

# The Role of Input Noise in Transcriptional Regulation

Gašper Tkačik<sup>1\*</sup>, Thomas Gregor<sup>1,2</sup>, William Bialek<sup>1,3</sup>

**1** Joseph Henry Laboratories of Physics, Lewis–Sigler Institute for Integrative Genomics, Princeton University, Princeton, New Jersey, United States of America, **2** Howard Hughes Medical Institute, Princeton University, Princeton, New Jersey, United States of America, **3** Princeton Center for Theoretical Physics, Princeton University, Princeton, New Jersey, United States of America

## Abstract

Gene expression levels fluctuate even under constant external conditions. Much emphasis has usually been placed on the components of this noise that are due to randomness in transcription and translation. Here we focus on the role of noise associated with the inputs to transcriptional regulation; in particular, we analyze the effects of random arrival times and binding of transcription factors to their target sites along the genome. This contribution to the total noise sets a fundamental physical limit to the reliability of genetic control, and has clear signatures, but we show that these are easily obscured by experimental limitations and even by conventional methods for plotting the variance vs. mean expression level. We argue that simple, universal models of noise dominated by transcription and translation are inconsistent with the embedding of gene expression in a network of regulatory interactions. Analysis of recent experiments on transcriptional control in the early *Drosophila* embryo shows that these results are quantitatively consistent with the predicted signatures of input noise, and we discuss the experiments needed to test the importance of input noise more generally.

**Citation:** Tkačik G, Gregor T, Bialek W (2008) The Role of Input Noise in Transcriptional Regulation. PLoS ONE 3(7): e2774. doi:10.1371/journal.pone.0002774

**Editor:** Rory Edward Morty, University of Giessen Lung Center, Germany

**Received:** January 28, 2008; **Accepted:** June 19, 2008; **Published:** July 23, 2008

**Copyright:** © 2008 Tkačik et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** NSF Grant PHY-0650617; NIH grants P50 GM071508, R01 GM077599; Burroughs Wellcome Program in Biological Dynamics

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: gtkacik@princeton.edu

## Introduction

A number of recent experiments have focused attention on noise in gene expression [1,2,3,4,5,6,7,8,9]. The study of noise in biological systems more generally has a long history, with two very different streams of thought. On the one hand, observations of noise in behavior at the cellular or even organismal level give us a window into mechanisms at a much more microscopic level. One of the classic examples of using noise to draw inferences about biological mechanism is the Luria–Delbrück experiment [10], which demonstrated the random character of mutations, but one can also point to early work on the nature of chemical transmission at synapses [11,12] and on the dynamics of ion channel proteins [13,14,15,16]. On the other hand, noise limits the reliability of biological function, and it is important to identify these limits. Examples include tracking the reliability of visual perception at low light levels down to the ability of the visual system to count single photons [17,18], the implications of channel noise for the reliability of neural coding [19,20,21], and the approach of bacterial chemotactic performance to the limits set by the random arrival of individual molecules at the cell surface [22].

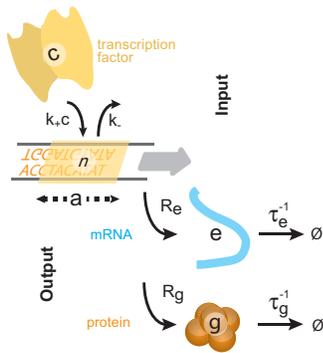
After demonstrating that one can observe noise in gene expression, most investigators have concentrated on the mechanistic implications of this noise. Working backward from the observation of protein concentrations, one can try to find the components of noise that derive from the translation of mRNA into protein, or the components that arise from noise in the transcription and degradation of the mRNA itself. At least in some organisms, a single mRNA transcript can give rise to many protein molecules, and this ‘burst’ both amplifies the fluctuations in the protein copy number and changes their statistics, so that even if the number of mRNA copies obeys the Poisson distribution the

number of protein molecules will not [23]; this discussion parallels the understanding that Poisson arrival of photons at the retina generates non-Poisson statistics of action potentials in retinal ganglion cells because each photon triggers a burst of spikes [24]. Recent large scale surveys of noise in eukaryotic transcription have suggested that the noise in most protein levels can be understood in terms of this picture, so that the fractional variance in the number of proteins expressed from gene *i* is universally given by

$$\eta_i^2 \equiv \frac{\sigma_{g_i}^2}{\langle g_i \rangle^2} = \frac{b}{\langle g_i \rangle}, \quad (1)$$

where  $\sigma_{g_i}^2 = \langle (g_i - \langle g_i \rangle)^2 \rangle$  is the variance in the protein copy number, and  $b \sim 10^3$  is the burst size, found to be approximately constant for all genes [9].

The mechanistic focus on noise in transcription vs translation perhaps misses the functional role of gene expression as part of a regulatory network. Almost all genes are subject to transcriptional regulation, and hence the expression level of a particular protein can be viewed as the cell’s response to the concentration of the relevant transcription factors. Seen in this way, transcription and translation are at the ‘output’ side of the response, and the binding of transcription factors to their targets along the genome is at the ‘input’ side (Fig. 1). Noise can arise at both the input and output, and while fluctuations in transcription factor concentration could be viewed as an extrinsic source of noise [1,25], there will be fluctuations in target site occupancy even at fixed transcription factor concentration [26,27,28]. There is a physical limit to how much the impact of these input fluctuations can be reduced, essentially because any system that responds to changes in the concentration ultimately is limited by the number of molecules



**Figure 1. A simple model for transcriptional regulation.** Transcription factor is present at an average concentration  $c$ , diffusing freely with diffusion constant  $D$ ; it can bind to the binding site of linear dimension  $a$  and the fractional occupancy of this site is  $n \in [0,1]$ . Binding occurs with a second order rate constant  $k_+$ , and unbinding occurs with a first order rate constant  $k_-$ . When the site is bound, the mRNA are transcribed at rate  $R_e$  and degraded with rate  $\tau_e^{-1}$ , resulting in a number of transcripts  $e$ . Proteins are translated from each mRNA molecule with rate  $R_g$  and degraded with rate  $\tau_g^{-1}$ , resulting in a copy number  $g$ . doi:10.1371/journal.pone.0002774.g001

that are carried by diffusion into the immediate vicinity of the receptor site. For realistic transcription factor concentrations and readout times this number can be small and therefore subject to considerable noise [26,22,29].

In this paper we revisit the relative contributions of input and output noise. Input noise has a clear signature, namely that its impact on the output protein concentration peaks at an intermediate value of the input transcription factor concentration. The analogous signature was essential, for example, in identifying the noise from random opening and closing of individual ion channels in neurons [30,31]. Perhaps surprisingly, we show that this signature is easily obscured in conventional ways of plotting the data on noise in gene expression. Recent experiments on the regulation of Hunchback expression by Bicoid in the early *Drosophila* embryo [32] are consistent with the predicted signature of input noise, and (although there are caveats) a quantitative analysis of these data supports a dominant contribution of diffusive input noise. We discuss what experiments would be required to test this conclusion more generally.

## Results

### Global consistency

We begin by asking if a universal noise model given by Eq (1) can be used simultaneously and in a consistent way for both transcription factors and their targets. Our goal for this section is to show that this is not the case, and that, furthermore, one can be led to models of the form of Eq (1) whenever the input noise is neglected in the analysis. These two findings will motivate the following sections in which we will reexamine the theoretical properties of the input noise and look for its signatures in experimental data of Ref [32].

Let us consider a gene  $g$  which is regulated by several transcription factors  $g_\mu, \mu = 1, \dots, K$ . In the simplest model, the dynamics for the expression of gene  $g$  can be written as:

$$\tau_g \frac{dg}{dt} = f(g_1, g_2, \dots, g_K) - g, \quad (2)$$

where  $f$  is the production function that depends on the concentrations of the regulators  $g_\mu$ , and  $-g$  accounts for the

gene's degradation with a time constant  $\tau_g$ . A more complete model would include the dynamics of mRNA and binding site occupancy. This level of detail is omitted here because it does not influence the arguments made in this section; in *Sources of noise* section, we will, however, treat all of these processes explicitly for the case of a single gene regulated by a single transcription factor.

