Bursty Expression Enhances Reliability of Messaging in Gene Networks

Leonardo R. Gama¹, Guilherme Giovanini², Gábor Balázsi^{3,4} and Alexandre F. Ramos^{1,2,*}

¹*Departamento de Radiologia e Oncologia & Instituto do Câncer do Estado de São Paulo, Faculdade de Medicina, Universidade de São Paulo, CEP 05403-911, São Paulo, SP, Brazil.*

²*Escola de Artes, Ciências e Humanidades, Universidade de São Paulo
Av. Arlindo Bettio, 1000, CEP 03828-000, São Paulo, SP, Brazil.*

³*The Louis and Beatrice Laufer Center for Physical and Quantitative Biology
Stony Brook University, Stony Brook, NY 11794, USA.*

⁴*Department of Biomedical Engineering, Stony Brook University
Stony Brook, NY 11794, USA.*

*Correspondence: alex.amos@usp.br

February 9, 2020

Abstract

The promoter state of a gene and its expression levels are modulated by the amounts of transcription factors interacting with its regulatory regions. Hence, one may interpret a gene network as a communicating system in which the state of the promoter of a gene (the *source*) is communicated by the amounts of transcription factors that it expresses (the *message*) to modulate the state of the promoter and expression levels of another gene (the *receptor*). The reliability of the gene network dynamics can be quantified by the Shannon's entropy of the message and the mutual information between the message and the promoter state. Here we consider a stochastic model for a binary gene and use its exact steady state solutions to calculate the entropy and mutual information. We show that a slow switching promoter having long and equally standing ON and OFF states maximizes the mutual information and reduces entropy. That is a bursty regime generating a high variance message governed by a bimodal probability distribution with peaks of the same height. Our results indicate that Shannon's theory can be a powerful framework for understanding how bursty gene expression conciliates with the striking spatio-temporal precision exhibited in pattern formation of developing organisms.

Keywords: Noise in gene networks; binary gene regulation; bursty gene expression; Shannon theory; stochastic modeling

1 Introduction

The remarkable spatio-temporal precision observed during development of *metazoa* is the result of well orchestrated gene networks regulating the expression of proteins that determine cell fate. That contrasts with the unavoidable fluctuations of intracellular molecular numbers caused by low copy numbers of reactants [1] as detected in expression of prokaryotic and eukaryotic cell's genes [2–17]. Hence, one may ask whether in the due course of evolution there emerged mechanisms equipping biological systems with

the capacity of modulating noise to exploit it functionally [18]. One possible example is bursty gene expression which appears to be an ubiquitous transcriptional regime occurring in developing *Drosophila* embryos [19–21]. During development, pattern formation is regulated by gene networks generating protein levels within proper ranges at proper instants and positions of the embryo [11, 19–25]. For those patterns to be achieved, the activation and inactivation of the promoters of the genes of those networks are modulated with sufficient precision. Hence, one may ask how the promoters, and transcription factors regulating the expression of these gene networks, contribute to the emergence of such a reliable dynamics.

Here we approach those questions considering a gene network as a communicating system in which a gene (the source) expresses transcription factors (the message) that regulate another gene (the receptor). That enables the application of information theory for investigating transduction by a gene network of input signal into functional output [26–28]. As opposed to the common consideration of gene activation/inactivation as completely random events, here we assign to them information content originating, for example, from upstream regulators. Therefore, we assume the information content of the input message to be the promoter state of the source, since the binding of transcription factors to the enhancer induces the state of the promoter of the receptor gene [29, 30]. The processing of all information arriving at the receptor is not considered here [31–34], as we are interested on the reliability of the information content of the message. Hence, we employ Shannon’s theory to compute the entropy of the message and the mutual information between the message and the promoter state of the source [35] by means of the probabilities given by the steady state exact solutions for the stochastic model for binary gene expression [36–38]. Since we are analysing the information content of the message we may interpret the number of products expressed by the binary gene as mRNA’s (which would result on the transcription model) or as transcription factors (which would result on a coupled transcription-translation model). Our results show that the gene having a slow switching promoter with ON and OFF states of the same duration maximizes mutual information between the message and the promoter state while reducing the message’s entropy. That reveals a new perspective to understand the origins of the precision observed in biological systems based on Shannon’s framework as a useful tool for investigating its mechanisms.

1.1 Preliminaries

Fig. 1 shows the effective set of reactions used to interpret a gene network as a communicating system. It depicts the symbols for: synthesis rate of the transcription factors when the promoter is ON, k ; transcription factors degradation rate, ρ ; promoter switching rate from ON to OFF state and vice-versa, h and f , respectively; the number of transcription factors available, n . No synthesis happens when the promoter is OFF. The curved lines indicate the interaction of the protein encoded in one gene with the enhancer of the receptor.

The average number of transcription factors (or expected message from the source) is denoted as usual by $\langle n \rangle$. The bursting (hyper-Poissonian) nature of the message can be measured by the Fano factor, $\mathcal{F}$, defined as the ratio of the variance to the mean (see Eq. 1). A Poissonian distribution has $\mathcal{F} = 1$ and a super Poissonian distribution has $\mathcal{F} = 1 + \delta$, where the increment δ is positive [2, 39]. We denote as bursty the expression of a gene for which its products are governed by super Poissonian distributions. In some cases, bursty expression can generate bimodal distributions, with numbers of protein products governed by distributions having two distinct peaks. The bimodal distributions are consistent with trajectories of the number of gene products characterized by a repetition of the following

Figure 1: Two elements of an arbitrary gene network are depicted as a communicating system. The continuation of the network is represented by the dashed lines. A source gene communicates with a receptor gene through transcription factors and the information is the promoter state of the source.

events, assuming that initially the promoter is OFF and there are no transcription factors : a burst towards a finite value $\langle n_\alpha \rangle$ after the promoter switching ON; small fluctuations around $\langle n_\alpha \rangle$ while the promoter remains ON; a sharp reduction towards zero after the promoter switching OFF; small fluctuations around zero while the promoter remains OFF; and the cycle repeats.

