Binary Expression Enhances Reliability of Messaging in Gene Networks

Leonardo R. Gama [1], Guilherme Giovanini [2] , 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, São Paulo CEP 05403-911, SP, Brazil

[2] Escola de Artes, Ciências e Humanidades, Universidade de São Paulo, Av. Arlindo Béttio, 1000, São Paulo CEP 03828-000, 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.ramos@usp.br

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 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 with long and equally standing ON and OFF states maximizes the mutual information and reduces entropy. That is a binary gene expression 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 environment caused by low copy numbers of reactants [1] as detected in expression of genes of prokaryotic [2] and eukaryotic cells [3] and high precision experiments demonstrating the pulsed transcription of developmental genes of social amoeba *Dictiostelium* [4], stochastic gene activation [5] and transcriptional bursting [6,7] in *Drosophila* embryos, or the bursty transcription of mammal genes [8]. Such features may lead one to ask whether in the due course of evolution there emerged mechanisms equipping biological systems with the capacity of modulating noise to exploit it functionally [9–12]. One example is the ubiquitous bursty gene expression [13,14] taking place during *Drosophila* embryogenesis when the gene networks govern multiple proteins to have proper amounts within proper spatial and temporal ranges during pattern formation [15–17]. For such patterns to be achieved, the activation and inactivation of the

promoters of the genes of developmental networks should be 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 the transduction of an input signal into a functional output by a gene network [18–20]. 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. Data for nascent transcripts in *Drosophila* embryos indicates that the position of activated promoters is in good agreement with that of their corresponding proteins during embryogenesis [13,14,21,22]. Therefore, we assume the information content of the input message to be the promoter state as the source, since the binding and unbinding of transcription factors to the enhancer induces the state of the promoter of the receptor gene [23,24]. The processing of all information arriving at the receptor is not considered here [25–28], as we are interested on the reliability of the information content of the message.

Shannon's information theory is employed to compute the entropy of the message and the mutual information between the message and the promoter state of the source [29]. Previous applications of information theory to biological problems employed a parameter-less analysis of experimental data to investigate optimal limits of information flow [19,30,31]. The absence of an underlying effective model, however, makes it difficult to biologically interpret the obtained results because of a lack of a mechanistic description of how optimal or near optimal information flow is achieved. The application of information theory to an exactly solvable stochastic model with a simple biological interpretation might help to fulfill that lacuna. Such an analysis may provide good guidance to synthetic biologists on designing gene circuits to behave as reliable (or unreliable) messaging devices. It can also be used as a basic building block useful to understand biological phenomena being described within the framework of information transmission [32,33], such as embryogenesis [34–36], carcinogenesis [37,38], or cell phenotype adaptation to fluctuating environments [39–41].

We investigate how unavoidable intracellular noise can be exploited by living organisms to produce reliable information exchange. We use the exact solutions of the stochastic binary model for gene expression formulated for coupled transcription-translation [42,43]. Although the binary model was later enhanced to describe transcription and translation as separate processes [44–46], its preliminary interpretation is sufficient to provide useful insights on the workings of a gene network. Our approximation implies, considering that, e.g., post-transcriptional regulation is absent, that translation delay is short in comparison with the relevant time scales governing the whole protein synthesis process, and that there is an approximately linear relation between the mRNA and protein numbers [47]. Experimental data for *Drosophila* embryos showing a good agreement between spatial patterns of transcripts and transcription factors [13,21] justify our approximation.

The remaining parts of this manuscript are organized as follows. Section 2 presents our picture of a gene network as a communicating system, and the equations used to compute the probabilities, noise, average values, entropy and mutual information for the binary gene. Our results are presented at Section 3 as a set of graphs relating the entropy and mutual information to gene expression noise, promoter switching speeds, average number of transcription factors and their corresponding probability distributions. A discussion of how our results relate to a non-exhaustive set of previous experimental or theoretical analysis related with our investigation is presented at Section 4. The results and analysis presented here contribute to the understanding of the detrimental role of noise for the proper development of an organism, or its benefits for the adaptation and robustness of living systems challenged by an ever changing environment.

2. Model and Methods

To set a reference value for a comparative analysis, we will evaluate the entropy of the message sent by a constitutive and by a binary gene at the steady state regime. The promoter of the constitutive gene is always ON while it switches between ON and OFF states on the latter [48]. We assume coupled transcription-translation for a clearer interpretation of our results. Our analysis focus on understanding: *i.* the entropy of a message sent by both the constitutive and binary sources; *ii.* the mutual information between the promoter state and the number of transcription factors sent by the binary source. These enable the investigation of the conditions for a binary source to send a reliable or non-reliable message about its state being ON or OFF. In this section, we present the formulas used to perform our calculation and their biological interpretation. The codes for performing the numerical calculations presented in this manuscript are available online (https://github.com/amphybio/stochastic-gene-expression).

2.1. Qualitative Model and Noise Quantification

Figure 1 shows the effective set of reactions that we 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 factor 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.
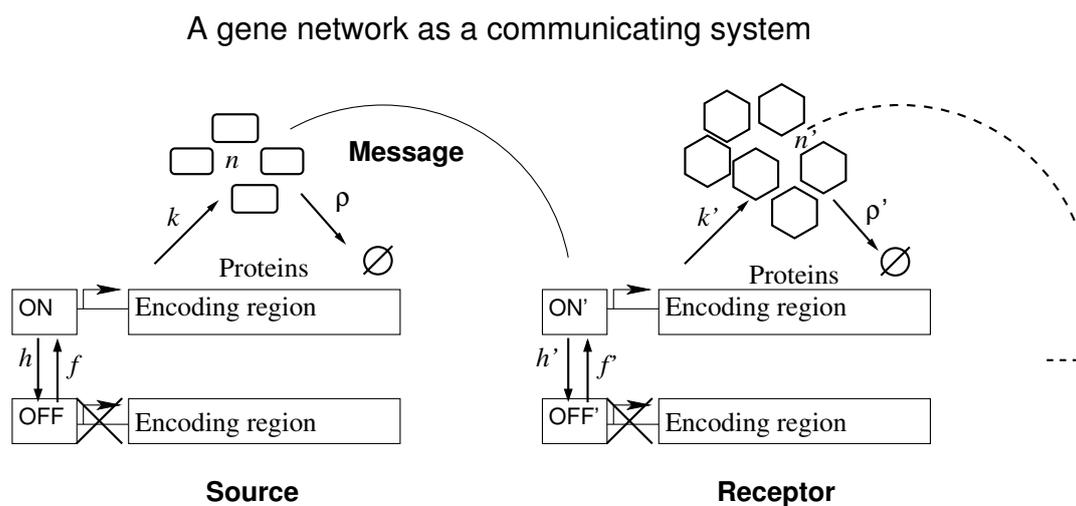
A gene network as a communicating system

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.