In steady state, the mean number of proteins of gene  $g$  in the cell will be a function of the mean copy numbers of all the relevant transcription factors. When the noise is small, we should be able to write:

$$\langle g \rangle = f(\langle g_1 \rangle, \langle g_2 \rangle, \dots, \langle g_K \rangle). \quad (3)$$

Even in steady state, however, transcription factor copy numbers  $g_\mu$  fluctuate around the mean value:  $g_\mu(t) = \langle g_\mu \rangle + \delta g_\mu(t)$ . When we speak about the 'noise' in TF levels, we are referring precisely to fluctuations  $\delta g_\mu$ , which can be characterized by two parameters: their strength, or variance,  $\sigma_{g_\mu}^2 = \langle (\delta g_\mu)^2 \rangle$ , and their correlation time  $\tau_{g_\mu}$  [5]. To compute the effect of these noise sources on the regulated gene  $g$ , we propagate them through the input/output relation  $f$  [23,6,33], so that

$$\langle (\delta g)^2 \rangle = \sum_{\mu=1, v=1}^K \frac{\partial f}{\partial g_\mu} \frac{\partial f}{\partial g_\nu} \Phi_{\mu\nu} \langle \delta g_\mu \delta g_\nu \rangle + \langle (\delta g)^2 \rangle_0, \quad (4)$$

where we include the noise  $\langle (\delta g)^2 \rangle_0$  due to stochastic transcription and translation of gene  $g$ ; this intrinsic contribution is present even at fixed transcription factor levels or in constitutively expressed genes.

Conceptually, Eq (4) looks much like computing a first order term in a Taylor expansion of Eq (3) around the steady state and squaring it to get  $\langle (\delta g)^2 \rangle$ , where the effect of the fluctuation  $\delta g_\mu$  in the regulator  $\mu$  on the output gene  $g$  is determined by the 'susceptibility'  $\partial f / \partial g_\mu$  [23]. One important difference between Eq (4) and the plain Taylor expansion around the steady state, however, is the presence of the noise filtering or noise averaging term  $\Phi_{\mu\nu}$ , which reminds us that Eq (3) is really a result of the *dynamical* system described by Eq (2), and that noise propagation depends not only on the noise magnitudes, but also fluctuation correlation times [23]. While it is possible to derive an expression for  $\Phi$  if the function  $f$  is given (and we shall do that in *Methods* for the model proposed in the next section), we highlight two important limiting cases here. If the dynamics of gene  $g$  is slow compared to that of its transcription factors, i.e.  $\tau_g \gg \tau_{g_\mu, g_\nu}$ , then  $\Phi_{\mu\nu} \rightarrow 0$  and gene  $g$  will "average away" upstream noise; in contrast, if transcription factors fluctuate more slowly or on a comparable timescale,  $\Phi_{\mu\nu}$  will be of order 1.

If the noise in gene expression is dominated by the processes of transcription and translation, and if the transcription factors are not regulating each other, then the correlations between fluctuations in the copy numbers of different proteins will be very small, so we expect that

$$\langle \delta g_\mu \delta g_\nu \rangle = \delta_{\mu\nu} \langle (\delta g_\mu)^2 \rangle, \quad (5)$$

where  $\delta_{\mu\nu}$  denotes a Kronecker delta symbol. Furthermore, it can be shown that the filtering function  $\Phi$  now reduces to

$$\Phi_\mu = \frac{\tau_{g_\mu}}{\tau_{g_\mu} + \tau_g}. \quad (6)$$

This allows us to simplify the propagation of noise in Eq (4) to give

$$\langle(\delta g)^2\rangle = \sum_{\mu=1}^K \left(\frac{\partial f}{\partial g_{\mu}}\right)^2 \Phi_{\mu} \langle(\delta g_{\mu})^2\rangle + \langle(\delta g)^2\rangle_0. \quad (7)$$

If, as in Eq (1), we express the noise in protein copy number as a fractional noise  $\eta$ , then this becomes

$$\eta^2 = \sum_{\mu=1}^K \left(\frac{\partial \log f}{\partial \log g_{\mu}}\right)^2 \Phi_{\mu} \eta_{\mu}^2 + \eta_0^2. \quad (8)$$

In particular, this means that there is a minimum level of noise,

$$\eta^2 \geq \sum_{\mu=1}^K \left(\frac{\partial \log f}{\partial \log g_{\mu}}\right)^2 \Phi_{\mu} \eta_{\mu}^2. \quad (9)$$

But if the fractional variance in protein copy number has a simple, universal relation to the mean copy number, as in Eq (1) [9], then this simplifies still further:

$$\frac{b}{\langle g \rangle} \geq \sum_{\mu=1}^K \left(\frac{\partial \log f}{\partial \log g_{\mu}}\right)^2 \Phi_{\mu} \frac{b}{\langle g_{\mu} \rangle} \quad (10)$$

$$\Rightarrow 1 \geq \sum_{\mu=1}^K \left(\frac{\partial \log f}{\partial \log g_{\mu}}\right)^2 \Phi_{\mu} \frac{\langle g \rangle}{\langle g_{\mu} \rangle}. \quad (11)$$

Since the proteins labeled by the indices  $\mu$  represent transcription factors, usually present at low concentrations, and the protein  $g$  is a regulated gene—such as a structural or metabolic protein—but not a transcription factor itself, one expects that  $\langle g \rangle / \langle g_{\mu} \rangle \ll 1$ . If the dynamics of gene expression for both the regulated gene and its transcription factors happen on the same timescale, we have  $\Phi_{\mu} = \tau_{g_{\mu}} / (\tau_{g_{\mu}} + \tau_g) \sim 1$ , and hence:

$$\sum_{\mu=1}^K \left(\frac{\partial \log f}{\partial \log g_{\mu}}\right)^2 \ll 1. \quad (12)$$

Since this inequality constrains the sum of squares of terms, each must be much smaller than one. This means that when we make a small change in the concentration of any transcription factor, the response of the regulated gene must be much less than proportional. In this sense, the assumption of a simple universal description for the level of noise in gene expression, Eq (1), leads us to the conclusion that transcriptional “regulation” can’t really be very effective, and this must be wrong. Notice that this problem is independent of the burst size  $b$ , and hence doesn’t depend on whether the noise is dominated by transcription or translation.

From Eq (12) we conclude that the original hypothesis about the simple noise model, Eq (1), should be re-examined. Alternatively, if this hypothesis were correct, either the downstream gene’s expression would have to change on a timescale which is much slower than those of its regulators (and consequently  $\Phi \rightarrow 0$ ), or there would have to exist subtle correlations among all the protein copy number fluctuations of the different transcription factors. However, since the gene expression machinery is shared between transcription factors and their targets, there is no particular reason to expect very different dynamics for various genes. If, conversely, the assumption of Eq (5) is wrong and the correlations between

protein levels provide consistency, they would have to take on a very special form—different transcription factors regulating a single gene would have to be correlated in a way that matches their impact on the expression of that gene—which seems implausible but would be very interesting if it were true.

### Sources of noise

The previous section points to severe difficulties for the universal noise model of Eq (1) when it is applied simultaneously to transcription factors and their target, as in Eq (2). To make this general problem more specific, we now focus on a simple case of a gene regulated by a *single* TF and work out carefully the noise in this small system.

Figure 1 makes clear that the concentration of a regulated gene product can fluctuate for many reasons. The processes of synthesis and degradation of the protein molecules themselves are discrete and stochastic, as are the synthesis and degradation of mRNA molecules; together these constitute the “output noise” which has been widely discussed. But if we are considering a gene whose transcription is regulated, we need a microscopic model for this process. For the case of a transcriptional activator, there are binding sites for the transcription factors upstream of the regulated gene, and when these sites are occupied transcription proceeds at some rate, but when the site is empty transcription is inhibited. Because there are only a small number of relevant binding sites (in the simplest case, just one), the occupancy of these sites must fluctuate, and this random switching is an additional source of noise. In addition, the binding of transcription factors to their target sites along the genome depends on the concentration in the immediate neighborhood of these sites, and this fluctuates as molecules diffuse into and out of the neighborhood.

All of the different processes described above and schematized in Fig. 1 can be worked out analytically using Langevin methods and the predictions of this analysis can be tested against detailed stochastic simulations. We computed the total noise variance of a regulated gene,  $\sigma_g^2 = \langle(\delta g)^2\rangle$ , by adding up the noise contributions from the relevant microscopic processes (see results below; for calculations see *Methods*, where we also make a connection with the derivations in the *Global consistency* section). Notice that variations in cell size, protein sorting in cell division, fluctuations in RNA polymerase and ribosome concentrations, and all other extrinsic contributions to the noise are neglected.

The variance in protein copy number  $\sigma_g^2$  can be written as a sum of three terms, which correspond to the output, switching, and diffusion noise (see Eq (58) in *Methods*). To set the scale, we have expressed the copy number as a fraction of its maximum possible mean value,  $g_0$ , which is reached at high concentrations of the transcriptional activator. In these units, we find

$$\left(\frac{\sigma_g}{g_0}\right)^2 = \frac{1 + R_g \tau_e}{g_0} \bar{g} + \frac{(1 - \bar{g})^2 \bar{g}}{k - \tau_g} + \frac{(1 - \bar{g})^2 \bar{g}^2}{\pi D a c \tau_g} \quad (13)$$

where  $\bar{g} = \langle g \rangle / g_0$  is the protein copy number expressed as a fraction of its maximal value,  $c$  is the concentration of the transcription factor, and other parameters are as explained in Fig. 1.