The information content of the message is quantified in the framework of Shannon's communication theory, where the variance of binary signals correlates with their information content. Hence, the average information content of the message is given by Shannon's entropy; the coupling between the message and the state of the promoter of the source is given by the mutual information; the average information content of the message given the promoter being ON is measured by the conditional entropy (see Subsec. 4.7). We assume that the stochastic binary gene expression model is at the steady state limit. Hence, the promoter states and the number of transcription factors will change accordingly with constant probabilities. The fraction of durations of the ON and OFF states of the promoter is governed by constant marginal probabilities p_α and p_β , respectively (see Eq. 7), for the promoter to be ON and OFF. For the promoter being exclusively ON the source would become constitutive (see Subsec. 4.1) with average number of transcription factors being N (see Eq. 3 or Eq. 7). The probability of finding n transcription factors and the promoter ON is denoted by α_n while ϕ_n indicates the marginal probability of finding n transcription factors (see Eq. 4 and Eq. 6).

The parameter ϵ is proportional to the ratio of the transcription factor's lifetime to the average period of a complete cycle of promoter switching (see Eq. 7) for a fixed p_α . For $\epsilon \ll 1$ the switching has the longer time scale and enables the average number of transcription factors to reach steady state value before the promoter switches from one state to another. For $\epsilon \gg 1$ we have the opposite condition and there will be multiple promoter switchings during the lifetime of a transcription factor.

Figure 2: Entropy (or mutual information) and Fano factor were obtained varying N for fixed p_α and ϵ . All curves have the same value for the average number of gene products ($\langle n \rangle = 50$). Graph A shows the entropy of the message sent by the gene versus the noise of the message. Graph B shows the mutual information between the message and the source state as a function of the bursting. Continuous lines refer to fixed values of ϵ and fixed values of p_α are indicated by the dashed ones.

2 Results

2.1 Bursty expression enables entropy reduction and mutual information increase

The horizontal axis of both graphs of Fig. 2 shows the bursting quantified by the Fano factor (see Eq. 1). The vertical axis of graphs A and B shows, respectively, the entropy of the message (see Eq. 11) and the mutual information between the message and the state of the promoter of the source (see Eq. 14). Each continuous line is associated to a specific value of the promoter switching rate, ϵ (see Eq. 7). The dashed curves are isolines for fixed values of the probability for the promoter to be ON, p_α . The color scheme for the curves is given at the caption. The black solid diamond gives the entropy of the message when it is sent by a Poissonian source which is usually employed on the description of a constitutive gene. Note that we have no mutual information for the constitutive gene because we assume that its promoter will be always ON. A condition of equivalence between the binary gene to the constitutive one can be devised for $\epsilon \gg N$ which corresponds to a switching cycle being much faster than the protein synthesis rate. In that case, the mutual information approaches zero.

Graph A shows that the constitutive gene sets a reference value for the entropy and the binary source at the small bursting regime ($\mathcal{F} \sim 1$) generating a message for which the entropy is greater than that sent by the constitutive gene. As bursting increases, however, the entropy of the message sent by the binary source reaches a maximum and then starts decreasing. Then the entropy of the message sent by the binary source becomes smaller than that of the constitutive source and decays towards arbitrarily small values. Besides, for a fixed value of $\mathcal{F}$, the message with minimal entropy is generated by a process having the smallest ϵ .

Graph B shows that the mutual information between the message and the source state reaches a maximum for values of the Fano factor being $\sim 1 + \langle n \rangle$ and $\epsilon \ll 1$. For a fixed value of $\mathcal{F}$ the mutual information becomes minimal for the greatest values of ϵ . As the switching rate becomes slower, *i.e.* $\epsilon \ll 1$, the mutual information is maximal for $p_\alpha \sim 1/2$. Furthermore, one may set an interval at which the mutual information has

Figure 3: The entropy (and mutual information) was obtained fixing p_α and N . All curves have the same average number of transcription factors ($\langle n \rangle = 50$). The horizontal axis of both graphs indicate the transcription factor's lifetime relative to the source's ON and OFF promoter switching duration. Graph A shows the entropy of the message and graph B shows the mutual information between the message and the source's promoter state. Each continuous line corresponds to a given value of p_α and the entropy value for a constitutive gene producing $\langle n \rangle = 50$ proteins is shown by the horizontal dashed black line. Values of ϵ used to construct Fig. 4, Fig. 5, and Fig. 6 appear in bold at the horizontal axis label.

higher, significant values, by arbitrarily fixing the probability for the promoter to be ON within a range $1/L \leq p_\alpha \leq 1 - 1/L$ such that the corresponding noise is $(L - 1)\langle n \rangle \geq \mathcal{F} \geq \langle n \rangle / (L - 1)$, where $L > 1$ and $\mathcal{F} \gg 1$ such that $(\mathcal{F} - 1)(1 + \epsilon) \sim \mathcal{F}$.

2.2 The slow switching genes generate reduced entropy and increased values for mutual information

The horizontal axis of both graphs of Fig. 3 indicates the values of the transcription factor's lifetime relative to the ON and OFF switching cycle of the source's promoter. The vertical axis of graphs A and B shows, respectively, the entropy of the message (see Eq. 11) and the mutual information between the message and the source state. To each continuous line we associate a specific value of the probability for the source to be ON, p_α (see Eq. 7). The dashed black curve indicates the entropy for the constitutive gene. This value is given as a reference since there is no switching on a constitutive promoter and, additionally, there is no mutual information between the promoter state and the message. The color scheme for the curves is given at the caption.

