Optimal auto–regulation to minimize first–passage time variability in protein level

Khem Raj Ghusinga[1], Pak–Wing Fok[2] and Abhyudai Singh[3]

***Abstract*— The timing of cellular events is inherently random because of the probabilistic nature of gene expression. Yet cells manage to have precise timing of important events. Here, we study how gene expression could possibly be regulated to precisely schedule timing of an event around a given time. Event timing is modeled as the first–passage time (FPT) for a protein's level to cross a critical threshold. Considering auto–regulation as a possible regulatory mechanism, we investigate what form of auto–regulation would lead to minimum stochasticity in FPT around a fixed time. We formulate a stochastic gene expression model and show that under certain assumptions, it reduces to a birth-death process. Our results show that when the death rate is zero, the objective is best achieved when all of the birth rates are equal. On the contrary, when the death rate is non–zero, the optimal birth rates are not equal. In terms of the gene expression model, these results illustrate that when protein does not degrade, stochasticity in FPT around a given time is minimized when there is no auto–regulation of its expression. However, when the protein degrades, some form of auto–regulation is required to achieve this. These results are consistent with experimental findings for the lysis time stochasticity in λ phage.**

I. Introduction

Gene expression, the process of synthesizing proteins from a gene via transcription and translation, is inherently a stochastic process, leading to cell–to–cell variability in the time evolution of protein levels even in a population of isogenic cells [1]–[3]. It also gives rise to randomness in the timing of cellular events which occur at critical protein thresholds after the onset of gene expression [4]–[8]. For instance, the time it takes for a bacteriophage λ virus to kill an infected *E. coli* cell (called the lysis time) is stochastic [7]. Its variability can be accounted for by randomness in the time a certain protein (called holin) takes to reach a critical level [9].

Despite the randomness in the protein levels and consequently in the event timings, cells typically manage to function robustly using various regulation mechanisms [10]–[14]. For example, lysis of a host cell by the bacteriophage λ virus is a precisely scheduled event even though synthesis of the lysis protein, holin, is stochastic [7], [15]. Furthermore, it has been suggested that an optimal lysis time exists that gives evolutionary advantage to the virus in terms of its fitness [16]–[20]. This implies that there could be regulation of expression of holin so that lysis happens around the optimal time, with minimum variability. Auto–regulation wherein a protein regulates its own transcription is one of the regulatory mechanism widely used by cells [14], [21].

In this work, we study how event timing could be regulated to minimize its stochasticity around a given time point. Event timing is modeled as the first–passage time (FPT) for the protein copy numbers to cross a critical threshold in the stochastic model of gene expression. Further, we analyze auto–regulation as a possible regulatory mechanism cells could employ to have precision in FPT. Ultimately, we seek the answer to: *what form of auto–regulatory feedback would minimize the stochasticity in FPT around a given time*?

Previously, we have dealt with similar a question assuming that the proteins are stable and produced in geometric bursts [22]. Here, we extend our analysis to the case where proteins are allowed to degrade. We start off with a standard two state model of gene expression. We make an additional assumption of the mean burst size being small so that the gene expression model can be analyzed by a birth-death process. Our analysis shows that when the death (protein degradation) rate is zero, the birth rates that minimize the stochasticity around a given time are equal. However, when the protein degradation (or death) term is introduced, these optimal birth rates are no longer equal. Connecting these observations back to the gene expression implies that when proteins are stable and do not degrade, the optimal auto–regulation is no feedback. However, when protein is unstable, auto–regulation is required to minimize the randomness in the timing around a given time.

This paper is organized as follows. In the section II, we mathematically prove that when the mean burst size is small, a gene expression model reduces to a birth-death process. In the next section, the distribution and first two moments of FPT for a birth-death process are calculated. Then, in section IV, we formulate the objective function that we aim to minimize. Section V discusses the optimal birth rates that would minimize the objective function. Lastly, we discuss the results and their implications. Also, the notations used in this work are summarized in Table I.

II. Gene expression as a birth-death process

In this section, we establish that under certain conditions a gene expression model converges to a birth-death process in terms of statistical properties of protein levels (refer to Fig. 1). Consider a gene expression model with auto–regulation. The transcription of mRNAs from the DNA is constitutive and takes places at a protein dependent rate $\hat{k}_x$ where the

[1]Khem Raj Ghusinga is with Department of Electrical Engineering,University of Delaware, Newark, DE 19716, USA khem@udel.edu

[2]Pak–Wing Fok is with the Faculty of Mathematical Sciences, University of Delaware, Newark, DE 19716, USA pakwing@udel.edu

[3]Abhyudai Singh is with the Faculty of Electrical and Computer Engineering, Biomedical Engineering and Mathematical Sciences, University of Delaware, Newark, DE 19716, USA absingh@udel.edu

