}{\gamma}, \frac{X}{b}\right]} \sum_{n=0}^{\infty} \frac{\left(\frac{s}{\gamma}\right)_n}{\left(1 + \frac{s+k}{\gamma}\right)_n} \frac{\Gamma(n+1, \frac{X}{b})}{n!} \right)} \quad (8)$$

We note that indeed  $\tilde{f}(X=0^+, 0, s) = k/(k+s)$  as expected and discussed above.

Although the functional forms of  $\tilde{f}(X, x > 0, s)$  and  $\tilde{f}(X, x=0, s)$  are different (in Eqs. (7) and (8) respectively), as  $x \rightarrow 0$ ,  $\tilde{f}(X, x \rightarrow 0, s)$  smoothly coincides with  $\tilde{f}(X, x=0, s)$ . We show this numerically, by plotting the ratio  $\frac{\tilde{f}(X, x > 0, s)}{\tilde{f}(X, x=0, s)}$ , as a function of  $x$  (see Fig. (1) below). Thus for all practical purposes, the function  $\tilde{f}(X, x \rightarrow 0, s)$  represents well the limit of vanishing protein concentration. Moreover since the function  $\tilde{f}(X, x > 0, s)$  in Eqs. (7) is simpler looking compared to  $\tilde{f}(X, x=0, s)$ , we have subsequently used the expression of  $\tilde{f}(X, x > 0, s)$  everywhere in the manuscript while discussing vanishingly small initial protein concentration.



**Figure 1.** For parameter values  $k = 4.5 \text{ min}^{-1}$ ,  $\gamma = \ln(2)/40 \text{ min}^{-1}$  and  $X/b = 10$ , the ratio  $\frac{\tilde{f}(X, x, s)}{\tilde{f}(X, 0, s)}$  is plotted as a function of  $x/b$ , for two different values of the Laplace variable:  $s = 0.1 \text{ min}^{-1}$  and  $s = 2 \text{ min}^{-1}$ . In both the cases the ratio approaches 1 smoothly as  $x = 0$ .

#### S4: The exact FPTD in the time domain for vanishing degradation and very large cell-doubling times ( $\gamma = 0$ ):

By writing explicitly the  ${}_1F_1$  functions, we have from Eq. (7),

$$\tilde{f}(X, x > 0, s) = \frac{\sum_{n=0}^{\infty} \frac{\left(\frac{s}{\gamma}\right)_n}{\left(1 + \frac{k+s}{\gamma}\right)_n} \frac{1}{n!} \left(\frac{x}{b}\right)^n}{\sum_{n=0}^{\infty} \frac{\left(\frac{s}{\gamma}\right)_n}{\left(1 + \frac{k+s}{\gamma}\right)_n} \frac{1}{n!} \left(\frac{X}{b}\right)^n}$$

which for  $\gamma = 0$ ,

$$\begin{aligned} \tilde{f}(X, x > 0, s) &= \frac{\sum_{n=0}^{\infty} \left(\frac{s}{k+s}\right)^n \frac{1}{n!} \left(\frac{x}{b}\right)^n}{\sum_{n=0}^{\infty} \left(\frac{s}{k+s}\right)^n \frac{1}{n!} \left(\frac{X}{b}\right)^n} \\ &= e^{-\frac{(X-x)s}{b(k+s)}} \\ &= e^{-\frac{(X-x)}{b}} e^{\frac{k(X-x)}{b(s+k)}} \\ &= e^{-\frac{(X-x)}{b}} \sum_{n=0}^{\infty} \left[\frac{k(X-x)}{b(s+k)}\right]^n \frac{1}{n!} \end{aligned}$$

Taking an inverse Laplace transform of the above expression we get,

$$\boxed{f(X, x > 0, t) \Big|_{\gamma=0} = e^{-\frac{(X-x)}{b}} \left[ \sum_{n=1}^{\infty} \frac{1}{n! (n-1)!} \left( \frac{(X-x)k}{b} \right)^n t^{n-1} \right] e^{-kt}} \quad (9)$$

## S5: Moments and cumulants:

For convenience of subsequent algebra, we rewrite Eq. (7) as,

$$\tilde{f}(X, x > 0, s) \Big|_{\gamma \neq 0} = \frac{g(s)}{h(s)} \quad (10)$$

where,  $g(s) = {}_1F_1\left[\frac{s}{\gamma}, 1 + \frac{k+s}{\gamma}, \frac{x}{b}\right] = \sum_{n=0}^{\infty} \frac{\left(\frac{s}{\gamma}\right)_n}{\left(1 + \frac{k+s}{\gamma}\right)_n} \frac{1}{n!} \left(\frac{x}{b}\right)^n$  and  $h(s) = {}_1F_1\left[\frac{s}{\gamma}, 1 + \frac{k+s}{\gamma}, \frac{X}{b}\right] = \sum_{n=0}^{\infty} \frac{\left(\frac{s}{\gamma}\right)_n}{\left(1 + \frac{k+s}{\gamma}\right)_n} \frac{1}{n!} \left(\frac{X}{b}\right)^n$ .

One of the important use of obtaining FPTD in the Laplace space is that the moments of any finite order may be calculated as derivatives of the Laplace transform at  $s = 0$ . Namely, the  $m$ th moment  $\langle t^m \rangle = (-1)^m \frac{d^m}{ds^m} \tilde{f}(X, x, s) \Big|_{s=0}$ . In this section, we calculate the first 3 moments.

### The mean first passage time $\langle t \rangle$ (first moment):

We see from Eq. (10) that

$$\boxed{\langle t \rangle = - \frac{\partial \tilde{f}(X, x, s)}{\partial s} \Big|_{s=0} = h^{(1)}(0) - g^{(1)}(0)} \quad (11)$$

since  $g(0) = h(0) = 1$ , where  $g^{(n)}(0) = \frac{\partial^n g(s)}{\partial s^n} \Big|_{s=0}$  and  $h^{(n)}(0) = \frac{\partial^n h(s)}{\partial s^n} \Big|_{s=0}$ . We have

$$h^{(1)}(s) = \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{X}{b}\right)^n \frac{\partial}{\partial s} \left[ \frac{\left(\frac{s}{\gamma}\right)_n}{\left(1 + \frac{k+s}{\gamma}\right)_n} \right]$$

and since for any function  $g(z)$ ,  $\frac{\partial g(z)}{\partial z} = g(z) \frac{\partial \ln(g(z))}{\partial z}$ , we have

$$h^{(1)}(s) = \frac{1}{\gamma} \sum_{n=1}^{\infty} \frac{1}{n!} \left(\frac{X}{b}\right)^n \frac{\left(\frac{s}{\gamma}\right)_n}{\left(1 + \frac{k+s}{\gamma}\right)_n} \sum_{m=0}^{n-1} \left[ \frac{1}{\frac{s}{\gamma} + m} - \frac{1}{1 + \frac{k+s}{\gamma} + m} \right] \quad (12)$$

In the limit of  $s = 0$ ,  $(s/\gamma)_n = 0$ . Collecting the surviving terms we have

$$h^{(1)}(0) = \frac{1}{\gamma} \sum_{n=1}^{\infty} \frac{1}{n} \left(\frac{X}{b}\right)^n \frac{1}{\left(1 + \frac{k}{\gamma}\right)_n} \quad (13)$$