Graph A shows that the entropy for a given p_α is minimal when the switching is slow. As the switching rate increases, the entropy reaches a maximum value (H_m) for an ϵ_m which is greater as the probability for the promoter to be ON is reduced. For $\epsilon > \epsilon_m$ and a given p_α , the entropy of the message decreases and asymptotically approaches that of a constitutive source. A source having $\epsilon \ll 1$ generates a message for which the entropy increases with p_α while the opposite happens for large values of ϵ . Furthermore, for $\epsilon \ll 1$ and $p_\alpha = 1/2$ the entropy is smaller than that of the constitutive source, but not minimal.

Graph B shows the mutual information as a function of ϵ with each line being related to a specific value of p_α . The similarity of the mutual information for complementary values of p_α as the switching rate increases relates to the bell-shaped curves (in semi-log scale) shown on Fig. 2 (graph B). For a given value of p_α the mutual information decreases monotonically with the increase of ϵ . For $\epsilon \ll 1$ the mutual information is maximal when $p_\alpha = 1/2$, and for $p_\alpha \ll 1$ the message will have minimal entropy and minimal mutual

information with the promoter state.

2.3 The distributions of the slow switching bursty regime with reduced entropy and maximal mutual information are bimodal

Figures 4-6 have four graphs each where: graph A shows the entropy on vertical axis and the average number of gene products ($\langle n \rangle$) on the horizontal; graph B shows the entropy conditional to the promoter being ON versus $\langle n_\alpha \rangle$ (see Eq. 8); graph C shows the mutual information versus $\langle n \rangle$; graph D shows the probability distribution of n . Each continuous line correspond to a given p_α with color code given at the key of *graph B*. The values of the entropy (vertical axis of *graph A*); entropy conditional to the source being ON (vertical axis of *graph B*); mutual information (vertical axis of *graph C*) are obtained fixing a given p_α and varying N to generate the values of $\langle n \rangle$ at the horizontal axis. The probability distributions of *graph D* are obtained by also fixing $\langle n \rangle = 50$, such that $N = 50/p_\alpha$. For each p_α we compute the values of ϕ_n (continuous lines) and α_n (dashed lines). In each set of graphs the value of ϵ is fixed and given at graph A's title. Results for the constitutive gene are shown by a black line on graphs A, B, and D.

Figure 4: The switching rate of the promoter of the source is 0.01. The vertical axis of graph's A, B, and C shows, respectively, entropy versus $\langle n \rangle$, entropy conditional to the source being ON versus $\langle n_\alpha \rangle$, and the mutual information versus $\langle n \rangle$. Graph D shows the probability distribution of n .

Fig. 4. Graph A shows that the entropy of the message increases monotonically with $\langle n \rangle$ for both the constitutive and binary sources. As p_α approaches one the entropy of the binary sources approaches that of a constitutive. A source having smaller values of p_α will

send messages with significantly reduced entropy which growth with $\langle n \rangle$ is much slower. *Graph B* shows the entropy of the message under the condition of the source being ON. In that case the conditional entropy of all messages approach that of a constitutive source for all values of p_α . That is observed by introducing the entropy for the constitutive source as a reference value (black curve). *Graph C* shows that the mutual information between the source's promoter state and the message increases monotonically until it reaches a limit at a value that depends on p_α . The mutual information for two complementary values of p_α approximate each other for a sufficiently large value of $\langle n \rangle$. For $p_\alpha = 1/2$ the mutual information reaches its maximum with its plateau being reached when $\langle n \rangle \sim 30$. *Graph D* shows the probability distribution for $\langle n \rangle = 50$. The black curve shows the probability distribution of n when emitted by a constitutive source. The distributions generated by the binary source are all bimodal, with one mode centered around zero and the other around N (where $N = \langle n \rangle / p_\alpha$). As the value of p_α approaches zero the second peak has its height lowered and its center displaced to the right. The opposite happens when p_α approaches unity with the taller peak approaching that of the constitutive source for the same mean.

Figure 5: The switching rate of the promoter of the source is 2. The vertical axis of graph's A, B, and C shows, respectively, entropy versus $\langle n \rangle$, entropy conditional to the source being ON versus $\langle n_\alpha \rangle$, and the mutual information versus $\langle n \rangle$. Graph D shows the probability distribution of n .

Fig. 5. *Graph A* shows that the entropy of the message increases monotonically with $\langle n \rangle$ for both the constitutive and the binary sources. For larger values of p_α , the entropy of the binary sources are greater than that of a constitutive. Entropy of the message from the binary source is a bit smaller than that of a constitutive for small values of p_α .

Graph B shows the entropy of the message given the promoter of the source being ON. In that case the conditional entropy of all messages is greater than that of a constitutive source. *Graph C* shows that the mutual information also increases monotonically until it reaches a limit value that depends on p_α . The mutual information for two complementary values of p_α approximate each other for a sufficiently large value of $\langle n \rangle$. For $p_\alpha = 1/2$ the mutual information reaches its maximum with its plateau being reached when $\langle n \rangle \sim 30$. Note, however, that the maximal value for the mutual information is smaller than that for $\epsilon = 0.01$. *Graph D* shows the probability distributions for the binary source. The distributions generated by the binary source have significant probability values spread along a large range of values of n . For $p_\alpha = 1/2$ the distribution is table shaped. As $p_\alpha \rightarrow 1$ the distribution forms a peak that gets closer to that of the distribution of a constitutive source while the distributions have a sector governed by a power law when p_α approaches zero.

