A COMMON REPRESSOR POOL RESULTS IN INDETERMINACY OF EXTRINSIC NOISE

Michail Stamatakis, Rhys Adams and Gábor Balázsi[*]
Department of Systems Biology, The University of Texas M. D. Anderson Cancer
Center, Houston, TX 77054, USA

Abstract

For over a decade, noise in gene expression has been the focus of experimental and theoretical studies. It is currently accepted that gene expression noise can be decomposed into extrinsic and intrinsic components, which have orthogonal contributions to the total noise. Intrinsic noise stems from the random occurrence of biochemical reactions and is inherent to gene expression. Extrinsic noise originates from fluctuations in the concentrations of regulatory components or random transitions in the cell's state and is imposed to the gene of interest by the intra- and extra-cellular environment. The basic assumption has been that extrinsic noise acts as a pure input on the gene of interest, which exerts no feedback on the extrinsic noise source. Consequently, multiple copies of a gene would be uniformly influenced by an extrinsic noise source. Here we report that this assumption falls short when multiple genes share a common pool of a regulatory molecule. Due to competitive utilization of the molecules existing in this pool, genes are no longer uniformly influenced by the extrinsic noise source. Rather, they exert negative feedbacks on each other. Thus, extrinsic noise calculated using the currently established method becomes ill-defined.

Keywords

Two promoter system, Intrinsic noise, Extrinsic noise, Indeterminacy, Chemical Master Equation, Negative feedback, Stochastic gene expression.

Introduction

For just over a decade, stochastic gene expression has been the focus of several experimental and theoretical studies. It is now widely accepted that noise in gene expression can be decomposed to extrinsic and intrinsic components, which have orthogonal contributions to the total noise, in the sense that their variances add up to the total variance (Swain et al. 2002). Intrinsic noise stems from the random occurrence of biochemical reactions and is inherent to gene expression. Extrinsic noise originates from fluctuations in the concentrations of regulatory components or random transitions in the cell's state and is imposed to the gene of interest by the intra- and extra-cellular environment.

This approach treats the gene expression process as a "noisy-machine" (Figure 1) which receives a noisy signal and performs a random transformation of that signal. Thus, the output of this machine will contain two noise components: the input (extrinsic) noise and the generated (intrinsic) noise. Considering two identical, non-interacting noisy machines, one would expect that their outputs have (i) perfectly correlated extrinsic components, since both "machines" receive the same input, (ii) perfectly uncorrelated intrinsic components, since the "machines" do not interact. Under these assumptions one can define the *extrinsic variables* as the noisy input variables to the "noisy-machine", and the *intrinsic variables* as the state variables of the "noisy-machine". Swain et. al (2002) give examples of extrinsic variables being the state of the cell or the contents of regulatory components, and examples of intrinsic

---

variables being the mRNA and protein contents for the gene of interest and demonstrate how one can calculate the extrinsic and intrinsic noise components in theory and experiment.

This approach is based on two assumptions, which we will refer to as follows: (i) The "pure-input assumption" stating that the gene of interest has no way to exert any feedback back to the extrinsic noise source; therefore, extrinsic noise acts as a pure-input to the process of gene expression. This assumption allows one to perform the partitioning of extrinsic and intrinsic variable as proposed by Swain et al. (2002). (ii) The "independent-genes assumption" stating that multiple identical copies of a gene do not influence one-another. This assumption allows one to calculate the intrinsic and extrinsic noise components by measuring the protein contents of two identical gene copies (otherwise, Eq. 9 in Swain et al. 2002 would not be valid because $P(\mathbf{E},\mathbf{I_1},\mathbf{I_2}) \neq P(\mathbf{E},\mathbf{I_1}) \cdot P(\mathbf{E},\mathbf{I_2})$).