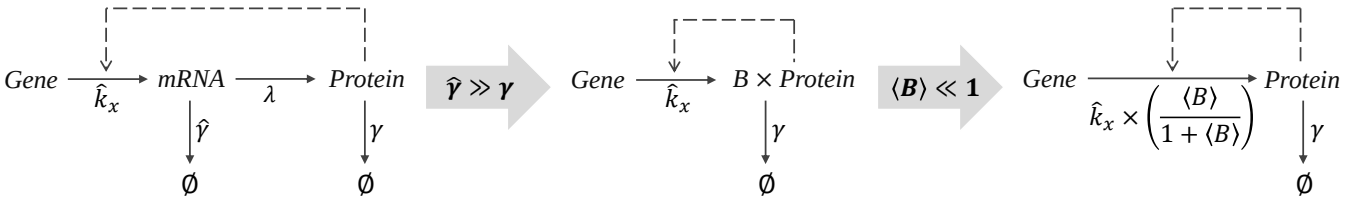

Fig. 1. A gene expression model converges to a birth-death process in terms of statistical properties of protein count. On the left, an auto–regulated gene expression model is shown wherein transcription takes place at a rate $\hat{k}_x$ while each mRNA is translated to proteins at a rate λ. mRNAs and proteins degrade with rates $\hat{\gamma}$ and γ respectively. In the limit when $\hat{\gamma} >> \gamma$, this simplifies to a burst model where each burst produces B (geometrically distributed) proteins. This has been depicted in the center. In the limit when $\langle B \rangle \ll 1$, this can be further simplified to a birth-death process as shown on the right. Here, each birth event, occurring at a rate $\hat{k}_x(\langle B \rangle / (1+\langle B \rangle))$, makes a protein while each death event takes place with rate γ.

subscript x is for protein level. Each mRNA is further translated to proteins at a rate λ. The rates of degradation of one molecule of mRNA and protein are respectively given by $\hat{\gamma}$ and γ. It has been established that when $\hat{\gamma} \gg \gamma$, mRNA dynamics can be ignored. In this case, each transcription event creates a random burst of protein molecules [23]–[26]. Below we describe this model in detail.

Let $x(t)$ denote the protein level at a time t. The probabilistic birth and death of proteins in the burst model are

$$\mathbb{P}(x(t+dt) = i+B | x(t) = i) = \hat{k}_i \, dt, \tag{1a}$$
$$\mathbb{P}(x(t+dt) = i-1 | x(t) = i) = i\gamma dt, \tag{1b}$$

where B is a random variable representing the protein burst size. The distribution of B is geometric with mean $\langle B \rangle := \lambda / \hat{\gamma}$ [27], [28]. Assuming μ to be the parameter of the geometric distribution, we have

$$\mathbb{P}(B \geq n) = (1-\mu)^n, \quad n \in \{0, 1, 2, \cdots\}. \tag{2}$$

where μ is related with $\langle B \rangle$ as $\mu = 1/(\langle B \rangle + 1)$.

TABLE I
SUMMARY OF NOTATIONS

$\hat{k}_x$	Transcription rate as a function of protein (x)
λ	Translation rate
$\hat{\gamma}$	mRNA degradation rate
γ	Protein degradation rate
$\mathbb{P}$	Probability
B	Protein burst size
μ	Parameter of burst distribution (geometric)
$p_i(t)$	$\mathbb{P}(x(t) = i)$ where $x(t)$ is protein count at time t
X	Threshold for FPT
k	Birth rate for the birth-death process
$S_i(t)$	Survival probability of the particle starting at i
$w_i(t)$	Density of FPT to reach X for a particle starting at i
$\tilde{S}_i(s)$	Laplace transform of $S_i(t)$
$\tilde{w}_i(s)$	Laplace transform of $w_i(t)$
$\langle . \rangle$	Expectation operator
τ_i	Mean exit time for birth-death process starting with initial protein count i. Also for birth-death process
η_i	Second order moment of FPT for birth-death process

To see under what conditions the statistical properties of $x(t)$ under this model converge to those of a birth-death process, we compare their master equation formulations. Let $p_i(t)$ be a shorthand notion for $\mathbb{P}(x(t) = i)$. The master equations for the gene expression model are given by

$$\frac{\partial p_0(t)}{\partial t} = -(1-\mu)\hat{k}_0 p_0(t) + \gamma p_1(t), \tag{3a}$$
$$\frac{\partial p_i(t)}{\partial t} = -\left((1-\mu)\hat{k}_i + i\gamma\right) p_i(t) + (i+1)\gamma p_{i+1}(t) + \sum_{n=0}^{i-1} \mu (1-\mu)^{i-n} \hat{k}_n p_n(t). \tag{3b}$$

When $\langle B \rangle \ll 1$, μ would be ≈ 1 and the terms $(1-\mu)^j, j \geq 2$ can be ignored. The master equations now simplify to

$$\frac{\partial p_0(t)}{\partial t} = -(1-\mu)\hat{k}_0 p_0(t) + \gamma p_1(t), \tag{4a}$$
$$\frac{\partial p_i(t)}{\partial t} = -\left((1-\mu)\hat{k}_i + i\gamma\right) p_i(t) + (i+1)\gamma p_{i+1}(t) + (1-\mu)\hat{k}_{i-1} p_{i-1}(t). \tag{4b}$$