A similar calculation gives

$$g^{(1)}(0) = \frac{1}{\gamma} \sum_{n=1}^{\infty} \frac{1}{n} \left(\frac{x}{b}\right)^n \frac{1}{\left(1 + \frac{k}{\gamma}\right)_n} \quad (14)$$

and thus finally

$$\langle t \rangle = \frac{1}{\gamma} \sum_{n=1}^{\infty} \frac{1}{n} \frac{1}{\left(1 + \frac{k}{\gamma}\right)_n} \left[ \left(\frac{X}{b}\right)^n - \left(\frac{x}{b}\right)^n \right] \quad (15)$$

**The second moment  $\langle t^2 \rangle$ :**

Proceeding as above, from Eq. (10),

$$\langle t^2 \rangle = \left. \frac{\partial^2 \tilde{f}(X, x, s)}{\partial s^2} \right|_{s=0} = g^{(2)}(0) - h^{(2)}(0) + 2h^{(1)}(0) \left( h^{(1)}(0) - g^{(1)}(0) \right) \quad (16)$$

where  $h^{(1)}(0)$  and  $g^{(1)}(0)$  is already evaluated (Eqs. 13 and 14). So, we now need to evaluate  $h^{(2)}(0)$  and  $g^{(2)}(0)$ . Differentiating Eq. (12) with respect to  $s$ :

$$\begin{aligned} h^{(2)}(s) = & \frac{1}{\gamma^2} \sum_{n=1}^{\infty} \frac{1}{n!} \left(\frac{X}{b}\right)^n \frac{\left(\frac{s}{\gamma}\right)_n}{\left(1 + \frac{k+s}{\gamma}\right)_n} \left( \left\{ \sum_{m=0}^{n-1} \left[ \frac{1}{\frac{s}{\gamma} + m} - \frac{1}{1 + \frac{k+s}{\gamma} + m} \right] \right\}^2 \right. \\ & \left. + \sum_{m=0}^{n-1} \left[ -\frac{1}{\left(\frac{s}{\gamma} + m\right)^2} + \frac{1}{\left(1 + \frac{k+s}{\gamma} + m\right)^2} \right] \right) \end{aligned} \quad (17)$$

Since for  $s = 0$ ,  $(s/\gamma)_n = 0$ , the non-zero terms yield

$$h^{(2)}(0) = \frac{1}{\gamma^2} \sum_{n=1}^{\infty} \frac{1}{n} \left(\frac{X}{b}\right)^n \frac{2}{\left(1 + \frac{k}{\gamma}\right)_n} \left[ -\frac{1}{1 + \frac{k}{\gamma}} + \sum_{m=1}^{n-1} \frac{1 + \frac{k}{\gamma}}{m(1 + m + \frac{k}{\gamma})} \right] \quad (18)$$

and similarly,

$$g^{(2)}(0) = \frac{1}{\gamma^2} \sum_{n=1}^{\infty} \frac{1}{n} \left(\frac{x}{b}\right)^n \frac{2}{\left(1 + \frac{k}{\gamma}\right)_n} \left[ -\frac{1}{1 + \frac{k}{\gamma}} + \sum_{m=1}^{n-1} \frac{1 + \frac{k}{\gamma}}{m(1 + m + \frac{k}{\gamma})} \right] \quad (19)$$

**The third moment  $\langle t^3 \rangle$ :**

From Eq. (10), we get

$$\begin{aligned} \langle t^3 \rangle = & - \left. \frac{\partial^3 \tilde{f}(X, x, s)}{\partial s^3} \right|_{s=0} \\ = & h^{(3)}(0) - g^{(3)}(0) + 3h^{(1)}(0) \left( g^{(2)}(0) - 2h^{(2)}(0) \right) + g^{(1)}(0) \left( 3h^{(2)}(0) - 6(h^{(1)}(0))^2 \right) + 6(h^{(1)}(0))^3 \end{aligned} \quad (20)$$

where  $h^{(1)}(0)$ ,  $h^{(2)}(0)$ ,  $g^{(1)}(0)$  and  $g^{(2)}(0)$  are given by Eqs. (13, 18, 14 and 19) respectively. So,  $h^{(3)}(0)$  and  $g^{(3)}(0)$  remains to be evaluated. Differentiating Eq. (17) with respect to  $s$ , we get

$$\begin{aligned}
h^{(3)}(s) = & \frac{1}{\gamma^3} \sum_{n=1}^{\infty} \frac{1}{n!} \left(\frac{X}{b}\right)^n \frac{\left(\frac{s}{\gamma}\right)_n}{\left(1 + \frac{k+s}{\gamma}\right)_n} \left( \left\{ \sum_{m=0}^{n-1} \left[ \frac{1}{\frac{s}{\gamma} + m} - \frac{1}{1 + \frac{k+s}{\gamma} + m} \right] \right\}^3 \right. \\
& + 2 \sum_{\substack{m=0 \\ l=0}}^{n-1} \left[ \frac{1}{\frac{s}{\gamma} + m} - \frac{1}{1 + \frac{k+s}{\gamma} + m} \right] \left[ -\frac{1}{\left(\frac{s}{\gamma} + l\right)^2} + \frac{1}{\left(1 + \frac{k+s}{\gamma} + l\right)^2} \right] \\
& + \sum_{\substack{m=0 \\ l=0}}^{n-1} \left[ \frac{1}{\frac{s}{\gamma} + m} - \frac{1}{1 + \frac{k+s}{\gamma} + m} \right] \left[ -\frac{1}{\left(\frac{s}{\gamma} + l\right)^2} + \frac{1}{\left(1 + \frac{k+s}{\gamma} + l\right)^2} \right] \\
& \left. - 2 \sum_{m=0}^{n-1} \left[ -\frac{1}{\left(\frac{s}{\gamma} + m\right)^3} + \frac{1}{\left(1 + \frac{k+s}{\gamma} + m\right)^3} \right] \right)
\end{aligned} \tag{21}$$

In the limit of  $s = 0$ , we have

$$\begin{aligned}
h^{(3)}(0) = & \frac{3}{\gamma^3} \sum_{n=1}^{\infty} \left(\frac{X}{b}\right)^n \frac{1}{n \left(1 + \frac{k}{\gamma}\right)_n} \left( \sum_{\substack{m=1 \\ l=1}}^{n-1} \frac{1 - \delta_{m,l}}{ml} + \sum_{\substack{m=0 \\ l=0}}^{n-1} \frac{1 + \delta_{m,l}}{\left(1 + \frac{k}{\gamma} + m\right)\left(1 + \frac{k}{\gamma} + l\right)} \right. \\
& \left. - 2 \sum_{\substack{m=1 \\ l=0}}^{n-1} \frac{1}{m\left(1 + \frac{k}{\gamma} + l\right)} \right)
\end{aligned} \tag{22}$$

and similarly,

$$\begin{aligned}
g^{(3)}(0) = & \frac{3}{\gamma^3} \sum_{n=1}^{\infty} \left(\frac{x}{b}\right)^n \frac{1}{n \left(1 + \frac{k}{\gamma}\right)_n} \left( \sum_{\substack{m=1 \\ l=1}}^{n-1} \frac{1 - \delta_{m,l}}{ml} + \sum_{\substack{m=0 \\ l=0}}^{n-1} \frac{1 + \delta_{m,l}}{\left(1 + \frac{k}{\gamma} + m\right)\left(1 + \frac{k}{\gamma} + l\right)} \right. \\
& \left. - 2 \sum_{\substack{m=1 \\ l=0}}^{n-1} \frac{1}{m\left(1 + \frac{k}{\gamma} + l\right)} \right)
\end{aligned} \tag{23}$$