The number of transcription factors is a random variable and the noise of the message can be computed using the Fano factor, defined as the ratio between the variance of n to its average value. The average number of transcription factors (or expected message from the source) is denoted by $\langle n \rangle$, and the Fano factor by $\mathcal{F}$, such that:

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

The Fano factor is a measure of how similar to 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 $\mathcal{F} > 1$.

2.2. 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

$$\phi_n = \frac{e^{-N}}{n!} N^n \,, \tag{2}$$

where

$$N = \frac{k}{\rho} \,, \tag{3}$$

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.

2.3. A 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 k (or 0) and degradation happens at rate ρ. There are two random variables, (m, n), where m indicates the state of the promoter as ON or OFF and n 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 for this model [42,43] we have:

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

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

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

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

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

ϕ_n indicates the marginal probability of finding n proteins, namely $\phi_n = \alpha_n + \beta_n$, and $M(a, b, x)$ 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 $(a)_n = a(a + 1) \ldots (a + n - 1)$ denotes the Pochhammer symbol [49].

2.3.1. Interpretation of the Parameters P_α, ϵ, and N.

The parameters p_α and p_β are the steady state marginal probabilities for the promoter to be at the ON and OFF states, namely $p_\alpha = \sum_{n=0}^{\infty} \alpha_n$ and $p_\beta = \sum_{n=0}^{\infty} \beta_n$ with $p_\alpha + p_\beta = 1$. p_α (or p_β) give the average fraction of a time interval T that the promoter will be ON (or OFF) while switching at random between these states.

For the promoter being exclusively ON the source would become constitutive (see Section 2.2) with average number of transcription factors being N (see Equation (3) or Equation (7)).

The parameter ϵ is the ratio of the promoter switching rate to the transcription factors degradation rate. It can also be interpreted as the ratio of the transcription factor's lifetime to the average period of a complete cycle of promoter switching (see Equation (7)) for a fixed p_α. For simplicity, we may assume $T_D = \frac{1}{\rho}$ as the protein lifetime, and $T_S = \frac{f+h}{fh}$ as the average time for the promoter to complete one switching cycle. Using the definitions at Equation (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 the protein's 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.

2.3.2. 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 or OFF (denoted by $\langle n_\alpha \rangle$ or $\langle n_\beta \rangle$) are

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

The mean number of transcription factors of the binary gene depends on the proportion of time that the promoter of the gene is ON. Hence, the average amount of transcription factors will range from zero to N accordingly with p_α. The conditional mean number of transcription factors is obtained using $\langle n_\alpha \rangle = \frac{1}{p_\alpha} \sum_{n=0}^{\infty} n \alpha_n$ and $\langle n_\beta \rangle = \frac{1}{p_\beta} \sum_{n=0}^{\infty} n \beta_n$ which can be obtained using the generating function technique [43,45]. The mean number of transcription factors can be written as $\langle n \rangle = p_\alpha \langle n_\alpha \rangle + p_\beta \langle n_\beta \rangle$.

For very small values of ϵ, $\langle n_\alpha \rangle \to N$ and $\langle n_\beta \rangle \to 0$. The binary gene behaves as a constitutive one when its promoter is ON and as a silenced one when its promoter is OFF. The average number is a quantity between these two values that does not provide a useful information about the instantaneous state of the system. However, measurement of the instantaneous number of transcription factors is useful to understand the promoter state.

For large values of ϵ we have $\langle n_\alpha \rangle$, $\langle n_\beta \rangle \sim \langle n \rangle$. The state of the promoter is indistinguishable by simply measuring the number of transcription factors. At that limit, the source approaches the behavior of a constitutive one.

2.3.3. The Variance and Bursting Noise of N on the Stochastic Model for a Binary Gene

The variance of n, denoted by σ^2, and the Fano factor ($\mathcal{F}$) when the transcription factors are produced from a binary gene are, respectively,

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

Since the discussion about the noise is similar for both σ^2 and $\mathcal{F}$ we focus on the second. Because of the promoter switching, the value of $\mathcal{F}$ is larger than that of a constitutive one by a quantity $\delta = N(1 - p_\alpha)/(1 + \epsilon)$.

The Fano factor is computed using the marginal probability ϕ_n, the probability of finding n transcription factors independently of the promoter state. The increment δ helps to understand the hyper-Poissonian nature of the message sent by the binary source. A Poissonian distribution describing the constitutive source has $\delta = 0$ while in a super-Poissonian distribution describing the message sent

by the binary source we have $\delta > 0$. The super-Poissonian distribution describes multiple qualitatively distinct gene expression regimes.

For $p_\alpha = 1$ the promoter is fully ON, the binary source becomes constitutive and $\delta = 0$. For $p_\alpha = 0$ we have a fully repressed gene. Additionally, when ϵ approaches infinity the noise of the message sent by the binary gene is similar to that of a constitutive one if N and p_α are finite constants. For (N, p_α) being finite constants and ϵ approaching zero we have $\delta \to N(1 - p_\alpha)$.

For instance, a bursty expression characterized by $k, h \gg \rho$ and $h \gg f$ or $N, \epsilon \gg 1$ and $p_\alpha \ll 1$ (see Equation (7)) has $\delta \approx k/h$ or $\delta \approx N/\epsilon$ which is called the burst size. This is a mathematical entity with interpretation as transcriptional or translational dependent on the meaning of the random variable n, which in here denotes the number of transcription factors. Hence, we interpret this limit as the translational bursting and δ as the translational burst size. Note, however, that in literature it has also been interpreted as the transcriptional bursting size [44,46,50–52]. Here, the burst size $\delta \sim k/h$ coincides with the noise increment in relation to the noise in the expression of the constitutive gene. In that case, the promoter remains OFF for a long interval, and switches ON for a brief time interval during which a burst takes place. Once the promoter switches OFF the amount of gene products decays exponentially until the next burst.

