Channel Capacity Computations for Unregulated and Autoregulated Gene Expression

Zahra Vahdat¹, Karol Nienaltowski², Zia Farooq², Michał Komorowski² and Abhyudai Singh³

Abstract—How living cells can reliably process biochemical cues in the presence of molecular noise is not fully understood. Here we investigate the fidelity of information transfer in the expression of a single gene. We use the established model of gene expression to examine how precisely the protein levels can be controlled by two distinct mechanisms: (i) the transcription rate of the gene, or (ii) the translation rate for the corresponding mRNA. The fidelity of gene expression is quantified with the information-theoretic notion of information capacity. Derived information capacity formulae reveal that transcriptional control generally provides a tangibly higher capacity as compared to the translational control. We next introduce negative feedback regulation in gene expression, where the protein directly inhibits its own transcription. While negative feedback reduces noise in the level of the protein for a given input signal, it also decreases the input-to-output sensitivity. Our results show that the combined effect of these two opposing forces is a reduced capacity in the presence of feedback. In summary, our analysis presents the first analytical quantification of information transfer in simple gene expression models, which provides insight into the fidelity of basic gene expression control mechanisms.

I. INTRODUCTION

Within the noisy environment of a living cell, biochemical species such as genes, RNAs, and proteins often occur at low molecular counts. These low copy numbers, coupled with the inherent random nature of biochemical processes, drive considerable stochastic fluctuations in copy numbers over time. Advances in single-cell technologies over the last decade have precisely quantified fluctuations in the level of a protein within single cells in diverse organisms [1]–[6]. There is also increasing appreciation of the beneficial and harmful roles played by these fluctuations with tangible consequences for biology and medicine [7]–[13].

As in engineering systems, cells use regulatory mechanisms to maintain the level of a protein around a set point. The simplest example of this is negative autoregulation, where a protein expressed from its gene inhibits its own transcription creating a feedback loop [14]–[23]. Such feedbacks have been shown to be ubiquitous across genes [24], and design of synthetic feedbacks for controlling gene expression is an intense area of current research [25]–[27]. In this contribution, we use information theory to study precision

Fig. 1. Schematic of the process of gene expression where the input signal may control either transcription or translation. Further, protein may inhibit its own transcription, which constitutes the negative feedback. The protein level is considered as the output of the system. We examine how precisely the input signal may control the output through transcription or translation, and with or without feedback.

with which different mechanisms can control gene expression levels. Specifically, we deploy the concept of information capacity that has recently been used to study information processing in diverse living systems [28]–[32]. Information capacity is expressed in bits, and an input-output system with the capacity of c bits can reliably discriminate 2^c input states from the output. Here, the main focus is to quantify the information capacity of the gene expression process for the input considered to be at different locations of the gene expression process, and also when the expression operates in open loop vs. feedback loop.

We modify a well-studied model of stochastic protein synthesis from its corresponding gene with the inclusion of an input signal, where the input either enhances the transcription rate of the gene, or the translation rate for the corresponding mRNA (Fig. 1). Analytical formulas for the channel capacity are derived as a function of the fold change in the input from its basal to its maximum value. While it is not surprising that the channel capacity monotonically increases with the input fold change, the location of the input leads to contrasting results. More specifically, input-driven transcription generally provides a much higher capacity as compared to input-driven translation, but we also identify pa-

¹Z. Vahdat is with the Department of Electrical and Computer Engineering, University of Delaware, Newark, DE USA 19716. zahravd@udel.edu

²K. Nienaltowski and Z. Farooq and M. Komorowski are with the Institute of Fundamental Technological Research, Polish Academy of Sciences.

³A. Singh is with the Department of Electrical and Computer Engineering, Biomedical Engineering, Mathematical Sciences, Center for Bioinformatics and Computational Biology, University of Delaware, Newark, DE USA 19716. absingh@udel.edu

parameter regimes where the opposite is true. We also consider the case of autoregulated gene expression where the protein inhibits its own transcription. Interestingly, the inclusion of a negative feedback reduces the capacity irrespective of where the input affects gene expression.

The paper is organized as follows. In Section II we formulate a stochastic model of gene expression without feedback, and quantify fluctuations in the output protein level for a given input signal. In Section III we derive formulas for the channel capacity and systematically compare scenarios where the input affects transcription vs. translation. Negative feedback is introduced in Section IV with capacity comparisons to the no feedback case. Finally, conclusions are presented in Section V.