Even though the decomposition of noise into extrinsic and intrinsic components is a useful tool for analyzing stochasticity in gene expression, such an analysis is ill-defined unless both of the aforementioned assumptions hold true. The aim of this paper is to demonstrate a case where these assumptions are violated and thus one cannot even partition the overall noise into extrinsic and intrinsic components. We will use mathematical modeling to a system similar to that used experimentally by (Elowitz et al. 2002), that comprises two reporter-gene variants under the influence of identical promoters repressed by LacI. We will show that the "pure-input assumption" does not hold, because the state of the gene (repressed or not) affects the free repressor content, and neither does the "independent-genes assumption", because if one gene is in the repressed state this prevents one repressor molecule from repressing the other gene. Thus, the competitive utilization of a common pool of repressor molecules results in mutual negative feedback interactions and subsequently in negatively correlated protein contents. As a matter of fact if one tries to calculate the extrinsic noise for this system, negative values with no physical meaning may be encountered.

The rest of the paper is organized as follows: we will first present a network of reactions that captures the salient features of the underlying biochemical interactions and will subsequently derive a stochastic model for the transitions of the two operators to the repressed and unrepressed states. We will show that the "pure input assumption" and the "independent genes assumption" only hold in the limit of infinitely large repressor pool, or otherwise meaningless negative extrinsic noise values may be encountered. Finally, we will discuss when it is meaningful to consider partitioning the noise into extrinsic and intrinsic components.
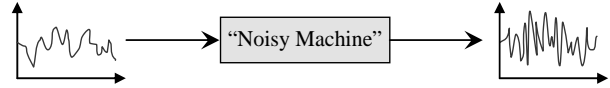

Figure 1: a gene can be viewed as a "noisy machine" that accepts a noisy signal as input, transforms it, and subsequently adds its own inherent noise. This process generates an output with two noise components: extrinsic, due to the input noise, and intrinsic due to the noise produced within the "machine".
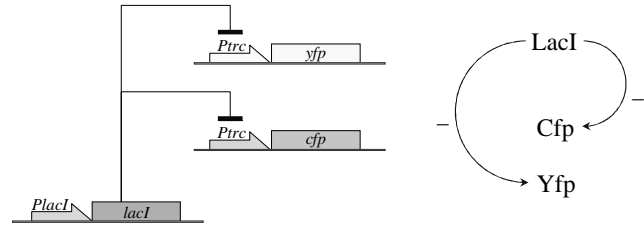

Figure 2: schematic representation of the interactions taken into account in the two promoter-reporter system. For species notation see Table 1. + and − denote positive and negative regulation respectively.

Table 1: Symbols used for the lac operon species

Symbol	Species denoted
Lac	LacI repressor
O_{Yfp}	yfp lacO operator
O_{Cfp}	cfp lacO operator
R_{Yfp}	yfp m-RNA
R_{Cfp}	cfp m-RNA
Yfp	Yfp
Cfp	Cfp
$\varnothing$	Generic source or sink

Competitive Repressor Utilization Model

The molecular mechanisms included in our model are depicted in Figure 2, the species notation is shown in Table 1 and the reaction network is summarized in Table 2. In constructing the reaction network we assume constitutive LacI production, single *lacO* operator sites per promoter, abundant RNA polymerases and ribosomes and 1st order degradation reactions. We ignore cell growth, division and DNA duplication effect as we want to concentrate on the effects of competitive utilization of a repressor pool. Thus, the stochastic model for this system is the following Chemical Master Equation which can be simulated with the Gillespie algorithm (Gillespie 1976; Gillespie 1977):

$$\frac{\partial P}{\partial t} = \left(\mathbb{E}_{Lac}^{-1} - 1 \right) \left(k_{Lac} \cdot P \right) +$$

$$\left(\mathbb{E}_{Lac} - 1 \right) \left(\lambda_{Lac} \cdot Lac \cdot P \right) +$$

$$\left(\mathbb{E}_{O_{Yfp}} \mathbb{E}_{Lac} \mathbb{E}_{O_{Yfp}Lac}^{-1} - 1 \right) \left(k_r \cdot O_{Yfp} \cdot Lac \cdot P \right) +$$