The binary expression, on the other hand, is characterized by $f + h \ll \rho$ or $\epsilon \ll 1$. The number of synthesized transcription factors is governed by a bimodal distribution which peaks are centered around 0 and N weighted by the probabilities p_β and p_α, respectively. During steady state regime, the trajectories of the number of transcription factors governed by a bimodal distribution can be described approximately as follows. Initially, we set the promoter OFF and no transcription factors, then there will be, at random instants: an increase of n towards a finite value $\langle n_\alpha \rangle \approx N$ after the promoter switching ON; fluctuations of n around $\langle n_\alpha \rangle$ corresponding to a Poissonian process while the promoter remains ON; a sharp reduction towards zero after the promoter switching OFF; then $n = 0$ while the promoter remains OFF; and the cycle repeats. Let us set $p_\alpha = 1/2$ and take the limit of $\epsilon \ll 1$ such that $\mathcal{F} \to 1 + N/2$. We set the promoter OFF and zero transcription factors at the initial instant. Once the promoter switches ON, n increases towards $\sim N$ as if it were a time-dependent Poissonian process because the promoter state remains stable for a sufficiently long time interval. For the case of a fast growth of n in comparison with the remaining time scales of the system, one might take it as a burst which size would not correspond to the increment in noise, δ, but to 2δ, and an analysis of the full trajectory would be necessary to understand this noise regime [22,52].

The two examples above points to the necessity of a more precise classification of noise and its biological implications. Information theory constitutes a useful toolbox for that purpose.

2.4. Analyzing the Information Content of the Message

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.

2.4.1. 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 with average value N the entropy is given by

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

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

2.4.2. 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:

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

where ϕ_n is given by Equation (6) and computed setting numerical values to its parameters given at Equation (7).

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

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

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

These formulas are obtained from the definition of the conditional entropy of X given $Y = y$, namely $H(X|Y = y) = -\sum_{x \in X} p(X|Y = y) \log p(X|Y = y)$, 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 Equations (4), (5) and (7).

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

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

where $H(n)$, $H_\alpha(n)$, and $H_\beta(n)$ are, respectively, given by Equations (11)–(13).

3. Results

We investigate the reliability of the information transfer considering the unavoidability of noise by characterizing the dynamics of the promoter and the distributions governing the message (gene product) generated by the source gene. Section 3.1 shows an analysis of the entropy and mutual information as a function of the Fano factor to motivate the hypothesis that noise properties are actively regulated by cells. Section 3.2 presents a characterization of the promoter switching dynamics to enable the condition of reduced entropy and increased mutual information of the message. Since the reliability of the signal is evaluated by means of the average signal, in Section 3.3 we investigate the behavior of the entropy and mutual information as functions of the average message. We also analyze the conditional entropy as a function of the average conditional message and determine the corresponding probability distributions for three distinctive switching dynamics: fast, slow, and intermediary.

Each subsection is structured with a set of graphs displaying the results obtained for a set of parameter values (see Equation (7)) indicated inside the graphs and on the figures' titles. We also present a description of the graphs and the data obtained using the corresponding equations (see Section 2). A final paragraph is dedicated to discussing and interpret the results using the qualitative picture at Figure 1.

3.1. Binary Expression Enables Entropy Reduction and Mutual Information Increase

The horizontal axis of both graphs of Figure 2 shows the bursting quantified by the Fano factor (see Equation (1)). The vertical axis of graphs A and B shows, respectively, the entropy of the message

(see Equation (11)) and the mutual information between the message and the state of the promoter of the source (see Equation (14)). Each continuous line is associated with a specific value of the promoter switching rate, ϵ (see Equation (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. Please 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 $\epsilon >> N$ which corresponds to a switching cycle being much faster than the protein synthesis rate. In that case, the mutual information approaches zero.
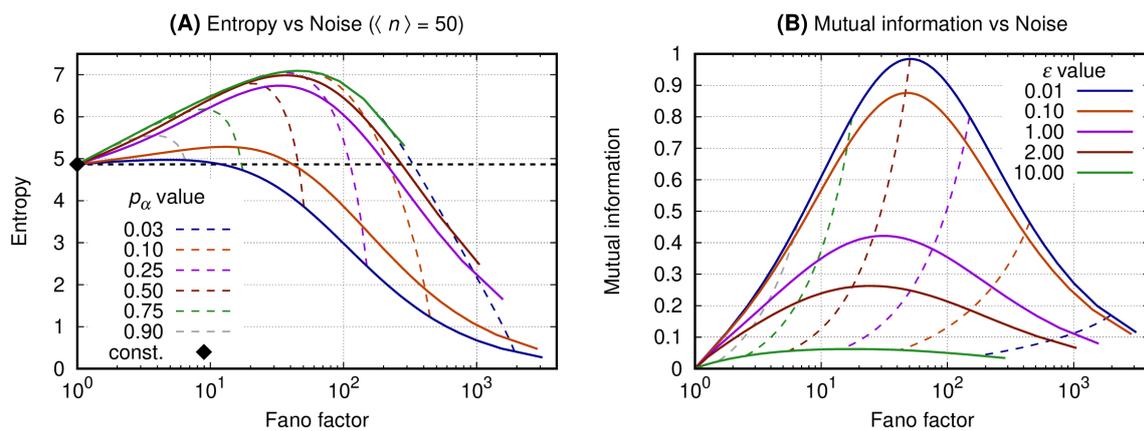
Figure 2. Entropy (or mutual information) and Fano factor were obtained varying N for fixed p_α and ϵ. All curves have the same value for the average number of gene products ($\langle n \rangle = 50$ such that for each curve we have $N = 50/p_\alpha$). Graph (**A**) shows the entropy of the message sent 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 noise. Continuous lines refer to fixed values of ϵ and fixed values of p_α are indicated by the dashed ones.