II. STOCHASTIC MODELS OF GENE EXPRESSION

Motivated by experimental data on gene activity in individual cells, the simplest mechanistic model of gene expression considers a protein being synthesized in stochastic bursts [33]–[42]. The burst events are assumed to arrive as per a Poisson process with rate (burst frequency) a . Each burst event leads to a several proteins being made instantaneously. Let $\mathbf{x}(t)$ denote the concentration of a given protein within a cell at time t . The burst size, i.e., the size of the jump in $\mathbf{x}(t)$, is an independent and identically distributed random variable following an exponential distribution with mean burst size b [43], [44]. Assuming that the protein has a long half-life with no active degradation, then in between two successive burst events the protein level is diluted as per

$$\frac{d\mathbf{x}}{dt} = -\gamma\mathbf{x}, \quad (1)$$

where γ is the cell growth rate. In essence, time evolution of $\mathbf{x}(t)$ is a Piecewise-Deterministic Markov Process (PDMP) with continuous exponential decay, interspersed with random burst events that increase protein levels at discrete times (as determined by the Poisson arrival process). Physiologically, the burst frequency a corresponding to the rate at which a gene is transcribed to produce mRNAs. Since mRNAs have short half-lives compared to the protein, each mRNA transcript quickly degrades after producing a burst of proteins [45], [46]. In this context, the burst size is the number of proteins made in the lifetime of an individual mRNA, and the average burst size b is directly related to the rate at which the mRNA is translated.

Having defined a PDMP model for bursty gene expression, the probability density function of the protein level at time t , $f(x, t)$, evolves as per the following continuous master equation

$$\frac{\partial f(x, t)}{\partial t} - \gamma \frac{\partial x f(x, t)}{\partial x} + \frac{\partial}{\partial x} \int_0^x a e^{-(x-y)/b} f(y, t) dy = 0 \quad (2)$$

[47]–[49]. Prior work has elegantly shown that solving this master equation at equilibrium ($\partial f(x, t)/\partial t = 0$), yields the steady-state distribution for the protein level

$$f(x; a, b) = \frac{1}{b^\gamma \Gamma(\frac{a}{\gamma})} x^{\frac{a}{\gamma}-1} e^{-x/b} \quad (3)$$

to be a gamma distribution, where Γ is the gamma function [47]. In the remainder of the paper, we set $\gamma = 1$ and a can be considered as the burst frequency normalized by the cell growth rate. The steady-state statistical moments of the protein level are

$$E\{x\} = ab, \quad E\{x^2\} = a(1+a)b^2, \quad (4)$$

where E denotes the expected value operation evaluated at steady-state. Interestingly, experimental measurements have found several proteins in *E. coli* and budding yeast to show statistical fluctuations consistent with a gamma distribution [43], [44].

In order to compute the channel capacity, we next consider an input signal z that activates gene expression in two different ways:

- 1) The input increases the transcription rate of the gene, and this corresponds to a monotonically increasing burst frequency $a(z)$ with respect to the input.
- 2) The input increases the translation rate of the mRNA, and this corresponds to a monotonically increasing average burst size $b(z)$ with respect to the input.

For simplicity, we assume a Michaelis–Menten form for these functions

$$b(z) = k_b + \frac{(k_{max} - k_b)z}{k + z}, \quad k_{max} > k_b > 0 \quad (5)$$

where k_b is the basal rate when there is no input ($z = 0$), k_{max} is the maximum rate ($z \rightarrow \infty$). The constant k represents the value of the input when the burst frequency reaches the halfway point, i.e., $b(z) = (k_{max} + k_b)/2$ when $z = k$, and thus k is inversely related to how rapidly $b(z)$ increases with increasing z to reach k_{max} . An identical functional form is assumed for $a(z)$. Having quantified statistical fluctuations in the output protein level as a function of the input signal, we next focus our attention on the channel capacity.