### The Cumulants:

The standard definitions of the  $CV^2$  and skewness are:

(i)

$$CV^2 = \frac{\text{Variance}}{\text{mean}^2} = \frac{\langle t^2 \rangle}{\langle t \rangle^2} - 1 \tag{24}$$

(ii)

$$\text{Skewness} = \frac{\langle (t - \langle t \rangle)^3 \rangle}{\text{Variance}^{3/2}} = \frac{\langle t^3 \rangle - 3\langle t^2 \rangle \langle t \rangle + 2\langle t \rangle^3}{(\langle t^2 \rangle - \langle t \rangle^2)^{3/2}} \quad (25)$$

## S6: Bootstrapping method:

Here we outline the method followed in “bootstrapping”, mentioned in the text.

1. For each mutant, out of around 150 lysis time samples available, 80 samples were chosen randomly, with replacement. The sample size (80) was chosen such that the error bars were shortest.
2. For the random set of 80 lysis times, the statistics (mean,  $CV^2$  and skewness) were calculated.
3. The above steps were repeated 1000 times. The number of repetition should be as large as possible. We found that beyond 1000 (for example 10000), the length of error bars did not change significantly, so we kept it 1000.
4. We then had a list of 1000 mean values, 1000  $CV^2$  values and 1000 skewness values. We arranged them in ascending order.
5. To get a confidence interval of  $x\%$ , we needed to ignore the  $(100 - x)\%$  extreme points. So, for 90% confidence interval, we discarded the left 5% and right 5% of the 1000 bootstrap values, that is, 50 data points from each of left and right. Hence the error bars were marked by the 50<sup>th</sup> and the 950<sup>th</sup> elements of the list of 1000 values.

## S7: Fitting of experimental data with our theory to extract the value of $k$ :

We can calculate experimental values of  $\langle t \rangle$  and  $CV^2$  from the 90 - 175 sample lysis times for any of the  $\lambda$ -mutants. The theoretical expression of FPTD in Laplace space, and hence all the moments depend on the four parameters-  $x/b$ ,  $k$ ,  $\gamma$  and  $X/b$ . We choose  $x/b$  to be a small value 0.01, as at the beginning of the lytic cycle the protein concentration is supposed to be vanishingly small. E. Coli cell-doubling time is  $\approx 40$  min (at 37° C) and holin degradation timescale is very large. So, we choose  $\gamma = \ln(2)/40 \text{ min}^{-1}$ . We hope that  $k$  is similar for all the mutants and they have distinct  $X/b$ .

We extract  $k$  by doing a best fit of the theoretical curve of  $CV^2$  versus  $\langle t \rangle$  to the experimental data. The procedure is as follows. We vary the value of  $k$  over a range  $1 \text{ min}^{-1} - 15 \text{ min}^{-1}$  in steps of  $0.1 \text{ min}^{-1}$ . For one such value of  $k$ , using Eq. (15) with upper limit of the sum on  $n$  being chosen 1000 (a large numerical value as a substitute for  $\infty$ ), we obtain numerically using Mathematica the value of  $X/b$ . Using this value of  $X/b$  and other parameter values as mentioned, using Eq. (16), we obtain a theoretical trial value of  $CV^2$  for that particular mutant. Having 20 such values for 20 mutants, we obtain a sum of squares of the differences of these

theoretical trial  $CV^2$  and the corresponding experimental  $CV^2$  of all the mutations. The value of  $k$  for which this sum of squares is a minimum, gives us the best fit for  $k$  — we obtained  $k = 4.5 \text{ min}^{-1}$ .

Once we have this best fitted  $k$  and the chosen values of  $\gamma$  and  $x/b$ , for any mutant we set the theoretical mean  $\langle t \rangle$  equal to the experimental mean FPT, and thus obtain the full time dependent FPTD by numerically inverting the exact FPTD in Laplace space (Eq. (7)).

## S8: Exponential tail and the characteristic time of the FPTD:

The FPTD which is the Laplace inverse transform of  $\tilde{f}(X, x, s)$  (Eq. (7)) has an asymptotic exponential tail, i.e. for large  $t$ ,  $f(X, x, t) \sim \exp(-t/\tau_c)$ . The quantity  $\tau_c$  is called the *characteristic* FPT. It may easily be obtained from a negative pole of  $\tilde{f}(X, x, s)$  as follows. Note that

$$\tilde{f}(X, x, s) = \frac{{}_1F_1\left[\frac{s}{\gamma}, 1 + \frac{k+s}{\gamma}, \frac{x}{b}\right]}{{}_1F_1\left[\frac{s}{\gamma}, 1 + \frac{k+s}{\gamma}, \frac{X}{b}\right]} = \frac{{}_1F_1\left[\frac{s}{\gamma}, 1 + \frac{k+s}{\gamma}, \frac{x}{b}\right]}{a_0 \prod_{i=1}^{\infty} (s + \alpha_i)} \quad (26)$$

where,  $-\alpha_i$ 's are the roots of the function  ${}_1F_1\left[\frac{s}{\gamma}, 1 + \frac{k+s}{\gamma}, \frac{X}{b}\right]$  and  $a_0$  is a constant dependent on the parameters  $k$ ,  $\gamma$  and  $X/b$ . Hence  $f(X, x, t) = \sum_{i=1}^{\infty} c_i e^{-\alpha_i t}$  and in the limit of large  $t$ ,  $f(X, x, t) = c_* e^{-\alpha_* t}$  where  $\alpha_*$  is the smallest of  $\{\alpha_i\}$ . Thus  $\tau_c = 1/\alpha_*$ .