Figure 6: The switching rate of the promoter of the source is 10. The vertical axis of graph's A, B, and C shows, respectively, entropy versus $\langle n \rangle$, entropy conditional to the source being ON versus $\langle n_\alpha \rangle$, and the mutual information versus $\langle n \rangle$. Graph D shows the probability distribution of n .

Fig. 6. *Graph A* shows that the entropy of the message increases monotonically with $\langle n \rangle$ for both the constitutive and the binary sources. The entropy of the message emitted by the binary source is larger than that of a constitutive gene for all values of p_α . *Graph B* shows the entropy of the message given the promoter being ON. The conditional entropy of all messages are greater than that of a constitutive source. *Graph C* shows that the mutual information also increases monotonically until it reaches a limit value that depends on p_α and that this maximal value is smaller than that for $\epsilon = 0.01$ or $\epsilon = 2$.

Graph *D* shows the probability distributions for the binary and the constitutive sources. The distributions generated by the binary source are spread along a large range of values of n . The distributions of n emitted by the binary source approach that of a constitutive gene as $p_\alpha \rightarrow 1$.

3 Discussion

This manuscript shows a functional role for the unavoidable biochemical stochasticity observed inside the cells. We demonstrate that under proper regulation expression bursting of the transcription factors produced from a source gene maximizes the information transmitted to a receptor. We consider that the information to be exchanged between the genes is their promoter's states since those are major determinants of gene expression levels. That is important for generating the internal conditions underpinning cell fate determination as observed during development, in which patterns of gene expression having striking spatio-temporal precision are observed.

The results presented at the Subsec. 2.1 show that one may increase the expression bursting of a binary gene to generate messages having arbitrarily small entropy. The reduction on the entropy and its maximal value depends both on the message's bursting noise and on the probabilities governing the promoter states. We complement that analysis by computing the mutual information between the message and the promoter state. We verify that it has a maximum when the probability for the promoter to be ON is 1/2 and the bursting noise is approximately the average number of gene products. Therefore, we also show that, besides promoter state probabilities, the bursting of gene expression must also be regulated for the maximization of the information content of the message sent by the binary source.

Eq. 1 shows that bursting can be regulated by three distinct parameters. The value of N is determined by the capacity of the cell of synthesizing transcription factors while p_α is given by the enhancer's state. Therefore, minimizing ϵ may be a useful complementary method for increasing level of bursting in the binary gene. Indeed, Subsec. 2.2 shows that we have reduced entropy and maximal mutual information for the slow switching genes. As expected, a very fast switching gene will generate a message having entropy approaching that of the constitutive gene having the same average expression levels.

The graphs of Subsec. 2.3 show that at maximal mutual information and reduced entropy, the distributions governing the message are bimodal. That implies on a slow switching and average duration of the ON or OFF states of the promoter being large enough for the number of transcription factors to reach steady state before the next switching. Hence, if we consider the conditional entropy given the promoter being ON, it will behave as a constitutive source. In other words, that regime permits the transcription factors to be a read out of the state of the promoter of the gene. As the parameter ϵ increases, the super Poissonian distributions change their shapes, *e.g.*, becoming uniform-like, negative-binomial like, or Poissonian-like. Those regimes will have decreasing maximal mutual information and reduced quality on the information content of the message.

Optimization methods have been employed previously to verify that the bimodal distributions governing the signal, here interpreted as the number of transcription factors, enables maximal mutual information between the source and the receptor's response [27, 40, 41]. That is done using a parameter-free approach useful for data analysis but not necessarily straightforward to be interpreted biologically and mechanistically. We use an approach based on an exactly solvable model with biologically interpretable, or observable, parameters. Hence, it can provide some guidance to synthetic biologists for de-

signing gene circuits behaving as reliable (or unreliable) messaging devices. Additionally, it is a basic building block that can help on further understanding noise related biological phenomena [42, 43], *e.g.*, occurring during development [44] or carcinogenesis [45, 46], in which information transmission plays a role.

The existence of a range of bursting regime for which mutual information is significant suggests that bursting is regulated in gene networks in which the promoter states are the information being exchanged between the source and receptor genes. There have been long term characterization of the enhancers as the gene structures responsible for regulating the promoter exposure for transcription and, consequently, average levels of gene products. Since gene expression bursting has also a functional role, one may wonder whether there exist a DNA structure surrounding a promoter that regulates its switching rates and, consequently, its bursting regimes, such that the gene network reliability is increased in contexts like development. Hence, it would be interesting to investigate experimentally the regulation of the promoter switching rates for multiple genes. Additionally, one might test the mutual information between a transcription factor and its target protein, or their promoters. Another interesting question is to test the information transfer between an external signal and genes involved in its processing. Our results also show bursty, non-constitutive promoters, have important biological role as elements of gene networks in which the important information is the promoter state. That contributes for the classification of noise function in gene expression besides self-repression as the mechanism responsible for noise reduction or homeostasis [47–50].

4 Materials and Methods

Here we assume that both the constitutive and the binary source are at the steady state regime. The source is interpreted as the gene encoding a transcription factor and the state of the source is determined as the promoter being ON or OFF. The constitutive source is always ON and the binary source switches at random between ON and OFF states. The regulation of a target gene depends on the interaction of the transcription factors with its enhancer. The binding of the transcription factors to an enhancer can be computed effectively as a function of the amounts of transcription factors around the DNA. Hence, the message emitted by the source is the number of transcription factors that are available for interaction with the enhancer of the receptor gene. The receptor gene promoter state is then determined by the action of the message on it, as either repressing or inducing transcription. The number of transcription factors denoted by n is a random variable and the noise on the message can be computed by the Fano factor, defined as the ratio between the variance of n to its average value. We denote the Fano factor by $\mathcal{F}$ such that:

$$\mathcal{F} = \frac{\langle n^2 \rangle - \langle n \rangle^2}{\langle n \rangle} . \quad (1)$$

The Fano factor is a measure of how similar to the Poissonian a probability distribution is. The Poisson distribution has $\mathcal{F} = 1$ while $\mathcal{F} < 1$ characterizes a sub-Poisson distribution. A super-Poissonian distribution has $\mathcal{F} > 1$.