III. CHANNEL CAPACITY FOR UNREGULATED GENE EXPRESSION

Channel capacity has been proposed initially to quantify the information transfer over a noisy communication channel [50]. The transfer was assumed to take place between the input, i.e., sender's end, and the output, i.e., the receiver's end. Due to noise, the input cannot be reconstructed from the output without error. Information capacity quantifies the overall fidelity of the transmission. Broadly speaking, it quantifies the $\log_2$ of the maximal number of different messages that can be sent over the channel, on average, per transmission. Here, we use the capacity to quantify the fidelity of gene expression control. We assume that an input signal impacts the output protein level via the burst size or the burst frequency. The capacity between the input and output quantities, therefore, is the fidelity of the control mechanism in terms of the overall number of distinct input values that can be discriminated from the output. The exact computation of the information capacity is usually problematic. Therefore, we use the approximation given by the asymptotic capacity. The asymptotic capacity calculates

the information capacity exactly for a large number of statistically independent copies of a given signaling system and then regresses back the capacity of a single copy [51].

Asymptotic capacity C_A is computed from the Fisher information matrix (FIM). For illustration, let's take an arbitrary probability distribution f with 2 parameters ($\theta_i, i = 1, 2$). Then the FIM is a 2×2 matrix with elements

$$[I(\theta_1, \theta_2)]_{i,j} = E \left\{ \frac{\partial \ln f(x; \theta_1, \theta_2)}{\partial \theta_i} \frac{\partial \ln f(x; \theta_1, \theta_2)}{\partial \theta_j} \right\}, \quad (6)$$

where E stands for expected value. Considering parameters θ_i depend on input signal z , $z \in [0, z_{max}]$, then the Fisher information for signal z is

$$F(z) = \begin{bmatrix} \frac{d\theta_1(z)}{dz} & \frac{d\theta_2(z)}{dz} \end{bmatrix} I(\theta_1(z), \theta_2(z)) \begin{bmatrix} \frac{d\theta_1(z)}{dz} \\ \frac{d\theta_2(z)}{dz} \end{bmatrix}, \quad (7)$$

and the general formulation for asymptotic channel capacity can be computed from the Fisher information using

$$C_A = \log_2 \left(\frac{1}{\sqrt{2\pi e}} \int_0^{z_{max}} \sqrt{|F(z)|} dz \right) \quad (8)$$

[51]. For the gamma distribution derived in (3), the Fisher information matrix is given by

$$I(a, b) = \begin{bmatrix} \psi_1(a) & \frac{1}{b} \\ \frac{1}{b} & \frac{a}{b^2} \end{bmatrix}, \quad (9)$$

where $\psi_1(\cdot)$ is Trigamma function (the logarithmic second derivative of gamma function),

$$\psi_1(a) = \frac{d^2 \ln \Gamma(a)}{da^2}. \quad (10)$$

Next we study specific cases when either a or b is modulated as a Michaelis–Menten function of the input z .

A. Input activates gene expression via the burst size

In this case the mean burst size varies with z as

$$b(z) = k_b + \frac{(k_{max} - k_b)z}{k + z} \quad (11)$$

and the burst frequency a is constant. Using (7) and (9) we calculate the Fisher information as

$$F(z) = \left(\frac{db(z)}{dz} \right)^2 \frac{a}{b(z)^2}, \quad (12)$$

that simplifies to

$$F(z) = \frac{ak^2(k_{max} - k_b)^2}{(k + z)^2(kk_b + k_{max}z)^2}. \quad (13)$$

Finally, we calculate the asymptotic channel capacity from (8)

$$C_A = \log_2 \left(\sqrt{\frac{a}{2\pi e}} \ln \left(\frac{kk_b + k_{max}z_{max}}{k_b(k + z_{max})} \right) \right), \quad (14)$$

which in the limit ($z_{max} \rightarrow \infty$) reduces to

$$C_A = \log_2 \left(\sqrt{\frac{a}{2\pi e}} \ln \frac{k_{max}}{k_b} \right). \quad (15)$$

Fig. 2. Channel capacity for unregulated stochastic gene expression when the input activates transcription via the burst frequency, or activates translation via the burst size. Channel capacities derived in (15) and (21) are plotted as function of k_{max} for different value of a assuming $k_b = 1$. The input affecting transcription generally provides a higher capacity. However, our results show that when a is large and k_{max} is small, then input-driven translation can provide a higher capacity.