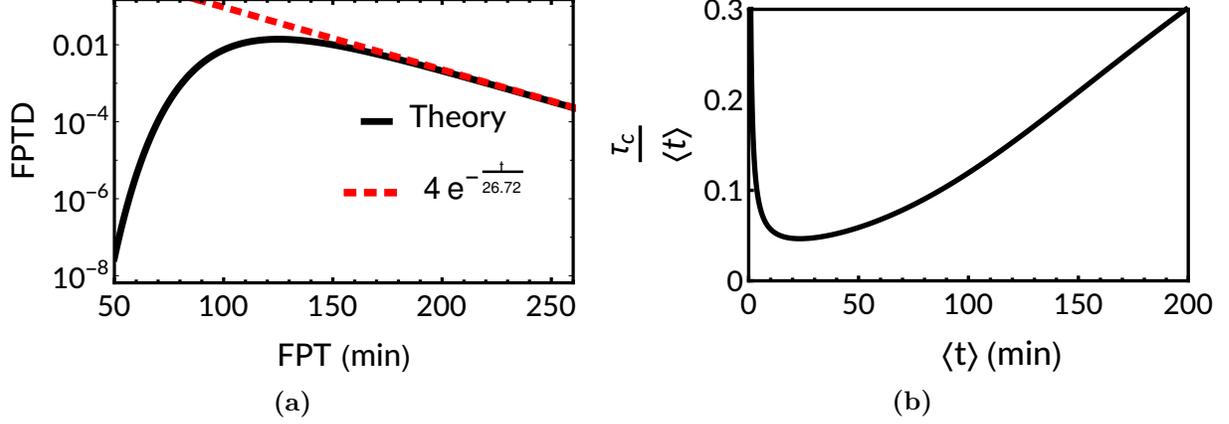
For the mutant JD426 with  $\langle t \rangle = 140 \text{ min}$ , we first obtain the value of  $X/b$  numerically solving Eq. (15) for  $\langle t \rangle = 140 \text{ min}$ ,  $k = 4.5 \text{ min}^{-1}$ ,  $\gamma = \ln(2)/40 \text{ min}^{-1}$  and  $x/b = 0.01$ . We find  $X/b = 241.9$ . We use this value to find  $\alpha_*$ , which is the negative of the root of  ${}_1F_1\left[\frac{s}{\gamma}, 1 + \frac{k+s}{\gamma}, \frac{X}{b}\right]$  nearest to 0. We get  $\alpha_* = 0.0374 \text{ min}^{-1}$  and the characteristic first passage time,  $\tau_c = 1/\alpha_* = 1/0.0374 = 26.72 \text{ min}$ . In Fig. (2a), we have plotted the exact theoretical FPTD along with the asymptotic exponential function,  $e^{-t/\tau_c}$ , in semi-log scale. The two curves merge for large  $t$  confirming the exponential tail and the calculated value of  $\tau_c$ .

Next we note that  $\tau_c$  is a measure of the spread of the FPTD and of fluctuations in FPT. We plot a scaled time  $\tau_c/\langle t \rangle$  against mean first passage time  $\langle t \rangle$  in Fig. (2b). Like  $CV^2$  and Skewness, we see that it has a non-monotonic shape. For larger mean FPT, the characteristic time gets larger indicating the broadening of the exponential tail.

## S9: Analytical argument for linear relationship between optimal mean FPT $t_m$ and $1/\gamma$ :

In Eq. (15) above for mean FPT, if we set initial concentration  $x \rightarrow 0$ ,

$$\langle t \rangle = \frac{1}{\gamma} \sum_{n=1}^{\infty} \frac{1}{n} \left(\frac{X}{b}\right)^n \frac{1}{\left(1 + \frac{k}{\gamma}\right)_n} \quad (27)$$



**Figure 2.** (a) For the mutant JD426, given the parameter values mentioned in the text, the exact theoretical FPTD (solid curve) is plotted (in semi-log scale) along with the asymptotic exponential function (dashed curve) with  $\tau_c = 26.72$  min, as a function of the the FPT. (b) Plot of scaled characteristic time  $\tau_c/\langle t \rangle$  versus mean FPT  $\langle t \rangle$ .

For threshold  $X$  less than the steady-state mean protein concentration,  $c_{ss} = kb/\gamma$ , in Eq. (27), the term  $\left(1 + \frac{k}{\gamma}\right)_n$  will be clearly greater than  $(X/b)^n$ . Moreover,  $1/n$  term makes the  $n^{th}$  term in the summation decrease fast so that only the first few terms actually contribute. Typically, values of  $k$  and  $\gamma$  are of the order  $10^0$  and  $10^{-2}$  respectively. So,  $k/\gamma$  will be of the order of  $10^2 \gg 1$ . Hence,

$$\left(1 + \frac{k}{\gamma}\right)_n \approx \left(\frac{k}{\gamma}\right)^n$$

Eq. (27) then becomes

$$\begin{aligned} \langle t \rangle &\approx \frac{1}{\gamma} \sum_{n=1}^{\infty} \frac{1}{n} \left(\frac{X}{b}\right)^n \frac{1}{\left(\frac{k}{\gamma}\right)^n} \\ &= -\frac{1}{\gamma} \ln\left(1 - \frac{X\gamma}{kb}\right) \end{aligned} \quad (28)$$

Now if one wants an “optimal  $\langle t \rangle = t_m$  to minimise fluctuations, one has to minimise the expression of  $CV^2$  (involving the moments in Section S5:) — that calculation does not give simple expressions. Yet one may expect  $t_m$  to correspond to some “optimal threshold concentration”  $X_{opt}$ , which on general dimensional grounds should be a fraction  $f$  of the steady state concentration  $c_{ss}$ . Thus,

$$X_{opt} = f \cdot c_{ss} = f \frac{kb}{\gamma} \quad (29)$$

and from Eqs. (28) and (29), putting  $X = X_{opt}$  we have optimal mean FPT  $\langle t \rangle = t_m$ , given by:

$$t_m = -\frac{1}{\gamma} \ln(1 - f) \quad (30)$$

Thus we have shown analytically that a linear relationship is expected between  $t_m$  and  $1/\gamma$ . We find the value of  $f \approx 0.55$  numerically.

## S10: Approximate Theory: Fokker-Planck equation and Langevin equation:

### Kramers-Moyal expansion:

Kramers–Moyal expansion [2] refers to a Taylor series expansion of the master equation, assuming small jumps  $y = x' - x$ . We first discuss the general procedure to derive a Fokker-Planck equation in this way, and then apply it to the problem of protein production. Given the transition rate from  $x$  to  $x'$  equal to  $W(x'|x)$ , the generic backward master equation is

$$\frac{\partial S(x, t)}{\partial t} = \int_{-\infty}^{\infty} dx' [W(x'|x)S(x', t) - W(x'|x)S(x, t)] \quad (31)$$

on substituting  $y = x' - x$  becomes

$$\frac{\partial S(x, t)}{\partial t} = \int_{-\infty}^{\infty} dy W(x + y|x)S(x + y, t) - \int_{-\infty}^{\infty} dy W(x + y|x)S(x, t) \quad (32)$$

Now, Taylor expanding Eq. (32) around  $x$

$$\begin{aligned} \frac{\partial S(x, t)}{\partial t} &= \int_{-\infty}^{\infty} dy W(x + y|x) \sum_{n=0}^{\infty} \left[ \frac{y^n}{n!} \frac{\partial^n S(x, t)}{\partial x^n} \right] - \int_{-\infty}^{\infty} dy W(x + y|x)S(x, t) \\ &= \sum_{n=1}^{\infty} \frac{\alpha^{(n)}(x, t)}{n!} \frac{\partial^n}{\partial x^n} [S(x, t)] \quad \text{where, } \alpha^{(n)}(x, t) = \int_{-\infty}^{\infty} dy W(x + y|x)y^n \end{aligned} \quad (33)$$

In our problem, the burst size is exponentially distributed, and the burst rate is  $k$ , such that

$$W(x + y|x) = \begin{cases} \frac{k}{b} e^{-\frac{y}{b}} & \text{for } y \geq 0 \\ 0 & \text{for } y < 0 \end{cases} \quad (34)$$

and we may get  $\alpha^{(0)}(x, t)$ ,  $\alpha^{(1)}(x, t)$  and  $\alpha^{(2)}(x, t)$  from Eq. (33) as

$$\alpha^{(0)}(x, t) = \int_0^{\infty} dy \frac{k}{b} e^{-\frac{y}{b}} = k \quad (35)$$

$$\alpha^{(1)}(x, t) = \int_0^{\infty} dy \frac{k}{b} e^{-\frac{y}{b}} y = kb \quad (36)$$

$$\alpha^{(2)}(x, t) = \int_0^{\infty} dy \frac{k}{b} e^{-\frac{y}{b}} y^2 = 2kb^2 \quad (37)$$

