Optimal feedback strength for noise suppression in auto-regulatory gene networks

Abhyudai Singh [1]

Department of Electrical and Computer Engineering,

University of California, Santa Barbara, CA

Joao P. Hespanha

Department of Electrical and Computer Engineering,

University of California, Santa Barbara, CA

[1]Corresponding author. Address: Department of Electrical and Computer Engineering, University of California, Santa Barbara, Santa Barbara, CA 93101, U.S.A., Tel.: (805)403-2533, Fax: (805) 893-3262

Abstract

Auto-regulatory feedback loops, where the protein expressed from a gene inhibits or activates its own expression are common gene network motifs within cells. In these networks, stochastic fluctuations in protein levels are attributed to two factors: intrinsic noise (i.e., the randomness associated with protein expression and degradation) and extrinsic noise (i.e., the noise caused by fluctuations in cellular components such as enzyme levels and gene-copy numbers). We present results that predict the level of both intrinsic and extrinsic noise in protein numbers as a function of quantities that can be experimentally determined and/or manipulated, such as the response time of the protein and the level of feedback strength. In particular, we show that for a fixed average number of protein molecules, decreasing response times leads to attenuation of both protein intrinsic and extrinsic noise, with the extrinsic noise being more sensitive to changes in the response time.

We further show that for auto-regulatory networks with negative feedback, the protein noise levels are always *minimized* at an optimal level of feedback strength. We provide an analytical expression for this highest level of noise suppression and the amount of feedback that achieves this minimal noise. These theoretical results are shown to be consistent and explain recent experimental observations. Finally, we illustrate how measuring changes in the protein noise levels as the feedback strength is manipulated can be used to determine the level of extrinsic noise in these gene networks.

*Key words:*Noise in gene expression; gene regulation; negative feedback; response time; feedback strength; extrinsic noise

1 Introduction

Gene expression and regulation is inherently a noisy process. The origins of this stochasticity lie in the probabilistic nature of transcription and translation and low copy numbers of RNAs and proteins within cells, which can lead to large statistical fluctuations in molecule numbers. Recent work (1–5) has provided considerable experimental evidence for these stochastic fluctuations and may explain for the large amounts of cell to cell variation observed in genetically identical cells exposed to the same environmental conditions (6, 7). Various gene network motifs within cells decrease/increase these stochastic fluctuations. A common such motif is an auto-regulatory gene network where the protein expressed from the gene inhibits/activates its own transcription (8, 9). Both theoretical and experimental studies have shown that negative feedback in these auto-regulatory gene networks reduces stochastic fluctuations in the protein population (10–14) where as positive feedback has the opposite effect (15, 16).

Auto-regulatory gene networks are characterized by their *transcriptional response* $g(\mathbf{x})$, which determines the transcription rate of the gene as a non-linear function g of the protein molecular count $\mathbf{x}$ within the cell. Monotonic decreasing and increasing functions $g(\mathbf{x})$ denote negative and positive feedback, respectively. The noise in the protein population is quantified by its *coefficient of variation* defined as the ratio of the standard deviation to the average number of protein molecules. Previous work has shown that this protein noise level is determined by a combination of two components (17, 18). The first is the intrinsic noise, which represents the stochastic fluc-

tuations in the protein population arising due to random protein formation and degradation events. The second component is the extrinsic noise, which corresponds to fluctuations in the protein numbers arising due to an exogenous noise source driving the auto-regulatory gene network, for example, fluctuations in gene copy numbers, enzyme levels and environmental stimuli. Table 1 provides a summary of the notations used for the different forms of noise in the protein population. Our goal is to understand how these different components of protein noise can be modulated by manipulating the *responsivity* of the auto-regulatory gene network, which is defined as follows: assuming $\mathbf{x}^*$ to be the steady-state average protein count, the *response time* T_r is the time taken for any initial perturbation about $\mathbf{x}^*$ to decay by 50% of its initial value. Negative and positive feedback in the auto-regulatory gene network, decreases and increases the response time, respectively, from the its value when there is no feedback (i.e., when the transcriptional response $g(\mathbf{x})$ is a constant and independent of $\mathbf{x}$).

We consider a simple model of gene expression where each expression event produces a random number of protein molecules according to an arbitrary probability distribution. Details on the stochastic formulation of this model are provided in Section 2. In Section 3, we determine the intrinsic noise in the protein population. Using a linear approximation for the transcriptional response

$$g(\mathbf{x}) \approx g(\mathbf{x}^*) + g'(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*), \tag{1}$$

where $\mathbf{x}^*$ is the steady-state average protein count, we show that the intrinsic

noise level is proportional to the ratio $T_r/\mathbf{x}^*$. Hence for a fixed $\mathbf{x}^*$, decreasing the protein's response time T_r attenuates the intrinsic noise whereas increasing the response time magnifies it. We also investigate the effects of non-linearities in the transcriptional response and show that a concave (convex) transcriptional response causes the noise in the protein to be smaller (larger) then what would be predicted by a linear transcriptional response as in equation Eq. 1.

We next quantify the extrinsic noise in the protein population. In Section 4, we derive analytical formulas that decompose the total noise in the protein into its extrinsic and intrinsic components. These formulas are simple generalizations of previous work, notably that of (18), where the number of protein molecules produced per expression event was not random but deterministic and equal to one. We show that for a given decrease in T_r through negative feedback, the extrinsic noise decreases by a much larger amount then does the intrinsic noise. Thus negative feedback is much more effective in reducing the extrinsic component of protein noise than its intrinsic component.

In Section 5, we use the above results to quantify noise in auto-regulatory gene networks that involve a common negative feedback with transcriptional response given by

$$g(\mathbf{x}) = g_0 \left(b + \frac{1-b}{1+(a\mathbf{x})^M} \right), \quad 0 \leq b < 1 \tag{2}$$

where $M \geq 1$ denotes the *hill coefficient* and g_0 corresponds to the transcription rate when there is no feedback (i.e., $a = 0$) (8, 19). The positive

constant b is less than one and is chosen such that the product $g_0 b$ represents the minimum level of transcription rate. The constant a characterizes the *feedback strength* and is determined by the binding affinity of the protein to the promoter of the gene. We shall perform a systematic analysis of how the protein noise level changes as the feedback strength a is increased from an initial value of zero. We first consider the situation where extrinsic noise is absent or negligible and intrinsic noise dominates the total noise in the protein population. In such a scenario, we show that if the hill coefficent is close to one, then the protein noise level actually increases as we increase the feedback strength. However, for hill coefficients larger than one, the protein noise level first decreases as we increae the feedback strength from zero and achieves a minimum value at some optimal level of feedback strength. Increasing the feedback strength above this optimal value causes an increase in the noise level. In summary, for hill coefficients M larger than one, we obtain a U-shaped noise profile as the feedback strength is increased. We quantify both the optimal level of feedback strength and the *limit of noise suppression*, which is defined as the ratio of the minimum possible noise in the protein population to the protein noise level when there is no feedback (i.e., $a = 0$). When the intrinsic noise dominates the total noise in the protein population, this limit is given by the simple expression

$$\sqrt{\frac{4M}{4M + (1-b)(M-1)^2}} \leq 1. \tag{3}$$

We next consider the situation where the extrinsic noise is not negligible and both extrinsic and intrinsic noise are present. We show that in this case,

irrespective of the value of the hill coefficient, the protein noise level always follows a U-shaped profile as the feedback strength is increased. This means that the noise level is minimized at some optimal value of feedback strength and decreasing or increasing feedback strength away from this optima will always causes an increase in the noise level. We again provide analytical formulae for the limit of noise suppression and show that in this case (when extrinsic noise is not small compared to the intrinsic noise), this ratio is much lower than what is given by equation Eq. 3, which corresponds to the situation where there is no extrinsic noise. In fact, determining how much the limit of noise suppression deviates from Eq. 3 can be used to estimate how much extrinsic noise is present in the gene network.

In Section 6 we validate our theoretical results by using experimental data from (20) for a synthetic auto-regulatory gene network. As predicted, we indeed see a U-shaped profile for the protein noise level as the feedback strength is experimentally manipulated. We also explain observations in (20) showing that for small levels of extrinsic noise, no U-shaped profile is observed, and instead, the protein noise level monotonically increase as the feedback strength is increased. Finally, we illustrate how the experimentally determined limit of noise suppression can be used to estimate the noise in the exogenous signal. Matching these estimates with independent measurements of noise associated with the plasmid population indeed confirmed that variability in plasmid numbers was the major source of extrinsic noise in this synthetic gene netwrok.

Table 1: A summary of the notation used in this paper. All estimates of noise, expect $CV_{int-quad}$, are based on a linear approximation of the transcriptional response.

CV_{tot}	Total noise in protein numbers
CV_{ext}	Extrinsic noise in protein numbers
CV_{int}	Intrinsic noise in protein numbers
CV_z	Noise in the exogenous signal driving the gene netwrok
$CV_{int-quad}$	Intrinsic noise in protein numbers based on a quadratic approximation for the transcriptional response.
CV_{tot-nr}	Total noise in protein numbers when there is no feedback
CV_{ext-nr}	Extrinsic noise in protein numbers when there is no feedback
CV_{int-nr}	Intrinsic noise in protein numbers when there is no feedback
$CV_{tot-min}$	Minimum possible total noise in protein numbers with optimal negative feedback
$CV_{ext-min}$	Minimum possible extrinsic noise in protein numbers with optimal negative feedback
$CV_{int-min}$	Minimum possible intrinsic noise in protein numbers with optimal negative feedback
a_{min}	Feedback strength where the total noise in the protein is minimum
$a_{int-min}$	Feedback strength where the intrinsic noise in the protein is minimum
$a_{ext-min}$	Feedback strength where the extrinsic noise in the protein is minimum
a_{Tmin}	Feedback strength where the protein's response time is minimum
T_r	Protein's response time
T_{nr}	Protein's response time when there is no feedback

2 Un-regulated gene expression

We consider a very simple model of gene expression where a gene expresses a protein X in bursts that occur at a rate K_x. Each expression event leads to the formation of $\mathbf{N}_x$ molecules of the protein X. Recent work suggests that the burst of proteins from each mRNA transcript follows a geometric distribution (21). Thus instead of assuming $\mathbf{N}_x$ to be a constant we assume it to be a random variable with mean N_x and variance V_x^2. We also assume that the protein decays at a constant rate d_x. Notice that our model omits the mRNA dynamics. This is a valid approximation as long as the protein's life time is much longer than the mRNA's life time, which is generally the case in gene-protein networks (22). Ignoring the mRNA dynamics leads to relatively simple expressions for the protein noise level, which help develop a qualitative understanding of how noise level changes in response to alterations of the gene network parameters.

In a stochastic formulation, gene expression and protein degradation are treated as probabilistic events with probabilities of occurring in an infinitesimal time interval $(t, t+dt]$ given by

$$\Pr\{\mathbf{x}(t+dt) = x + \mathbf{N}_x \mid \mathbf{x}(t) = x\} = K_x dt \tag{4a}$$

$$\Pr\{\mathbf{x}(t+dt) = x - 1 \mid \mathbf{x}(t) = x\} = d_x x dt, \tag{4b}$$

respectively, where $\mathbf{x}(t)$ denotes the number of molecules of protein X at time t.