B. Input activates gene expression via the burst frequency

In this case the burst frequency varies with z as

$$a(z) = k_b + \frac{(k_{max} - k_b)z}{k + z}, \quad (16)$$

and the mean burst size b is constant. Using (7), (9) and (16) yields the following Fisher information

$$F(z) = \left(\frac{da(z)}{dz} \right)^2 \psi_1(a(z)), \quad (17)$$

$$F(z) = \frac{k^2(k_{max} - k_b)^2}{(k + z)^4} \psi_1 \left(k_b + \frac{(k_{max} - k_b)z}{k + z} \right), \quad (18)$$

with the asymptotic channel capacity from (8) as

$$C_A = \log_2 \left(\frac{1}{\sqrt{2\pi e}} \int_0^{z_{max}} \frac{k(k_{max} - k_b)}{(k + z)^2} \times \sqrt{\psi_1 \left(k_b + \frac{(k_{max} - k_b)z}{k + z} \right)} dz \right). \quad (19)$$

The integration within (19) can be further simplified by doing a change of variable

$$y = k_b + \frac{(k_{max} - k_b)z}{k + z} \quad (20)$$

which in the limit $z_{max} \rightarrow \infty$ simplifies (19) to

$$C_A = \log_2 \left(\frac{1}{\sqrt{2\pi e}} \int_{k_b}^{k_{max}} \sqrt{\psi_1(y)} dy \right) \quad (21)$$

with ψ_1 being the Trigamma function defined in (10).

The channel capacities derived in (15) and (21) provide some intriguing insights:

- The capacities only depends on k_{max} and k_b , the maximum and minimum value of the input, respectively, and are invariant of the parameter k in (16) and (11).

$$F(z) = \frac{a_0 k^2 (k_{max} - k_b)^2}{(k+z)^2 (k k_b + k_{max} z)^3} (K(k+z) - \frac{a_0 (k k_b + k_{max} z) U^2(a_0, 0, \frac{K(k+z)}{k k_b + k_{max} z})}{U^2(a_0, 1, \frac{K(k+z)}{k k_b + k_{max} z})} + \frac{(k k_b + k_{max} z - K(k+z)) U(a_0, 0, \frac{K(k+z)}{k k_b + k_{max} z})}{U(a_0, 1, \frac{K(k+z)}{k k_b + k_{max} z})}). \quad (28)$$

- When the input drives the burst size b , then the capacity (15) monotonically increases with the burst frequency as in $\log_2 \sqrt{a}$.
- When the input drives the burst frequency a , then the capacity (19) becomes independent of the average burst size b , and hence, independent of the expression specific parameters.
- As illustrated in Fig. 2, for sufficiently small a the input in burst frequency (i.e., transcription) has higher capacity than input in burst size (i.e., translation) irrespective of k_{max} and k_b .
- For sufficiently large a , the input in burst size can provide higher capacity provided the ratio k_{max}/k_b is small enough (Fig. 2).

IV. CHANNEL CAPACITY FOR AUTOREGULATED GENE EXPRESSION

Negative feedback is incorporated in the gene expression process by modifying the burst frequency to be a monotonically decreasing function of the protein level x

$$a(x) = \frac{a_0}{1 + \frac{x}{K}}, \quad (22)$$

with two new positive parameters a_0 and K . Here a_0 is the burst frequency when there is no feedback ($K \rightarrow \infty$), and $1/K$ can be intuitively thought of as the feedback strength, i.e., lower values of K correspond to stronger feedback. In this case, the steady-state distribution of the protein concentration is given by

$$f(x; a_0, b, K) = c_0 e^{-x/b} x^{a_0-1} (1 + \frac{x}{K})^{-a_0}, \quad (23)$$

[52], where c_0 is the normalization constant

$$c_0 = \frac{1}{\int_0^\infty e^{-x'/b} x'^{a_0-1} (1 + \frac{x'}{K})^{-a_0} dx'} = \frac{K^{-a_0}}{\Gamma(a_0) U(a_0, 1, \frac{K}{b})}, \quad (24)$$

and U is the confluent hypergeometric function. This distribution yields the following steady-state first and second order moments

$$E\{x\} = \frac{abU(a_0, 0, \frac{K}{b})}{U(a_0, 1, \frac{K}{b})}, \quad (25)$$

$$E\{x^2\} = \frac{a_0(a_0+1)b^2U(a_0, -1, \frac{K}{b})}{U(a_0, 1, \frac{K}{b})}. \quad (26)$$

As before, we now consider two different cases of input-dependent expression, but now with negative feedback.

Fig. 3. Channel capacity for negative feedback regulation of burst frequency, and the input affects the mean burst size. The asymptotic channel capacity (C_A) is computed via (8) using (28) for parameters $a_0 = 90$, $k_b = 1$, $k = 1$. Decreasing values of K (i.e., increasing negative feedback strength) results in a reduced information capacity.