The first term in Eq (13) is the output noise and has a Poisson-like behavior, with variance proportional to the mean, but the proportionality constant differs from 1 by  $R_g \tau_e$ , i.e. the burst size or the number of proteins produced per mRNA [23]. This is just the simple model of Eq (1), with  $b = 1 + R_g \tau_e$ .

The second term in Eq (13) originates from binomial “switching” as the transcription factor binding site occupation fluctuates, and is most closely analogous to the noise from random opening

and closing of ion channels. This term will be small for unbinding rates  $k_-$  that are fast compared to the protein lifetime, but might be large for factors that take a long time to equilibrate or that form energetically stable complexes on their promoters.

The third term in Eq (13) arises because the diffusive flux of transcription factor molecules to the binding site fluctuates at low input concentration  $c$ ; in effect the receptor site “counts” the number of molecules arriving into its vicinity during a time window, and this number is of the order  $\sim Dac\tau_g$ . This argument is conceptually the same as that for the limits to chemoattractant detection in chemotaxis, as discussed by Berg and Purcell [22]. It can be shown that this is a theoretical noise floor that cannot be circumvented by using any sophisticated “binding site machinery” as long as this machinery is contained within a region of linear size  $a$  to which the molecules are transported by diffusion [26,29]. For example, cooperative binding to the promoter, or binding to a promoter with multiple internal states, will modify the binomial switching term. However, in both cases the input diffusion noise will remain unaffected, which is easily seen when this noise term is expressed as an effective noise in transcription factor concentration  $\sigma_c$  using the relation

$$\sigma_g = \left| \frac{\partial \bar{g}(c)}{\partial c} \right| \sigma_c. \tag{14}$$

Although cooperativity does not change the effective concentration noise due to diffusion, it does reduce the relative significance of the switching noise; see *Methods: Cooperative binding of transcription factors* and Ref [29]. Since we will discuss a system which is strongly cooperative, in much of what follows we neglect the switching noise term and focus on the output noise and diffusion noise. Then the generalization to multisite, cooperative regulation is straightforward (see *Methods: Cooperative binding of transcription factors*). We expect that cooperative effects among  $h$  transcription factors generate a sigmoidal dependence of expression on the transcription factor concentration, so that

$$\bar{g} = \frac{c^h}{c^h + K_d^h}, \tag{15}$$

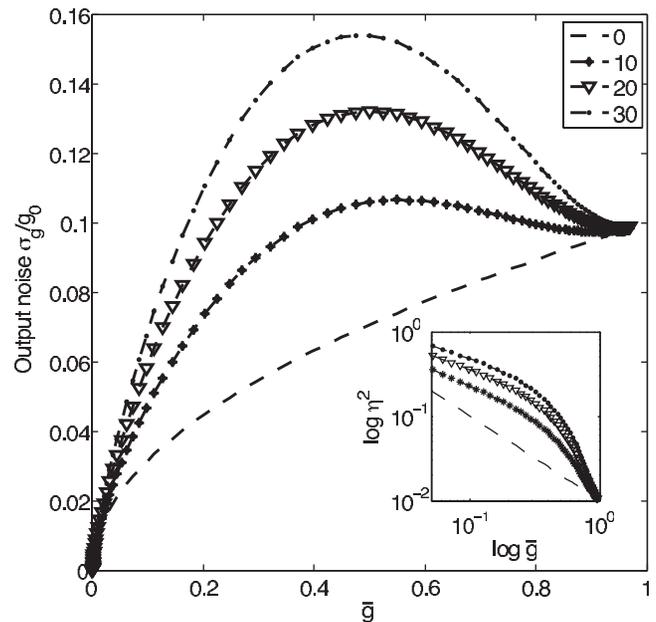
where  $h$  is called the Hill coefficient, and  $K_d$  is the concentration required for half maximal activation. We can invert this relationship to write the concentration  $c$ , which is relevant for the diffusive noise, as a function of the mean fractional expression level  $\bar{g}$ . Substituting back into Eq (13), and neglecting the switching noise, we obtain

$$\left( \frac{\sigma_g}{g_0} \right)^2 = \alpha \bar{g} + \beta \bar{g}^{2-1/h} (1-\bar{g})^{2+1/h}, \tag{16}$$

where  $\alpha$  and  $\beta$  are combinations of parameters that measure the strength of the output and diffusion noise, respectively. If we express the variance in fractional terms, this becomes

$$\eta_g^2 = \alpha \frac{1}{\bar{g}} + \beta \bar{g}^{-1/h} (1-\bar{g})^{2+1/h}. \tag{17}$$

The universal noise model of Eq (1) corresponds to  $\beta = 0$  (no input noise) and  $b = \alpha g_0$ . Figure 2 shows the predicted noise levels for different ratios of output to input noise ( $\beta/\alpha$ ) and illustrates how the monotonic relationship between the noise and the mean changes when input noise becomes the dominant contribution.



**Figure 2. Expression noise as a function of the mean.** The standard deviation of the protein concentration  $\sigma_g/g_0$  is plotted against the mean protein concentration  $\bar{g} = \langle g \rangle / g_0$ , from Eq (16) with  $h = 5$ . In all cases the output noise term has a strength  $\alpha = 0.01$ , and the different curves are indexed by the ratio of input noise to output noise  $\beta/\alpha = 0, 10, 20, 30$ . In the inset, we show the same results plotted as a fractional variance  $\eta_g^2$  vs the mean [Eq (17)], on a logarithmic scale. doi:10.1371/journal.pone.0002774.g002

For very highly cooperative, essentially switch-like systems, we can take the limit  $h \rightarrow \infty$  to obtain

$$\left( \frac{\sigma_g}{g_0} \right)^2 = \alpha \bar{g} + \beta \bar{g}^2 (1-\bar{g})^2 \tag{18}$$

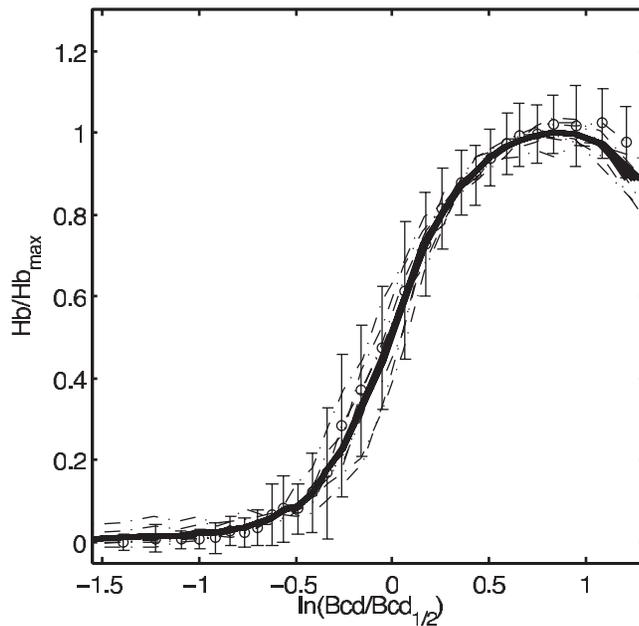
$$\eta_g^2 = \alpha \frac{1}{\bar{g}} + \beta (1-\bar{g})^2. \tag{19}$$

In particular, if we explore only expression levels well below the maximum ( $\bar{g} \ll 1$ ), then the diffusion noise just adds a constant  $\beta$  to the fractional variance. Because a constant contribution to  $\eta_g^2$  can be expected for global or some extrinsic noise sources [25,6,9], the diffusion noise in a highly cooperative system could be confused with either global or extrinsic noise.

### Signatures of input noise

Input noise arises from fluctuations in the occupancy of the transcription factor binding sites. These fluctuations must vanish at very high transcription factor concentrations, where all sites are fully occupied, or at very low concentrations, where the sites are never occupied. In the case of a transcriptional activator, full and zero occupancy correspond to maximal and minimal expression levels, respectively. Hence a key signature of input noise is a peak at some intermediate expression level as shown in Fig. 2.