A convenient way to model the time evolution of the number of molecules

$\mathbf{x}$ is through a Stochastic Hybrid System (SHS) characterized by trivial continuous dynamics

$$\dot{\mathbf{x}} = 0, \tag{5}$$

and two reset maps

$$\mathbf{x} \mapsto \phi_1(\mathbf{x}) = \mathbf{x} + \mathbf{N}_x, \quad \mathbf{x} \mapsto \phi_2(\mathbf{x}) = \mathbf{x} - 1 \tag{6}$$

with corresponding transition intensities given by

$$\lambda_1(\mathbf{x}) = K_x, \ \lambda_2(\mathbf{x}) = d_x \ \mathbf{x} \tag{7}$$

(23). In order to gauge the noise level in the protein population, we determine the time evolution of the first and second order moments of $\mathbf{x}$, i.e. the expected values $\mathbf{E}[\mathbf{x}]$ and $\mathbf{E}[\mathbf{x}^2]$. The moment dynamics can be obtained using the Dynkin's formula for the above SHS, according to which, for every differentiable function $\psi(\mathbf{x})$ we have that

$$\frac{d\mathbf{E}[\psi(\mathbf{x})]}{dt} = \mathbf{E}\left[\sum_{i=1}^{2} \left(\psi(\phi_i(\mathbf{x})) - \psi(\mathbf{x})\right) \lambda_i(\mathbf{x})\right] \tag{8}$$

(24, 25). Taking $\psi(\mathbf{x}) = \mathbf{x}$ and $\psi(\mathbf{x}) = \mathbf{x}^2$ in Eq. 8 we obtain the following moment dynamics

$$\frac{d\mathbf{E}[\mathbf{x}]}{dt} = N_x K_x - d_x \mathbf{E}[\mathbf{x}], \tag{9a}$$

$$\frac{d\mathbf{E}[\mathbf{x}^2]}{dt} = K_x(N_x^2 + V_x^2) + d_x \mathbf{E}[\mathbf{x}] + 2K_x N_x \mathbf{E}[\mathbf{x}] - 2d_x \mathbf{E}[\mathbf{x}^2]. \tag{9b}$$

As $t \to \infty$, the first and second order moments converge to constant steady-state values given by

$$\mathbf{x}^* := \lim_{t\to\infty} \mathbf{E}[\mathbf{x}(t)] = \frac{N_x K_x}{d_x} \tag{10a}$$

$$\mathbf{E}^*[\mathbf{x}^2] := \lim_{t\to\infty} \mathbf{E}[\mathbf{x}^2(t)] = \frac{K_x d_x N_x + 2K_x^2 N_x^2 + K_x d_x (N_x^2 + V_x^2)}{2d_x^2}. \tag{10b}$$

We quantify the noise in $\mathbf{x}(t)$ by its *coefficient of variation* defined as the ratio of the standard deviation in protein numbers to the average number of protein molecules. Using the above steady-state values we obtain

$$CV^2_{int-nr} = \frac{\mathbf{E}^*[\mathbf{x}^2] - \mathbf{x}^{*2}}{\mathbf{x}^{*2}} = \frac{d_x(N_x^2 + V_x^2 + N_x)}{2K_x N_x^2} = \frac{(N_x^2 + V_x^2 + N_x)}{2\mathbf{x}^* N_x}. \tag{11}$$

This quantity quantifies the noise in the protein X solely due to random gene expression and protein degradation, and is referred to as the *intrinsic noise* in the protein population when there is no regulation. Note that the noise in the protein increases with the variance V_x^2 in the number of protein molecules produced in each transcription event. A special case of equation Eq. 11 is obtained for $N_x = 1$ and $V_x = 0$, for which $\mathbf{x}(t)$ has a Poisson distribution and $CV^2_{int-nr} = 1/\mathbf{x}^*$. In the next section we examine what happens to this intrinsic noise when the gene expression rate is not a constant but a function of the number of molecules of the protein.

3 Auto-regulatory gene expression

Often the expressed protein binds to the promoter region of its own gene. In doing so it either recruits the enzyme RNA Polymerase to the promoter (which leads to an increase in gene expression) or blocks RNA Polymerase from binding to the promoter (which causes a decrease in gene expression). Such gene expression is referred to as an *auto-regulatory gene network*. We model this network by assuming that the rate of gene expression is no longer a constant and is instead a function $g(\mathbf{x})$ of the number of protein molecules $\mathbf{x}$. We refer to the function $g(\mathbf{x})$ as the *transcriptional response* of the network. This transcriptional response can be formally derived assuming that the rate of binding and dissociation between the protein and its promoter is much faster than the dynamics of protein production and degradation (8) or it can be determined directly from experiments. Monotonic decreasing and increasing functions $g(\mathbf{x})$ denote negative and positive feedback, respectively.

When an auto-regulation mechanism is present, the probabilities of gene expression and protein degradation events occurring in an infinitesimal time interval $(t, t+dt]$ are given by

$$\Pr\{\mathbf{x}(t+dt) = x + \mathbf{N}_x \mid \mathbf{x}(t) = x\} = g(x)dt \tag{12a}$$

$$\Pr\{\mathbf{x}(t+dt) = x - 1 \mid \mathbf{x}(t) = x\} = d_x x dt. \tag{12b}$$

To write the moment dynamics of $\mathbf{x}$ we first approximate $g(\mathbf{x})$ by a polyno-

mial in $\mathbf{x}$, which is done by expanding $g(\mathbf{x})$ as a Taylor series expansion

$$g(\mathbf{x}) = g(\mathbf{x}^*) + g'(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) + \frac{1}{2}g''(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*)^2 + \ldots, \tag{13}$$

about the steady-state average number of protein molecules $\mathbf{x}^*$.

3.1 Linear transcriptional response

We begin by ignoring quadratic and higher order terms in Eq. 13 which results in a linear transcriptional response

$$g(\mathbf{x}) \approx g(\mathbf{x}^*) + g'(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*). \tag{14}$$

This approximation is valid as long as the stochastic fluctuations in the protein do not leave the region in which $g(\mathbf{x})$ is approximately linear. As in Section 2, we model the time evolution of $\mathbf{x}$ through a Stochastic Hybrid System (SHS) but now the transition intensities are given by $\lambda_1(\mathbf{x}) = g(\mathbf{x}^*) + g'(\mathbf{x}^*)(\mathbf{x}-\mathbf{x}^*)$ and $\lambda_2(\mathbf{x}) = d_x\mathbf{x}$. Using the Dynkin's formula for this modified SHS we obtain the following dynamics for the mean $\mathbf{E}[\mathbf{x}]$:

$$\frac{d\mathbf{E}[\mathbf{x}]}{dt} = N_x g(\mathbf{x}^*) - d_x\mathbf{x}^* + (N_x g'(\mathbf{x}^*) - d_x)(\mathbf{E}[\mathbf{x}] - \mathbf{x}^*) \tag{15}$$

and the steady-state value $\mathbf{x}^*$ for the mean population $\mathbf{E}[\mathbf{x}]$ must satisfy

$$N_x g(\mathbf{x}^*) = d_x\mathbf{x}^*. \tag{16}$$

To be biologically meaningful, the average $\mathbf{E}[\mathbf{x}]$ must remain bounded which means that the linear system given by Eq. 15 must have a negative eigenvalue

$$\lambda := N_x g'(\mathbf{x}^*) - d_x < 0. \tag{17}$$

This eigenvalue λ can be expressed in terms of the response time T_r of the protein, a quantity that can be measured experimentally. The response time T_r is defined as the time taken for $\mathbf{E}[\mathbf{x}(t)] - \mathbf{x}^*$ to decay by 50% of its initial condition, i.e., $\mathbf{E}[\mathbf{x}(T_r)] - \mathbf{x}^* = (\mathbf{E}[\mathbf{x}(0)] - \mathbf{x}^*)/2$ and is given by

$$T_r = -\frac{\ln(2)}{\lambda} > 0, \qquad \lambda := N_x g'(\mathbf{x}^*) - d_x < 0. \tag{18}$$

Negative feedback, which correspond to $g'(\mathbf{x}^*) < 0$, decreases the response time from the value $T_{nr} = \ln(2)/d_x$ that corresponds to the absence of feedback (i.e., $g'(\mathbf{x}^*) = 0$). Positive feedback has an opposite effect.

We now compute the coefficient of variation of $\mathbf{x}(t)$ by writing the moment dynamics for the second order moment $\mathbf{E}[\mathbf{x}^2]$. Using Eq. 8, with $\psi(\mathbf{x}) = \mathbf{x}^2$ we obtain the following time derivative for $\mathbf{E}[\mathbf{x}^2]$:

$$\begin{aligned}\frac{d\mathbf{E}[\mathbf{x}^2]}{dt} =& [g(\mathbf{x}^*) - \mathbf{x}^* g'(\mathbf{x}^*)](N_x^2 + V_x^2) + d_x \mathbf{E}[\mathbf{x}] + 2[g(\mathbf{x}^*) - \mathbf{x}^* g'(\mathbf{x}^*)] N_x \mathbf{E}[\mathbf{x}] \\ & - 2d_x \mathbf{E}[\mathbf{x}^2] + g'(\mathbf{x}^*)(N_x^2 + V_x^2)\mathbf{E}[\mathbf{x}] + 2g'(\mathbf{x}^*) N_x \mathbf{E}[\mathbf{x}^2].\end{aligned} \tag{19}$$

Performing a steady-state analysis of the above equations and using Eq. 16

we obtain the following steady-state coefficient of variation

$$CV_{int} = \sqrt{\frac{d_x(N_x^2 + V_x^2 + N_x)}{2IN_x^2}}, \qquad I = g(\mathbf{x}^*) - \mathbf{x}^* g'(\mathbf{x}^*) \tag{20}$$

where I can be interpreted as the y-intercept of the tangent to the transcriptional response $g(\mathbf{x})$ at $\mathbf{x} = \mathbf{x}^*$ (see Figure 1). Using Eq. 16, Eq. 18 and Eq. 20 we can also relate the intrinsic noise to the response time T_r of the protein as

$$CV_{int} = \sqrt{\frac{T_r}{T_{nr}} \frac{N_x^2 + V_x^2 + N_x}{2\mathbf{x}^* N_x}} \tag{21}$$

where $T_{nr} = \ln(2)/d_x$ is the protein's response time when there is no regulation in gene expression (i.e., $g'(\mathbf{x}^*) = 0$ and the transcription rate is a constant as in Section 2). The formula in Eq. 21 shows that the intrinsic noise level in auto-regulatory gene networks is determined by three factors: the average number of protein molecules $\mathbf{x}^*$, the response time of the protein T_r and the gene expression burst characteristics, i.e. N_x and V_x^2. From Eq. 18 and Eq. 20 we also conclude that for a fixed $\mathbf{x}^*$, making the slope $g'(\mathbf{x}^*)$ more negative causes a decrease in the response time and leads to attenuation of intrinsic noise in the protein population. However, as we will see later, experimental manipulations that change the response time typically also alter $\mathbf{x}^*$, in which case, attenuation or magnification of intrinsic noise will depend on wether the ratio $T_r/\mathbf{x}^*$ in Eq. 21 decreases or increases, respectively.

When the number of proteins produced per mRNA follows a geometric

distribution (21), the variance V_x^2 is equal to $N_x^2 - N_x$. In this case Eq. 21 simplifies to

$$CV_{int} = \sqrt{\frac{T_r}{T_{nr}} \frac{N_x}{\mathbf{x}^*}} \tag{22}$$

which shows that for all other parameters fixed, the intrinsic noise increases as we increase the average number N_x of proteins produced per gene expression event, which is consistent with other similar theoretical and experimental observations (19, 26).