Graph A shows that the constitutive gene sets a reference value for the entropy for a fixed $\langle n \rangle$. The binary source at the small noise regime ($\mathcal{F} \sim 1$) generates a message which entropy is greater than that sent by the constitutive gene. As the noise 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. Moreover, for a fixed value of $\mathcal{F}$, the message with minimal entropy is generated by a process with 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 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 $(\mathcal{F} - 1)(1 + \epsilon) \sim \mathcal{F}$.

The results presented here show that one may increase the expression noise of a binary gene to generate messages with arbitrarily small entropy. The reduction of the entropy and of its maximal value depends both on the message's noise and on the probabilities governing the promoter states. Hence, reducing the entropy of the message is not sufficient for optimizing its reliability, so an analysis of the mutual information between the message and the promoter state is necessary. Indeed, the mutual information has a maximum when the probability for the promoter to be ON is $1/2$ and the noise

increment δ is approximately the average number of gene products. Therefore, besides promoter state probabilities, the noise in gene expression must also be regulated for the maximization of the information content of the message sent by the binary source.

3.2. The Slow Switching Genes Generate Reduced Entropy and Increased Values for Mutual Information

The horizontal axis of both graphs of Figure 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 Equation (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 Equation (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.
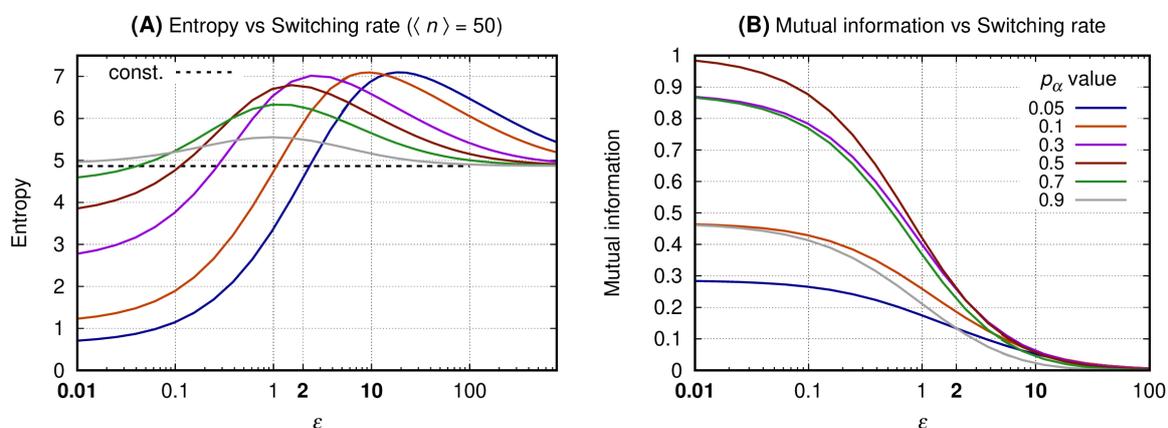

Figure 3. The entropy (and mutual information) was obtained fixing p_α and N. All curves have the same average number of transcription factors ($\langle n \rangle = 50$ such that for each line we have $N = 50/p_\alpha$). The horizontal axis of both graphs indicates the transcription factor's lifetime 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 Figures 4–6 appear in bold at the horizontal axis label.

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 with $\epsilon \ll 1$ 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 Figure 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.

Equation (9) shows that the increment δ on the noise of the message sent by the binary gene 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 method for increasing δ. Indeed, here we have obtained reduced entropy and maximal mutual information for the slow switching genes. As expected, a very fast switching gene will generate a message with entropy approaching that of the constitutive gene with the same average expression levels. An additional regime is the bursting that takes place when we have $\epsilon, N \gg 1$ and $p_\alpha \ll 1$ [52]. This regime leads to a strong reduction on the entropy of the message but its mutual information is low. Hence, it would not be reliable in the context of a gene network in which the information content of the message is the promoter state.

3.3. 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 Equation (8)); graph C shows the mutual information versus $\langle n \rangle$; graph D shows the probability distribution of n. Each continuous line corresponds 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.
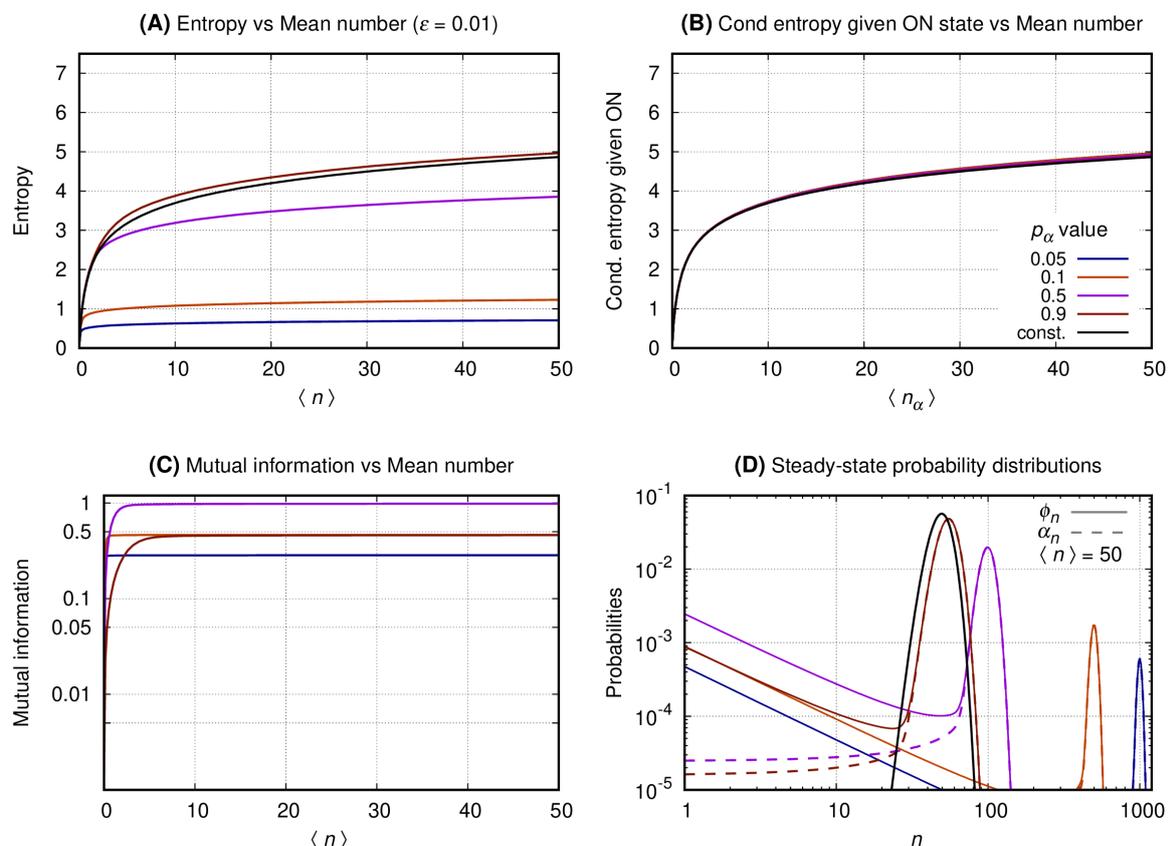