The claim that many genes have expression noise levels which fit the universal output noise model of Eq (1) would seem to contradict the prediction of a peak in the noise as a function of the mean. But if we plot the predictions of the model with input noise as a fractional variance vs mean, the prominent peak disappears

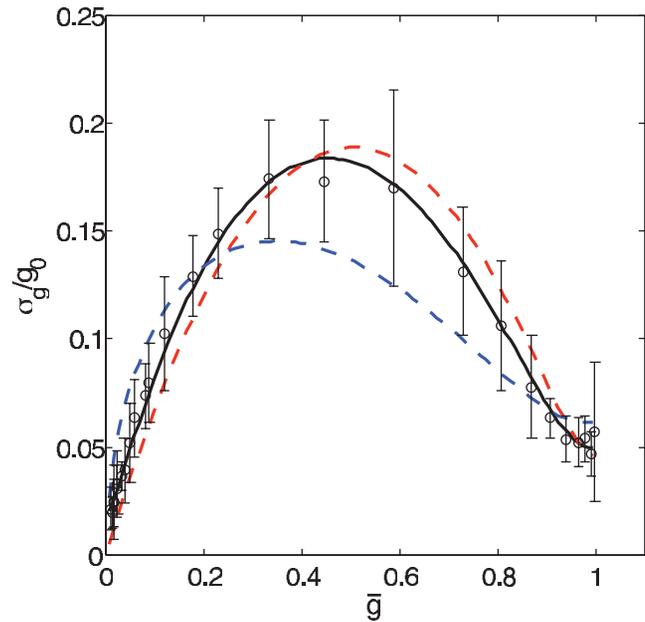


**Figure 3. The input-output relation for Bicoid regulation of Hunchback expression, redrawn from Ref [32] (cf. Fig. 4 of the reference).** Dashed curves show mean expression levels in different embryos, thick black line is the mean across all embryos, and points with error bars show the mean and standard deviation of Hb expression at a given Bcd concentration in one embryo.  
doi:10.1371/journal.pone.0002774.g003

(inset to Fig. 2). In fact, over a large dynamic range, the input noise seems just to increase the magnitude of the fractional variance while not making a substantial change in the slope of  $\log(\eta_g^2)$  vs  $\log(\langle g \rangle)$ . Confronted with real data on a system with significant input noise, we could thus fit much of those data with the universal output noise model but with a larger value of  $b$ . There is, of course, a difference between input and output noise, even when plotted as  $\log(\eta_g^2)$  vs  $\log(\langle g \rangle)$ , namely a rapid drop in noise level as we approach maximal expression. But this effect is confined to a narrow range, essentially a factor of two in mean expression level. As we discuss below, there is a variety of reasons why this might not have been seen in the data of Ref [9].

Recent experiments on the precision of gene expression in the early *Drosophila* embryo provide us with an opportunity to search for the signatures of input noise [32]. The embryo contains a spatial gradient of the protein Bicoid (Bcd), translated from anteriorly located maternal mRNA, and this protein is a transcription factor which activates, among other genes, *hunchback*. Looking along the anterior-posterior axis of the embryo one thus has an array of nuclei that experience a graded range of transcription factor concentrations. Using antibody staining and image processing methods, it is thus possible to collect thousands of points on a scatter plot of input (Bicoid concentration) vs. output (Hunchback protein concentration); since even in a single embryo there are many nuclei that have the same Bcd concentration, one can examine both the mean Hunchback (Hb) response and its variance. In Fig. 3 we replot the data from Ref [32] (cf. Fig. 4 of the reference).

The mean response of Hb to Bcd is fit reasonably well by Eq (15) with a Hill coefficient  $h=5$  [32], and in Fig. 4 we show the noise in this response as a function of the mean. The peak of expression noise near half maximal expression—the signature of input noise—is clearly visible. More quantitatively, we find that



**Figure 4. Standard deviation of Hunchback expression as a function of the mean (points with error bars), replotted from Ref [32].** The black line is a fit of combined output and diffusion noise contributions, from Eq (16) with  $h=5$ , and the dashed red line is with  $h \rightarrow \infty$ , from Eq (18). In contrast, the dashed blue line is the best fit of combined output and switching noise contributions, i.e.  $(\sigma_g/g_0)^2 = \alpha\bar{g} + (1-\bar{g})^2\bar{g}$ . Although both diffusion and switching noise produce a peak at intermediate expression levels, the shapes of the peaks are distinguishable, and the data favor the diffusion noise model.  
doi:10.1371/journal.pone.0002774.g004

the data are well fit by Eq (16) with the contribution from output noise ( $\alpha \approx 1/380$ ) much smaller than that from input noise ( $\beta \approx 1/2$ ). We also consider the same model with  $h \rightarrow \infty$ , and this fully switch-like model, although formally still within error bars, systematically deviates from the data. Finally we consider a model in which diffusion noise is absent, but we include the switching noise from Eq (13), which generalizes to the case of cooperative binding (see *Methods: Cooperative binding of transcription factors*). Interestingly, this model has the same number of parameters as the diffusion noise model, but does a significantly poorer job of fitting the data. While the fit can be improved further by adding a small background to the noise, we emphasize that Eq (16) correctly captures the non-trivial shape of the noise curve with only two parameters. Because input noise falls to zero at maximal expression, the sole remaining noise at that point is the output noise, and this uniquely determines the parameter  $\alpha$ . The strength of the input noise ( $\beta$ ) then is determined by the height of the noise peak, and there is no further room for adjustment. The *shape* of the peak is predicted by the theory with no additional parameters, and the different curves in Fig. 4 demonstrate that the data can distinguish among various functional forms for the peak.

Are the parameters  $\alpha$  and  $\beta$  that fit the Bcd/Hb data biologically reasonable? The fact that diffusive noise dominates at intermediate levels of expression ( $\beta \gg \alpha$ ) means that the Hunchback expression level provides a readout of Bcd concentration with a reliability that is close to the physical limit set by diffusional noise, as was argued in Ref [32] based on the magnitude of the noise level and estimates of the relevant microscopic parameters that determine  $\beta$  ( $\beta \approx h^2/\pi Da K_d \tau_g$ , see the *Methods*). The dominance of diffusive noise over switching noise

presumably is related to the high cooperativity of the Bcd/Hb input/output relation [29].

The parameter  $\alpha = (1+b)/g_0$  measures the strength of the output noise and thus depends on the absolute number of Hb molecules and on the number of proteins produced per mRNA transcript. If this burst size is in the range  $b = R_g \tau_e \sim 1-10$ , then our fit predicts the maximum expression level of Hb corresponds to  $g_0 = 700-4000$  molecules in the nucleus. Given the volume of the nuclei at this stage of development ( $\sim 140 \mu\text{m}^3$ ; see Refs [32,34]), this is a concentration of 8–48 nM. Although we don't have independent measurements of the absolute Hunchback concentration, this is reasonable for transcription factors, which typically act in the nanoMolar range [35,36,37,38,39,40], and can be compared with the maximal nuclear concentration of Bcd, which is  $55 \pm 3$  nM [32]. Larger burst sizes would predict larger maximal expression levels, or conversely measurements of absolute expression levels might give suggestions about the burst size for translation in the early *Drosophila* embryo.

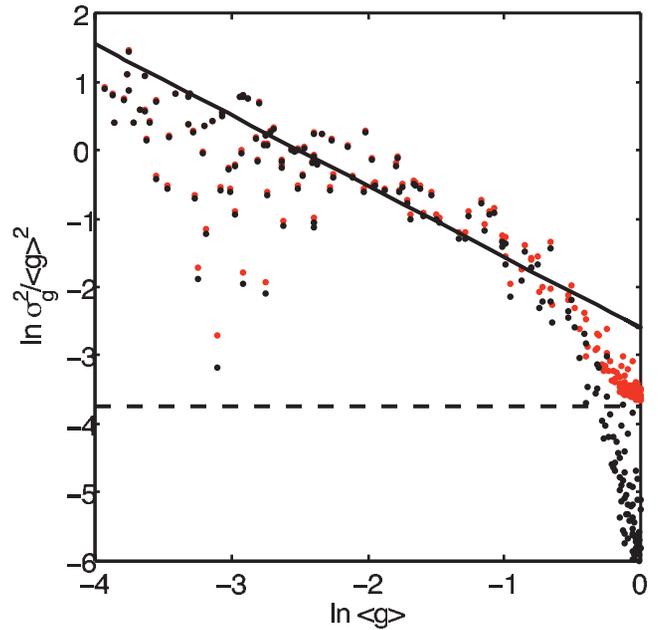
## Discussion

In the process of transcriptional regulation, the (output) expression level of regulated genes acts as a sensor for the (input) concentration of transcription factors. The performance of this sensor, and hence the regulatory power of the system, is limited by noise. While changes in the parameters of the transcriptional and translational apparatus can change the level of output noise, the input noise is determined by the physical properties of the transcription factor and its interactions with the target sites along the genome. Ultimately, there is a lower bound on this input noise level set by the noise in random arrival of the transcription factors at their targets, in much the same way that any imaging process ultimately is limited by the random arrival of photons.

Input and output noise seem to be so different that it is hard to imagine that they could be confused experimentally. Some of the difficulty, however, can be illustrated by plotting the results from the Bcd/Hb experiments of Ref [32] in the form which has become conventional in the study of gene expression noise, as a fractional variance vs mean expression level (Fig. 5). The signature of input noise, so clear in Fig. 4, now is confined to a narrow range ( $\sim \times 2$ ) near maximal expression. In contrast, over more than a decade of expression levels the noise level is a good fit to  $\eta_g^2 \propto \langle g \rangle^{-\zeta}$ , with  $\zeta = 1.04$  being very similar to the prediction of the universal noise model ( $\zeta = 1$ ) in Eq (1). The departures from power-law behavior are easily obscured by global noise sources, experimental error, or by technical limitations that lead to the exclusion of data at the very highest expression levels, as in Ref [9].