An important feature of equation Eq. 22 is that it relates the noise in the protein to parameters that can be experimentally determined. In particular, $N_x = L_x/d_r$ where L_x is the translation rate of the mRNA and d_r is the mRNA degradation rate, and the response times can be measured by tracking the time evolution of the number of molecules within the cell. For example, in (27) an auto-regulatory gene network was designed where the protein repressed its own transcription. The protein was fluorescently tagged which allowed one to compute the time evolution of the average number of protein molecules in the cell. Figure 2 plots this time evolution with and without negative feedback in the gene. The promoter strength was appropriately adjusted such that the steady-state population of the protein was the same in both cases. The figure shows that with negative feedback it takes about $T_r = .21$ time unit for the protein count to reach half of its steady-state average protein count $\mathbf{x}^*$. The response time when there is no feedback is $T_{nr} = 1$ time unit, which is five times larger than T_r. We can then conclude from Eq. 22 that for this network, the presence of negative

feedback reduces the intrinsic noise levels in the protein population by a factor of $\sqrt{5} \approx 2.2$.

3.2 Effect of nonlinearities

We now examine the effects of non-linear quadratic terms in $g(\mathbf{x})$. Towards that end we approximate $g(\mathbf{x})$ as

$$g(\mathbf{x}) = g(\mathbf{x}^*) + g'(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) + \frac{1}{2}g''(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*)^2 \tag{23}$$

where $\mathbf{x}^*$ is given by Eq. 16 and ignore cubic and higher order terms in $\mathbf{x} - \mathbf{x}^*$. Referring the reader to Appendix A for more details, the intrinsic noise in the protein population can now be approximated by

$$CV^2_{int-quad} = \frac{CV^2_{int}}{1 + \frac{N_x \mathbf{x}^* g''(\mathbf{x}^*) CV^2_{int}}{2\lambda}}, \qquad \lambda := N_x g'(\mathbf{x}^*) - d_x < 0 \tag{24}$$

where CV_{int} is the intrinsic noise when $g(\mathbf{x})$ is linear (as computed in Section 3.1) and is given by Eq. 20. To obtain Eq. 24 we assumed that $CV^2_{int-quad}$ is much smaller than one and the distribution of the protein population is symmetrically distributed about its mean. The above result shows two important points. Firstly, a transcriptional response which is concave at $\mathbf{x}^*$ ($g''(\mathbf{x}^*) < 0$) results in smaller intrinsic noise then as predicted by CV_{int}, where as a convex response ($g''(\mathbf{x}^*) > 0$) has the opposite effect. Secondly, as long as CV_{int} and the non-linearity in the transcriptional

response are small in the sense that

$$\left|\frac{N_x \mathbf{x}^* g''(\mathbf{x}^*) CV_{int}^2}{2\lambda}\right| \ll 1, \tag{25}$$

linearzing the transcriptional response will yield a good approximation for the intinsic noise in the protein population.

4 Extrinsic and intrinsic contributions to noise

We now consider extrinsic noise in the protein population arising due to an exogenous noise source driving the auto-regulatory gene network. Towards that end, we now consider a transcriptional response $g(\mathbf{x}, \mathbf{z})$ that depends on a noisy exogenous signal $\mathbf{z}$. The transcriptional response $g(\mathbf{x}, \mathbf{z})$ may take different forms. For example, if the gene is encoded on a low-copy plasmid, then fluctuations in the number of copies of the plasmid are known to be a major source of extrinsic noise (20). In this case, the transcriptional response takes the form $\mathbf{z}g(\mathbf{x})$, where $\mathbf{z}$ represents the number of copies of the plasmid. Alternatively $\mathbf{z}$ could represent the number of molecules of the RNA polymerase, in which case, the transcriptional response for an auto-regulatory gene network with negative feedback would be

$$g(\mathbf{x}, \mathbf{z}) = \frac{k_0 \mathbf{z}}{1 + k_1 \mathbf{z} + k_2 \mathbf{x}} \tag{26}$$

where k_0, k_1 and k_2 are positive constants (13).

We model the stochastic fluctuations in $\mathbf{z}$ by a birth-death process. In particular, the probabilities of formation and degradation of $\mathbf{z}$ in the in-

finitesimal time interval $(t, t+dt]$ are given by

$$\Pr\{\mathbf{z}(t+dt) = z + \mathbf{N}_z \mid \mathbf{z}(t) = z\} = K_z dt \tag{27a}$$

$$\Pr\{\mathbf{z}(t+dt) = z - 1 \mid \mathbf{z}(t) = z\} = d_z z dt \tag{27b}$$

where K_z and d_z represent the production and degradation rate of $\mathbf{z}$, respectively, and $\mathbf{N}_z$ is a random variable with mean N_z and variance V_z^2. In the sequel we refer to $T_z = \ln(2)/d_z$ as the response time of the exogenous signal. Following steps such as those outlined in Section 2, we can conclude from Eq. 10 and Eq. 11 that the steady-state average level and the coefficient of variation of $\mathbf{z}$ are given by

$$\mathbf{z}^* = \frac{N_z K_z}{d_z} \tag{28}$$

and

$$CV_z = \sqrt{\frac{(N_z^2 + V_z^2 + N_z)}{2\mathbf{z}^* N_z}}, \tag{29}$$

respectively. The quantity CV_z represents the amount of noise that enters the auto-regulatory gene network through the exogenous signal $\mathbf{z}$. Assuming that the stochastic fluctuations in $\mathbf{x}$ and $\mathbf{z}$ around their respective means $\mathbf{x}^*$ and $\mathbf{z}^*$ are sufficiently small, we approximate the transcriptional response as

$$g(\mathbf{x}, \mathbf{z}) \approx g(\mathbf{x}^*, \mathbf{z}^*) + \frac{dg(\mathbf{x}, \mathbf{z}^*)}{d\mathbf{x}}|_{\mathbf{x}=\mathbf{x}^*}(\mathbf{x} - \mathbf{x}^*) + \frac{dg(\mathbf{x}^*, \mathbf{z})}{d\mathbf{z}}|_{\mathbf{z}=\mathbf{z}^*}(\mathbf{z} - \mathbf{z}^*) \tag{30}$$

and ignore quadratic and higher order terms in $\mathbf{x} - \mathbf{x}^*$ and $\mathbf{z} - \mathbf{z}^*$. In the sequel, we refer to $g(\mathbf{x}, \mathbf{z}^*)$, the transcriptional response when there is no noise in the exogenous signal by $g(\mathbf{x})$. Details are presented in Appendix B, where we show that for this linearized transcriptional response, $\mathbf{x}^*$ is the solution to Eq. 16 and the total protein noise CV_{tot} is given by

$$CV_{tot}^2 = CV_{int}^2 + CV_{ext}^2 \tag{31}$$

where CV_{int} is the intrinsic noise and

$$CV_{ext} = \frac{T_r}{T_{nr}}\sqrt{\frac{T_z}{T_z + T_r}}SCV_z, \quad S = \frac{\mathbf{z}^*}{g(\mathbf{x}^*, \mathbf{z}^*)}\frac{dg(\mathbf{x}^*, \mathbf{z})}{d\mathbf{z}}|_{\mathbf{z}=\mathbf{z}^*}, \quad T_z = \frac{\ln(2)}{d_z} \tag{32a}$$

represents the extrinsic noise in the protein population. Note that signals $\mathbf{z}$ with small response times T_z result in smaller values of CV_{ext} because rapid fluctuations in the exogenous signal are averaged out by the dynamics of the gene network. Typically, only those exogenous signals that have response times much larger than the protein's response time, contribute significantly to the extrinsic component of protein noise.

The extrinsic noise CV_{ext} is a monotonically increasing function of the protein response time T_r, which in turn is determined by the slope of the transcriptional response $g(\mathbf{x})$ at $\mathbf{x} = \mathbf{x}^*$ (see Eq. 18). This is in contrast to the intrinsic noise CV_{int} which is determined by the y-intercept of the tangent to the transcriptional response $g(\mathbf{x})$ at $\mathbf{x} = \mathbf{x}^*$ (see Eq. 20). Another important difference is that unlike CV_{int}, CV_{ext} does not depend on gene

expression parameters such as the average number N_x and variance V_x^2 of proteins produced per gene expression event.

We now contrast how rapidly intrinsic and extrinsic noise attenuate as the response time T_r is decreased. We first express CV_{ext} as a function of the extrinsic noise level CV_{ext-nr} that would be observed in the absence of feedback:

$$CV_{ext} = \frac{T_r}{T_{nr}} \sqrt{\frac{T_{nr} + T_z}{T_r + T_z}} CV_{ext-nr}, \tag{33}$$

where

$$CV_{ext-nr} = \sqrt{\frac{T_z}{T_z + T_{nr}}} SCV_z. \tag{34}$$

From Eq. 33, we conclude that the five fold decrease in the response time (i.e., $T_r = T_{nr}/5$) that we had observed in Figure 2 corresponds to a reduction of CV_{ext} by a factor of 3.9 compared to CV_{ext-nr} when $T_z \approx T_{nr}$ or a reduction by a factor of 5 when $T_z \gg T_{nr}$. Recall from the previous section that a five fold decrease in the response time leads to a reduction of intrinsic noise level in the protein by a factor of only 2.2 (assuming that $\mathbf{x}^*$ is kept fixed). This illustrates an important point: negative feedback is much more effective in reducing the extrinsic component of protein noise than its intrinsic component.

5 Auto-regulatory gene networks with negative feedback

In this section we quantify protein noise levels and limits of noise suppression for auto-regulatory gene networks involving a common form of transcriptional response given by

$$g(\mathbf{x}) = g_0 \left(b + \frac{1-b}{1+(a\mathbf{x})^M} \right), \quad 0 \leq b < 1 \tag{35}$$

where $M \geq 1$ denotes the hill coefficient and g_0 is the transcription rate when there is no feedback ($a = 0$) (8, 19). The constant b is chosen such that the product $g_0 b$ represents the minimum level of transcription rate (also called the basal level of transcription rate) that is achieved when the number of protein molecules is very large ($\mathbf{x} \to \infty$). Typically, the constant b is either zero or much smaller than one. The constant a is the feedback strength that depends on the binding affinity of the protein to the promoter, with lower binding affinities corresponding to smaller values of a.

For the above transcriptional response we conclude from Eq. 16 that the equilibrium $\mathbf{x}^*$ is the unique solution to

$$N_x g(\mathbf{x}^*) = N_x g_0 \left(b + \frac{1-b}{1+(a\mathbf{x})^M} \right) = d_x \mathbf{x}^* \tag{36}$$

and monotonically decreases as we increase a. The response time T_r in Eq.

18 is given by

$$T_r = \frac{T_{nr}(1+(a\mathbf{x}^*)^M)(1+b(a\mathbf{x}^*)^M)}{1+[1+M-b(M-1)](a\mathbf{x}^*)^M+b(a\mathbf{x}^*)^{2M}}, \qquad T_{nr} = \frac{\ln(2)}{d_x} \tag{37}$$

which starts by decreasing as we increase the feedback strength a from an initial value of $a = 0$. It achieves a minimum value of

$$T_{rmin} = T_{nr}\frac{1+\sqrt{b}}{1+M-\sqrt{b}(M-1)} \tag{38}$$

for

$$a = a_{Tmin} = \frac{b^{-\frac{M+1}{2M}}d}{g_0 N} \tag{39}$$

and then increases as we increase a beyond a_{Tmin}. Note that when $b = 0$, then $a_{Tmin} = \infty$ and T_r always decreases as we increase the feedback strength a with the asymptote

$$\lim_{a\to\infty} T_r = \frac{T_{nr}}{M+1}. \tag{40}$$

In the sections below we investigate how the different components of the noise and the total noise in the protein numbers change as the feedback strength varies.