The backward master equation (Eq. (1)), including protein degradation, for our problem is

$$\frac{\partial S(X, x, t)}{\partial t} = k \int_x^X dx' [\nu(x' - x) - \delta(x' - x)] S(X, x', t) - \gamma x \frac{\partial S(X, x, t)}{\partial x}$$

and putting Eqs. (35), (36) and (37) in Eq. (33), and truncating the Kramers-Moyal expansion up to the second-order term we get the approximate backward Fokker-Planck equation as follows:

$$\boxed{\frac{\partial S(X, x, t)}{\partial t} = (kb - \gamma x) \frac{\partial S(X, x, t)}{\partial x} + kb^2 \frac{\partial^2 S(X, x, t)}{\partial x^2}} \quad (38)$$

### The Langevin equation in forward formalism:

The forward Fokker-Planck equation corresponding to Eq. (38) follows as a standard result [2,5]:

$$\boxed{\frac{\partial p(c, t')}{\partial t'} = -kb \frac{\partial p(c, t')}{\partial c} + kb^2 \frac{\partial^2 p(c, t')}{\partial c^2} + \frac{\partial}{\partial c} [\gamma c p(c, t')]} \quad (39)$$

which may be also obtained as the Kramers-Moyal expansion of the forward Master equation Eq. (2) for the protein concentration  $p(c, t')$  developed in [3]:

$$\frac{\partial p(c, t')}{\partial t'} = k \int_0^c dc' [\nu(c - c') - \delta(c - c')] p(c', t') + \gamma \frac{\partial [c p(c, t')]}{\partial c}$$

Under Ito's convention [2,5], the forward Fokker-Planck equation Eq. (39) corresponds to the following Langevin equation:

$$\boxed{\frac{dc}{dt'} = kb - \gamma c + \eta(t')} \quad (40)$$

where,  $\eta(t')$  is a Gaussian noise with zero mean,  $\langle \eta(t') \rangle = 0$ , and delta correlation,  $\langle \eta(t'_1) \eta(t'_2) \rangle = 2D \delta(t'_1 - t'_2)$ , and the diffusion constant needs to be identified as  $D = kb^2$ . The term  $kb$  represents production of protein molecules with a fixed burst size  $b$ . The second term,  $-\gamma c$ , represents decay of protein concentration.

## S11: Exact FPTD in Laplace space, for the approximate theory:

We may exactly solve the Eq. (38) as follows. Defining the Laplace transform  $\tilde{S}(X, x, s) = \int_0^\infty e^{-st} S(X, x, t)$ , the Laplace transform of Eq. (38) is

$$s\tilde{S}(X, x, s) - 1 = (kb - \gamma x) \frac{\partial \tilde{S}(X, x, s)}{\partial x} + kb^2 \frac{\partial^2 \tilde{S}(X, x, s)}{\partial x^2} \quad (41)$$

which leads to

$$\frac{\partial^2 \tilde{S}(X, x, s)}{\partial x^2} + \left( \frac{kb - \gamma x}{kb^2} \right) \frac{\partial \tilde{S}(X, x, s)}{\partial x} - \frac{s}{kb^2} \tilde{S}(X, x, s) = -\frac{1}{kb^2}$$

Performing a linear transformation,  $\left( \frac{kb - \gamma x}{kb^2} \right) = -2v$ , with  $dx = \frac{2kb^2}{\gamma} dv$  we get

$$\frac{\gamma^2}{4k^2b^4} \frac{\partial^2 \tilde{S}}{\partial v^2} - \frac{2v\gamma}{2kb^2} \frac{\partial \tilde{S}}{\partial v} - \frac{s}{kb^2} \tilde{S} = -\frac{1}{kb^2}$$

and then a scale transformation,  $v\sqrt{\frac{2kb^2}{\gamma}} = u$ ,  $dv = \sqrt{\frac{\gamma}{2kb^2}} du$  leads to

$$\frac{\partial^2 \tilde{S}}{\partial u^2} - 2u \frac{\partial \tilde{S}}{\partial u} - 2 \frac{s}{\gamma} \tilde{S} = -\frac{2}{\gamma} \quad (42)$$

The homogeneous part of Eq. (42) is a Hermite equation [6], and a particular solution of the inhomogeneous equation is  $1/s$ . Using the relations between the Hermite function and confluent hypergeometric function of the first kind  ${}_1F_1$  and that of Tricomi's  $U$  [6], we write the solution of Eq. (42) as:

$$\tilde{S}(X, x, s) = \frac{1}{s} + \alpha_1 F_1 \left[ \frac{s}{2\gamma}, \frac{1}{2}, \frac{(bk - \gamma x)^2}{2b^2 \gamma k} \right] + \beta U \left[ \frac{s}{2\gamma}, \frac{1}{2}, \frac{(bk - \gamma x)^2}{2b^2 \gamma k} \right] \quad (43)$$

We note that as  $k$  tends to 0, the term  $\frac{(bk - \gamma x)^2}{2b^2 \gamma k} \rightarrow \infty$ , and in such a limit asymptotically,  ${}_1F_1 \left[ \frac{s}{2\gamma}, \frac{1}{2}, \frac{(bk - \gamma x)^2}{2b^2 \gamma k} \right]$  tends to  $\infty$  while  $U \left[ \frac{s}{2\gamma}, \frac{1}{2}, \frac{(bk - \gamma x)^2}{2b^2 \gamma k} \right]$  tends to 0 [6]. This would imply that  $\tilde{S}(X, x, s)$  will be unbounded as  $k \rightarrow 0$  for every value of the Laplace variable  $s$ . Consequently  $S(X, x, t)$  would be unbounded, which cannot be, as it belongs to the range  $[0, 1]$ . So, we discard the  ${}_1F_1$  solution, i.e., we set  $\alpha = 0$ .

Using the boundary condition,  $S(X, X, t) = 0$ , we fix  $\beta$  and get

$$\tilde{S}(X, x, s) \Big|_{\text{approx}} = \frac{1}{s} \left( 1 - \frac{U \left[ \frac{s}{2\gamma}, \frac{1}{2}, \frac{(bk - \gamma x)^2}{2b^2 \gamma k} \right]}{U \left[ \frac{s}{2\gamma}, \frac{1}{2}, \frac{(bk - \gamma X)^2}{2b^2 \gamma k} \right]} \right) \quad (44)$$

Using the relation  $\tilde{f}(X, x, s) = 1 - s\tilde{S}(X, x, s)$ , we get the exact FPTD in the Laplace space for this approximate theory:

$$\tilde{f}(X, x, s) \Big|_{\text{approx}} = \frac{U \left[ \frac{s}{2\gamma}, \frac{1}{2}, \frac{(bk - \gamma x)^2}{2b^2 \gamma k} \right]}{U \left[ \frac{s}{2\gamma}, \frac{1}{2}, \frac{(bk - \gamma X)^2}{2b^2 \gamma k} \right]} \quad (45)$$

## S12: Exact moments, of the approximate theory:

Like Section (S5:) above, we proceed by defining the FPTD (Eq. (45)) of the approximate theory as

$$\tilde{f}(X, x, s)_{\text{approx}} = \frac{g(s)}{h(s)} \quad (46)$$

where,  $g(s) = U \left[ \frac{s}{2\gamma}, \frac{1}{2}, \frac{(bk - \gamma x)^2}{2b^2 \gamma k} \right]$  and  $h(s) = U \left[ \frac{s}{2\gamma}, \frac{1}{2}, \frac{(bk - \gamma X)^2}{2b^2 \gamma k} \right]$ .