The present analysis of the Bicoid/Hunchback data shows that the signatures of the input noise are surprisingly subtle. In this system, the behavior near half maximal expression is biologically most relevant, since it is at this very concentration that the system has to 'decide' where to draw a fairly sharp gene expression boundary as one of the first steps in constructing a spatial pattern. While in other systems the behavior of noise at the intermediate gene expression levels might not be of so high a biological importance for the organism, it is only in this region that different sources of noise are qualitatively distinguishable, as is clear from Fig. 5. Thus, unless we have independent experiments to measure some of the parameters of the system, we need experimental access to the full range of expression levels and hence, implicitly, to the full dynamic range of transcription factor concentrations, if we want to disentangle input and output noise.

The early *Drosophila* embryo is an attractive model system precisely because the organism itself generates a broad range of



**Figure 5. Logarithmic plot of fractional variance vs the mean expression level for Hunchback, replotted from Ref [32].** Each black point represents the noise level measured across nuclei that experience the same Bcd concentration within one embryo, and results are collected from nine embryos. The solid line shows a fit to  $\eta_g^2 \propto \langle g \rangle^{-\zeta}$  in the region below half maximal mean expression; we find a good fit, with  $\zeta = 1.04$ , despite the fact that these data show a clear signature of input noise when plotted in Fig. 4. Dashed line indicates the global noise floor suggested in Ref [9], and red points show the raw data with this variance added. Although the input noise still appears as a drop in fractional noise level near maximal mean expression, this now is quite subtle and easily obscured by experimental errors. doi:10.1371/journal.pone.0002774.g005

transcription factor concentrations, and conveniently arranges these different samples along the major axes of the embryo. A caveat to our analysis is the possibility that, in the concentration range of interest, Hunchback is controlled by factors other than Bicoid [41]. While it explains the systematic variation of Hunchback with Bicoid through the sigmoidal input/output relation, our model would treat any other regulatory input influence as an 'extrinsic' noise source. Indeed, judging by the qualitative shape of the noise vs mean expression curve in Fig. 4 alone and recalling the analysis of Elowitz et al [1], one could attribute the whole peak in the noise at intermediate levels of expression to some (unknown) extrinsic noise source rather than to the noise in Bicoid concentration. The existing measurements cannot definitely rule out this possibility, yet there are two strong reasons to consider it unlikely.

First, Bicoid is a known and probably the dominant factor controlling Hunchback expression at midpoint of the anterior-posterior axis. With an absolute measurement of Bicoid concentration – the equivalent of which did not exist in case of Elowitz et al [1] – the theory predicts that we should see a non-negligible contribution to the noise peaking at intermediate expression levels. For an unknown factor  $x$  to generate an effect of similar size, the factor would have to have a susceptibility (i.e.  $\partial \bar{g} / \partial x$ ) at the major axis midpoint comparable in strength to Bicoid's [Eq (2)]. One would thus need to justify a presence of such a strong alternative regulator at that position *as well as* the absence of the predicted noise contribution from Bicoid at the same position.

Second, we emphasize that the matching between noise measurements and theoretical predictions is not just qualitative. Concretely, the detailed *shape* of the noise vs the mean expression in Fig. 4 can discriminate between various regulatory scenarios: Eq (16) shows the non-trivial dependence of the noise on the mean expression in the case of diffusion input noise, in particular its scaling as  $\bar{g}^{-1/h}(1-\bar{g})^{2+1/h}$ . These exponents will be different for a non-diffusional extrinsic noise contribution, yet the data of Ref [32] give support to the above fit, using the measured Hill slope of  $h=5$ . We note that despite these arguments it is possible that some, but presumably not the majority, of the noise at intermediate expressions is contributed by unknown extrinsic factors.

Another caveat is that since we don't directly control the transcription factor concentration, we have to measure it. In particular, in order to measure the variance of the output (Hunchback, in the present discussion) we have to find many nuclei that all have the same input transcription factor (Bicoid) concentration. Because the mean output is a steep function of the input, errors in the measurement of transcription factor concentration can simulate the effects of input noise, as discussed in Ref [32]. Thus, a complete analysis of input and output noise requires not only access to a wide range of transcription factor concentrations, but rather precise measurements of these concentrations.

Why are the different sources of noise so easily confused? If noise is dominated by randomness in a single step of the translation process, then the number of protein molecules will obey the Poisson distribution, and the variance in copy number will be equal to the mean. But if we can't actually turn measurements of protein level into molecule counts, then all we can say is that the variance will be *proportional* to the mean. If the dominant noise source is a single step in transcription, then the number of mRNA transcripts will obey the Poisson distribution, and the variance of protein copy numbers still will be proportional to the mean, but the proportionality constant will be enhanced by the burst size. The same reasoning, however, can be pushed further back: if, far from maximal expression, the dominant source of noise is the infrequent binding of a transcriptional activator (or dissociation of a repressor) to its target site, then the variance in protein copy number still will be proportional to the mean. Thus, the proportionality of variance to mean implies that there is some single rare event that dominates the noise, and by itself doesn't distinguish the nature of this event.

If noise is dominated by regulatory events, then the number of mRNA transcripts should be drawn from a distribution broader than Poisson. In effect the idea of bursting, which amplifies protein relative to mRNA number variance, applies here too, amplifying the variance of transcript number above the expectations from the Poisson distribution. Transcriptional bursting has in fact been observed directly [7], although it is not clear whether this arises from fluctuations in transcription factor binding or from other sources.

## Summary

The main findings of this paper are as follows:

- Universal and simple noise model of Eq (1), in which the noise scales in inverse proportion to the mean expression for all genes, is inconsistent with the embedding of gene expression in a regulatory network. One possible cause of this inconsistency stems from assigning all of the noise to the processes on the 'output' side of transcriptional regulation, and ignoring the contribution of input noise. In particular, previous theoretical work has shown that the contribution of diffusive input noise to

the observed variance of gene expression could be significant [26].

- Despite appearing straightforward in theory, in real data the signatures of input noise can be surprisingly subtle to detect: inability to probe the whole physiological range of transcription factor concentrations, experimental noise, the difficulties in measuring absolute concentrations of both the regulator and controlled gene and established ways of plotting and analyzing data can act, among others, as confounding factors when noise measurements are used to learn about microscopic processes of transcriptional regulation.
- We argue that the Bicoid/Hunchback system of the fruit fly *Drosophila* provides an example in which input noise is dominant, and furthermore, that the detailed form of the variance vs mean quantitatively supports a dominant role for diffusion rather than switching noise.

Although there are caveats, these conclusions are consistent with the idea that, as with other critical biological processes [18,22,42,43], the regulation of gene expression can operate with a precision limited by fundamental physical principles.

## Methods

### Langevin derivation of input noise

We consider a simplified model of regulated gene expression, as schematized in Fig. 1:

$$\partial_t c = D\nabla^2 c(\mathbf{x}, t) - \dot{n}\delta(\mathbf{x} - \mathbf{x}_0) + \mathfrak{S} - \mathfrak{D} \quad (20)$$

$$\dot{n} = k_+ c(\mathbf{x}_0, t)(1 - n) - k_- n + \xi_n \quad (21)$$

$$\dot{e} = R_e n - \tau_e^{-1} e + \xi_e \quad (22)$$

$$\dot{g} = R_g e - \tau_g^{-1} g + \xi_g. \quad (23)$$

Equation (20) describes the diffusion of the transcription factor that can be absorbed to or released from a binding site on the DNA located at  $\mathbf{x}_0$ ;  $\delta(\mathbf{x} - \mathbf{x}_0)$  denotes the Dirac-delta function. These transcription factors are produced at sources  $\mathfrak{S}$  and degraded at sinks  $\mathfrak{D}$ , which can both be spatially distributed and can also contribute to the noise in  $c$ . Equation (21) describes the dynamics of the binding site occupancy; binding occurs with a second order rate constant  $k_+$  and unbinding with a first order rate constant  $k_-$ , and the dissociation constant of the site is  $K_d = k_-/k_+$ . The Langevin term  $\xi_n$  induces stochastic (binomial) switching between occupied and empty states of the site. Equations (22) and (23) describe the production and degradation of mRNA and protein, respectively, and include Langevin noise terms associated with these birth and death processes.

This seems a good place to note that, while conventional, the assumption that transcription and translation are simple one step processes seems a bit strong. We hope to return to this point at another time.

Our goal is to compute the variance in protein copy number,  $\sigma_g^2(\bar{c})$ . For simplicity we will assume that the transcription factors are present at a fixed total number in the cell and that they do not decay,  $\mathfrak{S} = \mathfrak{D} = 0$ . We will see that even with this simplification, where the overall concentration of transcription factors does not

fluctuate, we still get an interesting noise contribution from the randomness associated with diffusion in Eq (20).