5.1 Suppression of intrinsic noise in the protein

We first investigate the intrinsic component of noise given by equation Eq. 21 for this specific transcriptional response. Substituting Eq. 35 in Eq. 21,

and using Eq. 36 and Eq. 37, we conclude that the intrinsic noise CV_{int} in the protein is given by

$$CV_{int} = \sqrt{\frac{T_r}{T_{nr}} \frac{1 + (a\mathbf{x}^*)^M}{1 + b(a\mathbf{x}^*)^M}} CV_{int-nr} \tag{41a}$$

$$= \sqrt{\frac{[1 + (a\mathbf{x}^*)^M]^2}{1 + [1 + M - b(M-1)](a\mathbf{x}^*)^M + b(a\mathbf{x}^*)^{2M}}} CV_{int-nr} \tag{41b}$$

where

$$CV_{int-nr} = \sqrt{\frac{d_x(N_x^2 + V_x^2 + N_x)}{2g_0 N_x^2}} \tag{42}$$

is the intrinsic noise in the protein when there is no feedback (i.e. $a = 0$). Our goal is to understand how CV_{int}^2 varies with the hill coefficient M and the feedback strength a. Straightforward calculus shows that the above intrinsic noise is smallest when the feedback strength is equal to

$$a_{int-min} = \frac{d}{Ng_0} \frac{2M}{M + 1 + b(M-1)} \left(\frac{M-1}{M+1}\right)^{\frac{1}{M}} \tag{43}$$

and the corresponding minimum intrinsic noise $CV_{int-min}$ is given by

$$CV_{int-min} = \sqrt{\frac{4M}{4M + (1-b)(M-1)^2}} CV_{int-nr} \leq CV_{int-nr}. \tag{44}$$

When $M = 1$, then $a_{int-min} = 0$ and $CV_{int-min} = CV_{int-nr}$, i.e., the intrinsic noise level is minimum when there in no feedback. In this particular case, increasing a causes CV_{int} to monotonically increase (see Figure 3). This happens because as we increase a from zero, both T_r and $\mathbf{x}^*$ decrease

in Eq. 21. However, as $\mathbf{x}^*$ decreases at a faster rate than T_r, their ratio $T_r/\mathbf{x}^*$ increases, and hence, the intrinsic noise increases as we increases as the feedback strength a. When $M > 1$, the intrinsic noise first decreases when we increase a from zero and achieves a minimum at some optimal value $a = a_{int-min} > 0$. Increasing a beyond $a_{int-min}$ causes an increase in the intrinsic noise level (see Figure 3).

From Eq. 44, the quantity

$$\frac{CV_{int-min}}{CV_{int-nr}} = \sqrt{\frac{4M}{4M + (1-b)(M-1)^2}} \tag{45}$$

represents the limit of suppression of intrinsic noise in the protein from CV_{int-nr}. Notice that this limit decreases as we decrease the basal level of transcription (i.e., decrease b) and achieves a minimum at $b = 0$. When $b \ll 1$, this limit of intrinsic noise suppression is given by

$$\frac{CV_{int-min}}{CV_{int-nr}} = \frac{\sqrt{4M}}{M+1} \tag{46}$$

and is completely determined by the hill coefficient M, with larger values of M causing more reduction in the protein intrinsic noise. This is consistent with results in the literature, which show that a large hill coefficient is more effective in reducing stochastic fluctuations in the protein (11, 19, 29). For example, when $b = 0$, and $M = 2$ there can be at most a $1-\sqrt{4M}/(M+1) = 5.7\%$ reduction in intrinsic noise from CV_{int-nr} where as for $M = 4$ we can have a 20% reduction.

In summary, depending on the hill coefficient, the protein intrinsic noise

levels can either monotonically increase or exhibit a U-shaped curve as the feedback strength is increased. Moreover, large hill coefficients are much more effective in reducing noise.

5.2 Suppression of extrinsic noise in the protein

We now investigate the extrinsic component of protein noise CV_{ext}. As CV_{ext} is a monotonically increasing function of the response time T_r, it will be minimum when the response time is the smallest. Recall from Eq. 38, Eq. 39 that the response time T_r achieves its minimum value T_{rmin} when the feedback strength is equal to $a = a_{Tmin}$. Hence we conclude that level of extrinsic noise is minimum when feedback strength is equal to

$$a_{ext-min} = \frac{b^{-\frac{M+1}{2M}} d}{g_0 N}, \tag{47}$$

and from Eq. 33, this minimum level $CV_{ext-min}$ is given by

$$\frac{CV_{ext-min}}{CV_{ext-nr}} = \sqrt{\frac{T_{nr}+T_z}{T_{rmin}+T_z}}\frac{T_{rmin}}{T_{nr}}, \quad T_{rmin} = \frac{1+\sqrt{b}}{1+M-\sqrt{b}(M-1)}T_{nr}. \tag{48}$$

The above expression provides the limit of extrinsic noise suppression and reduces to

$$\frac{CV_{ext-min}}{CV_{ext-nr}} = \begin{cases} \sqrt{\frac{T_{nr}+T_z}{[T_{nr}+T_z(M+1)](M+1)}} & \text{when} \quad b \ll 1 \\ \frac{1}{M+1} & \text{when} \quad b \ll 1 \text{ and } T_z \gg T_{nr}. \end{cases} \tag{49}$$

As we increase M these limits decrease at a much faster rate then the limit of intrinsic noise suppression for the same value of b (compare with right-hand-side of Eq. 46). For example, when $T_z \approx T_{nr}$ and $b \ll 1$, for $M = 2$ we have a maximum reduction in extrinsic noise of $1 - \sqrt{2/[(M+1)(M+2)]} \approx 42\%$ whereas for $M = 4$ we have a reduction of 74%. These reductions are much larger than the maximum reductions of 5.7% and 20% in the protein intrinsic noise level for the same values of M and b (see previous section). This reinforces the earlier point that negative feedback is much more efficient in reducing the extrinsic component of the noise than its intrinsic component.

5.3 Suppression of total noise in the protein

Finally, we investigate how the total noise in the protein population varies with the feedback strength. As derived in Section 4, the total protein noise level is given by

$$CV_{tot}^2 = CV_{int}^2 + CV_{ext}^2, \tag{50}$$

which using Eq. 32, Eq. 37 and Eq. 41 can be written as

$$CV_{tot}^2 = CV_{int-nr}^2 \frac{T_r}{T_{nr}} \frac{1 + (a\mathbf{x}^*)^M}{1 + b(a\mathbf{x}^*)^M} + S^2 CV_z^2 \left(\frac{T_r}{T_{nr}}\right)^2 \frac{T_z}{T_z + T_r} \tag{51a}$$

$$T_r = \frac{T_{nr}(1 + (a\mathbf{x}^*)^M)(1 + b(a\mathbf{x}^*)^M)}{1 + [1 + M - b(M-1)](a\mathbf{x}^*)^M + b(a\mathbf{x}^*)^{2M}}. \tag{51b}$$

Now, for all $M \geq 1$ and $CV_z > 0$ we have that

$$\frac{dCV_{tot}^2}{da}|_{a=0} = -(1-b)\left[(CV_{int-nr}^2)(M-1) + CV_z^2 M T_z \frac{(2T_z + T_{nr})}{(T_z + T_{nr})^2}\right] < 0, \tag{52}$$

which means that in the presence of extrinsic noise, the total protein noise level will always decrease as we increase the feedback strength from $a = 0$, irrespective of the value of the hill coefficient, but eventually will start to increase for sufficiently large values of a past an optimal feedback strength a_{min}. In summary, in the presence of extrinsic noise, the total noise in the protein is always minimized at some optimal feedback strength and decreasing or increasing feedback strength away from this optima will always causes an increase in the noise level. This point is shown in Figure 4 which plots CV_{tot}/CV_{tot-nr} as a function of a when the hill coefficient in one and CV_{tot-nr} given by

$$CV_{tot-nr}^2 = CV_{int-nr}^2 + S^2 CV_z^2 \frac{T_z}{T_z + T_{nr}}, \tag{53}$$

represents the protein noise level when there is no feedback. In the absence of extrinsic noise ($CV_z = 0$), CV_{tot}/CV_{tot-nr} monotonically increases as the feedback strength is increased. However, in the presence of extrinsic noise, it follows a U-shaped profile and is minimized at some $a = a_{min} > 0$.

When the hill coefficient is larger than one ($M > 1$), then even in the absence of any extrinsic noise ($CV_z = 0$), the protein noise level will show a U-shaped profile as the feedback strength is altered (see Section 5.1). In

particular, for $CV_z = 0$, we conclude from Eq. 45 that the minimum value of CV_{tot}/CV_{tot-nr}, i.e., the limit of noise suppression is given by

$$\frac{CV_{tot-min}}{CV_{tot-nr}} = \sqrt{\frac{4M}{4M + (1-b)(M-1)^2}} \tag{54}$$

and is attained when the feedback strength is equal to

$$a_{min} = \frac{d}{Ng_0} \frac{2M}{M+1+b(M-1)} \left(\frac{M-1}{M+1}\right)^{\frac{1}{M}}. \tag{55}$$

As shown in Figure 4 (for $M = 1$) and Figure 5 (for $M = 2$), when we now increase CV_z away from zero, this limit of noise suppression decreases and is much lower than as predicted by Eq. 54. On the other hand the optimal feedback strength a_{min} where the protein noise is minimum, increases and is much higher than Eq. 55. As we further increase the noise CV_z in the exogenous signal, both $CV_{tot-min}/CV_{tot-nr}$ and a_{min} approach Eq. 48 and Eq. 47, respectively, which correspond to the scenario where extrinsic noise dominates the total noise in protein numbers.

In Appendix C, we provide formulas that predict both the minimum level of noise $CV_{tot-min}$ and the optimal feedback strength when both intrinsic and extrinsic noise are present but neither dominates the total noise in the protein population. As we will illustrate later, an important application of these formulas is that one can directly compute the noise in the exogenous signal from the experimentally obtained value of $CV_{tot-min}$.

6 Experimental verification

We now validate our theoretical results with recent experimental measurements of protein noise levels that were obtained as the feedback strength was changed via experimental manipulation. In (20) a synthetic auto-regulatory gene network is constructed where the protein inhibits its own transcription. The feedback strength is altered by adding a compound aTc that binds to the protein and the resulting complex has a significantly smaller binding affinity to the promoter. As the feedback strength is directly related to the binding affinity of the protein to its promote, increasing the concentration of aTc corresponds to decreasing the feedback strength a. The gene is encoded on a low-copy plasmid with high variability in plasmid population contributing to large levels of extrinsic noise in the protein population. Based on our theoretical analysis the protein noise level should show a U-shaped profile as the feedback strength is changed. In particular, at low values of a (i.e., high levels of aTc), increasing a (i.e., deceasing aTc) should lead to a decrease in protein noise levels. However, at high values of a (i.e., low levels of aTc), increasing a (i.e., deceasing aTc) should increase the protein noise levels. Such a U-shaped profile is indeed what is experimentally observed and the protein noise level is minimized at an optimal level of feedback strength (see bottom left plot of (20)).

In (20), the results from detailed stochastic simulations of the auto-regulatory gene network are also reported. The authors observe in simulation that both in the absence of any extrinsic noise or when the extrinsic noise form only the enzyme RNA polymerase is included, instead of seeing

a U-shaped profile, the protein noise level monotonically increased as the feedback strength is increased (i.e., aTc concentration is decreased). Our theoretical results fully explain this phenomenon: As for this synthetic gene network the hill coefficient is one ($M = 1$), our analysis in Section 5.1 shows that the intrinsic noise level will always increases if the feedback strength is increased. As the extrinsic noise associated with fluctuations in RNA polymerase numbers is very small [we calculate $CV_{RNA\ polymearse} \approx 0.02$ using Eq. 29 and the reaction rates provided in Table I of (20)], in both the above cases the protein noise is dominated by the intrinsic noise which always increases with the feedback strength, and hence, no U-shaped profile should be observed.