Mean first passage (first moment,  $\langle t \rangle$ ):

$$\boxed{\langle t \rangle = - \left. \frac{\partial \tilde{f}(X, x, s)}{\partial s} \right|_{s=0} = h^{(1)}(0) - g^{(1)}(0)} \quad \text{where, } g^{(n)}(0) = \left. \frac{\partial^n g(s)}{\partial s^n} \right|_{s=0} \quad \text{and } h^{(n)}(0) = \left. \frac{\partial^n h(s)}{\partial s^n} \right|_{s=0} \quad (47)$$

The Tricomi's confluent hypergeometric function,  $U$ , can be written in terms of confluent hypergeometric function (of first kind),  ${}_1F_1$ , as [6]:

$$U \left[ \frac{s}{2\gamma}, \frac{1}{2}, z \right] = \frac{\Gamma(\frac{1}{2})}{\Gamma(\frac{s}{2\gamma} + \frac{1}{2})} {}_1F_1 \left[ \frac{s}{2\gamma}, \frac{1}{2}, z \right] + \frac{\Gamma(-\frac{1}{2})}{\Gamma(\frac{s}{2\gamma})} z^{\frac{1}{2}} {}_1F_1 \left[ \frac{s}{2\gamma} + \frac{1}{2}, \frac{3}{2}, z \right] \quad (48)$$

$$\begin{aligned} \frac{\partial U}{\partial s} = & \Gamma \left( \frac{1}{2} \right) \left[ - \frac{1}{\Gamma(\frac{s}{2\gamma} + \frac{1}{2})^2} \Gamma \left( \frac{s}{2\gamma} + \frac{1}{2} \right) \psi \left( \frac{s}{2\gamma} + \frac{1}{2} \right) {}_1F_1 \left[ \frac{s}{2\gamma}, \frac{1}{2}, z \right] \frac{1}{2\gamma} \right. \\ & \left. + \frac{1}{\Gamma(\frac{s}{2\gamma} + \frac{1}{2})} {}_1F_1^{(1)} \left[ \frac{s}{2\gamma}, \frac{1}{2}, z \right] \right] + \Gamma \left( -\frac{1}{2} \right) z^{\frac{1}{2}} \left[ - \frac{1}{\Gamma(\frac{s}{2\gamma})^2} \Gamma \left( \frac{s}{2\gamma} \right) \psi \left( \frac{s}{2\gamma} \right) \frac{1}{2\gamma} {}_1F_1 \left[ \frac{s}{2\gamma} + \frac{1}{2}, \frac{3}{2}, z \right] \right. \\ & \left. + \frac{1}{\Gamma(\frac{s}{2\gamma})} {}_1F_1^{(1)} \left[ \frac{s}{2\gamma} + \frac{1}{2}, \frac{3}{2}, z \right] \right] \end{aligned} \quad (49)$$

In Eq. (49), the superscripts,  $(n)$ , denotes  $n^{\text{th}}$  derivatives and  $\psi(x)$  is the polygamma function defined as  $\psi(x) = \frac{\partial \ln \Gamma(x)}{\partial x}$  [6]. We note three results (i)-(iii) below.

(i)

$${}_1F_1^{(1)} \left[ \frac{s}{2\gamma}, \frac{1}{2}, z \right] = \frac{1}{2\gamma} \sum_{n=1}^{\infty} \left[ \frac{\left( \frac{s}{2\gamma} \right)_n}{\left( \frac{1}{2} \right)_n} \frac{z^n}{n!} \sum_{p=0}^{n-1} \frac{1}{\frac{s}{2\gamma} + p} \right] \quad (50)$$

which in the limit of  $s = 0$ ,

$$= \frac{1}{2\gamma} \sum_{n=1}^{\infty} \frac{z^n}{\left( \frac{1}{2} \right)_n n} \quad (51)$$

(ii) It is know from [6]:

$$\left. \frac{\psi(x)}{\Gamma(x)} \right|_{x=0} = -1 \quad (52)$$

(iii) It is known that

$$\left. \Gamma(x) \right|_{x=0} = -\infty \quad (53)$$

Using Eqs. (51), (52) and (53) in Eq. (49), we get

$$\frac{\partial U}{\partial s} \Big|_{s=0} = \frac{1}{2\gamma} \left[ -\psi\left(\frac{1}{2}\right) + \sum_{n=1}^{\infty} \frac{z^n}{\left(\frac{1}{2}\right)_n n} + \Gamma\left(-\frac{1}{2}\right) z^{\frac{1}{2}} {}_1F_1\left[\frac{1}{2}, \frac{3}{2}, z\right] \right] \quad (54)$$

Note  $g^{(1)}(0)$  and  $h^{(1)}(0)$  are given by:

$$g^{(1)}(0) = \frac{\partial U}{\partial s} \Big|_{z=\frac{(bk-\gamma x)^2}{2b^2\gamma k}}^{s=0} \quad \text{and} \quad h^{(1)}(0) = \frac{\partial U}{\partial s} \Big|_{z=\frac{(bk-\gamma X)^2}{2b^2\gamma k}}^{s=0} \quad (55)$$

So, using Eq. (54) in Eq. (47), we get the mean first passage time as

$$\begin{aligned} \langle t \rangle &= -\frac{\partial U}{\partial s} \Big|_{z=\frac{(bk-\gamma x)^2}{2b^2\gamma k}}^{s=0} + \frac{\partial U}{\partial s} \Big|_{z=\frac{(bk-\gamma X)^2}{2b^2\gamma k}}^{s=0} \quad (56) \\ &= \frac{1}{2\gamma} \left[ \sum_{n=1}^{\infty} \frac{1}{\left(\frac{1}{2}\right)_n n} \left( \left( \frac{(bk-\gamma X)^2}{2b^2\gamma k} \right)^n - \left( \frac{(bk-\gamma x)^2}{2b^2\gamma k} \right)^n \right) \right. \\ &\quad \left. - 2\sqrt{\pi} \left( \left( \frac{(bk-\gamma X)^2}{2b^2\gamma k} \right)^{\frac{1}{2}} {}_1F_1\left[\frac{1}{2}, \frac{3}{2}, \frac{(bk-\gamma X)^2}{2b^2\gamma k}\right] - \left( \frac{(bk-\gamma x)^2}{2b^2\gamma k} \right)^{\frac{1}{2}} {}_1F_1\left[\frac{1}{2}, \frac{3}{2}, \frac{(bk-\gamma x)^2}{2b^2\gamma k}\right] \right) \right] \quad (57) \end{aligned}$$