Our basic strategy is to find the steady state solution of the model, and then linearize around this to compute the response of the variables  $\{n, e, g\}$  to the various Langevin forces  $\{\xi_n, \xi_e, \xi_g\}$ . In the linear approximation, the steady states are also the mean values:

$$c = \bar{c} \tag{24}$$

$$\langle n \rangle = \frac{k_+ \bar{c}}{k_+ \bar{c} + k_-} = \frac{\bar{c}}{\bar{c} + K_d} \tag{25}$$

$$\langle e \rangle = R_e \tau_e \langle n \rangle \tag{26}$$

$$\langle g \rangle = R_g \tau_g \langle e \rangle = g_0 \langle n \rangle, \tag{27}$$

where  $g_0 = R_e \tau_e R_g \tau_g$  is the maximum mean expression level. Notice that what we have called  $\bar{g} = \langle g \rangle / g_0$  in the text is just the mean occupancy,  $\langle n \rangle$ , of the transcription factor binding site.

Small departures from steady state are written in a Fourier representation:

$$c(\mathbf{x}, t) = \bar{c} + \int \frac{d\omega}{2\pi} \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{x}} e^{-i\omega t} \delta\tilde{c}(\mathbf{k}, \omega) \tag{28}$$

$$n = \langle n \rangle + \int \frac{d\omega}{2\pi} e^{-i\omega t} \delta\hat{n}(\omega) \tag{29}$$

$$e = \langle e \rangle + \int \frac{d\omega}{2\pi} e^{-i\omega t} \delta\hat{e}(\omega) \tag{30}$$

$$g = \langle g \rangle + \int \frac{d\omega}{2\pi} e^{-i\omega t} \delta\hat{g}(\omega). \tag{31}$$

Similarly, each of the Langevin terms is written in its Fourier representation,

$$\xi_\mu = \int \frac{d\omega}{2\pi} e^{-i\omega t} \hat{\xi}_\mu(\omega), \tag{32}$$

where  $\mu = n, e, g$ .

As a first step we use the Fourier representation to solve Eq (20) for  $\delta c(\mathbf{x}_0, t)$  that we need to substitute into Eq (21) for the binding site occupancy:

$$\delta c(\mathbf{x}_0, t) = \int \frac{d\omega}{2\pi} e^{-i\omega t} \delta\tilde{c}(\mathbf{x}_0, \omega) \tag{33}$$

$$\delta c(\mathbf{x}_0, \omega) = i\omega \delta\hat{n}(\omega) \int \frac{d^3k}{(2\pi)^3} \frac{1}{-\omega + D|\mathbf{k}|^2} \tag{34}$$

$$= \frac{i\omega \delta\hat{n}(\omega)}{\pi D a}. \tag{35}$$

The integral over  $\mathbf{k}$  in Eq (34) is divergent at large  $|\mathbf{k}|$  (ultraviolet). This arises, as explained in Ref [26], because we started with the assumption that the binding reaction occurs at a point—the delta function in Eq (20). In fact our description needs to be coarse grained on a scale corresponding to the size of the binding site, so we introduce a cutoff so that  $|\mathbf{k}| \leq k_{max} = 2\pi/a$ , where  $a$  is the linear size of the binding site.

Linearizing Eq (21) for the dynamics of the site occupancy, we have

$$\begin{aligned} -i\omega \delta\hat{n}(\omega) = \\ -(k_+ \bar{c} + k_-) \delta\hat{n}(\omega) + k_+ (1 - \langle n \rangle) \delta\tilde{c}(\mathbf{x}_0, \omega) + \hat{\xi}_n(\omega). \end{aligned} \tag{36}$$

Substituting our result for  $\delta\tilde{c}(\mathbf{x}_0, \omega)$  from Eq (35), we find

$$\begin{aligned} -i\omega \delta\hat{n}(\omega) = \\ -(k_+ \bar{c} + k_-) \delta\hat{n}(\omega) + k_+ (1 - \langle n \rangle) \frac{i\omega \delta\hat{n}(\omega)}{\pi D a} + \hat{\xi}_n(\omega) \end{aligned} \tag{37}$$

$$\begin{aligned} -i\omega \left[ 1 + \frac{k_+ (1 - \langle n \rangle)}{\pi D a} \right] \delta\hat{n}(\omega) = \\ -(k_+ \bar{c} + k_-) \delta\hat{n}(\omega) + \hat{\xi}_n(\omega) \end{aligned} \tag{38}$$

$$\delta\hat{n}(\omega) = \frac{\hat{\xi}_n(\omega)}{-i\omega(1 + \Sigma) + (k_+ \bar{c} + k_-)} \tag{39}$$

where  $\Sigma = k_+(1 - \langle n \rangle) / (\pi D a)$ . The linearization of Eqs (22) and (23) takes the form

$$-i\omega \delta\hat{e}(\omega) = \frac{1}{\tau_e} \delta\hat{e}(\omega) + R_e \delta\hat{n}(\omega) + \hat{\xi}_e(\omega) \tag{40}$$

$$-i\omega \delta\hat{g}(\omega) = -\frac{1}{\tau_g} \delta\hat{g}(\omega) + R_g \delta\hat{e}(\omega) + \hat{\xi}_g(\omega) \tag{41}$$

Each Langevin term is independent, and each frequency component  $\omega$  is correlated only with the component at  $-\omega$ , defining the noise power spectrum  $\langle \hat{\xi}_\mu(\omega) \hat{\xi}_\mu(-\omega) \rangle = 2\pi \delta(\omega - \omega') N_\mu(\omega)$  for  $\mu = n, e, g$ , where  $\delta(\omega - \omega')$  is the Dirac-delta function. Solving the three linear equations, Eqs (39–41), we can find the power spectrum of the protein copy number fluctuations,

$$\begin{aligned} S_g(\omega) = \frac{N_g}{\omega^2 + 1/\tau_g^2} + R_g^2 \frac{N_e}{(\omega^2 + 1/\tau_g^2)(\omega^2 + 1/\tau_e^2)} + \\ + R_g^2 R_e^2 \frac{N_n}{(\omega^2 + 1/\tau_g^2)(\omega^2 + 1/\tau_e^2) [(1 + \Sigma)^2 \omega^2 + 1/\tau_c^2]}, \end{aligned} \tag{42}$$

where  $1/\tau_c = k_+ \bar{c} + k_-$ . This form has a very intuitive interpretation: each Langevin term represents a noise source; as this noise propagates from the point where it enters the dynamical system to the output, it is subjected both to gain of each successive stage (prefactors  $R$ ), and to filtering by factors of  $\mathfrak{F}_\tau = (\omega^2 + 1/\tau^2)^{-1}$ .

The total variance in protein copy number is given by an integral over the spectrum,

$$\langle (\delta g)^2 \rangle \equiv \sigma_g^2 = \int \frac{d\omega}{2\pi} S_g(\omega), \quad (43) \quad \frac{1}{\tau_c} \equiv k_+ \bar{c} + k_- \quad (52)$$

and the noise power spectra of the Langevin terms associated with the mRNA and protein dynamics have the simple forms  $N_e(\omega) = 2R_e \langle n \rangle$  and  $N_g(\omega) = 2R_g \langle e \rangle$ , respectively. The spectrum  $N_n(\omega)$  is more subtle. One way to derive it is to realize that since there is only one binding site and this site is either occupied or empty, the total variance of  $\delta n$  must be given by the binomial formula,

$$\langle (\delta n)^2 \rangle = \langle n \rangle (1 - \langle n \rangle). \quad (44)$$

Starting with Eq (39) and the analog of Eq (43), we can use this condition to set the magnitude of  $N_n$ . Alternatively, we can use the fact that binding and unbinding come to equilibrium, and hence the fluctuations in  $n$  are a form of thermal noise, like Brownian motion or Johnson noise. The spectrum  $N_n$  is therefore determined by the fluctuation-dissipation theorem [26]. The result is that

$$N_n = \frac{2}{\tau_c} (1 + \Sigma) \langle n \rangle (1 - \langle n \rangle). \quad (45)$$

For simplicity we consider the case where the protein lifetime  $\tau_g$  is long compared with all other time scales in the problem. Then we can approximate Eq (42) as

$$S_g(\omega) \approx \frac{1}{\omega^2 + 1/\tau_g^2} \left[ N_g + (R_g \tau_e)^2 N_e + (R_g \tau_e R_e \tau_c)^2 N_n \right]. \quad (46)$$

Substituting the forms of the individual noise spectra  $N_i$  and doing the integral over  $\omega$  [Eq (43)], we find the variance in protein copy number

$$\begin{aligned} \sigma_g^2 = & \tau_g \left[ R_g \langle e \rangle + (R_g \tau_e)^2 R_e \langle n \rangle \right] \\ & + \frac{\tau_g}{\tau_c} (R_g \tau_e R_e \tau_c)^2 (1 + \Sigma) \langle n \rangle (1 - \langle n \rangle). \end{aligned} \quad (47)$$

We notice that the first term in this equation is  $R_g \tau_g \langle e \rangle$ , which is just the mean number of proteins  $\langle g \rangle$  from Eq (27). The second term is

$$\tau_g (R_g \tau_e)^2 R_e \langle n \rangle = R_g \tau_g (R_e \tau_e \langle n \rangle) (R_g \tau_e) \quad (48)$$

$$= R_g \tau_g \langle e \rangle (R_g \tau_e) \quad (49)$$

$$= R_g \tau_e \langle g \rangle. \quad (50)$$

Thus, the first two terms together contribute  $(1 + R_g \tau_e) \langle g \rangle$  to the variance, and this corresponds to the output noise term in Eq (16).