As mentioned earlier, our results also allow us to predict the level of noise in the exogenous signal that drives the synthetic auto-regulated gene network. Assuming the source of extrinsic noise to be the plasmid population, and using the experimentally obtained minimal noise level of approximately 0.4, we estimate in Appendix D that

$$CV_{plasmid} \approx 0.63. \tag{56}$$

Independent measurements of plasmid noise [using Eq. 29 and the reaction rates provided in Table I of (20)] show that $CV_{plasmid}$ is equal to 0.51 which is slightly smaller than as given by Eq. 56. This confirms that variability in plasmid numbers is indeed the major source of extrinsic noise in the protein population. The fact that the estimate in Eq. 56 is larger than the actual plasmid noise suggests that variability in other cellular components also

make minor contributions to the extrinsic noise.

In summary, the experimental results of (20) provide an experimental verification of our theoretical predictions. They also indicate that measuring changes in the protein noise level as a function of the feedback strength could be useful in determining the level of noise in the exogenous signal.

7 Discussion

Auto-regulatory gene networks where the protein inhibits/activates its own transcription are common motifs occurring within living cell. These networks are characterized by their transcriptional response $g(\mathbf{x})$ which provides information on how the transcription rate of the gene varies as a function of the number of protein molecules $\mathbf{x}$ present in the cell.

7.1 Noise dependence on the shape of the transcriptional response

We developed a full understanding of how the protein noise levels are related to the functional form of the transcriptional response. Using a linear approximation for g($\mathbf{x}$), we showed that the extrinsic noise levels are determined by the slope $g'(\mathbf{x}^*)$ of the transcriptional response at $\mathbf{x}^*$, with more negative values of the slope (i.e., more stable equilibriums $\mathbf{x}^*$) leading to smaller levels of extrinsic noise. On the other hand, the intrinsic noise levels are determined by $I = g(\mathbf{x}^*) - \mathbf{x}^* g'(\mathbf{x}^*)$ which is the y-intercept of the tangent to the transcriptional response at $\mathbf{x} = \mathbf{x}^*$ (as shown in Figure 1), and larger values of I lead to smaller levels of intrinsic noise. Consequently, given two

hypothetical transcriptional responses $g_1(\mathbf{x}) = 1$ and $g_2(\mathbf{x}) = 1 - \mathbf{x}/2$, the response $g_2(\mathbf{x})$ will give lower levels of extrinsic noise. However, since both transcriptional responses have the same intercept I equal to one, they both yield the same level of intrinsic noise in the protein population.

We considered deviations from a linear transcriptional response and showed that concave responses have better noise suppression properties compared to linear and convex responses. We also derived appropriate conditions under which a linear $g(\mathbf{x})$ will accurately predict the protein noise level.

7.2 Intrinsic and extrinsic noise

Analytical formulas that relate the noise levels to the response time of the protein show key differences between extrinsic and intrinsic noise: as one decreases the protein response time T_r through feedback, the levels of extrinsic noise decreases much more than those of intrinsic noise. This leads to an important conclusion that negative feedback is much more effective in reducing the extrinsic component of protein noise than its intrinsic component, which is consistent with other theoretical and experimental studies (18, 30, 31). Another difference is that unlike intrinsic noise, the extrinsic noise is independent of the average burst size N_x and the variance V_x^2 in the number of protein molecules produced in each transcription event. Recent work (32) has suggested that many genes operate with very low values of N_x, which could be an adaptation to reduce the intrinsic noise but will have no effect on the extrinsic noise levels.

7.3 U-shaped protein noise profile

We investigated how protein noise levels change as we vary the feedback strength for a biologically meaningfull class of auto-regulatory gene networks with negative feedback and characterized by the transcriptional response

$$g(\mathbf{x}) = g_0 \left(b + \frac{1-b}{1+(a\mathbf{x})^M} \right), \quad 0 \leq b < 1. \tag{57}$$

Our main result shows that the total noise level in the protein population is minimized at an optimal level of feedback strength. Recall from Section 5 that increasing the feedback strength causes a decrease in the average number of protein molecules, which results in an increase in the intrinsic noise level. On the other hand increasing the feedback strength causes the protein response time to decrease which attenuates both the intrinsic and extrinsic noise. The net result of these two opposing effects is a U-shaped profile where increasing feedback strength first causes the noise level to decrease and then increase. This U-shaped profile was shown to be in good agreement with experimental data for a synthetic auto-regulatory gene network. We also identified a scenario where noise is minimum when there is no feedback and any amount of negative feedback will always increase the noise: the case where intrinsic noise dominates the total noise in the protein population and the hill coefficient is close to one. This explained the observation that when the source of the extrinsic noise was removed, the U-shaped profile vanished, and instead, the noise level monotonically increased with the feedback strength. However, if the synthetic gene network was characterized by a hill coefficient much larger than one, then our theory predicts that

even in the absence of extrinsic noise a U-shaped profile should have been observed. This remains to be experimentally verified.

7.4 Limit of noise suppression

We characterized the smallest level of noise that is inherent to this type of auto-regulation by the limit of noise suppression. This limit, which can be interpreted as the depth of the U-shape profile, was defined as the ratio of the minimum possible noise in the protein population with feedback to the protein noise level when there is no feedback (i.e., $a = 0$). For auto-regulatory networks with small basal level of transcription (i.e. $b \approx 0$) this limit is given by

$$\frac{\sqrt{4M}}{M+1} \tag{58}$$

when the intrinsic noise dominates the total noise in the protein (see equation Eq. 46). As we now increase the amount of extrinsic noise this limit decreases and asymptotically approaches a values of

$$\sqrt{\frac{T_{nr}+T_z}{[T_{nr}+T_z(M+1)](M+1)}}, \tag{59}$$

which corresponds to the situation where extrinsic noise dominates the total noise in the protein (see equation Eq. 48). We also showed that the optimal level of feedback strength where the protein noise level is minimum monotonically increases with increasing levels of extrinsic noise in the protein population. Thus if negative feedback loops indeed function to min-

imize protein noise, then networks with larger contributions from extrinsic noise compared to intrinsic noise, should operate at higher levels of feedback strength, i.e., higher binding affinities between the protein and its promoter.

The above results can be used to quantify the level of extrinsic noise in the protein population. This is useful for synthetic and natural auto-regulatory gene networks where the feedback strength can be manipulated. As illustrated, noise in the exogenous signal can then directly be estimated from the minimum possible protein noise. Matching these estimates with independent measurements of noise in the exogenous signal can also be done to confirm that a particular noise source is indeed the major contributor of extrinsic noise in the protein population.

7.5 Positive feedback loops

Our analysis can also be used for auto-regulatory gene networks with positive feedback. These networks are characterized by a transcriptional response $g(\mathbf{x})$ as in Eq. 57 but for a constant b larger than one. A similar analysis reveals that instead of being minimized, the protein noise levels are maximized at the optimal level of feedback strength. Thus protein noise levels follow an inverted U-shape profile as the feedback strength is increased. Moreover, for positive feedback loops characterized by a hill coefficient of one (such as the Tat protein in the HIV gene network (33, 34)), and when intrinsic noise dominates the total noise, the noise level monotonically decrease with increasing feedback strength. Such differences in noise profiles can again be exploited to determine the level of extrinsic noise in these gene networks.

In summary, we have developed results relating the noise levels to the feedback strength in auto-regulatory gene networks. We have show that for negative feedback loops, protein noise levels are always minimized at an optimal level of feedback strength. The noise resulting from these optimal levels of feedback characterizes the smallest level of noise that can be achieved in these networks through the use of negative feedback. Our results have implications for the design of synthetic auto-regulatory gene networks with minimal protein noise. They also raise the question whether these widely occurring auto-regulatory gene networks (for example, over 40% of known Escherichia coli transcription factors negatively regulate their own transcription (27)) have naturally evolved to operate at the optimal feedback strength. If these networks have indeed evolved to operate at the optimal point then any mutation or experimental manipulation that changes the feedback strength should always cause the protein noise levels to increase.

8 Acknowledgements

We thank Mustafa Khammash for many discussions of the work reported in this paper. This material is based upon work supported by the Institute for Collaborative Biotechnologies through grant DAAD19-03-D-0004 from the U.S. Army Research Office and by the National Science Foundation under Grant No. CCR-0311084.

A Appendix : Effects of nonlinearities in $g(\mathbf{x})$

For the transcriptional response given by Eq. 23 the time evolution of the moments $\mathbf{E}[\mathbf{x}]$, $\mathbf{E}[\mathbf{x}^2]$ are given by

$$\begin{aligned}\frac{d\mathbf{E}[\mathbf{x}]}{dt} =& g(\mathbf{x}^*)N_x - d_x\mathbf{E}[\mathbf{x}] + g'(\mathbf{x}^*)N_x\mathbf{E}[\mathbf{x}] + \frac{g''(\mathbf{x}^*)}{2}N_x\mathbf{E}[\mathbf{x}^2] - g'(\mathbf{x}^*)N_x\mathbf{x}^* \\ &- g''(\mathbf{x}^*)N_x\mathbf{E}[\mathbf{x}]\mathbf{x}^* + \frac{g''(\mathbf{x}^*)}{2}N_x\mathbf{x}^{*2} \qquad (60a)\\ \frac{d\mathbf{E}[\mathbf{x}^2]}{dt} =& g(\mathbf{x}^*)(N_x^2+V_x^2) + dr\mathbf{E}[\mathbf{x}] + 2g(\mathbf{x}^*)N_x\mathbf{E}[\mathbf{x}] + g'(\mathbf{x}^*)(N_x^2+V_x^2)\mathbf{E}[\mathbf{x}] \\ &- 2d_x\mathbf{E}[\mathbf{x}^2] + 2g'(\mathbf{x}^*)N_x\mathbf{E}[\mathbf{x}^2] + \frac{g''(\mathbf{x}^*)}{2}(N_x^2+V_x^2)\mathbf{E}[\mathbf{x}^2] \\ &+ g''(\mathbf{x}^*)N_x\mathbf{E}[\mathbf{x}^3] - g'(\mathbf{x}^*)(N_x^2+V_x^2)\mathbf{x}^* - 2g'(\mathbf{x}^*)N_x\mathbf{E}[\mathbf{x}]\mathbf{x}^* \\ &- g''(\mathbf{x}^*)(N_x^2+V_x^2)\mathbf{E}[\mathbf{x}]\mathbf{x}^* - 2g''(\mathbf{x}^*)N_x\mathbf{E}[\mathbf{x}^2]\mathbf{x}^* \\ &+ \frac{g''(\mathbf{x}^*)}{2}(N_x^2+V_x^2)\mathbf{x}^{*2} + g''(\mathbf{x}^*)N_x\mathbf{E}[\mathbf{x}]\mathbf{x}^{*2} \qquad (60b)\end{aligned}$$

which can be written more compactly as

$$\begin{bmatrix} \frac{d\mathbf{E}[\mathbf{x}]}{dt} \\ \frac{d\mathbf{E}[\mathbf{x}^2]}{dt} \end{bmatrix} = \mathbf{a} + \mathbf{A}\begin{bmatrix} \mathbf{E}[\mathbf{x}] \\ \mathbf{E}[\mathbf{x}^2] \end{bmatrix} + \mathbf{B}\ \mathbf{E}[\mathbf{x}^3], \qquad (61)$$

for some vector $\mathbf{a}$ and matrices $\mathbf{A}$ and $\mathbf{B}$. The above moment dynamics is not closed as the time derivative of $\mathbf{E}[\mathbf{x}]$ and $\mathbf{E}[\mathbf{x}^2]$ depends on $\mathbf{E}[\mathbf{x}]$, $\mathbf{E}[\mathbf{x}^2]$ and $\mathbf{E}[\mathbf{x}^3]$. We use moment closure techniques to approximate $\mathbf{E}[\mathbf{x}^3]$ as a nonlinear function of $\mathbf{E}[\mathbf{x}]$ and $\mathbf{E}[\mathbf{x}^2]$ and close the above set of differential equations. A standard assumption at this point is to assume that the third cumulant of the distribution is zero, which will be valid approximation as long as the distribution is symmetrically distributed about its mean.