**The second moment  $\langle t^2 \rangle$ :**

$$\langle t^2 \rangle = \frac{\partial^2 \tilde{f}(X, x, s)}{\partial s^2} \Big|_{s=0} = g^{(2)}(0) - h^{(2)}(0) + 2h^{(1)}(0) \left( h^{(1)}(0) - g^{(1)}(0) \right) \quad (58)$$

Differentiating Eq. (49) with respect to  $s$ :

$$\begin{aligned} \frac{\partial^2 U}{\partial s^2} &= \Gamma\left(\frac{1}{2}\right) \left[ -\left( \frac{\psi\left(\frac{s}{2\gamma} + \frac{1}{2}\right)}{\Gamma\left(\frac{s}{2\gamma} + \frac{1}{2}\right)} \right)^{(1)} {}_1F_1\left[\frac{s}{2\gamma}, \frac{1}{2}, z\right] \frac{1}{2\gamma} - \frac{\psi\left(\frac{s}{2\gamma} + \frac{1}{2}\right)}{\Gamma\left(\frac{s}{2\gamma} + \frac{1}{2}\right)} \frac{1}{2\gamma} {}_1F_1^{(1)}\left[\frac{s}{2\gamma}, \frac{1}{2}, z\right] \right. \\ &\quad \left. + \frac{{}_1F_1^{(2)}\left[\frac{s}{2\gamma}, \frac{1}{2}, z\right] - \psi\left(\frac{s}{2\gamma} + \frac{1}{2}\right) {}_1F_1^{(1)}\left[\frac{s}{2\gamma}, \frac{1}{2}, z\right] \frac{1}{2\gamma}}{\Gamma\left(\frac{s}{2\gamma} + \frac{1}{2}\right)} \right] \\ &\quad + \Gamma\left(-\frac{1}{2}\right) z^{\frac{1}{2}} \left[ -\left( \frac{\psi\left(\frac{s}{2\gamma}\right)}{\Gamma\left(\frac{s}{2\gamma}\right)} \right)^{(1)} {}_1F_1\left[\frac{s}{2\gamma} + \frac{1}{2}, \frac{3}{2}, z\right] \frac{1}{2\gamma} - \frac{\psi\left(\frac{s}{2\gamma}\right)}{\Gamma\left(\frac{s}{2\gamma}\right)} \frac{1}{2\gamma} {}_1F_1^{(1)}\left[\frac{s}{2\gamma} + \frac{1}{2}, \frac{3}{2}, z\right] \right. \\ &\quad \left. + \frac{{}_1F_1^{(2)}\left[\frac{s}{2\gamma} + \frac{1}{2}, \frac{3}{2}, z\right] - \psi\left(\frac{s}{2\gamma}\right) {}_1F_1^{(1)}\left[\frac{s}{2\gamma} + \frac{1}{2}, \frac{3}{2}, z\right] \frac{1}{2\gamma}}{\Gamma\left(\frac{s}{2\gamma}\right)} \right] \quad (59) \end{aligned}$$

(i) From Eq. (50),

$${}_1F_1^{(2)} \left[ \frac{s}{2\gamma}, \frac{1}{2}, z \right] = \frac{1}{(2\gamma)^2} \sum_{n=1}^{\infty} \frac{z^n}{n! \left(\frac{1}{2}\right)_n} \left(\frac{s}{2\gamma}\right)_n \left[ \sum_{p=0}^{n-1} \frac{1}{\left(\frac{s}{2\gamma} + p\right) \left(\frac{s}{2\gamma} + m\right)} - \sum_{p=0}^{n-1} \frac{1}{\left(\frac{s}{2\gamma} + p\right)^2} \right]$$

which in the limit of  $s = 0$ ,

$$= \frac{1}{(2\gamma)^2} \sum_{n=1}^{\infty} \frac{2z^n}{n \left(\frac{1}{2}\right)_n} \sum_{m=1}^{n-1} \frac{1}{m} \quad (60)$$

(ii)

$${}_1F_1^{(1)} \left[ \frac{s}{2\gamma} + \frac{1}{2}, \frac{3}{2}, z \right] = \frac{1}{2\gamma} \sum_{n=1}^{\infty} \frac{\left(\frac{s}{2\gamma} + \frac{1}{2}\right)_n}{\left(\frac{1}{2}\right)_n} \frac{z^n}{n!} \sum_{m=0}^{n-1} \frac{1}{\frac{s}{2\gamma} + \frac{1}{2} + m}$$

which in the limit of  $s = 0$ ,

$$= \frac{1}{2\gamma} \sum_{n=1}^{\infty} \frac{z^n}{n!} \sum_{m=0}^{n-1} \frac{1}{\frac{1}{2} + m} \quad (61)$$

(iii) From (6):

$$\left. \frac{\partial}{\partial x} \left( \frac{\psi(x)}{\Gamma(x)} \right) \right|_{x=0} = -2E \quad (62)$$

where  $E$  is Euler's constant,  $E = 0.57721\dots$

Put Eqs. (51) - (53) and (60) - (62) in Eq. (59),

$$\begin{aligned} \left. \frac{\partial^2 U}{\partial s^2} \right|_{s=0} &= \frac{1}{(2\gamma)^2} \left[ -\sqrt{\pi} \left( \frac{\psi\left(\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)} \right)^{(1)} - 2\psi\left(\frac{1}{2}\right) \sum_{n=1}^{\infty} \frac{1}{\left(\frac{1}{2}\right)_n} \frac{z^n}{n} + \sum_{n=1}^{\infty} \frac{2z^n}{n \left(\frac{1}{2}\right)_n} \sum_{m=1}^{n-1} \frac{1}{m} \right] \\ &\quad - \frac{\sqrt{\pi} z^{\frac{1}{2}}}{\gamma^2} \left[ E_1 F_1 \left[ \frac{1}{2}, \frac{3}{2}, z \right] + \sum_{n=1}^{\infty} \frac{z^n}{n!} \sum_{m=0}^{\infty} \frac{1}{\frac{1}{2} + m} \right] \end{aligned} \quad (63)$$

From Eq. (63) we can get the expressions of  $g^{(2)}(0)$  and  $h^{(2)}(0)$ :

$$g^{(2)}(0) = \left. \frac{\partial^2 U}{\partial s^2} \right|_{\substack{s=0 \\ z = \frac{(bk-\gamma x)^2}{2b^2\gamma k}}} \quad \text{and} \quad h^{(2)}(0) = \left. \frac{\partial^2 U}{\partial s^2} \right|_{\substack{s=0 \\ z = \frac{(bk-\gamma X)^2}{2b^2\gamma k}}}$$

Substituting these expressions along with the previously obtained expressions of  $g^{(1)}(0)$  and  $h^{(1)}(0)$  (Eq. (55)) into Eq. (58), we will finally get the expression for  $\langle t^2 \rangle$