The third term in Eq (47) contains the contribution of input noise to the variance in protein copy number. To simplify this term we note that the steady state of Eq (21) is equivalent to

$$k_+ \bar{c} (1 - \langle n \rangle) = k_- \langle n \rangle. \quad (51)$$

Thus we can write

$$= k_- \left[ \frac{\langle n \rangle}{1 - \langle n \rangle} + 1 \right] = \frac{k_-}{1 - \langle n \rangle}. \quad (53)$$

The term we are interested in is

$$\begin{aligned} & \frac{\tau_g}{\tau_c} (R_g \tau_e R_e \tau_c)^2 (1 + \Sigma) \langle n \rangle (1 - \langle n \rangle) \\ & = (R_g \tau_g R_e \tau_e)^2 \frac{\tau_c}{\tau_g} (1 + \Sigma) \langle n \rangle (1 - \langle n \rangle) \end{aligned} \quad (54)$$

$$= g_0^2 \frac{1}{k_- \tau_g} (1 + \Sigma) \langle n \rangle (1 - \langle n \rangle)^2 \quad (55)$$

$$= g_0^2 \frac{\langle n \rangle (1 - \langle n \rangle)^2}{k_- \tau_g} + g_0^2 \frac{1}{k_- \tau_g} \frac{k_+ (1 - \langle n \rangle)}{\pi D a} \langle n \rangle (1 - \langle n \rangle)^2 \quad (56)$$

$$= g_0^2 \frac{\langle n \rangle (1 - \langle n \rangle)^2}{k_- \tau_g} + g_0^2 \frac{\langle n \rangle^2 (1 - \langle n \rangle)^2}{\pi D a \bar{c} \tau_g}, \quad (57)$$

where in the last step we once again use Eq (51) to rewrite the ratio  $k_+/k_-$  in terms of  $\langle n \rangle$ . Using Eq (50) and Eq (57) together to simplify the output and input noise terms in the equation for total noise, Eq (47), we finally find

$$\left( \frac{\sigma_g}{g_0} \right)^2 = \frac{1 + R_g \tau_e}{g_0} \bar{g} + \frac{(1 - \bar{g})^2 \bar{g}}{k_- \tau_g} + \frac{(1 - \bar{g})^2 \bar{g}^2}{\pi D a \bar{c} \tau_g} \quad (58)$$

which is Eq (13) in the main text.

To establish a connection between this result and the discussion in the *Global consistency* section, in particular with Eq (4), we first point out that the derivation presented here is more detailed, as it includes the dynamics of mRNA and binding site occupancy in addition to the dynamics of the protein levels  $g$ . This allows us to capture all the relevant noise sources – especially the input noise that we claim is important – and compute their variances (for example, without including binding site occupancy as an independent dynamical variable, we would not be able to compute the switching noise contribution). It is nevertheless instructive to rewrite the output and diffusion input noise terms of Eq (58) as follows:

$$\sigma_g^2 = g_0 (1 + R_g \tau_e) \bar{g} + g_0^2 \frac{(1 - \bar{g})^2 \bar{g}^2}{\pi D a \bar{c} \tau_g} \quad (59)$$

$$= \sigma_{g,0}^2 + \left( \frac{\partial \langle g \rangle}{\partial \bar{c}} \right)^2 \times \frac{\tau_D}{\tau_g} \times \sigma_c^2, \quad (60)$$

where the correlation time of diffusion fluctuations  $\tau_D = a^2/\pi D$  is approximately the time that a molecule needs to clear out of a region of dimension  $a$  by diffusion, and  $\sigma_c^2 = \bar{c}/a^3$  is the variance of

the concentration in a small volume  $V = a^3$ . To see why the latter is true, consider a small volume  $V$  that on average contains  $\bar{N} = \bar{c}V$  molecules. The variance in the number of molecules is Poisson, and therefore  $\sigma_N^2 = \bar{N} = \bar{c}V$ . But  $\sigma_c = \sigma_N/V$ , and hence  $\sigma_c^2 = \bar{c}/V$ .

After these rearrangements, Eq (60) is clearly of the form presented in Eq (4), if the only regulator of gene  $g$  is the transcription factor  $c$ . Then the filtering term  $\Phi = \tau_D/(\tau_D + \tau_g) \approx \tau_D/\tau_g$ , because diffusion fluctuations happen on timescales much shorter than  $\tau_g$  (which is usually of the order of at least a minute); and the susceptibility,  $\partial\langle g \rangle/\partial\bar{c}$ , can be computed from the steady state  $\langle g \rangle = g_0\bar{c} = g_0\bar{c}/(\bar{c} + K_d)$ . To summarize: in this section we computed the noise arising from a particular set of microscopic processes, summarized by Eqs (20–23), using the Langevin formalism, and have concluded by showing how the result fits into a general framework of noise propagation of Eq (2).

### Cooperative binding of transcription factors

To generalize this analysis of noise to cooperative interactions among transcription factors it is useful to think more intuitively about the two terms in Eq (57), corresponding to switching and diffusion noise. Consider first the switching noise.

We are looking at a binary variable  $n$  such that the number of proteins is  $g_0n$ . The total variance in  $n$  must be  $\langle(\delta n)^2\rangle = \langle n \rangle(1 - \langle n \rangle)$  [Eq (44)]. This noise fluctuates on a time scale  $\tau_c$ , so during the lifetime of the protein we see  $N_s = \tau_g/\tau_c$  independent samples. The current protein concentration is effectively an average over these samples, so the effective variance is reduced to

$$\langle(\delta n)^2\rangle_{\text{eff}} = \frac{1}{N_s} \langle n \rangle(1 - \langle n \rangle) = \frac{\tau_c}{\tau_g} \langle n \rangle(1 - \langle n \rangle). \quad (61)$$

Except for the factor of  $g_0$  that converts  $n$  into  $g$ , this is the first term in Eq (57).

Now if  $h$  transcription factors bind cooperatively, we can still have two states, one in which transcription is possible and one in which it is blocked. For the case of activation, which we are considering here, the active state corresponds to all binding sites being filled, and so the rate at which the system leaves this state,  $k_-$ , should not depend on the concentration of the transcription factors. The rate at which the system enters the active state does depend on concentration, but this does not matter, because with only two states we must always have an analog of Eq (51), which allows us to eliminate the “on rate” in favor of  $k_-$  and  $\langle n \rangle$ . The conclusion is that the first term in Eq (57), corresponding to switching noise, is unchanged by cooperativity as long as the system is still well approximated as having just two states of transcriptional activity that depend on the potentially many more states of binding site occupancy.

For the diffusion noise term we use the ideas of Refs [22,26,29]. Diffusion noise should be thought of as an effective noise in the measurement of the concentration  $c$ , with a variance

$$\frac{\sigma_c^2}{\bar{c}^2} \sim \frac{1}{\pi D a \bar{c} \tau_g}, \quad (62)$$

where again we identify the protein lifetime as the time over which the system averages. For the system with a single binding site,

$$\langle n \rangle = \frac{\bar{c}}{\bar{c} + K_d}, \quad (63)$$

so that

$$\frac{\partial\langle n \rangle}{\partial c} = \frac{1}{\bar{c}} \langle n \rangle(1 - \langle n \rangle). \quad (64)$$

The noise in concentration, together with this sensitivity of  $n$  to changes in the concentration, should contribute a noise variance

$$\langle(\delta n)^2\rangle_{\text{eff}} = \left| \frac{\partial\langle n \rangle}{\partial c} \right|^2 \sigma_c^2 = \frac{\langle n \rangle^2(1 - \langle n \rangle)^2}{\pi D a \bar{c} \tau_g}. \quad (65)$$

This is (up to the factor of  $g_0$ ) the second term in Eq (57). Now the generalization to cooperative interactions is straightforward. If we have

$$\langle n \rangle = \frac{\bar{c}^h}{\bar{c}^h + K_d^h}, \quad (66)$$

then

$$\frac{\partial\langle n \rangle}{\partial c} = \frac{h}{\bar{c}} \langle n \rangle(1 - \langle n \rangle). \quad (67)$$

Since the effective noise in concentration is unchanged [29], the only effect of cooperativity is to multiply the second term in Eq (57) by a factor of  $h^2$ .