4.1 A stochastic model for a constitutive gene

We model a constitutive gene as a Poissonian birth and death process. The random variable n indicates the number of transcription factors available for regulating a receptor gene. The probability of finding n transcription factors is indicated by ϕ_n . The synthesis

and degradation rates are denoted by k and ρ , respectively. The probability ϕ_n is given by

$$\phi_n = \frac{e^{-N}}{n!} N^n, \quad (2)$$

where

$$N = \frac{k}{\rho}, \quad (3)$$

is the average number of transcription factors produced from a constitutive gene at the steady state regime. The variance of the Poissonian process is also N .

4.2 The stochastic model for the binary gene

We model the binary gene as two Poissonian birth and death processes coupled by a random telegraph process. The random telegraph process describes the switching of the promoter between ON and OFF states. When the promoter is ON (or OFF) proteins are synthesized at a rate denoted by k (or 0). The proteins degradation rate is denoted by ρ . There are two random variables, (m, n) , where m indicates the state of the promoter as ON or OFF and n indicates the number of proteins. The probability of finding the promoter ON (or OFF) and n proteins is indicated by α_n (or β_n). Using the exact solutions [36–38] for this model we have:

$$\alpha_n = p_\alpha \frac{N^n}{n!} \frac{(1 + \epsilon p_\alpha)_n}{(1 + \epsilon)_n} M(1 + \epsilon p_\alpha + n, 1 + \epsilon + n, -N), \quad (4)$$

$$\beta_n = (1 - p_\alpha) \frac{N^n}{n!} \frac{(\epsilon p_\alpha)_n}{(1 + \epsilon)_n} M(\epsilon p_\alpha + n, 1 + \epsilon + n, -N), \quad (5)$$

$$\phi_n = \frac{N^n}{n!} \frac{(\epsilon p_\alpha)_n}{(\epsilon)_n} M(\epsilon p_\alpha + n, \epsilon + n, -N). \quad (6)$$

where the constants p_α , ϵ and N are defined as:

$$p_\alpha = \frac{f}{f + h}, \quad p_\beta = \frac{h}{f + h}, \quad \epsilon = \frac{f + h}{\rho}, \quad N = \frac{k}{\rho}, \quad (7)$$

and ϕ_n indicates the marginal probability of finding n proteins, namely $\phi_n = \alpha_n + \beta_n$.

$M(a, b, x)$ denotes the KummerM function, defined as a power series:

$$M(a, b, x) = \sum_{n=0}^{\infty} \frac{(a)_n}{(b)_n} \frac{x^n}{n!},$$

where the $(a)_n = a(a + 1) \dots (a + n - 1)$ denotes the Pochhammer symbol [51].

4.3 Interpretation of the parameters p_α , ϵ , and N .

The parameters p_α and p_β is the steady state probability for the promoter to be at the ON and OFF state, with $p_\alpha + p_\beta = 1$. p_α (or p_β) give the fraction of a time interval T that the promoter will be ON (or OFF) while switching at random between its possible states.

The parameter ϵ is the ratio of the promoter switching rate to the transcription factors degradation rate. It can be interpreted as a ratio of the two time scales of the stochastic processes being coupled. For simplicity, we assume $T_D = \frac{1}{\rho}$ as the proteins

lifetime, and $T_S = \frac{f+h}{fh}$ as the average time for the promoter to complete one switching cycle. Using the definitions at Eq. 7 we obtain

$$\epsilon = \frac{1}{p_\alpha p_\beta} \frac{T_D}{T_S} .$$

When $\epsilon \ll 1$ the switching cycle is long and proteins lifetime is shorter than the time spent by the promoter at ON or OFF state if $p_\alpha \sim 1/2$. When $\epsilon \gg 1$ the switching cycle is very short compared to the transcription factors lifetime and multiple switchings will happen while a transcription factor remains functional.

The parameter N is the average number of transcription factors if the binary gene remains ON all the time as if it was a constitutive gene.

4.4 The mean number and the conditional mean number of the stochastic model for a binary gene

The average number of transcription factors synthesized from the binary gene ($\langle n \rangle$), and the average number of transcription factors given that the promoter is ON (denoted by $\langle n_\alpha \rangle$) are

$$\langle n \rangle = N p_\alpha , \quad \langle n_\alpha \rangle = N \frac{1 + p_\alpha \epsilon}{1 + \epsilon} . \quad (8)$$

The mean number of transcription factors of the binary gene are dependent on the proportion of time that the promoter of the gene is ON. That indicates that the average amount of transcription factors will range from zero to N accordingly with p_α . The conditional mean number of transcription factors given the promoter being ON is obtained from $\langle n_\alpha \rangle = \frac{1}{p_\alpha} \sum_{n=0}^{\infty} n \alpha_n$. For very small values of ϵ , $\langle n_\alpha \rangle$ approaches N , which is the average for a Poissonian process. For large values of ϵ we have $\langle n_\alpha \rangle \sim \langle n \rangle$.

4.5 The variance and bursting noise of n on the stochastic model for a binary gene

The variance on n , denoted by σ^2 , and the Fano factor ($\mathcal{F}$) are, respectively,

$$\sigma^2 = \langle n \rangle \left(1 + \frac{N(1-p_\alpha)}{1+\epsilon} \right) , \quad \mathcal{F} = 1 + \frac{N(1-p_\alpha)}{1+\epsilon} . \quad (9)$$

We interpret the variance and the Fano factor of n synthesized from a binary gene in comparison to that of a constitutive gene. Since the analysis is similar for both quantities, we focus on the Fano factor of n produced by the binary source is greater than that of the constitutive gene by a quantity $N(1-p_\alpha)/(1+\epsilon)$. This increment on $\mathcal{F}$ is caused by the promoter switching. For $p_\alpha = 1$ the promoter is fully ON, the binary source becomes constitutive and the increment on $\mathcal{F}$ is null. For $p_\alpha = 0$ we have a fully repressed gene. Additionally, when ϵ approaches infinity the bursting noise on the message sent by a binary source gene is similar to that of a constitutive gene for finite values of N and p_α . For fixed values of N and p_α we have maximal bursting for ϵ approaching zero.

4.6 Entropy for the constitutive source

The Shannon's entropy of a probability distribution ϕ_n is denoted by $H(n)$ and defined as $H(n) = -\sum_{n=0}^{\infty} \phi_n \log \phi_n$, with the logarithm taken to base 2. For a Poissonian distribution having average value N the entropy is given by

$$H(n) = N - N \log(N) + e^{-N} \sum_{n=0}^{\infty} \frac{N^n}{n!} \log(n!) . \quad (10)$$

The entropy of the constitutive gene will be considered here as a reference value to be compared with that of the binary one.

4.7 Entropy, conditional entropy and mutual information for the binary source

Since it is not trivial to find a closed form to compute the logarithm of a KummerM function, we obtained the entropy and mutual information for the binary gene by computing their formulas numerically. The entropy of the message sent by the binary gene is:

$$H(n) = - \sum_{n=0}^{\infty} \phi_n \log \phi_n , \quad (11)$$

where ϕ_n is given by Eq. 6.

The conditional entropy of n given the promoter being ON or OFF is denoted by $H_{\alpha}(n)$ or $H_{\beta}(n)$, respectively, and they are written as:

$$H_{\alpha}(n) = \log(p_{\alpha}) - \frac{1}{p_{\alpha}} \sum_{n=0}^{\infty} \alpha_n \log(\alpha_n) , \quad (12)$$

$$H_{\beta}(n) = \log(p_{\beta}) - \frac{1}{p_{\beta}} \sum_{n=0}^{\infty} \beta_n \log(\beta_n) . \quad (13)$$

These formulas are obtained from the definition of the conditional entropy of X given $Y = y$, $H(X|Y = y) = - \sum_{x \in X} p(X|Y = y) \log p(X|Y = y)$, where we assume X being the number of transcription factors and Y the promoter state. Here α_n , β_n , and (p_{α}, p_{β}) are, respectively, given by Eq. 4, Eq. 5, and Eq. 7.

The mutual information between the message n and the source state being ON or OFF is given by

$$I(n; m) = H(n) - p_{\alpha} H_{\alpha}(n) - p_{\beta} H_{\beta}(n) , \quad (14)$$

where $H(n)$, $H_{\alpha}(n)$, and $H_{\beta}(n)$ are, respectively, given by Eq. 11, Eq. 12, and Eq. 13.

Author Contributions: Formal analysis, LRG, GG, GB, and AFR; Investigation, LRG, GG, and AFR; Visualization, LRG, GG, and AFR; Writing - original draft, AFR; Writing - review & editing LRG, GG, GB, and AFR; Funding acquisition, AFR.

Funding: We thank CAPES (88881.062174/2014-01) and FAPESP (2012/24962-3) for previous funding. LRG thanks Dean of Research of the University of São Paulo for support from the "Use of Intelligent Systems" call (18.5.245.86.7). GB is supported by the National Institutes of Health, NIGMS MIRA Program (R35 GM122561) and by the Laufer Center for Physical and Quantitative Biology.

Acknowledgments: AFR is thankful to Yuri Suhov for introduction to Information Theory. We are thankful to Alan U. Sabino for useful comments.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