Referring the reader to (35–37) for further details we approximate $\mathbf{E}[\mathbf{x}^3]$ as

$$\mathbf{E}[\mathbf{x}^3] \approx 3\mathbf{E}[\mathbf{x}]\mathbf{E}[\mathbf{x}^2] - 2\mathbf{E}[\mathbf{x}]^3. \tag{62}$$

Denoting the steady-state values of the moments $\mathbf{E}[\mathbf{x}]$ and $\mathbf{E}[\mathbf{x}^2]$ by $\mathbf{x}_q^*$ and $\mathbf{E}^*[\mathbf{x}^2]$, respectively, we have from Eq. 61 and Eq. 62 that

$$0 = \mathbf{a} + \mathbf{A}\begin{bmatrix} \mathbf{x}_q^* \\ \mathbf{E}^*[\mathbf{x}^2] \end{bmatrix} + \mathbf{B}\left(3\mathbf{x}_q^*\mathbf{E}^*[\mathbf{x}^2] - 2{\mathbf{x}^*}_q^3\right). \tag{63}$$

Analytically solving for these steady-state moments from Eq. 63 is not an easy task and we use perturbation methods to compute approximate steady-states. This is done by writing $\mathbf{x}_q^*$ as a perturbation about $\mathbf{x}^*$ and $\mathbf{E}^*[\mathbf{x}^2]$ as a perturbation about ${\mathbf{x}^*}_q^2$, as follows

$$\mathbf{x}_q^* := \mathbf{x}^*(1+\epsilon_1), \qquad \mathbf{E}^*[\mathbf{x}^2] := {\mathbf{x}^*}_q^2(1+\epsilon_2). \tag{64}$$

Assuming

$$\epsilon_1^2 \ll 1, \quad \epsilon_2^2 \ll 1, \quad |\epsilon_1\epsilon_2| \ll 1 \tag{65}$$

we have from Eq. 64

$$\mathbf{E}^*[\mathbf{x}^2] \approx {\mathbf{x}^*}^2(1+2\epsilon_1+\epsilon_2) \tag{66a}$$

$$3\mathbf{x}_q^*\mathbf{E}^*[\mathbf{x}^2] - 2{\mathbf{x}^*}_q^3 \approx {\mathbf{x}^*}^3(1+3\epsilon_1+3\epsilon_2). \tag{66b}$$

Substituting Eq. 66 in Eq. 63 we obtain

$$0 = \hat{\mathbf{a}} + \hat{\mathbf{A}} \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix} \tag{67}$$

for some vector $\hat{\mathbf{a}}$ and matrix $\hat{\mathbf{A}}$. Solving for ϵ_1 and ϵ_2 from Eq. 67 we have

$$\epsilon_1 = \frac{1}{1 + \frac{N_x \mathbf{x}^* g''(\mathbf{x}^*) CV_{int}^2}{2\lambda}} - 1, \tag{68a}$$

$$CV_{int-quad}^2 = \frac{\mathbf{E}^*[\mathbf{x}^2] - \mathbf{x}_q^{*2}}{\mathbf{x}_q^{*2}} = \epsilon_2 = \frac{CV_{int}^2}{1 + \frac{N_x \mathbf{x}^* g''(\mathbf{x}^*) CV_{int}^2}{2\lambda}} \tag{68b}$$

where CV_{int}^2 is given by Eq. 20.

B Appendix : Extrinsic and intrinsic contributions of noise

We model the time evolution of the number of molecules $\mathbf{x}$ and $\mathbf{z}$ through a Stochastic Hybrid System (SHS) with state $\mathbf{y} = [\mathbf{z}, \mathbf{x}]^T$ characterized by trivial continuous dynamics

$$\dot{\mathbf{y}} = \begin{bmatrix} \dot{\mathbf{z}} \\ \dot{\mathbf{x}} \end{bmatrix} = 0 \tag{69}$$

and four reset maps

$$\mathbf{y} \mapsto \phi_1(\mathbf{y}) = \begin{bmatrix} \mathbf{z} + \mathbf{N}_z \\ \mathbf{x} \end{bmatrix}, \quad \mathbf{y} \mapsto \phi_2(\mathbf{y}) = \begin{bmatrix} \mathbf{z} - 1 \\ \mathbf{x} \end{bmatrix} \tag{70}$$

$$\mathbf{y} \mapsto \phi_3(\mathbf{y}) = \begin{bmatrix} \mathbf{z} \\ \mathbf{x} + \mathbf{N}_x \end{bmatrix}, \quad \mathbf{y} \mapsto \phi_4(\mathbf{y}) = \begin{bmatrix} \mathbf{z} \\ \mathbf{x} - 1 \end{bmatrix} \tag{71}$$

with corresponding transition intensities given by

$$\lambda_1(\mathbf{y}) = K_z, \ \lambda_2(\mathbf{y}) = d_z\mathbf{z}, \tag{72}$$

$$\lambda_3(\mathbf{y}) = g(\mathbf{x}^*, \mathbf{z}^*) + \frac{dg(\mathbf{x}, \mathbf{z}^*)}{d\mathbf{x}}|_{\mathbf{x}=\mathbf{x}^*}(\mathbf{x} - \mathbf{x}^*) + \frac{dg(\mathbf{x}^*, \mathbf{z})}{d\mathbf{z}}|_{\mathbf{z}=\mathbf{z}^*}(\mathbf{z} - \mathbf{z}^*), \tag{73}$$

$$\lambda_4(\mathbf{y}) = d_x\mathbf{x}. \tag{74}$$

Using Dynkin's equations Eq. 8 we have that the time evolution of all the first and second order moments of $\mathbf{y}$ is given by

$$\begin{bmatrix} \frac{d\mathbf{E}[\mathbf{z}]}{dt} \\ \frac{\mathbf{E}[\mathbf{x}]}{dt} \\ \frac{\mathbf{E}[\mathbf{z}^2]}{dt} \\ \frac{\mathbf{E}[\mathbf{x}^2]}{dt} \\ \frac{\mathbf{E}[\mathbf{zx}]}{dt} \end{bmatrix} = \bar{a} + \bar{A} \begin{bmatrix} d\mathbf{E}[\mathbf{z}] \\ \mathbf{E}[\mathbf{x}] \\ \mathbf{E}[\mathbf{z}^2] \\ \mathbf{E}[\mathbf{x}^2] \\ \mathbf{E}[\mathbf{zx}] \end{bmatrix} \tag{75}$$

for some vector $\bar{a}$ and matrix $\bar{A}$. A steady-state analysis of Eq. 75 shows that the steady-state average number molecules of the protein, $\mathbf{x}^*$, is given as the solution to Eq. 16 and the total noise level in the protein population is given by Eq. 31

C Appendix : Limit of noise suppression

It is not easy to derive an explicit expression for the minimum protein noise, $CV_{tot-min}$. However, for the biologically meaningful case of

$$b \ll 1, \quad T_z \gg T_{nr}, \tag{76}$$

analytical formulations for both $CV_{tot-min}$ and the optimal level of feedback strength which achieves this minimum noise are possible. When these assumptions are true we have from Eq. 51 that the total protein noise level is given by

$$CV_{tot}^2 = CV_{int-nr}^2 \frac{\left(1+(a\mathbf{x}^*)^M\right)^2}{1+(M+1)(a\mathbf{x}^*)^M} + S^2 CV_z^2 \left(\frac{1+(a\mathbf{x}^*)^M}{1+(M+1)(a\mathbf{x}^*)^M}\right)^2. \tag{77}$$

Straightforward calculus shows that the right-hand-side of Eq. 77 is minimum when

$$(a\mathbf{x}^*)^M = \frac{M-2+\sqrt{M}\sqrt{8S^2CV_z^2 + MCV_{int-nr}^2}/CV_{int-nr}}{2(M+1)}, \tag{78}$$

which implies from Eq. 77 and Eq. 36 that

$$CV^2_{tot-min} = \frac{2L}{(1+M)^2\left(\sqrt{M}CV_{int-nr} + \sqrt{8S^2CV_z^2 + MCV^2_{int-nr}}\right)^2}$$

$$\begin{aligned} L &= 4S^4CV_z^4 + 19MS^2CV_z^2CV^2_{int-nr} + 4M^2CV^4_{int-nr} \\ &+ 5\sqrt{M}S^2CV_z^2CV_{int-nr}\sqrt{8S^2CV_z^2 + MCV^2_{int-nr}} \\ &+ 4\sqrt{M^3}CV^3_{int-nr}\sqrt{8S^2CV_z^2 + MCV^2_{int-nr}} \end{aligned} \tag{79}$$

and

$$a_{min} = \frac{d\sqrt{M}}{2g_0N(M+1)}\left(3\sqrt{M} + \sqrt{8S^2CV_z^2 + MCV^2_{int-nr}}/CV_{int-nr}\right)P \tag{80a}$$

$$P = \left(\frac{M-2+\sqrt{M}\sqrt{8S^2CV_z^2 + MCV^2_{int-nr}}/CV_{int-nr}}{2(M+1)}\right)^{\frac{1}{M}}. \tag{80b}$$

Given estimates of $CV_{tot-min}$ and CV_{int-nr}, one can compute CV_z from Eq. 79. We also conclude from Eq. 53 that when $T_z \gg T_{nr}$, we have

$$CV^2_{tot-nr} = CV^2_{int-nr} + S^2CV_z^2. \tag{81}$$

Hence, given experimental measurements of the noise CV_{tot-nr} when there is no feedback and the minimal noise $CV_{tot-min}$ in the protein population, one can also determine SCV_z by simultaneously solving equations Eq. 79 and Eq. 81.

In addition to the above assumptions (i.e., $b \ll 1$ and $T_z \gg T_{nr}$) if we

also have that

$$CV^2_{int-nr} \ll S^2 CV^2_z, \tag{82}$$

then equations Eq. 79 reduce to

$$CV^2_{tot-min} \approx \frac{S^2 CV^2_z}{(1+M)^2} + \frac{5SCV_z CV_{int-nr}\sqrt{M}}{\sqrt{2}(1+M)^2}. \tag{83}$$

D Appendix : Estimating the noise in the exogenous signal

Assuming the source of extrinsic noise to be the plasmid population, we have from Section 4 that $g(\mathbf{x}, \mathbf{z}) = \mathbf{z}g(\mathbf{x})$ and therefore

$$S = \frac{\mathbf{z}^*}{g(\mathbf{x}^*, \mathbf{z}^*)} \frac{dg(\mathbf{x}^*, \mathbf{z})}{d\mathbf{z}}|_{\mathbf{z}=\mathbf{z}^*} = 1. \tag{84}$$

For this synthetic auto-regulatory gene network $b = 0$ and we calculate from the reaction rates provided in Table I of (20),

$$CV^2_{int-nr} \approx 0.008, \quad T_{nr}/T_z \approx 0.1. \tag{85}$$

Using $CV_{tot-min} \approx 0.4$ in Eq. 83 we obtain $CV_z = CV_{plasmid}$ to be approximately 0.63.