Next, consider a birth-death process described by

$$\mathbb{P}(x(t+dt) = i+1 | x(t) = i) = k_i \, dt, \tag{5a}$$
$$\mathbb{P}(x(t+dt) = i-1 | x(t) = i) = i\gamma dt, \tag{5b}$$

where k_i denotes the birth rate when protein count is i. The death rate for each protein is denoted by γ. The corresponding master equation is

$$\frac{\partial p_0(t)}{\partial t} = -k_0 p_0(t) + \gamma p_1(t), \tag{6a}$$
$$\frac{\partial p_i(t)}{\partial t} = -(k_i + i\gamma) p_i(t) + (i+1)\gamma p_{i+1}(t) + k_{i-1} p_{i-1}(t). \tag{6b}$$

Comparing the master equation formulations in (4a) and (4b) with those in (6a) and (6b), it can be concluded that when $\langle B \rangle \ll 1$, a gene expression model converges to a birth-death process in terms of statistical properties of protein count. Notice that the birth rates in the birth-death model are proportional to the transcription rates in the gene expression model. More specifically, $k_i = \hat{k}_i(1-\mu) = \hat{k}_i(\langle B \rangle / (1+\langle B \rangle))$. Next, we derive the distribution and the moments of first–passage time (FPT) of a birth-death process.

III. First–passage time for a birth-death process

The first-passage time calculations for a birth-death process have been previously dealt with in the literature (see, for instance, [29]–[31]).Here, we describe the calculations which are specific to the problem at hand as a convenience to the reader.

The first–passage time (FPT) for the protein level $x(t)$ to cross a threshold X is defined as

$$FPT := \inf\{t : x(t) \geq X\}. \tag{7}$$

To calculate the distribution of FPT, let us consider the birth-death process with absorbing boundary at X. The process can be imagined as a particle hopping on an integer lattice with forward and backward rates as birth and death rates respectively. The particle's position at any given time represents the protein count at that time. When the particle reaches the protein count X, the process terminates and the FPT is recorded. This is shown in Fig. 2.

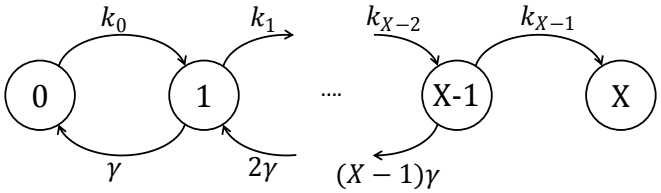

Fig. 2. Illustration of FPT calculation for a birth-death process. The birth-death process is depicted as particle hopping on integer lattice. The sites $\{0,1,..X\}$ denote the protein count. Forward hopping rates are protein dependent birth rates k_i while backward hopping rates are death rates $i\gamma$. The process terminates when the particle reaches X and FPT is recorded.

Let $\mathbb{P}(j,t|i,\hat{t})$ be the probability that the protein count (the particle's position) is j at time t given that it was i at time $\hat{t}$. The backward equation describing $\mathbb{P}(j,t|i,\hat{t})$ is

$$\begin{aligned}\frac{\partial \mathbb{P}(j,t|i,\hat{t})}{\partial t} &= k_i \mathbb{P}(j,t|i+1,\hat{t}) + i\gamma \mathbb{P}(j,t|i-1,\hat{t}) \\ &\quad - (i\gamma + k_i)\mathbb{P}(j,t|i,\hat{t}), \quad 1 \leq i \leq X-1, \qquad (8a)\\ \frac{\partial \mathbb{P}(j,t|0,\hat{t})}{\partial t} &= k_0 \mathbb{P}(j,t|1,\hat{t}) - k_0 \mathbb{P}(j,t|0,\hat{t}), \qquad (8b)\\ \mathbb{P}(j,t|X,\hat{t}) &= 0. \qquad (8c)\end{aligned}$$

The initial conditions are $\mathbb{P}(j,\hat{t}|i,\hat{t}) = \delta_{ij}$, where δ_{ij} is the Kronecker delta function. Let us define $S_i(t)$ as the survival probability of a particle that starts at protein count i given that it gets absorbed when it reaches the protein count X. Then, $S_i(t) := \sum_{j=1}^{X} \mathbb{P}(j,t|i,\hat{t}=0)$, and

$$\begin{aligned}\frac{\partial S_i}{\partial t} &= k_i S_{i+1}(t) + i\gamma S_{i+1}(t) - (i\gamma + k_i)S_i(t), \qquad (9a)\\ \frac{\partial S_0}{\partial t} &= k_0 S_1(t) - k_0 S_0(t), \qquad (9b)\\ S_X(t) &= 0. \qquad (9c)\end{aligned}$$

Also, $S_i(0) = 1$ for $0 \leq i \leq X$. Taking Laplace transforms

$$\begin{aligned}-1 + s\tilde{S}_i(s) &= k_i \tilde{S}_{i+1}(s) + i\gamma \tilde{S}_{i-1}(s) - (i\gamma + k_i)\tilde{S}_i(s), \quad (10a)\\ -1 + s\tilde{S}_0(s) &= k_0 \tilde{S}_1(s) - k_0 \tilde{S}_0(s), \qquad (10b)\\ \tilde{S}_X(s) &= 0. \qquad (10c)\end{aligned}$$