$$\left(\mathbb{E}_{O_{Yfp}}^{-1} \mathbb{E}_{Lac}^{-1} \mathbb{E}_{O_{Yfp}Lac} - 1 \right) \left(k_{-r} \cdot O_{Yfp}Lac \cdot P \right) +$$

$$\left(\mathbb{E}_{R_{Yfp}}^{-1} - 1 \right) \left(k_m \cdot O_{Yfp} \cdot P \right) +$$

$$\left(\mathbb{E}_{R_{Yfp}} - 1 \right) \left(\lambda_{R_{Yfp}} \cdot R_{Yfp} \cdot P \right) +$$

$$\left(\mathbb{E}_{Yfp}^{-1} - 1 \right) \left(k_m \cdot R_{Yfp} \cdot P \right) +$$

$$\left(\mathbb{E}_{Yfp} - 1 \right) \left(\lambda_{Yfp} \cdot Yfp \cdot P \right) +$$

$$\left(\mathbb{E}_{O_{Cfp}} \mathbb{E}_{Lac} \mathbb{E}_{O_{Cfp}Lac}^{-1} - 1 \right) \left(k_r \cdot O_{Cfp} \cdot Lac \cdot P \right) +$$

$$\left(\mathbb{E}_{O_{Cfp}}^{-1} \mathbb{E}_{Lac}^{-1} \mathbb{E}_{O_{Cfp}Lac} - 1 \right) \left(k_{-r} \cdot O_{Cfp}Lac \cdot P \right) +$$

$$\left(\mathbb{E}_{R_{Cfp}}^{-1} - 1 \right) \left(k_m \cdot O_{Cfp} \cdot P \right) +$$

$$\left(\mathbb{E}_{R_{Cfp}} - 1 \right) \left(\lambda_{R_{Cfp}} \cdot R_{Cfp} \cdot P \right) +$$

$$\left(\mathbb{E}_{Cfp}^{-1} - 1 \right) \left(k_m \cdot R_{Cfp} \cdot P \right) + \tag{1}$$

$$\left(\mathbb{E}_{Cfp} - 1 \right) \left(\lambda_{Cfp} \cdot Cfp \cdot P \right)$$

where we have made use of the step operator as follows (van Kampen 1992):

$$\mathbb{E}_m^p f \left(k, l, m, n, ... \right) = f \left(k, l, m + p, n, ... \right) \tag{2}$$

The state vector with the numbers of molecules is $\mathbf{x} = (Lac, O_{Yfp}, O_{Yfp}Lac, O_{Cfp}, O_{Cfp}Lac, R_{Yfp}, Yfp, R_{Cfp}, Cfp)$.

Let us now focus on the operator states under the assumption of fast operator fluctuations. Noting that O_{Yfp} and O_{Cfp} are zero or unity and operator sites are conserved, let us define:

$$O_{Yfp} + O_{Yfp}Lac = O_{Yfp,T} = 1$$
$$O_{Cfp} + O_{Cfp}Lac = O_{Cfp,T} = 1 \tag{3}$$

Furthermore, the operator fluctuations between the free and the repressed state do not change the overall repressor content, which is defined as:

$$Lac_T = Lac + O_{Yfp}Lac + O_{Cfp}Lac \tag{4}$$

We can further apply the summing operator:

$$\sum_{R_{Yfp} \geq 0} \sum_{Yfp \geq 0} \sum_{R_{Cfp} \geq 0} \sum_{Cfp \geq 0} \bullet \tag{5}$$

to Eq. (1) to eliminate RNA and protein species.