- [1] Delbrück, M. Statistical fluctuations in autocatalytic reactions. *J. Chem. Phys.* **1940**, *8*, 120–124. doi:10.1063/1.1750549.
- [2] Thattai, M.; van Oudenaarden, A. Intrinsic noise in gene regulatory networks. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 8614–8619. doi:10.1073/pnas.151588598.
- [3] Elowitz, M.B.; Levine, A.J.; Siggia, E.D.; Swain, P.S. Stochastic gene expression in a single cell. *Science* **2002**, *297*, 1183–1186. doi:10.1126/science.1070919.
- [4] Simpson, M.L.; Cox, C.D.; Sayler, G.S. Frequency domain analysis of noise in autoregulated gene circuits. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 4551–4556. doi:10.1073/pnas.0736140100.
- [5] Blake, W.J.; Kærn, M.; Cantor, C.R.; Collins, J.J. Noise in eukaryotic gene expression. *Nature* **2003**, *422*, 633–637. doi:10.1038/nature01546.
- [6] Raser, J.M.; O’Shea, E.K. Noise in gene expression: origins, consequences, and control. *Science* **2005**, *309*, 2010–2013. doi:10.1126/science.1105891.
- [7] Arias, A.M.; Hayward, P. Filtering transcriptional noise during development: concepts and mechanisms. *Nat. Rev. Genet.* **2006**, *7*, 34–44. doi:10.1038/nrg1750.
- [8] Hooshangi, S.; Weiss, R. The effect of negative feedback on noise propagation in transcriptional gene networks. *Chaos* **2006**, *16*, 026108. doi:10.1063/1.2208927.
- [9] Chubb, J.R.; Treck, T.; Shenoy, S.M.; Singer, R.H. Transcriptional Pulsing of a Developmental Gene. *Curr. Biol.* **2006**, *16*, 1018–1025. doi:10.1016/j.cub.2006.03.092.
- [10] Boettiger, A.N.; Levine, M. Synchronous and Stochastic Patterns of Gene Activation in the Drosophila Embryo. *Science* **2009**, *325*, 471–473. doi:10.1126/science.1173976.
- [11] Paré, A.; Lemons, D.; Kosman, D.; Beaver, W.; Freund, Y.; McGinnis, W. Visualization of individual Scr mRNAs during Drosophila embryogenesis yields evidence for transcriptional bursting. *Curr. Biol.* **2009**, *19*, 2037–2042. doi:10.1016/j.cub.2009.10.028.
- [12] Crudu, A.; Debussche, A.; Radulescu, O. Hybrid stochastic simplifications for multiscale gene networks. *BMC Syst. Biol.* **2009**, *3*, 89:1–89:25. doi:10.1186/1752-0509-3-89.
- [13] Suter, D.M.; Molina, N.; Gatfield, D.; Schneider, K.; Schibler, U.; Naef, F. Mammalian genes are transcribed with widely different bursting kinetics. *Science* **2011**, *332*, 472–474. doi:10.1126/science.1198817.
- [14] Munsky, B.; Neuert, G.; van Oudenaarden, A. Using Gene Expression Noise to Understand Gene Regulation. *Science* **2012**, *336*, 183–187. doi:10.1126/science.1216379.
- [15] Chalancon, G.; Ravarani, C.N.; Balaji, S.; Martinez-Arias, A.; Aravind, L.; Jothi, R.; Babu, M.M. Interplay between gene expression noise and regulatory network architecture. *Trends Genet.* **2012**, *28*, 221–232. doi:10.1016/j.tig.2012.01.006.

- [16] Sanchez, A.; Golding, I. Genetic Determinants and Cellular Constraints in Noisy Gene Expression. *Science* **2013**, *342*, 1188–1193. doi:10.1126/science.1242975.
- [17] Little, S.C.; Tikhonov, M.; Gregor, T. Precise Developmental Gene Expression Arises from Globally Stochastic Transcriptional Activity. *Cell* **2013**, *154*, 789–800. doi:10.1016/j.cell.2013.07.025.
- [18] Balázsi, G.; van Oudenaarden, A.; Collins, J.J. Cellular decision making and biological noise: from microbes to mammals. *Cell* **2011**, *144*, 910–925. doi:10.1016/j.cell.2011.01.030.
- [19] Bothma, J.P.; Garcia, H.G.; Esposito, E.; Schlissel, G.; Gregor, T.; Levine, M. Dynamic regulation of eve stripe 2 expression reveals transcriptional bursts in living *Drosophila* embryos. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 10598–10603. doi:10.1073/pnas.1410022111.
- [20] Fukaya, T.; Lim, B.; Levine, M. Enhancer Control of Transcriptional Bursting. *Cell* **2016**, *166*, 358–368. doi:10.1016/j.cell.2016.05.025.
- [21] Zoller, B.; Little, S.C.; Gregor, T. Diverse spatial expression patterns emerge from unified kinetics of transcriptional bursting. *Cell* **2018**, *175*, 835–847. doi:10.1016/j.cell.2018.09.056.
- [22] Garcia, H.G.; Tikhonov, M.; Lin, A.; Gregor, T. Quantitative live imaging of transcription in *Drosophila* embryos links polymerase activity to macroscopic patterns. *Curr. Biol.* **2013**, *23*, 2140–2145. doi:10.1016/j.cub.2013.08.054.
- [23] Holloway, D.M.; Spirov, A.V. Transcriptional bursting in *Drosophila* development: Stochastic dynamics of eve stripe 2 expression. *PLOS ONE* **2017**, *12*, e0176228. doi:10.1371/journal.pone.0176228.
- [24] Faló-Sanjuan, J.; Bray, S.J. Decoding the Notch signal. *Dev. Growth Differ.* **2019**, *62*, 4–14. doi:10.1111/dgd.12644.
- [25] Lammers, N.C.; Galstyan, V.; Reimer, A.; Medin, S.A.; Wiggins, C.H.; Garcia, H.G. Multimodal transcriptional control of pattern formation in embryonic development. *Proc. Natl. Acad. Sci. U.S.A.* **2020**, *117*, 836–847. doi:10.1073/pnas.1912500117.
- [26] Ziv, E.; Nemenman, I.; Wiggins, C.H. Optimal Signal Processing in Small Stochastic Biochemical Networks. *PLOS ONE* **2007**, *2*, e1077. doi:10.1371/journal.pone.0001077.
- [27] Tkačik, G.; G, C.; Jr., C.; Bialek, W. Information flow and optimization in transcriptional regulation. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 12265–12270. doi:10.1073/pnas.0806077105.
- [28] Bowsheer, C.G.; Swain, P.S. Identifying sources of variation and the flow of information in biochemical networks. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, E1320–E1328. doi:10.1073/pnas.1119407109.
- [29] Reinitz, J.; Hou, S.; Sharp, D.H. Transcriptional control in *Drosophila*. *Complexus* **2003**, *1*, 54–64. doi:10.1159/000070462.