Let $w_i(t)$ denote the first–passage time probability distribution, then $w_i(t)dt = S_i(t) - S_i(t+dt)$ or $w_i(t) = -\frac{\partial S_i}{\partial t}$. This implies $\tilde{w}_i(s) = -s\tilde{S}_i(s) + 1$. Therefore, we could write the Laplace transformed survival density function as

$$\begin{aligned}s\tilde{w}_i(s) &= k_i \tilde{w}_{i+1}(s) + i\gamma \tilde{w}_{i+1}(s) - (i\gamma + k_i)\tilde{w}_i(s), \qquad (11a)\\ s\tilde{w}_0(s) &= k_0 \tilde{w}_1(s) - k_0 \tilde{w}_0(s), \qquad (11b)\\ \tilde{w}_X(s) &= 1. \qquad (11c)\end{aligned}$$

The mean FPT and second moment of FPT can now be derived. The results are presented in a theorem–proof format.

Theorem 1: The first two moments of the FPT for a birth-death process with birth rates k_i and death rate γ are given by following expressions

$$\langle FPT \rangle = \sum_{l=0}^{X-1} \sum_{j=0}^{l} \gamma^{l-j} \left(\frac{k_0 k_1 ... k_j}{k_0 k_1 ... k_l} \right) \frac{l!}{j!} \frac{1}{k_j}. \tag{12}$$

$$\left\langle FPT^2 \right\rangle = 2 \sum_{l=0}^{X-1} \sum_{j=0}^{l} \gamma^{l-j} \left(\frac{k_0 k_1 ... k_j}{k_0 k_1 ... k_l} \right) \frac{l!}{j!} \frac{\tau_j}{k_j}, \tag{13a}$$

where

$$\tau_i = \sum_{l=i}^{X-1} \sum_{j=0}^{l} \gamma^{l-j} \left(\frac{k_0 k_1 ... k_j}{k_0 k_1 ... k_l} \right) \frac{l!}{j!} \frac{1}{k_j}. \tag{13b}$$

Proof: The proof is divided in two parts. The first part deals with derivation of expression for mean FPT while the second one with that of the second order moment of FPT.

(1) Calculation of mean FPT

Let us denote the mean exit time starting the site i as $\tau_i := \int_0^\infty t w_i(t)\, dt = \tilde{w}_i'(s=0)$. Note that $\langle FPT \rangle = \tau_0$. Using (11a)–(11c), one may write

$$\begin{aligned}k_0(\tau_1 - \tau_0) &= -1, \qquad (14a)\\ k_i(\tau_{i+1} - \tau_i) - i\gamma(\tau_i - \tau_{i-1}) &= -1, \qquad (14b)\\ \tau_X &= 0. \qquad (14c)\end{aligned}$$

Let $a_i := \tau_{i+1} - \tau_i$. Note that (14a) is just (14b) evaluated at $i = 0$. Thus, (14a)–(14c) transform to

$$\begin{aligned}a_0 &= -\frac{1}{k_0}, \qquad (15a)\\ k_i a_i - i\gamma a_{i-1} &= -1, \quad 1 \leq i \leq X-1, \qquad (15b)\\ a_{X-1} &= \tau_X - \tau_{X-1} = -\tau_{X-1}. \qquad (15c)\end{aligned}$$

Let's solve (15b). Multiplying both sides by f_i (to be determined)

$$f_i a_i - \frac{\gamma}{k_i} i f_i a_{i-1} = -\frac{f_i}{k_i}. \tag{16}$$

Let $r_i = \frac{\gamma}{k_i}$. Choosing f_i so that

$$f_{i-1} = \frac{\gamma}{k_i} i f_i = r_i i f_i \implies f_i = \frac{C}{r_i r_{i-1} ... r_1 r_0 i!}. \tag{17a}$$

Substituting this expression of f_i in (16) yields

$$f_i a_i - f_{i-1} a_{i-1} = -\frac{f_i}{k_i},\ 1 \le i \le X-1, \tag{18a}$$

$$\implies f_i a_i = -\sum_{j=0}^{i} \frac{f_j}{k_j} + C'', \tag{18b}$$

$$\implies a_i = -\frac{1}{f_i}\sum_{j=0}^{i} \frac{f_j}{k_j} + \frac{C''}{f_i}, \tag{18c}$$

$$= -\sum_{j=0}^{i} \frac{r_0 r_1 \ldots r_i}{r_0 r_1 \ldots r_j} \frac{i!}{j!}\frac{1}{k_j} + \frac{C''}{C} r_0 r_1 \ldots r_i i!. \tag{18d}$$