A. Input activates gene expression via the burst size

Considering the input activates gene expression by enhancing mRNA translation, we model the average burst size as a function of input z as per (11), while the burst frequency is given by (22). Following (6), (7), and (23) the Fisher information $F(z)$ is calculated in Appendix A and is given by (28). The channel capacity (C_A) is computed via (8) using (28). Given the complexity of the formula we investigate it numerically. Our results show that decreasing values of K , which corresponds to increasing negative feedback strength, results in a reduced capacity (Fig. 3).

B. Input activates gene expression via the burst frequency

Next, we consider the input signal affecting the burst frequency. Since both the input and the feedback come at the same location, we modify the burst frequency to

$$a(x, z) = \frac{a_0(z)}{1 + \frac{x}{K}}, \quad a_0(z) = k_b + \frac{(k_{max} - k_b)z}{k + z}. \quad (29)$$

The mean burst size is now assumed to be a constant b . The Fisher information for input signal z is given by (30) (see Appendix B for details), where $U_{(1,0,0)}$ and $U_{(2,0,0)}$ are the first and second derivative of the confluent hypergeometric function with respect to $a_0(z)$, respectively. Calculating C_A from (8) again shows that adding negative feedback reduces the information capacity (Fig. 4).

V. CONCLUSION

We have systematically investigated fidelity of information transfer in a simple yet mechanistic model of stochastic

Fig. 4. Channel capacity for negative feedback regulation of burst frequency, and the input also affects the burst frequency. The asymptotic channel capacity (C_A) is computed via (8) using (30) for parameters $b = 5$, $k_b = 1$, $k = 1$. As in Fig. 3, decreasing values of K (i.e., increasing negative feedback strength) result in a reduced information capacity.

gene expression. In the absence of any feedback loops, our results provide simple analytical formulas connecting the information capacity to properties of the input signal (the maximum and minimum values of the input), and expression parameters. A key finding from this analysis is that transcriptional control of gene expression generally leads to a higher capacity as compared to translational control (Fig. 2). However, if the burst frequency is high enough, then translational control can provide a higher capacity for small fold changes in the input (Fig. 2).

These results were extended to consider feedback in burst frequency. In the feedback case, the formulas are quite involved and we have investigated them numerically. It is quite possible that formulas can be further simplified in limits of strong ($K \rightarrow 0$) or weak feedback ($K \rightarrow \infty$) and we will investigate this in our future work. Our results lead to an intriguing find that in spite of feedback loops suppressing noise in protein levels for a given input value [16], [53], they actually reduce the information capacities (Figs. 3 and 4). Intuitively, inclusion of feedback makes the output less sensitive to changes in the input, and the overall effect of noise buffering and reduced input-to-output sensitivity is a lower information capacity. While our analysis is restricted to a simple non-cooperative feedback, it will be interesting to see if more complex forms of feedback can lead to a higher capacity.

Another direction of future work is to consider feedbacks in the burst size and/or the protein decay rate. Our recent collaborative work has derived probability density functions for such control mechanisms [49], and we plan to investigate channel capacities across feedback architectures. Prior work has shown the negative feedback is much more efficient in suppressing external disturbances in protein production, rather than noise arising from the inherent bursty nature of expression [16]. We plan to expand models presented here to incorporate external disturbances by making parameters

itself random processes that biologically corresponds to fluctuations in the abundance of enzymes involved in gene expression. With multiple noise mechanism affecting the protein level, we will study information capacity both with and without feedback.

ACKNOWLEDGMENT

AS is supported by the National Science Foundation grant ECCS-1711548 and ARO grant W911NF-19-1-0243. ZF, KN, and MK were supported by the Foundation for Polish Science within the First TEAM (First TEAM/2017-3/21) programme co-financed by the European Union under the European Regional Development Fund.

VI. APPENDIX

A. Calculating Fisher information when an input activates gene expression via the burst size

Fisher information is calculated by

$$F(z) = \left(\frac{\partial b(z)}{\partial z} \right)^2 E_x \left\{ \left(\frac{\partial \ln f(x; a_0, b(z), K)}{\partial b(z)} \right)^2 \right\}. \quad (31)$$

Same as (35) we have

$$E_x \left\{ \left(\frac{\partial \ln f(x; a_0, b(z), K)}{\partial b(z)} \right)^2 \right\} = -E_x \left\{ \frac{\partial^2 \ln f(x; a_0, b(z), K)}{\partial^2 b(z)} \right\}, \quad (32)$$

where

$$E_x \left\{ \frac{\partial^2 \ln f(x; a_0, b(z), K)}{\partial b^2(z)} \right\} = -\frac{a_0 K + 2E\{x\}}{b^3(z)} + \frac{a_0^2 U^2 \left(a_0, 0, \frac{K}{b(z)} \right)}{U^2 \left(a_0, 1, \frac{K}{b(z)} \right) b^2(z)} + \frac{a_0 (b(z) + K) U \left(a_0, 0, \frac{K}{b(z)} \right)}{U \left(a_0, 1, \frac{K}{b(z)} \right) b^3(z)}. \quad (33)$$

$E\{x\}$ is calculated in (25). Finally using the above equations we get $F(z)$ in (28).

B. Calculating Fisher information when an input activates gene expression via the burst frequency

Since a_0 is a function of input z , $a_0(z)$ changes as per (16), with constant burst size, following (7) Fisher information is as follows

$$F(z) = \left(\frac{\partial a_0(z)}{\partial z} \right)^2 E_x \left\{ \left(\frac{\partial \ln f(x; a_0(z), b, K)}{\partial a_0(z)} \right)^2 \right\}. \quad (34)$$

As $\frac{\partial^2 \ln f(x; a_0(z), b, K)}{\partial a_0^2(z)}$ exists,

$$E_x \left\{ \left(\frac{\partial \ln f(x; a_0(z), b, K)}{\partial a_0(z)} \right)^2 \right\} = -E_x \left\{ \frac{\partial^2 \ln f(x; a_0(z), b, K)}{\partial a_0^2(z)} \right\}, \quad (35)$$

where

$$\frac{\partial^2 \ln f(x; a_0(z), b, K)}{\partial a_0^2(z)} = -\psi_1(a_0(z)) + \frac{U_{(1,0,0)}^2(a_0(z), 1, \frac{K}{b}) - U(a_0(z), 1, \frac{f}{b}) U_{(2,0,0)}(a_0(z), 1, \frac{K}{b})}{U(a_0(z), 1, \frac{K}{b})^2}. \quad (36)$$

Finally, $F(z)$ can be calculated by using (34), (35) and (36).