Figure 4. The switching rate of the promoter of the source is 0.01. The vertical axis of Graph's (**A**–**C**) shows, respectively, entropy versus $\langle n \rangle$, entropy conditional to the source being ON versus $\langle n_\alpha \rangle$, and the mutual information versus $\langle n \rangle$.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

Figure 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 with 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 approaches 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 approximates unity with the taller peak approaching that of the constitutive source for the same mean.

Figure 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 entropies 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 $p_\alpha \to 1$ 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. *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 entropies 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 $p_\alpha \to 1$.
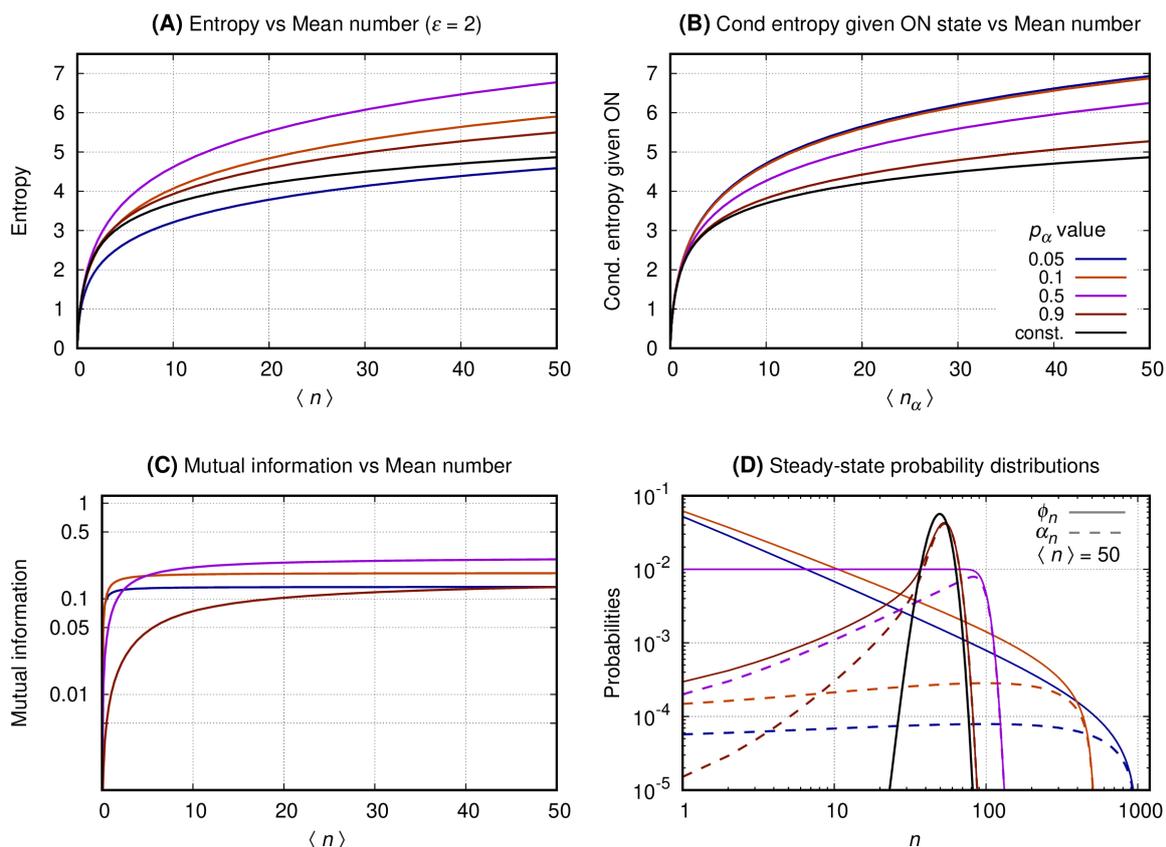
Figure 5. The switching rate of the promoter of the source is 2. The Graph's (**A**–**C**) shows, respectively, entropy versus $\langle n \rangle$, entropy conditional to the source being ON versus $\langle n_\alpha \rangle$, and the mutual information versus $\langle n \rangle$. Graph (**D**) shows the probability distribution of n.