- [30] Janssens, H.; Hou, S.; Jaeger, J.; Kim, A.; Myasnikova, E.; Sharp, D.; Reinitz, J. Quantitative and predictive model of transcriptional control of the *Drosophila melanogaster* even-skipped gene. *Nat. Genet.* **2006**, *38*, 1159–1165. doi:10.1038/ng1886.
- [31] Kim, A.R.; Martinez, C.; Ionides, J.; Ramos, A.F.; Ludwig, M.Z.; Ogawa, N.; Sharp, D.H.; Reinitz, J. Rearrangements of 2.5 Kilobases of Noncoding DNA from the *Drosophila* even-skipped Locus Define Predictive Rules of Genomic cis-Regulatory Logic. *PLoS Genet.* **2013**, *9*, e1003243. doi:10.1371/journal.pgen.1003243.
- [32] Barr, K.A.; Martinez, C.; Moran, J.R.; Kim, A.R.; Ramos, A.F.; Reinitz, J. Synthetic enhancer design by in silico compensatory evolution reveals flexibility and constraint in cis-regulation. *BMC Syst. Biol.* **2017**, *11*, 116:1–116:15. doi:10.1186/s12918-017-0485-2.
- [33] Granados, A.A.; Pietsch, J.M.J.; Cepeda-Humerez, S.A.; Farquhar, I.L.; Tkačik, G.; Swain, P.S. Distributed and dynamic intracellular organization of extracellular information. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115*, 6088–6093. doi:10.1073/pnas.1716659115.
- [34] Liu, Y.; Barr, K.; Reinitz, J. Fully Interpretable Deep Learning Model of Transcriptional Control. *bioRxiv* **2019**, pp. 1–11. doi:10.1101/655639.
- [35] Shannon, C.E. A mathematical theory of communication. *The Bell System Technical Journal* **1948**, *27*, 379–423. doi:10.1002/j.1538-7305.1948.tb01338.x.
- [36] Peccoud, J.; Ycart, B. Markovian modelling of gene product synthesis. *Theor. Popul. Biol.* **1995**, *48*, 222–234. doi:10.1006/tpbi.1995.1027.
- [37] Innocentini, G.C.P.; Hornos, J.E.M. Modeling stochastic gene expression under repression. *J. Math. Biol.* **2007**, *55*, 413–431. doi:10.1007/s00285-007-0090-x.
- [38] Ramos, A.F.; Innocentini, G.C.P.; Forger, F.M.; Hornos, J.E.M. Symmetry in biology: from genetic code to stochastic gene regulation. *IET Syst. Biol.* **2010**, *4*, 311–329. doi:10.1049/iet-syb.2010.0058.
- [39] Shahrezaei, V.; Swain, P.S. Analytical distributions for stochastic gene expression. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 17256–17261. doi:10.1073/pnas.0803850105.
- [40] Ashall, L.; Horton, C.A.; Nelson, D.E.; Paszek, P.; Harper, C.V.; Sillitoe, K.; Ryan, S.; Spiller, D.G.; Unitt, J.F.; Broomhead, D.S.; Kell, D.B.; Rand, D.A.; See, V.; White, M.R.H. Pulsatile Stimulation Determines Timing and Specificity of NF- κ B-Dependent Transcription. *Science* **2009**, *324*, 242–246. doi:10.1126/science.1164860.
- [41] Cheong, R.; Rhee, A.; Wang, C.; Nemenman, I.; Levchenko, A. Information Transduction Capacity of Noisy Biochemical Signaling Networks. *Science* **2011**, *334*, 354–358. doi:10.1126/science.1204553.
- [42] Levchenko, A.; Nemenman, I. Cellular noise and information transmission. *Curr. Opin. Biotechnol.* **2014**, *28*, 156–164. doi:10.1016/j.copbio.2014.05.002.

- [43] Ghorbani, M.; Jonckheere, E.A.; Bogdan, P. Gene Expression Is Not Random: Scaling, Long-Range Cross-Dependence, and Fractal Characteristics of Gene Regulatory Networks. *Front. Physiol.* **2018**, *9*, 1446:1–1446:12. doi:10.3389/fphys.2018.01446.
- [44] Gregor, T.; Tank, D.W.; Wieschaus, E.F.; Bialek, W. Probing the limits to positional information. *Cell* **2007**, *130*, 153–164. doi:10.1016/j.cell.2007.05.025.
- [45] West, J.; Bianconi, G.; Severini, S.; Teschendorff, A.E. Differential network entropy reveals cancer system hallmarks. *Sci. Rep.* **2012**, *2*, 802:1–802:8. doi:10.1038/srep00802.
- [46] Gatenby, R.A.; Frieden, B.R. The critical roles of information and nonequilibrium thermodynamics in evolution of living systems. *Bull. Math. Biol.* **2013**, *75*, 589–601. doi:10.1007/s11538-013-9821-x.
- [47] Savageau, M.A. Comparison of classical and autogenous systems of regulations in inducible operons. *Nature* **1974**, *252*, 546–549. doi:10.1038/252546a0.
- [48] Becskei, A.; Serrano, L. Engineering stability in gene networks by autoregulation. *Nature* **2000**, *405*, 590–593. doi:10.1038/35014651.
- [49] Nevozhay, D.; Adams, R.M.; Murphy, K.F.; Josic, K.; Balázsi, G. Negative autoregulation linearizes the dose-response and suppresses the heterogeneity of gene expression. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 5123–5128. doi:10.1073/pnas.0809901106.
- [50] Ramos, A.F.; Hornos, J.E.M.; Reinitz, J. Gene regulation and noise reduction by coupling of stochastic processes. *Phys. Rev. E* **2015**, *91*, 020701(R). doi:10.1103/physreve.91.020701.
- [51] Abramowitz, M. *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*, 10th ed.; Vol. 55, *National Bureau of Standards Applied Mathematics Series*, U.S. Department of Commerce, National Bureau of Standards: Washington, DC, USA, 1972.