Thus, in the expression [Eq (16)] for the variance of protein copy number, cooperativity has no effect on the switching noise but actually increases the diffusion noise by a factor of  $h^2$ . When written as a function of the mean copy number and the transcription factor concentration, this leaves the functional form of the variance fixed, only changing the coefficients. The overall effect is to make the contribution of diffusion noise more important. One way to say this is that, when we refer the noise in copy number back to the input, cooperativity causes the equivalent concentration noise to become closer to the limit Eq (62) set by diffusive noise [29].

Reference [32] also considers the possibility that noise is reduced by averaging among neighboring nuclei. This does not change the form of any of the noise terms, but does change the microscopic interpretation of the coefficients  $\alpha$  and  $\beta$ . For example, averaging for a time  $\tau_g$  over  $N$  nuclei is equivalent to having one nucleus with an averaging time  $N\tau_g$ .

### Acknowledgments

We thank CG Callan, JB Kinney, P Mehta, RR de Ruyter van Steveninck, S Setayeshgar, DW Tank, and EF Wieschaus for helpful discussions.

### Author Contributions

Conceived and designed the experiments: GT TG WB. Performed the experiments: TG. Analyzed the data: GT TG WB. Wrote the paper: GT TG WB.

### References

1. Elowitz MB, Levine AJ, Siggia ED, Swain PD (2002) Stochastic gene expression in a single cell. *Science* 297: 1183–86.
2. Ozbudak E, Thattai M, Kurtser I, Grossman AD, van Oudenaarden A (2002) Regulation of noise in the expression of a single gene. *Nature Gen* 31: 69–73.
3. Blake WJ, Kaern M, Cantor CR, Collins JJ (2003) Noise in eukaryotic gene expression. *Nature* 422: 633–637.
4. Raser JM, O’Shea EK (2004) Control of stochasticity in eukaryotic gene expression. *Science* 304: 1811–1814.

5. Rosenfeld N, Young JW, Alon U, Swain PW, Elowitz MB (2005) Gene regulation at the single cell level. *Science* 307: 1962–1965.
6. Pedraza JM, van Oudenaarden A (2005) Noise propagation in gene networks. *Science* 207: 1965–1969.
7. Golding I, Paulsson J, Zawilski SM, Cox EC (2005) Real-time kinetics of gene activity in individual bacteria. *Cell* 123: 1025–1036.
8. Newman JR, Ghaemmaghami S, Ihmels J, Breslow DK, Noble M, et al. (2006) Single-cell proteomic analysis of *S cerevisiae* reveals the architecture of biological noise. *Nature* 441: 840–846.
9. Bar-Even A, Paulsson J, Maheshri N, Carmi M, O’Shea E, et al. (2006) Noise in protein expression scales with natural protein abundance. *Nature* 38: 636–643.
10. Luria SE, Delbrück M (1943) Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* 28: 491–511.
11. Fatt P, Katz B (1950) Some observations on biological noise. *Nature* 166: 597–598.
12. Fatt P, Katz B (1952) Spontaneous subthreshold activity at motor nerve endings. *J Physiol (Lond)* 117: 109–128.
13. Lecar H, Nossal R (1971) Theory of threshold fluctuations in nerves. I: Relationships between electrical noise and fluctuations in axon firing. *Biophys J* 11: 1048–1067.
14. Lecar H, Nossal R (1971) Theory of threshold fluctuations in nerves. II: Analysis of various sources of membrane noise. *Biophys J* 11: 1068–1084.
15. Stevens CF (1972) Inferences about membrane properties from electrical noise measurement. *Biophys J* 12: 1028–1048.
16. Conti F, de Felice LJ, Wanke E (1975) Potassium and sodium ion current noise in the membrane of the squid giant axon. *J Physiol (Lond)* 248: 45–82.
17. Hecht S, Shlaer S, Pirenne MH (1942) Energy, quanta, and vision. *J Gen Physiol* 25: 819–840.
18. Barlow HB (1981) Critical limiting factors in the design of the eye and visual cortex. *Proc R Soc Lond Ser B* 212: 1–34.
19. Verveen AA, Derksen HE (1965) Fluctuation in membrane potential of axons and the problem of coding. *Kybernetik* 2: 152–160.
20. Verveen AA, Derksen HE (1968) Fluctuation phenomena in nerve membrane. *Proc IEEE* 56: 906–916.
21. Schneidman E, Freedman B, Segev I (1998) Ion channel stochasticity may be critical in determining the reliability and precision of spike timing. *Neural Comp* 10: 1679–1704.
22. Berg HC, Purcell EM (1977) Physics of chemoreception. *Biohys J* 20: 193–219.
23. Paulsson J (2004) Summing up the noise in gene networks. *Nature* 427: 415–418.
24. Barlow HB, Levick WR, Yoon M (1971) Responses to single quanta of light in the retinal ganglion cells of the cat. *Vision Res Suppl* 3: 87–101.
25. Swain PS, Elowitz MB, Siggia ED (2002) Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc Natl Acad Sci (USA)* 99: 12795–12800.
26. Bialek W, Setayeshgar S (2005) Physical limits to biochemical signaling. *Proc Natl Acad Sci (USA)* 102: 10040–10045. arXiv.org/physics/0301001.
27. Walczak AM, Sasai M, Wolynes PG (2005) Self-consistent proteomic field theory of stochastic gene switches. *Biophys J* 88: 828–850. arXiv.org:q-bio.MN/0407041.
28. van Zon JS, Morelli MJ, Tanase-Nicola S, ten Wolde PR (2006) Diffusion of transcription factors can drastically enhance the noise in gene expression. *Biophys J* 91: 4350–4367. arXiv.org:q-bio.MN/0604005.
29. Bialek W, Setayeshgar S (2006) Cooperativity, sensitivity and noise in biochemical signaling. arXiv.org:q-bio.MN/0601001.
30. Sigworth FJ (1977) Sodium channels in nerve apparently have two conductance states. *Nature* 270: 265–267.
31. Sigworth FJ (1980) The variance of sodium current fluctuations at the node of Ranvier. *J Physiol (Lond)* 307: 97–129.
32. Gregor T, Tank DW, Wieschaus EF, Bialek W (2007) Probing the limits to positional information. *Cell* 130: 153–164.
33. Hooshang S, Thiberge S, Weiss R (2005) Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proc Natl Acad Sci (USA)* 102: 3581–3586.
34. Gregor T, Wieschaus EF, McGregor AP, Bialek W, Tank DW (2007) Stability and nuclear dynamics of the Bicoid morphogen gradient. *Cell* 130: 141–152.
35. Ptashne M () *A Genetic Switch*. Second edition: Phage  $\lambda$  and Higher Organisms Cell Press: Cambridge, 1992).
36. Pedone PV, Ghirlando R, Clore GM, Gronenbron AM, Felsenfeld G, et al. (1996) The single Cys2–His2 zinc finger domain of the GAGA protein flanked by basic residues is sufficient for high-affinity specific DNA binding. *Proc Natl Acad Sci (USA)* 93: 2822–2826.
37. Ma X, Yuan D, Diepold K, Scarborough T, Ma J (1996) The *Drosophila* morphogenetic protein Bicoid binds DNA cooperatively. *Development* 122: 1195–1206.
38. Burz DS, Rivera-Pomar R, Jäckle H, Hanes SD (1998) Cooperative DNA-binding by Bicoid provides a mechanism for threshold dependent gene activation in the *Drosophila* embryo. *EMBO J* 17: 5998–6009.
39. Winston RL, Millar DP, Gottesfeld JM, Kent SB (1999) Characterization of the DNA binding properties of the bHLH domain of Deadpan to single and tandem sites. *Biochemistry* 38: 5138–5146.
40. Zhao C, York A, Yang F, Forsthoefel DJ, Dave V, et al. (2002) The activity of the *Drosophila* morphogenetic protein Bicoid is inhibited by a domain located outside its homeodomain. *Development* 129: 1669–1680.
41. Reinitz J (2007) Developmental biology: A ten per cent solution. *Nature* 448: 420–421.
42. Bialek W (1987) Physical limits to sensation and perception. *Ann Rev Biophys Biophys Chem* 16: 455–478.
43. Bialek W (2002) Thinking about the brain. In *Physics of Biomolecules and Cells: Les Houches Session LXXXV Flyvbjerg H, Jülicher F, Ormos P, David F, eds. pp 485–577, EDP Sciences, Les Ulis; Springer-Verlag: Berlin, 2002, arXiv.org: physics/0205030.*