The graphs of Section 3.3 show that at maximal mutual information and reduced entropy, the distributions governing the message are bimodal. That implies the 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 is 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.
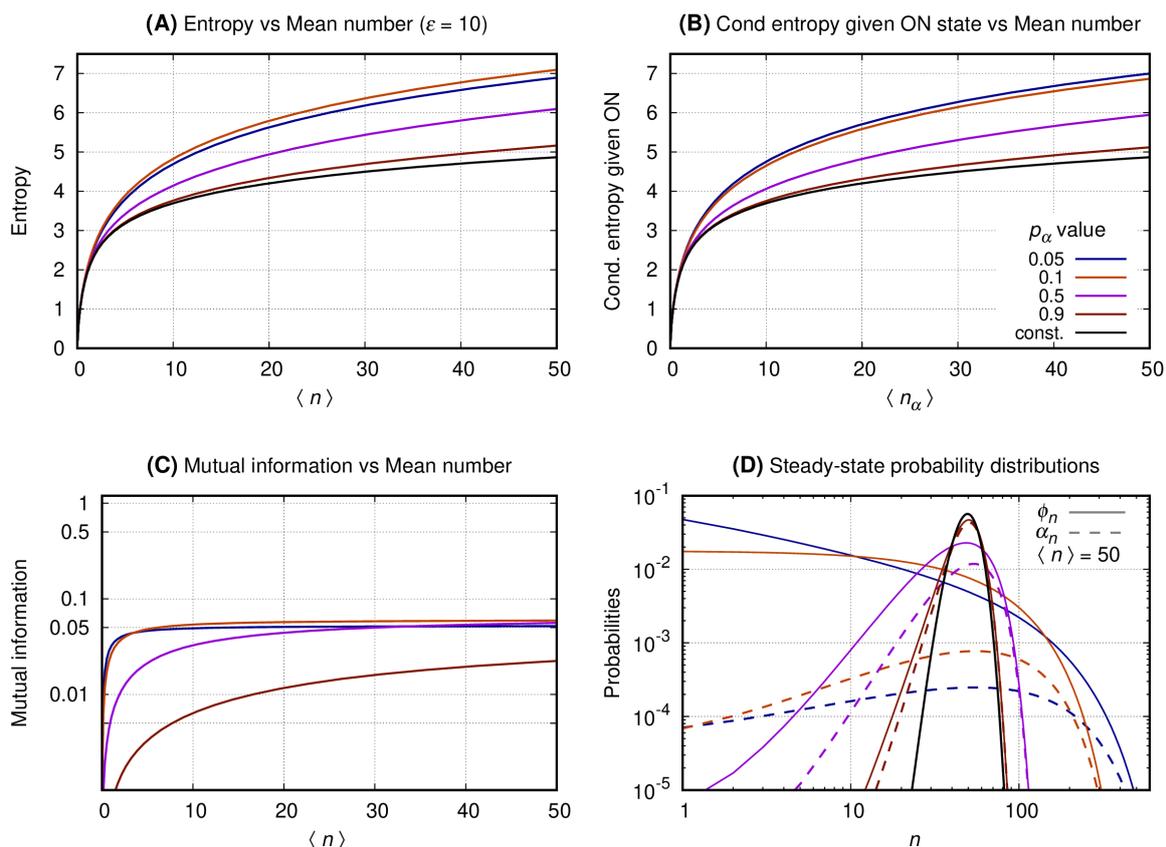
Figure 6. The switching rate of the promoter of the source is 10. The vertical axis of Graph's (**A**–**C**) shows, respectively, entropy versus $\langle n \rangle$, entropy conditional to the source being ON versus $\langle n_\alpha \rangle$, and the mutual information versus $\langle n \rangle$. Graph (**D**) shows the probability distribution of n.

4. Discussion

The advent of fluorescence techniques generating precise molecular level data has encouraged the use of theoretical approaches on the search for general principles underpinning biological processes. Here we contribute to these efforts by means of an information theoretic analysis of the well-known stochastic binary model for gene expression. The bimodal probability distribution obtained from this model, characterized by large noise, optimizes the mutual information (see Figure 2 or [19,30,31]). Because our results are derived using an exactly solvable master equation modeling a small set of effective reactions, they provide guidance for the interpretation and description of experimental results. Though we are not claiming that we have a model for expression of the *Drosophila* embryo developmental gene *hunchback*, we use its expression as an example. The histogram of the levels of Hunchback protein along the AP axis of a *Drosophila* embryo at cycle 14 is also bimodal, and characterize a near optimal information flow in transcriptional regulation [19]. Expression of *hunchback* is modulated by self-activation, repression by Knirps and Kruppel transcription factors, and activation by the maternal protein Bicoid [53]. In a "first order" approximation, however, one may assume Bicoid as a major regulator of *hunchback* gene such that with some precision [53], its promoter state is a readout of the concentration of Bicoid—low concentration of Bicoid implies on the promoter being OFF and vice-versa [36]. In that case, the bimodality of the distribution of Hunchback suggests that its promoter has a slow switching dynamics and, since the state of the promoter is regulated by multiple transcription factors [54], we assume that their binding and unbinding holds its promoter at either ON and OFF states for a relatively long interval. This incipient analysis is important to be completed because of the similar kinetics of the gap genes despite their distinctive spatial patterns [14]. Furthermore, the number of elongating transcripts in a gene with promoter controlled

by distinct developmental enhancers of *Drosophila* exhibit trajectories consistent with the bimodal distributions [22]. Hence, mapping the biological systems in which binary expression occurs may provide useful general principles to understand the functioning of gene networks in the context of development.

A major challenge in Biology is to classify noise in gene expression accordingly with the role of its products in a functional network or a cell. For example, self-repression has long been recognized as the mechanism responsible for noise reduction or homeostasis in gene expression [55–58]. Our results indicate that in a gene network in which the relevant information is the promoter state, the binary expression is the proper regime to increase the reliability and information content of the message for a wide range of the noise amplitudes. Another noise regime of gene expression is the bursting, which generates a message with minimal entropy and low mutual information (Figure 2). The bursts have been reported in many recent publications, e.g., see Table 1 [59] for a compilation showing average values of burst size and frequency among $\sim$ 4k and $\sim$ 2k genes of, respectively, mouse and human fibroblasts cells [8,60–64], or two genes in yeast [51]. The recurrence of observations of burstings indicates the necessity of additional analysis using more general models [59,65] to understand the flow of information in gene networks with bursty components. One may consider that the bursty genes are participants of networks in which the relevant information is not the promoter state, and they may even not need to operate at a near optimal regime of promoter state information transmission. These three examples, self-repression, binary and bursty expression, lead us to *conjecture* that evolution of living systems do not necessarily select or avoid selecting gene networks for maximal information flow, but for near optimal operation under the necessity of sufficiently reliable function and near minimal energy consumption.