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] M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain, "Stochastic gene expression in a single cell," *Science*, vol. 297, pp. 1183–1186, 2002.
- [3] A. Bar-Even, J. Paulsson, N. Maheshri, M. Carmi, E. O'Shea, Y. Pilpel, and N. Barkai, "Noise in protein expression scales with natural protein abundance," *Nature Genetics*, vol. 38, pp. 636–643, 2006.
- [4] W. J. Blake, M. Kaern, C. R. Cantor, and J. J. Collins, "Noise in eukaryotic gene expression," *Nature*, vol. 422, pp. 633–637, 2003.
- [5] W. J. Blake, G. Balazsi, M. A. Kohanski, F. J. Isaacs, K. F. Murphy, Y. Kuang, C. R. Cantor, D. R. Walt, and J. J. Collins, "Phenotypic consequences of promoter-mediated transcriptional noise," *Molecular Cell*, vol. 24, pp. 853–865, 2006.
- [6] J. R. S. Newman, S. Ghaemmaghami, J. Ihmels, D. K. Breslow, M. Noble, J. L. DeRisi, and J. S. Weissman, "Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise," *Nature Genetics*, vol. 441, pp. 840–846, 2006.
- [7] A. Eldar and M. B. Elowitz, "Functional roles for noise in genetic circuits," *Nature*, vol. 467, pp. 167–173, 2010.
- [8] H. Maamar, A. Raj, and D. Dubnau, "Noise in gene expression determines cell fate in *Bacillus subtilis*," *Science*, vol. 317, pp. 526–529, 2007.
- [9] G. Balázsi, A. van Oudenaarden, and J. J. Collins, "Cellular decision making and biological noise: From microbes to mammals," *Cell*, vol. 144, pp. 910–925, 2014.
- [10] H. H. Chang, M. Hemberg, M. Barahona, D. E. Ingber, and S. Huang, "Transcriptome-wide noise controls lineage choice in mammalian progenitor cells," *Nature*, vol. 453, pp. 544–547, 2008.
- [11] J. Veening, W. K. Smits, and O. P. Kuipers, "Bistability, epigenetics, and bet-hedging in bacteria," *Annual Review of Microbiology*, vol. 62, pp. 193–210, 2008.
- [12] T. M. Norman, N. D. Lord, J. Paulsson, and R. Losick, "Stochastic switching of cell fate in microbes," *Annual Review of Microbiology*, vol. 69, pp. 381–403, 2015.
- [13] N. Balaban, J. Merrin, R. Chait, L. Kowalik, and S. Leibler, "Bacterial persistence as a phenotypic switch," *Science*, vol. 305, pp. 1622–1625, 2004.
- [14] A. Becskei and L. Serrano, "Engineering stability in gene networks by autoregulation," *Nature*, vol. 405, pp. 590–593, 2000.
- [15] Y. Dublanche, K. Michalodimitrakis, N. Kummerer, M. Foglierini, and L. Serrano, "Noise in transcription negative feedback loops: simulation and experimental analysis," *Molecular Systems Biology*, vol. 2, p. 41, 2006.
- [16] A. Singh and J. P. Hespanha, "Optimal feedback strength for noise suppression in autoregulatory gene networks," *Biophysical Journal*, vol. 96, pp. 4013–4023, 2009.
- [17] D. Orrell and H. Bolouri, "Control of internal and external noise in genetic regulatory networks," *Journal of Theoretical Biology*, vol. 230, pp. 301–312, 2004.
- [18] A. Borri, P. Palumbo, and A. Singh, "The impact of negative feedback in metabolic noise propagation," *IET Systems Biology*, pp. 179–186, 2016.
- [19] I. Lestas, G. Vinnicombe, and J. Paulsson, "Fundamental limits on the suppression of molecular fluctuations," *Nature*, vol. 467, pp. 174–178, 2010.
- [20] 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.
- [21] S. Kumar and A. J. Lopez, "Negative feedback regulation among splicing factors encoded by *rbp1* and *rbp1*-like in *Drosophila*," *The EMBO Journal*, vol. 24, pp. 2646–2655, 2005.
- [22] A. Singh and J. P. Hespanha, "Evolution of autoregulation in the presence of noise," *IET Systems Biology*, vol. 3, pp. 368–378, 2009.
- [23] N. Rosenfeld, M. B. Elowitz, and U. Alon, "Negative autoregulation speeds the response times of transcription networks," *Journal of molecular biology*, vol. 323, no. 5, pp. 785–793, 2002.
- [24] U. Alon, "Network motifs: theory and experimental approaches," *Nature Reviews Genetics*, vol. 8, pp. 450–461, 2007.
- [25] A. Miliadis-Argeitis, S. Summers, J. Stewart-Ornstein, I. Zuleta, D. Pincus, H. El-Samad, M. Khammash, and J. Lygeros, "In silico feedback for in vivo regulation of a gene expression circuit," *Nature Biotechnology*, vol. 29, 2011.
- [26] Y. Qian and D. D. Vecchio, "Realizing 'integral control' in living cells: how to overcome leaky integration due to dilution?" *Journal of The Royal Society Interface*, vol. 15, 2018.
- [27] S. K. Aoki, G. Lillacci, A. Gupta, A. Baumschlager, D. Schweingruber, and M. Khammash, "A universal biomolecular integral feedback controller for robust perfect adaptation," *Nature*, p. 1, 2019.