Since $a_0 = -1/k_0 \implies C''/C = 0$. Therefore

$$\tau_{i+1} - \tau_i = -\sum_{j=0}^{i} \frac{r_0 r_1 \ldots r_i}{r_0 r_1 \ldots r_j} \frac{i!}{j!}, \tag{19a}$$

$$\implies \tau_i = -\sum_{l=0}^{i-1}\sum_{j=0}^{l} \frac{r_0 r_1 \ldots r_i}{r_0 r_1 \ldots r_j} \frac{i!}{j!}\frac{1}{k_j} + C''', \tag{19b}$$

$$\implies \tau_0 = C'''. \tag{19c}$$

C''' can be determined by $-a_{X-1} = \tau_{X-1}$ as follows

$$\tau_{X-1} = -\sum_{l=0}^{X-2}\sum_{j=0}^{l} \frac{r_0 r_1 \ldots r_i}{r_0 r_1 \ldots r_j} \frac{i!}{j!}\frac{1}{k_j} + C''' \tag{20a}$$

$$= -a_{X-1} = -\sum_{j=0}^{X-2} \frac{r_0 r_1 \ldots r_{X-1}}{r_0 r_1 \ldots r_j} \frac{(X-1)!}{j!}\frac{1}{k_j}. \tag{20b}$$

Let $\beta(l) := \sum_{j=0}^{l} \frac{r_0 r_1 \ldots r_i}{r_0 r_1 \ldots r_j} \frac{l!}{j!}\frac{1}{k_j}$. This implies

$$C''' = \beta(X-1) + \sum_{l=0}^{X-2} \beta(l) = \sum_{l=0}^{X-1} \beta(l), \tag{21a}$$

$$\implies \tau_0 = \sum_{l=0}^{X-1}\sum_{j=0}^{l} \frac{r_0 r_1 \ldots r_l}{r_0 r_1 \ldots r_j} \frac{l!}{j!}\frac{1}{k_j} = \sum_{l=0}^{X-1}\sum_{j=0}^{l} \gamma^{l-j} \frac{k_0 k_1 \ldots k_j}{k_0 k_1 \ldots k_l} \frac{l!}{j!}\frac{1}{k_j}, \tag{21b}$$

and

$$\tau_i = \sum_{l=0}^{X-1} \beta(l) - \sum_{l=0}^{i-1} \beta(l) = \sum_{l=i}^{X-1} \beta(l) \tag{22a}$$

$$= \sum_{l=i}^{X-1}\sum_{j=0}^{l} \frac{r_0 r_1 \ldots r_l}{r_0 r_1 \ldots r_j} \frac{l!}{j!}\frac{1}{k_j} = \sum_{l=i}^{X-1}\sum_{j=0}^{l} \gamma^{l-j} \frac{k_0 k_1 \ldots k_j}{k_0 k_1 \ldots k_l} \frac{l!}{j!}\frac{1}{k_j}. \tag{22b}$$

(2) Calculation of second order moment of FPT

Let us denote the second order moments of FPT distribution for a particle starting at site i by $\eta_i := \int_0^\infty t^2 w_i(t)\,dt = \tilde{w}_i''(s=0)$. Note that $\langle FPT^2 \rangle = \eta_0$. Using (11a)–(11c)

$$k_0(\eta_1 - \eta_0) = -2\tau_0, \tag{23a}$$

$$k_i \eta_{i+1} + i\gamma\eta_{i-1} - (i\gamma + k_i)\eta_i = -2\tau_i, \tag{23b}$$

$$\eta_X = 0. \tag{23c}$$

Defining $A_i := \eta_{i+1} - \eta_i$, we have

$$k_0 A_0 = -2\tau_0, \tag{24a}$$

$$k_i A_i - i\gamma A_{i-1} = -2\tau_i, \tag{24b}$$

$$A_{X-1} = -\eta_{X-1}. \tag{24c}$$

Multiplying both sides of (24b) by the integrating factor f_i defined in (17a) and then dividing by k_i gives

$$f_i A_i - f_{i-1} A_{i-1} = -\frac{2f_i}{k_i}\tau_i, \tag{25a}$$

$$\implies A_i = -\frac{2}{f_i}\sum_{j=0}^{i} \frac{f_j}{k_j}\tau_j + \frac{D}{f_i}. \tag{25b}$$

Since $A_0 = -2\tau_0/k_0$, we have $D = 0$. Expression of A_i now becomes

$$A_i = \eta_{i+1} - \eta_i = -2\sum_{j=0}^{i} \frac{f_j}{f_i}\frac{\tau_j}{k_j}, \tag{26a}$$

$$\implies \eta_i = \sum_{k=0}^{i-1} A_k + D' \implies \eta_0 = D'. \tag{26b}$$

Using $\eta_{X-1} = -A_{X-1}$

$$\sum_{l=0}^{X-2} A_l + D' = -A_{X-1} \implies \sum_{l=0}^{X-1} A_l = -D', \tag{27a}$$

$$\implies \eta_0 = -\sum_{l=0}^{X-1} A_l = 2\sum_{l=0}^{X-1}\sum_{j=0}^{l} \frac{f_j}{f_l}\frac{\tau_j}{k_j} \tag{27b}$$

$$= 2\sum_{l=0}^{X-1}\sum_{j=0}^{l} \gamma^{l-j} \frac{k_0 k_1 \ldots k_j}{k_0 k_1 \ldots k_l} \frac{l!}{j!}\frac{\tau_j}{k_j}. \tag{27c}$$

This concludes the proof. ∎

So far we have developed the formulas for the moments of FPT for a birth-death process. In the following section, we discuss the optimal birth rates such that the variability in the FPT around a fixed time is minimized.