Table 2: Reaction Network

(i)	$\emptyset \xrightarrow{k_{lac}}$ Lac	
(ii)	$O_{Yfp} + Lac \xrightarrow{k_r} O_{Yfp}Lac$	
(iii)	$O_{Yfp}Lac \xrightarrow{k_{-r}} O_{Yfp} + Lac$	
(iv)	$O_{Yfp} \xrightarrow{k_m} O_{Yfp} + R_{Yfp}$	
(v)	$R_{Yfp} \xrightarrow{k_p} R_{Yfp} + Yfp$	
(vi)	$O_{Cfp} + Lac \xrightarrow{k_r} O_{Cfp}Lac$	
(vii)	$O_{Cfp}Lac \xrightarrow{k_{-r}} O_{Cfp} + Lac$	
(viii)	$O_{Cfp} \xrightarrow{k_m} O_{Cfp} + R_{Cfp}$	
(ix)	$R_{Cfp} \xrightarrow{k_p} R_{Cfp} + Cfp$	
(x)	$Lac \xrightarrow{\lambda_{Lac}} \emptyset$	
(xi)	$R_{Yfp} \xrightarrow{\lambda_m} \emptyset$	
(xii)	$Yfp \xrightarrow{\lambda_p} \emptyset$	
(xiii)	$R_{Cfp} \xrightarrow{\lambda_m} \emptyset$	
(xiv)	$Cfp \xrightarrow{\lambda_p} \emptyset$	

Finally, assuming fast operator fluctuations:

$$k_r = \frac{\kappa_r}{\varepsilon} \qquad \text{and} \qquad k_{-r} = \frac{\kappa_{-r}}{\varepsilon} \tag{6}$$

allows us to introduce asymptotic expansions for the following probability mass functions:

$$\widehat{P} \left(O_{Yfp}, O_{Cfp} \mid Lac_T; t \right) = \widehat{P}_0 \left(O_{Yfp}, O_{Cfp} \mid Lac_T; t \right) + \varepsilon \cdot \widehat{P}_1 \left(O_{Yfp}, O_{Cfp} \mid Lac_T; t \right) + ... \tag{7}$$

$$\widehat{P} \left(Lac_T; t \right) = \widehat{P}_0 \left(Lac_T; t \right) + \varepsilon \cdot \widehat{P}_1 \left(Lac_T; t \right) + ...$$

Then from Eq. (1) using Eq. (3) − (7) the $\mathcal{O}(\varepsilon^{-1})$ expression is:

$$0 = \left(\mathbb{E}_{O_{Yfp}} - 1 \right) \left(\kappa_r \cdot O_{Yfp} \cdot \left(Lac_T \right. \right.$$
$$\left. \left. - \left(O_{Yfp,T} - O_{Yfp} \right) - \left(O_{Cfp,T} - O_{Cfp} \right) \right) \cdot \widehat{P}_0 \right)$$
$$\left(\mathbb{E}_{O_{Yfp}}^{-1} - 1 \right) \left(\kappa_{-r} \cdot \left(O_{Yfp,T} - O_{Yfp} \right) \cdot \widehat{P}_0 \right) +$$
$$\left(\mathbb{E}_{O_{Cfp}} - 1 \right) \left(\kappa_r \cdot O_{Cfp} \cdot \left(Lac_T \right. \right. \tag{8}$$
$$\left. \left. - \left(O_{Yfp,T} - O_{Yfp} \right) - \left(O_{Cfp,T} - O_{Cfp} \right) \right) \cdot \widehat{P}_0 \right)$$
$$\left(\mathbb{E}_{O_{Cfp}}^{-1} - 1 \right) \left(\kappa_{-r} \cdot \left(O_{Cfp,T} - O_{Cfp} \right) \cdot \widehat{P}_0 \right)$$

where the new state vector yielding $\hat{P}_0$ is $\mathbf{x} = (Lac_T, O_{Yfp}, O_{Cfp})$.

By solving Eq. (8) one obtains the extrinsic and intrinsic noise contributions given $Lac_T = q$:

$$\eta^2_{\substack{oper|Lac_T \\ int}} = \frac{\left\langle \left(O_{Yfp} - O_{Cfp}\right)^2\right\rangle}{2\cdot\left\langle O_{Yfp}\right\rangle\cdot\left\langle O_{Cfp}\right\rangle} = \tag{9}$$
$$\frac{q\cdot K_r}{\left(q\cdot K_r + 1\right)^2}\cdot\left(q\cdot(q-1)\cdot K_r^2 + 2\cdot q\cdot K_r + 1\right)$$

$$\eta^2_{\substack{oper|Lac_T \\ ext}} = \frac{\left\langle O_{Yfp}\cdot O_{Cfp}\right\rangle - \left\langle O_{Yfp}\right\rangle\cdot\left\langle O_{Cfp}\right\rangle}{\left\langle O_{Yfp}\right\rangle\cdot\left\langle O_{Cfp}\right\rangle} = \tag{10}$$
$$-\frac{q\cdot K_r^2}{\left(q\cdot K_r + 1\right)^2}$$