- [28] G. Tkačik, C. G. Callan Jr, and W. Bialek, "Information capacity of genetic regulatory elements," *Physical Review E*, vol. 78, no. 1, p. 011910, 2008.
- [29] S. Hormoz, "Cross talk and interference enhance information capacity of a signaling pathway," *Biophysical Journal*, vol. 104, no. 5, pp. 1170–1180, 2013.
- [30] C. G. Bowsher and P. S. Swain, "Environmental sensing, information transfer, and cellular decision-making," *Current opinion in biotechnology*, vol. 28, pp. 149–155, 2014.
- [31] E. Ziv, I. Nemenman, and C. H. Wiggins, "Optimal signal processing in small stochastic biochemical networks," *PloS one*, vol. 2, no. 10, p. e1077, 2007.
- [32] R. Cheong, A. Rhee, C. J. Wang, I. Nemenman, and A. Levchenko, "Information transduction capacity of noisy biochemical signaling networks," *science*, vol. 334, no. 6054, pp. 354–358, 2011.
- [33] D. M. Suter, N. Molina, D. Gutfeld, K. Schneider, U. Schibler, and F. Naef, "Mammalian genes are transcribed with widely different bursting kinetics," *Science*, vol. 332, pp. 472–474, 2011.
- [34] R. D. Dar, B. S. Razoosky, A. Singh, T. V. Trimeloni, J. M. McCollum, C. D. Cox, M. L. Simpson, and L. S. Weinberger, "Transcriptional burst frequency and burst size are equally modulated across the human genome," *Proceedings of the National Academy of Sciences*, vol. 109, pp. 17454–17459, 2012.
- [35] T. Fukaya, B. Lim, and M. Levine, "Enhancer control of transcriptional bursting," *Cell*, vol. 166, pp. 358–368, 2015.
- [36] C. R. Bartman, S. C. Hsu, C. C.-S. Hsiung, A. Raj, and G. A. Blobel, "Enhancer regulation of transcriptional bursting parameters revealed by forced chromatin looping," *Molecular Cell*, vol. 62, pp. 237–247, 2016.
- [37] A. M. Corrigan, E. Tunnacliffe, D. Cannon, and J. R. Chubb, "A continuum model of transcriptional bursting," *eLife*, vol. 5, p. e13051, 2016.
- [38] S. Chong, C. Chen, H. Ge, and X. S. Xie, "Mechanism of transcriptional bursting in bacteria," *Cell*, vol. 158, pp. 314–326, 2014.
- [39] A. Singh, B. Razoosky, C. D. Cox, M. L. Simpson, and L. S. Weinberger, "Transcriptional bursting from the HIV-1 promoter is a significant source of stochastic noise in HIV-1 gene expression," *Biophysical Journal*, vol. 98, pp. L32–L34, 2010.
- [40] R. D. Dar, S. M. Shaffer, A. Singh, B. S. Razoosky, M. L. Simpson, A. Raj, and L. S. Weinberger, "Transcriptional bursting explains the noise-versus-mean relationship in mRNA and protein levels," *PLOS ONE*, vol. 11, p. e0158298, 2016.
- [41] I. Golding, J. Paulsson, S. Zawilski, and E. Cox, "Real-time kinetics of gene activity in individual bacteria," *Cell*, vol. 123, pp. 1025–1036, 2005.
- [42] A. Singh, B. S. Razoosky, R. D. Dar, and L. S. Weinberger, "Dynamics of protein noise can distinguish between alternate sources of gene-expression variability," *Molecular Systems Biology*, vol. 8, p. 607, 2012.
- [43] 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.
- [44] L. Cai and N. F. X. S. Xie, "Stochastic protein expression in individual cells at the single molecule level," *Nature*, vol. 440, pp. 358–362, Sep. 2006.
- [45] M. Thattai and A. van Oudenaarden, "Intrinsic noise in gene regulatory networks," *Proceedings of the National Academy of Sciences*, vol. 98, pp. 8614–8619, 2001.
- [46] V. Shahrezaei and P. S. Swain, "Analytical distributions for stochastic gene expression," *Proceedings of the National Academy of Sciences*, vol. 105, pp. 17256–17261, 2008.
- [47] N. Friedman, L. Cai, and X. Xie, "Linking stochastic dynamics to population distribution: an analytical framework of gene expression," *Physical Review Letters*, vol. 97, p. 168302, 2006.
- [48] N. Van Kampen, *Stochastic processes in physics and chemistry*. Elsevier, 2011.
- [49] P. Bokes and A. Singh, "Gene expression noise is affected differentially by feedback in burst frequency and burst size," *Journal of Mathematical Biology*, vol. 74, pp. 1483–1509, 2017.

- [50] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.
- [51] T. Jetka, K. Nienaltowski, S. Filippi, M. P. Stumpf, and M. Komorowski, "An information-theoretic framework for deciphering pleiotropic and noisy biochemical signaling," *Nature communications*, vol. 9, no. 1, p. 4591, 2018.
- [52] 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, no. 16, p. 168302, 2006.
- [53] A. Singh, "Negative feedback through mRNA provides the best control of gene-expression noise," *IEEE Transactions on Nanobioscience*, vol. 10, pp. 194–200, 2011.