IV. Optimal birth rates for a birth-death process

Our aim is to investigate what form of auto–regulation in gene expression would lead to minimum variability in the FPT around a given value, say t^*. We hypothesize following two possible objective functions for that purpose

1) Minimizing variance of FPT subject to $\langle FPT \rangle = t^*$.
2) Minimizing the expected mean squared error of FPT from the point t^*, i.e., $\left\langle (FPT - t^*)^2 \right\rangle$.

To incorporate both these hypotheses, we consider the following objective function (denoted by ϕ)

$$\phi := \mathrm{Var}(FPT) + \frac{1}{\varepsilon}\left(t^* - \langle FPT \rangle\right)^2, \tag{28}$$

where $\varepsilon > 0$ is a tuning parameter. Eq. (28) covers a wide class of cost functions involving the variance of FPT and penalizes the distance of mean FPT from the point t^*. Observe that when $\varepsilon = 1$, (28) is the objective function for 2). However, it is not immediately obvious how the first hypothesis of minimizing variance of FPT for fixed mean can be covered by such a choice of ϕ. We will revisit this issue while discussing the optimal birth rates when $\gamma = 0$.

A. Optimal birth rates when $\gamma = 0$

When $\gamma = 0$, the process reduces to a pure birth process. Using Theorem 1, we can write

$$\langle FPT \rangle = \sum_{i=0}^{X-1} \frac{1}{k_i}, \tag{29a}$$

$$\mathrm{Var}(FPT) = \langle FPT^2 \rangle - \langle FPT \rangle^2 = \sum_{i=0}^{X-1} \frac{1}{k_i^2}. \tag{29b}$$

Thus, the objective function ϕ in (28) can be written as

$$\phi(k_0, k_1, ..., k_{X-1}) = \sum_{i=0}^{X-1} \frac{1}{k_i^2} + \frac{1}{\varepsilon} \left(t^* - \sum_{i=0}^{X-1} \frac{1}{k_i} \right)^2. \tag{30}$$

Note that for $0 \le i \le X-1$

$$\frac{\partial \phi}{\partial k_i} = -\frac{2}{k_i^3} + \frac{2}{\varepsilon} \left(t^* - \sum_{j=0}^{X-1} \frac{1}{k_j} \right) \frac{1}{k_i^2}. \tag{31}$$

To find the minimum of the function ϕ, we solve the system of following X equations for k_i's

$$\frac{\partial \phi}{\partial k_i} = 0, \;\; 0 \le i \le X-1. \tag{32a}$$

We restrict ourselves to $k_i > 0$ since k_i's represent birth rates. The system of equations to be solved now reduces to

$$-\frac{1}{k_i} + \frac{1}{\varepsilon} \left(t^* - \sum_{j=0}^{X-1} \frac{1}{k_j} \right) = 0, \;\; 0 \le i \le X-1. \tag{32b}$$

The solution is given by the following:

$$k_i = \frac{X + \varepsilon}{t^*}, \quad 0 \le i \le X-1 \tag{33}$$

Further, to ascertain that above solution is indeed a minimum, we have checked that the Hessian matrix consisting of second order derivatives of the objective function ϕ is positive definite. The calculations have been omitted due to space constraints.

This calculation implies that when the death rate is zero, the optimal birth rates have to be equal, given by (33). The parameter ε changes the value of the optimal birth rates. In particular, for $\varepsilon \to 0$, we have $k_i = X/t^*$. This solution is same as what we get by minimizing the variance of FPT subject to $\langle FPT \rangle = t^*$ using the Lagrange multiplier method, which substantiates our choice of the objective function in (28). Also, when $\varepsilon = 1$, the optimal birth rates are $k_i = (X+1)/t^*$. Next, we introduce a death rate in the process and investigate the optimal birth rates that minimize the objective function under consideration.

Remark: Note that if one wants to minimize the function $(t^* - \langle FPT \rangle)^2$, the solution is ill-posed as there are infinitely many k_i that make the objective function equal to zero. However, the problem can be regularized by minimizing $(t^* - \langle FPT \rangle)^2 + \varepsilon \mathrm{Var}(FPT)$, $\varepsilon \to 0$ instead. In this case, we are choosing the unique set of k_i's which also minimize the variance of FPT. This is another reason behind choosing the objective function ϕ described in (28).