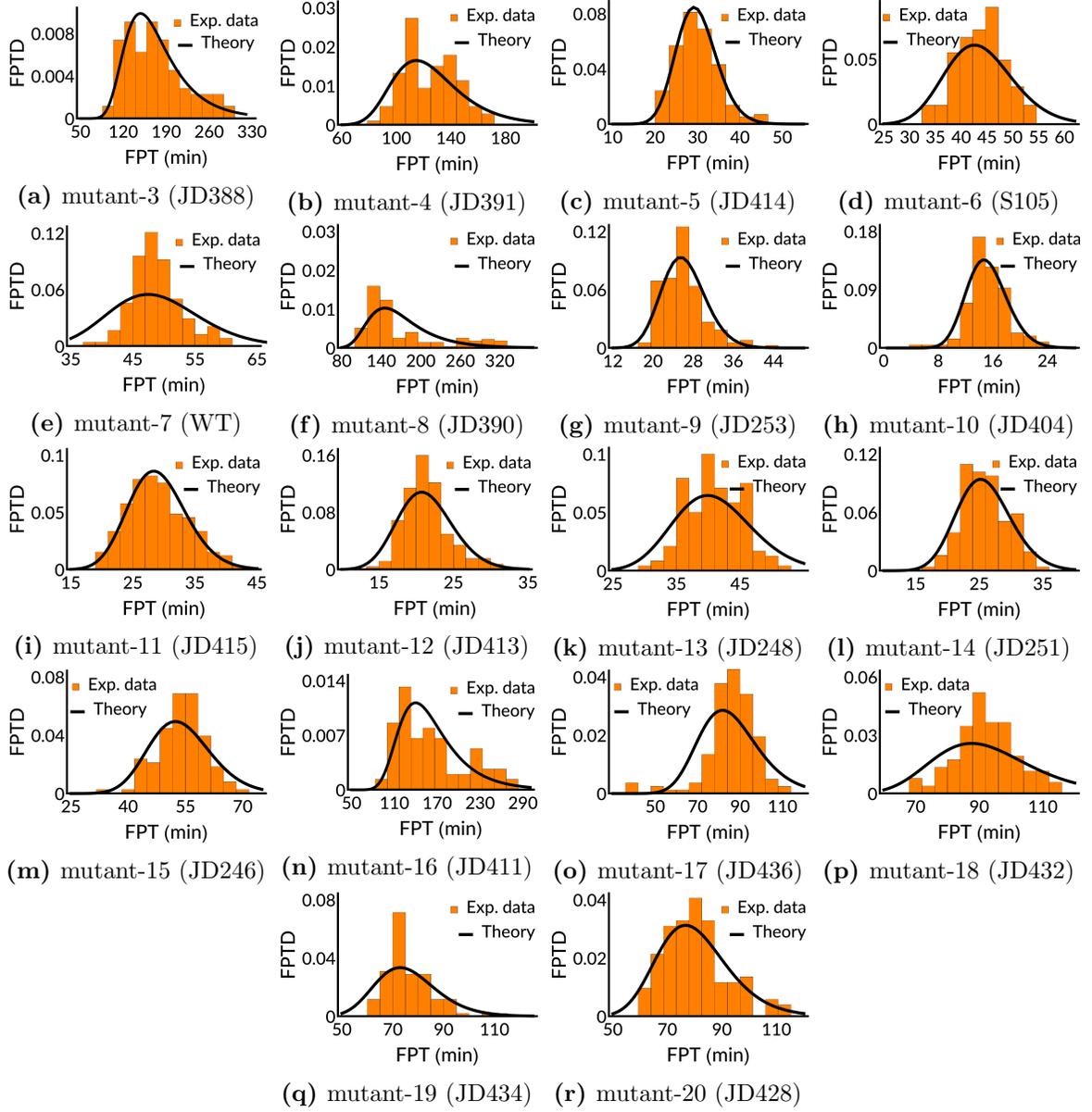
**Third moment,  $\langle t^3 \rangle$ :**

$$\begin{aligned} \langle t^3 \rangle &= - \left. \frac{\partial^3 \tilde{f}(X, x, s)}{\partial s^3} \right|_{s=0} \\ &= h^{(3)}(0) - g^{(3)}(0) + 3h^{(1)}(0) \left( g^{(2)}(0) - 2h^{(2)}(0) \right) + g^{(1)}(0) \left( 3h^{(2)}(0) - 6(h^{(1)}(0))^2 \right) + 6(h^{(1)}(0))^3 \end{aligned} \quad (64)$$

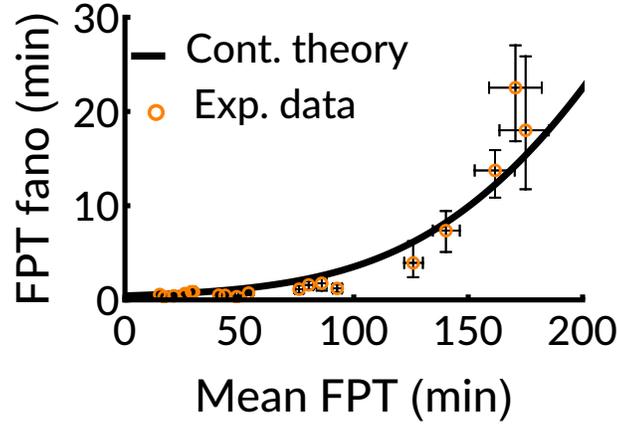
For third moment, we need to evaluate two more terms,  $g^{(3)}(0)$  and  $h^{(3)}(0)$ :

$$g^{(3)}(0) = \left. \frac{\partial^3 U}{\partial s^3} \right|_{z = \frac{(bk - \gamma x)^2}{2b^2 \gamma k}}^{s=0} \quad \text{and} \quad h^{(3)}(0) = \left. \frac{\partial^3 U}{\partial s^3} \right|_{z = \frac{(bk - \gamma X)^2}{2b^2 \gamma k}}^{s=0}$$

$\frac{\partial^3 U}{\partial s^3}$  can be evaluated by taking one more derivative of Eq. (59) with respect to  $s$ . We skip the calculation as it gets lengthy. The derivatives are easily evaluated numerically by Mathematica which we use for our plots of Skewness in the main manuscript.



**Figure 3.** The FPTD from experiment and theory for 18 different mutants. The plots for the remaining two mutants: mutant-1 (JD405) and mutant-2 (JD426) are in the main manuscript (Fig. (2)).



**Figure 4.** FPT Fano factor against  $\langle t \rangle$  for 20 mutants (symbols) and exact continuous theory (solid line) with  $k = 4.5 \text{ min}^{-1}$ ,  $x/b = 0.01$  and  $\gamma = \ln(2)/40 \text{ min}^{-1}$ . Error bars are 90% confidence intervals, obtained after bootstrapping 1000 replicates.

## References

- [1] S. Kannoly, T. Gao, S. Dey, N. Wang, A. Singh, and J. J. Dennehy, “Optimum threshold minimizes noise in timing of intracellular events,” *iScience*, vol. 23, p. 101186, 2020.
- [2] C. W. Gardiner, *Handbook of stochastic methods*, vol. 3. springer Berlin, 1985.
- [3] N. Friedman, L. Cai, and X. S. Xie, “Linking stochastic dynamics to population distribution: an analytical framework of gene expression,” *Phys. Rev. Lett.*, vol. 97, no. 16, p. 168302, 2006.
- [4] G. F. Simmons, *Differential equations with applications and historical notes*. CRC Press, 2016.
- [5] N. G. Van Kampen, “Itô versus stratonovich,” *Journal of Statistical Physics*, vol. 24, no. 1, pp. 175–187, 1981.
- [6] M. Abramowitz, I. A. Stegun, and R. H. Romer, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. American Association of Physics Teachers, 1988.