Information transfer in biological systems also play a role on phenotypic switching for adaptation to constantly changing environments. When the environment fluctuates, two strategies (and their combinations) are possible [40]: (*i*) sensing and response; and (*ii*) blind, random switching. In case (*i*) we a assume a cell which survival in a given environment depends on the presence of a specific protein. The occurrence of the environment leads to the expression of the specific protein by activation of its corresponding promoter. The optimal response regime would require the protein numbers to be a readout of the promoter state and, for a fluctuating environment, the trajectories of the protein numbers would be similar to those described for bimodal distributions (see Section 2.3.3). The duration of the high concentrations of the responding protein would be similar to duration of the demanding environment. Note, however, that the chemicals responsible for activating the responding gene may have a specific production regime optimized for detection of environmental changes. In case (*ii*) there is a random switching that does not respond to the environment [39,41]. How do the cells implement the optimal switching rates? For instance, they could just have a randomly switching gene with arbitrary, optimal ON→OFF and OFF→ON rates. However, the phenotype comes from the protein level, not from the gene. So, if the randomly switching gene's ON/OFF states are not faithfully followed by the protein, the phenotypic switching will not be optimal, even though the gene switching is optimal. So, for optimal random phenotypic switching, there should be optimal information transfer from the gene's state to the protein levels.

Our results may also help unravel why transcription factors are present in lower copy numbers inside the cells when compared with, e.g., glycolytic or elongating proteins (see the online book http://book.bionumbers.org/what-are-the-copy-numbers-of-transcription-factors/ related to BioNumbers webpage—[66]. Assume that the transcription factors are the carriers of information about the promoter state and that their interaction with the regulatory regions of the receptor gene depends on their quantity. One may wonder whether there is a threshold above which the amount of transcription factors is an unnecessary burden for the cell because of lack of additional information and their near certainty of binding to the DNA. Indeed, Section 3.3 shows that the mutual information between the message and promoter state saturates for a finite value of the average number of transcription factors while entropy grows monotonically. It would be interesting to connect this

result with the functioning of specific gene networks. As previously discussed, gene networks may or may not function transmitting information about promoter states and further analysis is necessary to understand additional possibilities relating a gene network function and its noise regime.

For example, 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. Assuming that all gene expression noise regimes have a functional role, one may wonder whether there exist a DNA structure surrounding a promoter that regulates its switching rates and, consequently, whether its expression will be binary, bursty, constitutive-like accordingly with its context necessities. That would be interesting to investigate experimentally to relate an expression regime with its functional role. Additionally, one might also 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.

References

1. Delbrück, M. Statistical fluctuations in autocatalytic reactions. *J. Chem. Phys.* **1940**, *8*, 120–124, doi:10.1063/1.1750549. [CrossRef]

2. 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. [CrossRef]

3. 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. [CrossRef]

4. Chubb, J.R.; Trcek, 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. [CrossRef] [PubMed]

5. 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. [CrossRef] [PubMed]

6. 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. [CrossRef]

7. 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. [CrossRef]

8. 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. [CrossRef]

9. 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. [CrossRef]

10. 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. [CrossRef]

11. 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. [CrossRef] [PubMed]

12. 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. [CrossRef] [PubMed]

13. 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. USA* **2014**, *111*, 10598–10603, doi:10.1073/pnas.1410022111. [CrossRef]

14. 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. [CrossRef] [PubMed]

15. 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. [CrossRef] [PubMed]

16. Falo-Sanjuan, J.; Bray, S.J. Decoding the Notch signal. *Dev. Growth Differ.* **2019**, *62*, 4–14, doi:10.1111/dgd.12644. [CrossRef] [PubMed]

17. 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. USA* **2020**, *117*, 836–847, doi:10.1073/pnas.1912500117. [CrossRef]

18. 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. [CrossRef] [PubMed]

19. Tkačik, G.; Callan, C.G., Jr.; Bialek, W. Information flow and optimization in transcriptional regulation. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 12265–12270, doi:10.1073/pnas.0806077105. [CrossRef]

20. Bowsher, C.G.; Swain, P.S. Identifying sources of variation and the flow of information in biochemical networks. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E1320–E1328, doi:10.1073/pnas.1119407109. [CrossRef]

21. 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. [CrossRef] [PubMed]

22. Fukaya, T.; Lim, B.; Levine, M. Enhancer Control of Transcriptional Bursting. *Cell* **2016**, *166*, 358–368, doi:10.1016/j.cell.2016.05.025. [CrossRef] [PubMed]

23. Reinitz, J.; Hou, S.; Sharp, D.H. Transcriptional control in Drosophila. *Complexus* **2003**, *1*, 54–64, doi:10.1159/000070462. [CrossRef]

24. 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. [CrossRef] [PubMed]

25. 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. [CrossRef]

26. 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. [CrossRef] [PubMed]

27. 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. USA* **2018**, *115*, 6088–6093, doi:10.1073/pnas.1716659115. [CrossRef]

28. Liu, Y.; Barr, K.; Reinitz, J. Fully Interpretable Deep Learning Model of Transcriptional Control. *bioRxiv* **2019**, 1–11, doi:10.1101/655639. [CrossRef]

29. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423, doi:10.1002/j.1538-7305.1948.tb01338.x. [CrossRef]

30. 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.; et al. Pulsatile Stimulation Determines Timing and Specificity of NF-κB-Dependent Transcription. *Science* **2009**, *324*, 242–246, doi:10.1126/science.1164860. [CrossRef]

31. 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. [CrossRef]

32. Levchenko, A.; Nemenman, I. Cellular noise and information transmission. *Curr. Opin. Biotechnol.* **2014**, *28*, 156–164, doi:10.1016/j.copbio.2014.05.002. [CrossRef]

33. 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. [CrossRef] [PubMed]

34. Reinitz, J.; Mjolsness, E.; Sharp, D.H. Cooperative control of positional information in *Drosophila* by *bicoid* and maternal *hunchback*. *J. Exp. Zool.* **1995**, *271*, 47–56. [CrossRef]

35. Jaeger, J.; Surkova, S.; Blagov, M.; Janssens, H.; Kosman, D.; Kozlov, K.N.; Myasnikova, E.; Vanario-Alonso, C.E.; Samsonova, M.; Sharp, D.H.; et al. Dynamic control of positional information in the early *Drosophila* embryo. *Nature* **2004**, *430*, 368–371. [CrossRef]

36. 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. [CrossRef] [PubMed]

37. 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. [CrossRef] [PubMed]

38. 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. [CrossRef] [PubMed]

39. Balaban, N.; Merrin, J.; Chait, R.; Kowalik, L.; Leibler, S. Bacterial persistence as a phenotypic switch. *Science* **2004**, *305*, 1622–1625.