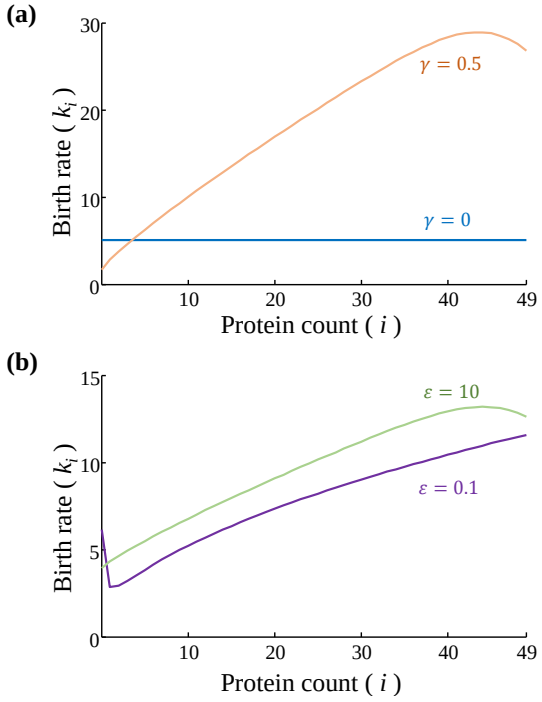

Fig. 3. Optimal birth rates for a birth-death process. The birth rates are assumed to follow a Hill function mentioned in (34). Part (a) shows optimal birth rates when degradation rate is allowed to change while keeping the other parameters constant ($X = 50, t^* = 10, \varepsilon = 1$). In part (b), optimal birth rates are plotted when the tuning parameter ε is changed (other parameters: $X = 50, t^* = 10, \gamma = 0.1$).

B. Optimal birth rates when $\gamma > 0$

We have seen that the optimal birth rates for a birth-death process with zero death rate are equal. However, this is not true anymore once the death term is introduced. For example, if we take a simple case of $X = 3, \gamma = 1, \varepsilon = 1$ and solve for the optimal birth rates, the answer is $k_0 = 4.0097, k_1 = 8.5745, k_2 = 4.6059$.

In Fig. 3, we present the results for $X = 50$. In order to reduce the number of parameters to be optimized and keep the results biologically relevant, we confined the optimization space to a mixture of negative and positive feedbacks implemented using the Hill function. Particularly, we used the following form for birth rate at protein count i

$$k_i = \frac{k_{\max}(1 + (ic_1)^{H_1})}{a + (ic_2)^{H_2}}, \quad H_1 \le H_2, \tag{34}$$

and optimized over the parameters $k_{\max}, c_1, H_1, a, c_2, H_2$.

In 3(a), we have allowed the death rate γ to change while keeping other parameters as $t^* = 10, \varepsilon = 1$. We see that when $\gamma = 0$, the optimal birth rates are equal. In fact, they are equal to $X/t^* = 5$, validating the result in (33). Further, when $\gamma = 0.5$, the optimal birth rates are no longer equal. Rather, they first increase and then decrease, indicating auto–regulation involving both positive and negative feedbacks.

In part Fig. 3(b), we have changed the tuning parameter ε while other parameters are kept as $t^* = 10, \gamma = 0.1$. It can be seen that in both cases, the optimal birth rates represent a mix

feedback mechanism involving both positive and negative feedbacks.

V. Discussion

The inherent stochastic nature of gene expression is a driving force behind randomness in how a protein's level evolves with time. Therefore, timing of cellular events, which typically depends upon levels of important proteins, is also stochastic. In this work, our objective is to understand how cells regulate event timing. More specifically, motivated by precision in the lysis time of a λ phage, we studied events whose stochasticity is minimized around a fixed time. We modeled event timing as a first–passage time problem. Also, we considered auto–regulation as a possible regulatory mechanism and studied what form of auto–regulation would lead to minimum stochasticity in FPT around a fixed time. Our results show that when protein does not degrade appreciably and the mean burst size is very small, the best way to ensure tight regulation of event timing would be to have no auto–regulation. However, when protein is unstable, some form of auto–regulation is required to optimize timing of events.

These results can be connected to phage λ lysis time. The existence of optimal lysis time suggests that there could be auto–regulation in expression of holin. While the proteins responsible for lysis (holin) are stable [32], the effective mean burst size is also small [9]. Thus, the best strategy for the phage would be to have no auto–regulation in the expression of holin. This interpretation is consistent with experimental observations [33], [34]. It should be noted that there could be other types of regulation schemes that can minimize stochasticity in FPT around a fixed time. For instance, when a gene expression is considered in bursts with no protein decay, the stochasticity decreases as the mean burst size is reduced [9], [22]. In fact, analysis in [9] suggests that the holin-antiholin system of λ phage could be a mechanism to regulate the effective mean burst size and subsequently the stochasticity.

ACKNOWLEDGMENT

AS is supported by the National Science Foundation Grant DMS-1312926, University of Delaware Research Foundation (UDRF) and Oak Ridge Associated Universities (ORAU).

References

[1] J. M. Raser and E. K. O'Shea, "Noise in gene expression: origins, consequences, and control," *Science*, vol. 309, pp. 2010–2013, 2005.

[2] A. Raj and A. van Oudenaarden, "Nature, nurture, or chance: stochastic gene expression and its consequences," *Cell*, vol. 135, pp. 216–226, 2008.

[3] A. Singh and M. Soltani, "Quantifying intrinsic and extrinsic variability in stochastic gene expression models," *PloS One*, vol. 8, p. e84301, 2013.

[4] H. H. McAdams and A. Arkin, "Stochastic mechanisms in gene expression," *Proceedings of the National Academy of Sciences*, vol. 94, pp. 814–819, 1997.

[5] A. Amir, O. Kobiler, A. Rokney, A. B. Oppenheim, and J. Stavans, "Noise in timing and precision of gene activities in a genetic cascade," *Molecular Systems Biology*, vol. 3, 2007.

[6] J. M. Pedraza and J. Paulsson, "Random timing in signaling cascades," *Molecular Systems Biology*, vol. 3, 2007.