Immediately apparent is the fact that the operator "extrinsic noise squared" is *always* calculated to be negative, which has no physical meaning. This happens because of the negative correlations resulting from the competitive utilization of the common repressor pool. Furthermore, it is interesting to investigate the limiting behavior of the two noise components for high repressor contents ($q \rightarrow \infty$) and for very strong repression ($K_r \rightarrow \infty$):

$$\lim_{q\rightarrow\infty}\eta^2_{\substack{oper|Lac_T \\ ext}} = 0$$
$$\tag{11}$$
$$\lim_{K_r\rightarrow\infty}\eta^2_{\substack{oper|Lac_T \\ ext}} = -\frac{1}{q}$$

Non-Competitive Repressor Utilization

In order to show that indeed it is the competitive repressor utilization that results in the negative feedback between genes and the paradox of negative extrinsic noise values, we considered the case where genes do not alter the free repressor concentration. Thus, reactions (ii) and (iii) are substituted with the following:

$$O_{Yfp} + Lac \xrightarrow{k_r} O_{Yfp}Lac + Lac \tag{ii}$$

$$O_{Yfp}Lac \xrightarrow{k_{-r}} O_{Yfp} \tag{iii}$$
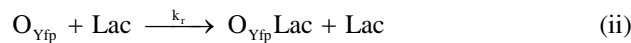
Performing the analysis as previously described, the intrinsic and extrinsic noise values, given the total Lac, can now be calculated:

$$\eta^2_{\substack{oper|Lac_T \\ int}} = q\cdot K_r \tag{12}$$

$$\eta^2_{\substack{oper|Lac_T \\ ext}} = 0 \tag{13}$$

Evidently, the extrinsic noise is always zero in this case due to the fact that the free repressor concentration remains constant.

Conclusions

Starting from a reaction network for the two promoter system, used by Elowitz et al. (2002) to quantify the contributions of extrinsic versus intrinsic noise, we focused on the operator states and calculated the corresponding noise values. We found that due to the competitive repressor utilization, the extrinsic noise value will always be negative if the total repressor is assumed constant. This unrealistic situation is due to the negative correlations that result from the fact that a LacI molecule bound to the *yfp* operator is unavailable for repressing the other operator (*cfp*). These negative correlations essentially result in an underestimation of the extrinsic noise.

Our analysis shows that these effects become less pronounced and finally disappear in the limit of abundant repressor contents (infinitely many molecules).

Hence, the presumption that noise can be partitioned into extrinsic and intrinsic components is subject to limitations. In particular, the species' concentrations that act as inputs to the genes of interest must be high enough so that depletion effects are negligible. Otherwise, the calculated (or experimentally measured) extrinsic and intrinsic noise values may not be meaningful.

References

Elowitz, M. B., A. J. Levine, E. D. Siggia and P. S. Swain (2002). "Stochastic gene expression in a single cell." Science **297**(5584): 1183-1186.

Gillespie, D. T. (1976). "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions." Journal of Computational Physics **22**(4): 403-434.

Gillespie, D. T. (1977). "Exact stochastic simulation of coupled chemical reactions." The Journal of Physical Chemistry **81**(25): 2340-2361.

Swain, P. S., M. B. Elowitz and E. D. Siggia (2002). "Intrinsic and extrinsic contributions to stochasticity in gene expression." Proceedings of the National Academy of Sciences of the United States of America **99**(20): 12795-12800.

van Kampen, N. G. (1992). Stochastic processes in physics and chemistry. New York, Amsterdam, North-Holland-Personal-Library.