40. Kussell, E.; Leibler, S. Phenotypic diversity, population growth, and information in fluctuating environments. *Science* **2005**, *309*, 2075–2078. [CrossRef]

41. Belete, M.K.; Balázsi, G. Optimality and adaptation of phenotypically switching cells in fluctuating environments. *Phys. Rev. E* **2015**, *92*, 062716. [CrossRef] [PubMed]

42. Peccoud, J.; Ycart, B. Markovian modelling of gene product synthesis. *Theor. Popul. Biol.* **1995**, *48*, 222–234, doi:10.1006/tpbi.1995.1027. [CrossRef]

43. 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. [CrossRef] [PubMed]

44. Raj, A.; Peskin, C.S.; Tranchina, D.; Vargas, D.Y.; Tyagi, S. Stochastic mRNA Synthesis in Mammalian Cells. *PLoS Biol.* **2006**, *4*, 1707–1719. [CrossRef] [PubMed]

45. 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. [CrossRef] [PubMed]

46. Shahrezaei, V.; Swain, P.S. Analytical distributions for stochastic gene expression. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 17256–17261, doi:10.1073/pnas.0803850105. [CrossRef]

47. Paulsson, J. Models of stochastic gene expression. *Phys. Life Rev.* **2005**, *2*, 157–175. [CrossRef]

48. 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. [CrossRef]

49. Abramowitz, M.; Stegun, I.A. (Eds.) *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*, 10th ed.; National Bureau of Standards Applied Mathematics Series; U.S. Department of Commerce, National Bureau of Standards: Washington, DC, USA, 1972; Volume 55.

50. Thattai, M.; van Oudenaarden, A. Intrinsic noise in gene regulatory networks. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 8614–8619, doi:10.1073/pnas.151588598. [CrossRef]

51. Zenklusen, D.; Larson, D.R.; Singer, R.H. Single-RNA counting reveals alternative modes of gene expression in yeast. *Nat. Struct. Mol. Biol.* **2008**, *15*, 1263. [CrossRef]

52. Yvinec, R.; da Silva, L.G.S.; Prata, G.N.; Reinitz, J.; Ferreira Ramos, A. Bursting on a two state stochastic model for gene transcription in Drosophila embryos. *bioRxiv* **2017**, doi:10.1101/107979. [CrossRef]

53. Manu, S.S.; Spirov, A.V.; Gursky, V.V.; Janssens, H.; Kim, A.R.; Radulescu, O.; Vanario-Alonso, C.E.; Sharp, D.H.; Samsonova, M.; Reinitz, J. Canalization of gene expression in the Drosophila blastoderm by gap gene cross regulation. *Plos Biol.* **2009**, *7*, e1000049. [CrossRef] [PubMed]

54. Bothma, J.P.; Garcia, H.G.; Ng, S.; Perry, M.W.; Gregor, T.; Levine, M. Enhancer additivity and non-additivity are determined by enhancer strength in the *Drosophila* embryo. *eLife* **2015**, *4*, e07956, doi:10.7554/eLife.07956. [CrossRef]

55. Savageau, M.A. Comparison of classical and autogenous systems of regulations in inducible operons. *Nature* **1974**, *252*, 546–549, doi:10.1038/252546a0. [CrossRef] [PubMed]

56. Becskei, A.; Serrano, L. Engineering stability in gene networks by autoregulation. *Nature* **2000**, *405*, 590–593, doi:10.1038/35014651. [CrossRef] [PubMed]

57. 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. USA* **2009**, *106*, 5123–5128, doi:10.1073/pnas.0809901106. [CrossRef]

58. 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. [CrossRef]

59. Cao, Z.; Grima, R. Analytical distributions for detailed models of stochastic gene expression in eukaryotic cells. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 4682–4692, doi:10.1073/pnas.1910888117. [CrossRef]

60. Ochiai, H.; Sugawara, T.; Sakuma, T.; Yamamoto, T. Stochastic promoter activation affects Nanog expression variability in mouse embryonic stem cells. *Sci. Rep.* **2014**, *4*, 1–9. [CrossRef]

61. Senecal, A.; Munsky, B.; Proux, F.; Ly, N.; Braye, F.E.; Zimmer, C.; Mueller, F.; Darzacq, X. Transcription factors modulate c-Fos transcriptional bursts. *Cell Rep.* **2014**, *8*, 75–83. [CrossRef]

62. Halpern, K.B.; Tanami, S.; Landen, S.; Chapal, M.; Szlak, L.; Hutzler, A.; Nizhberg, A.; Itzkovitz, S. Bursty gene expression in the intact mammalian liver. *Mol. Cell* **2015**, *58*, 147–156. [CrossRef]

63. Skinner, S.O.; Xu, H.; Nagarkar-Jaiswal, S.; Freire, P.R.; Zwaka, T.P.; Golding, I. Single-cell analysis of transcription kinetics across the cell cycle. *Elife* **2016**, *5*, e12175. [CrossRef] [PubMed]

64. Larsson, A.J.; Johnsson, P.; Hagemann-Jensen, M.; Hartmanis, L.; Faridani, O.R.; Reinius, B.; Segerstolpe, Å.; Rivera, C.M.; Ren, B.; Sandberg, R. Genomic encoding of transcriptional burst kinetics. *Nature* **2019**, *565*, 251–254. [CrossRef] [PubMed]

65. Xu, H.; Skinner, S.O.; Sokac, A.M.; Golding, I. Stochastic kinetics of nascent RNA. *Phys. Rev. Lett.* **2016**, *117*, 128101. [CrossRef] [PubMed]

66. Milo, R.; Jorgensen, P.; Moran, U.; Weber, G.; Springer, M. BioNumbers—the database of key numbers in molecular and cell biology. *Nucleic Acids Res.* **2009**, *38*, D750–D753, doi:10.1093/nar/gkp889. [CrossRef]