[7] J. J. Dennehy and N. Wang, "Factors influencing lysis time stochasticity in bacteriophage λ," *BMC Microbiology*, vol. 11, p. 174, 2011.

[8] E. Yurkovsky and I. Nachman, "Event timing at the single–cell level," *Briefings in Functional Genomics*, vol. 12, pp. 90–98, 2013.

[9] A. Singh and J. J. Dennehy, "Stochastic holin expression can account for lysis time variation in the bacteriophage λ," *Journal of The Royal Society Interface*, vol. 11, p. 20140140, 2014.

[10] U. Alon, "Network motifs: theory and experimental approaches," *Nature Reviews Genetics*, vol. 8, pp. 450–461, 2007.

[11] Y. Tao, X. Zheng, and Y. Sun, "Effect of feedback regulation on stochastic gene expression," *Journal of Theoretical Biology*, vol. 247, pp. 827–836, 2007.

[12] A. Singh and J. P. Hespanha, "Optimal feedback strength for noise suppression in autoregulatory gene networks," *Biophysical Journal*, vol. 96, pp. 4013–4023, 2009.

[13] A. Singh, "Negative feedback through mrna provides the best control of gene-expression noise," *IEEE Transactions on NanoBioscience*, vol. 10, pp. 194–200, 2011.

[14] A. Singh and J. P. Hespanha, "Evolution of gene auto-regulation in the presence of noise," *Systems Biology, IET*, vol. 3, pp. 368–378, 2009.

[15] R. White, S. Chiba, T. Pang, J. S. Dewey, C. G. Savva, A. Holzenburg, K. Pogliano, and R. Young, "Holin triggering in real time," *Proceedings of the National Academy of Sciences*, vol. 108, pp. 798–803, 2011.

[16] I.-N. Wang, D. E. Dykhuizen, and L. B. Slobodkin, "The evolution of phage lysis timing," *Evolutionary Ecology*, vol. 10, pp. 545–558, 1996.

[17] I.-N. Wang, "Lysis timing and bacteriophage fitness," *Genetics*, vol. 172, pp. 17–26, 2006.

[18] R. H. Heineman and J. J. Bull, "Testing optimality with experimental evolution: lysis time in a bacteriophage," *Evolution*, vol. 61, pp. 1695–1709, 2007.

[19] Y. Shao and I.-N. Wang, "Bacteriophage adsorption rate and optimal lysis time," *Genetics*, vol. 180, pp. 471–482, 2008.

[20] J. A. Bonachela and S. A. Levin, "Evolutionary comparison between viral lysis rate and latent period," *Journal of Theoretical Biology*, vol. 345, pp. 32–42, 2014.

[21] A. Becskei and L. Serrano, "Engineering stability in gene networks by autoregulation," *Nature*, vol. 405, pp. 590–593, 2000.

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

[23] J. Paulsson, "Models of stochastic gene expression," *Physics of Life Reviews*, vol. 2, pp. 157–175, 2005.

[24] N. Friedman, L. Cai, and X. S. Xie, "Linking stochastic dynamics to population distribution: an analytical framework of gene expression," *Physical Review Letters*, vol. 97, p. 168302, 2006.

[25] V. Shahrezaei and P. S. Swain, "Analytical distributions for stochastic gene expression," *Proceedings of the National Academy of Sciences*, vol. 105, pp. 17 256–17 261, 2008.

[26] O. G. Berg, "A model for the statistical fluctuations of protein numbers in a microbial population," *Journal of Theoretical Biology*, vol. 71, pp. 587–603, 1978.

[27] J. Yu, J. Xiao, X. Ren, K. Lao, and X. S. Xie, "Probing gene expression in live cells, one protein molecule at a time," *Science*, vol. 311, pp. 1600–1603, 2006.

[28] V. Elgart, T. Jia, A. T. Fenley, and R. Kulkarni, "Connecting protein and mrna burst distributions for stochastic models of gene expression," *Physical biology*, vol. 8, p. 046001, 2011.

[29] O. Jouini and Y. Dallery, "Moments of first passage times in general birth–death processes," *Mathematical Methods of Operations Research*, vol. 68, pp. 49–76, 2008.

[30] L. J. Allen, *An Introduction to Stochastic Processes with Applications to Biology*. Prentice Hall, 2003.

[31] C. Doering, K. Sargsyan, and L. Sander, "Extinction times for birth-death processes: Exact results, continuum asymptotics, and the failure of the fokker–planck approximation," *Multiscale Modeling & Simulation*, vol. 3, pp. 283–299, 2005.

[32] Y. Shao and N. Wang, "Effect of late promoter activity on bacteriophage λ fitness," *Genetics*, vol. 181, pp. 1467–1475, 2009.

[33] I.-N. Wang, D. L. Smith, and R. Young, "Holins: the protein clocks of bacteriophage infections," *Annual Reviews in Microbiology*, vol. 54, pp. 799–825, 2000.

[34] C.-Y. Chang, K. Nam, and R. Young, "S gene expression and the timing of lysis by bacteriophage lambda." *Journal of bacteriology*, vol. 177, pp. 3283–3294, 1995.