References

1. Walters, M. C., S. Fiering, J. Eidemiller, W. Magis, M. Groudine, and D. I. K. Martin, 1995. Enhancers increase the probability but not the level of gene expression. *Proceedings of the National Academy of Sciences U.S.A* 92:7125–7129.

2. Arkin, A., J. Ross, and H. H. McAdams, 1998. Stochastic Kinetic Analysis of Developmental Pathway Bifurcation in Phage λ-Infected *Escherichia coli* Cells. *Genetics* 149:1633–1648.

3. Blake, W. J., M. Krn, C. R. Cantor, and J. J. Collins, 2003. Noise in eukaryotic gene expression. *Nature* 422:633–637.

4. Elowitz, M. B., A. J. Levine, E. D. Siggia, and P. S. Swain, 2002. Stochastic Gene Expression in a Single Cell. *Science* 297:1183–1186.

5. Raser, J. M., and E. K. O'Shea, 2005. Noise in Gene Expression: Origins, Consequences, and Control. *Science* 309:2010 – 2013.

6. Spudich, J. L., and D. E. K. Jr, 1976. Non-genetic individuality: chance in the single cell. *Nature (London)* 262:467–471.

7. McAdams, H. H., and A. P. Arkin, 1997. Stochastic mechanisms in gene expression. *Proceedings of the National Academy of Sciences U.S.A* 94:814–819.

8. Alon, U., 2007. Network motifs: theory and experimental approaches. *Nature Reviews Genetics* 8:450–461.

9. Wall, M. E., W. S. Hlavacek, and M. A. Savageau, 2003. Design Principles for Regulator Gene Expression in a Repressible Gene Circuit. *J. of Molecular Biology* 332:861–876.

10. Savageau, M. A., 1974. Comparison of classical and autogenous systems of regulation in inducible operons. *Nature* 252:546–549.

11. Orrell, D., and H. Bolouri, 2004. Control of internal and external noise in genetic regulatory networks. *J. of Theoretical Biology* 230:301–312.

12. Tao, Y., X. Zheng, and Y. Sun, 2007. Effect of feedback regulation on stochastic gene expression. *J. of Theoretical Biology* 247:827–836.

13. Becskei, A., and L. Serrano, 2000. Engineering stability in gene networks by autoregulation. *Nature* 405:590–593.

14. Tomioka, R., H. Kimura, T. J. Kobayashi, and K. Aihara, 2004. Multivariate analysis of noise in genetic regulatory networks. *J. of Theoretical Biology* 229:501–521.

15. Hasty, J., J. Pradines, M. Dolnik, and J. J. Collins, 2000. Noise-based switches and amplifiers for gene expression. *PNAS* 97:2075–2080.

16. Brandman, O., J. E. Ferrell, R. Li, and T. Meyer, 2005. Interlinked Fast and Slow Positive Feedback Loops Drive Reliable Cell Decisions. *Science* 310:496 – 498.

17. Swain, P. S., M. B. Elowitz, and E. D. Siggia, 2002. Intrinsic and extrinsic contributions to stochasticity in gene expression. *PNAS* 99:12795–12800.

18. Paulsson, J., 2004. Summing up the noise in gene networks. *Nature* 427:415–418.

19. Thattai, M., and A. van Oudenaarden, 2001. Intrinsic noise in gene regulatory networks. *Proceedings of the National Academy of Sciences U.S.A* 98:8614–8619.

20. Dublanche, Y., K. Michalodimitrakis, N. Kummerer, M. Foglierini, and L. Serrano, 2006. Noise in transcription negative feedback loops: simulation and experimental analysis. *Molecular Systems Biology* .

21. Longo, D., and J. Hasty, 2006. Imaging gene expression: Tiny signals make a big noise. *Nature Chemical Biology* 2:181–182.

22. Paulsson, J., 2005. Model of stochastic gene expression. *Physics of Life Reviews* 2:157–175.

23. Hespanha, J. P., and A. Singh, 2005. Stochastic Models for Chemically reacting Systems Using Polynomial Stochastic Hybrid Systems. *Int. J. of Robust and Nonlinear Control* 15:669–689.

24. Davis, M. H. A., 1993. Markov models and Optimization. Chapman and Hall.

25. Hespanha, J. P., 2004. Stochastic Hybrid Systems: Applications to Communication Networks. *In* R. Alur, and G. J. Pappas, editors, Hybrid Systems: Computation and Control, Springer-Verlag, Berlin, number 2993 in Lect. Notes in Comput. Science, 387–401.

26. Ozbudak, E. M., M. Thattai, I. Kurtser, A. D. Grossman, and A. van Oudenaarden, 2002. Regulation of noise in the expression of a single gene. *Nature Genetics* 31:69–73.

27. Rosenfeld, N., M. B. Elowitz, and U. Alon., 2002. Negative autoregulation speeds the response times of transcription networks. *J. Molecular Biology* 323:785–793.

28. Tsimring, L. S., D. Volfson, and J. Hasty, 2006. Stochastically driven genetic circuits. *CHAOS* 16.

29. Singh, A., and J. P. Hespanha, 2006. Stochastic Analysis of Gene Regulatory Networks Using Moment Closure. *In* Proc. of the 2007 Amer. Control Conference, New York, NY.

30. Hooshangi, S., and R. Weiss, 2006. The effect of negative feedback on noise propagation in transcriptional gene networks. *CHAOS* 16.

31. Shahrezaei, V., J. F. Ollivier, and P. S. Swain, 2008. Colored extrinsic fluctuations and stochastic gene expression. *Molecular Systems Biology* 4.

32. Fraser, H. B., A. E. Hirsh, G. Giaever, J. Kumm, and M. B. Eisen, 2004. Noise Minimization in Eukaryotic Gene Expression. *PLoS Biology* 2.

33. Weinberger, L. S., and T. Shenk, 2007. An HIV Feedback Resistor: Auto-Regulatory Circuit Deactivator and Noise Buffer. *PLoS Biology* 5.

34. Weinberger, L. S., R. D. Dar, and M. L. Simpson, 2008. Transient-mediated fate determination in a transcriptional circuit of HIV. *Nature Genetics* 40:466 – 470.

35. Gmez-Uribe, C. A., and G. C. Verghese, 2007. Mass fluctuation kinetics: Capturing stochastic effects in systems of chemical reactions through coupled mean-variance computations. *J. of Chemical Physics* 126.

36. Singh, A., and J. P. Hespanha, 2006. Lognormal Moment Closures for Biochemical Reactions. *In* Proc. of the 45th Conf. on Decision and Control, San Diego.

37. Nasell, I., 2003. Moment closure and the stochastic logistic model. *Theoretical Population Biology* 63:159–168.

Figure Legends

Figure 1.

A graphical interpretation of the quantity $I = g(\mathbf{x}^*) - \mathbf{x}^* g'(\mathbf{x}^*)$ in Eq. 20 for any arbitrary transcriptional response $g(\mathbf{x})$: I is the intercept of the tangent to the transcriptional response $g(\mathbf{x})$ at $(\mathbf{x}^*, g(\mathbf{x}^*))$ with the y-axis.

Figure 2.

Time evolution of the average number of protein molecules. 1) corresponds to the case when these is negative feedback (i.e., protein repressed its own transcription) and 2) corresponds to the case when these is there is gene expression with no negative feedback. The solid and dashed lines represent experimentally measured and fitted approximations to the time evolution of the average number of protein molecules, respectively. This figure was taken from (8).

Figure 3.

Intrinsic noise CV_{int} in the protein as a function of the feedback strength a and hill coefficient M. CV_{int} is normalized by CV_{int-nr}, the intrinsic noise in the protein when there is no feedback. Other parameters taken as $g_0 = 1$, $b = 0$, $N_x = 1$, $V_x = 0$ and $d_x = 0.01$.

Figure 4.

Total noise CV_{tot} as a function of the feedback strength a when the hill coefficient is one ($M = 1$) for different values of noise CV_z in the exogenous

signal. CV_{tot} is normalized by CV_{tot-nr}, the total noise in the protein when there is no feedback. Other parameters are taken as $g_0 = 1$,$N_x = 4$, $V_x^2 = N_x^2 - N_x$, $b = 0$, $S = 1$, $d_x = 0.04$. The response time T_z is assumed be much larger than T_{nr}.

Figure 5.

Total noise CV_{tot} as a function of the feedback strength a when the hill coefficient is two ($M = 2$) for different values of noise CV_z in the exogenous signal. CV_{tot} is normalized by CV_{tot-nr}, the total noise in the protein when there is no feedback. Other parameters are taken as $g_0 = 1$, $N_x = 4$, $V_x^2 = N_x^2 - N_x$, $b = 0$, $S = 1$, $d_x = 0.04$. The response time T_z is assumed be much larger than T_{nr}.

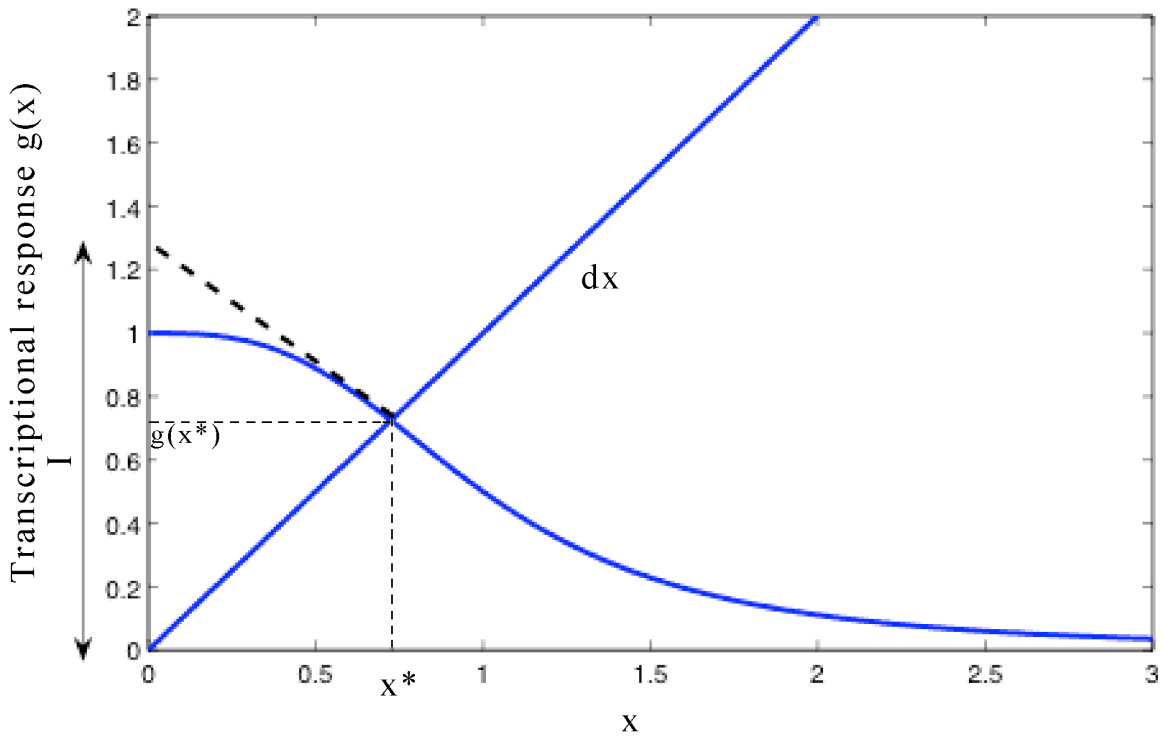
Transcriptional response g(x)
dx
g(x*)
I
x*
x
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
0.5
1.5
2.5
3

Figure 1:

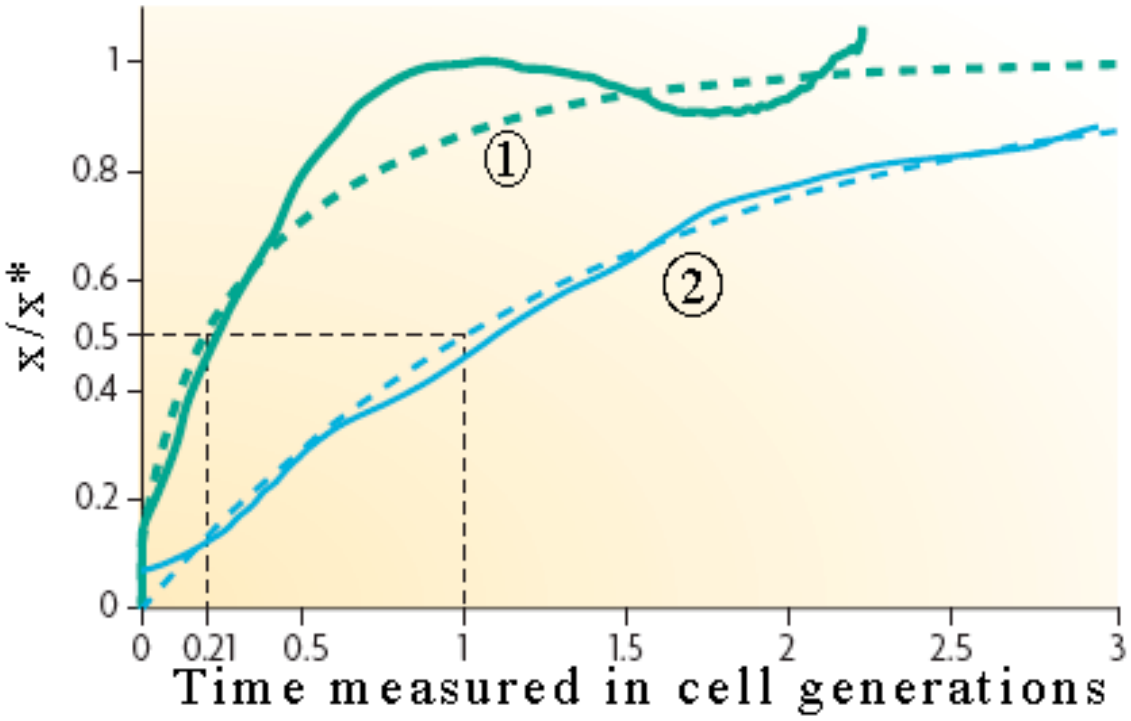
x / x*
1
0.8
0.6
0.5
0.4
0.2
0
0
0.21
0.5
1
1.5
2
2.5
3
Time measured in cell generations
①
②

Figure 2:

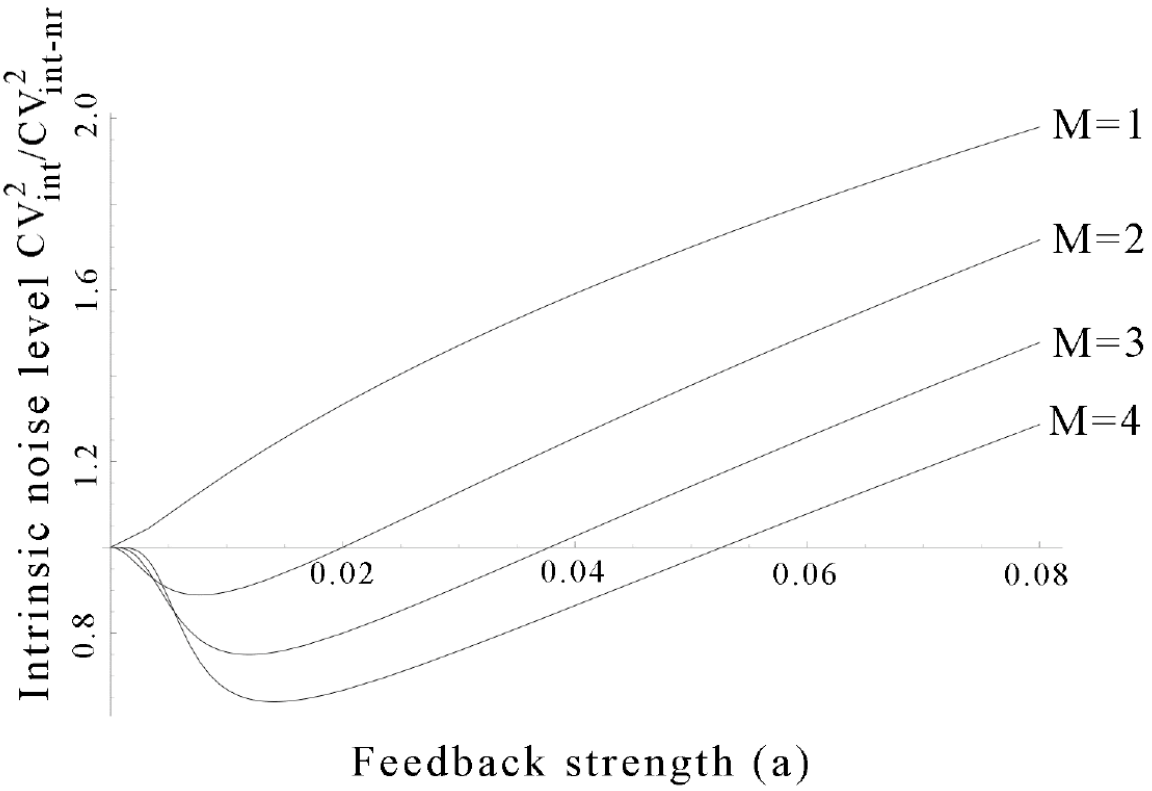
Intrinsic noise level $CV^2_{int}/CV^2_{int\text{-}nr}$
2.0
1.6
1.2
0.8
0.02
0.04
0.06
0.08
M=1
M=2
M=3
M=4
Feedback strength (a)

Figure 3:

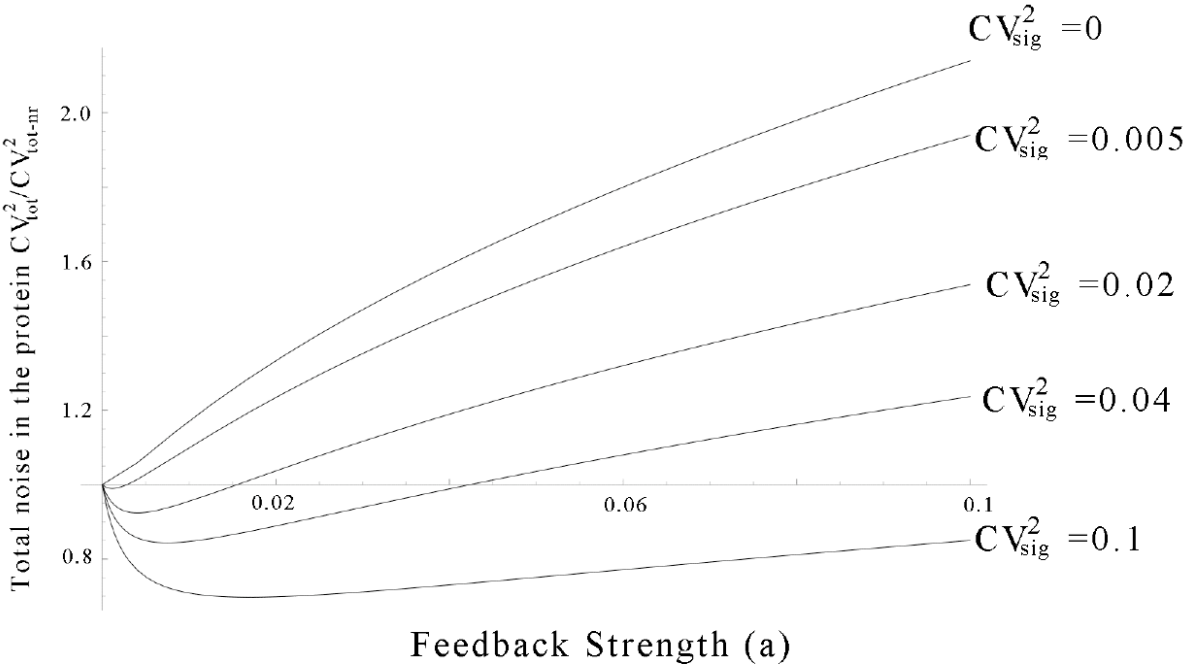
CV^2_{sig} =0
CV^2_{sig} =0.005
CV^2_{sig} =0.02
CV^2_{sig} =0.04
CV^2_{sig} =0.1
Total noise in the protein CV^2_{tot}/CV^2_{tot-nr}
2.0
1.6
1.2
0.8
0.02
0.06
0.1
Feedback Strength (a)

Figure 4:

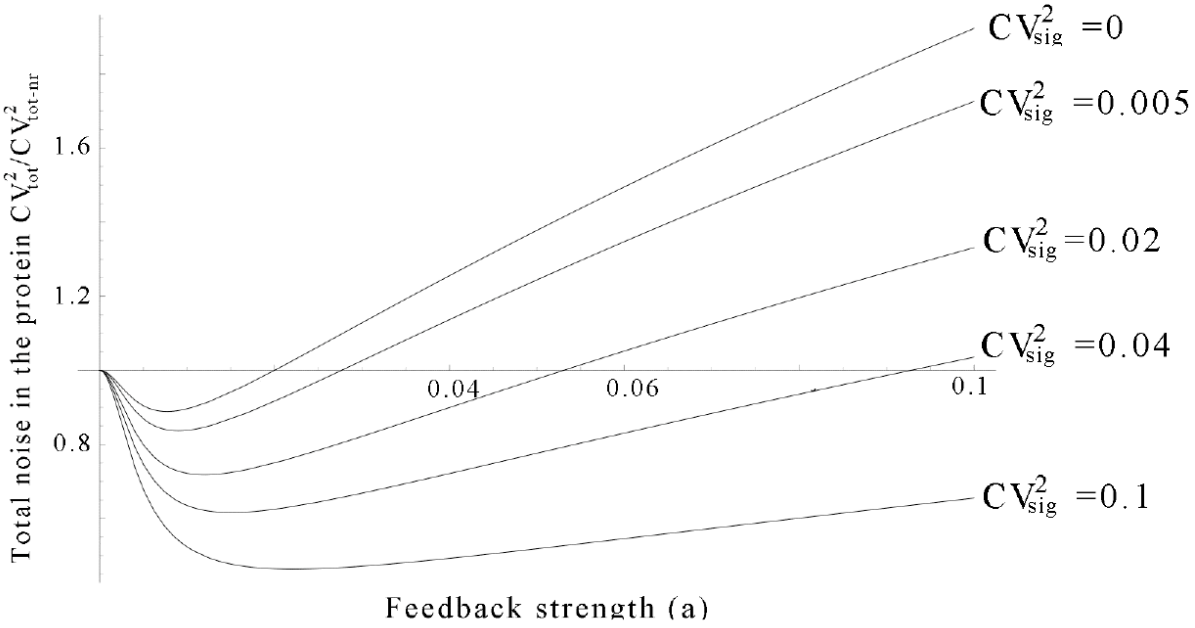
Total noise in the protein CV^2_{tot}/CV^2_{tot-nr}
1.6
1.2
0.8
0.04
0.06
0.1
Feedback strength (a)
CV^2_{sig} =0
CV^2_{sig} =0.005
CV^2_{sig} =0.02
CV^2_{sig} =0.04
CV^2_{sig} =0.1

Figure 5: