

Hill kinetics as a noise filter: The role of transcription factor autoregulation in gene cascades.

Electronic supplementary material.

Anna Ochab-Marcinek,^{*†} Jakub Jędrak,[†] and Marcin Tabaka^{†‡}

August 9, 2017

Contents

S1	Dependence of Hill function on external signal	1
S2	Intrinsic noise of the target gene	4
S3	Extrema of the transcription rate distribution	5
S4	Extrema of the TF number distribution	8
S5	Normalization constant for the protein number distribution of an autoregulated gene	8
S6	Simulations	10
S7	Time scales and parameters of the system	10
S8	Response of the cascade in absence of self-regulation	13
S9	Comparison between deterministic and stochastic model for Figs. 2B1 and 2B2 in main text	17
S10	Dependence of distribution shape on leakage	20
S11	Example of a trimodal transcription rate distribution	23
S12	Zoomed view of Fig. 5D from the main manuscript	24
S13	Coefficients of variation and Fano factors of the distributions	24
S14	Precision of response of the target gene to external signal	26
S15	Upstream gene positively autoregulated, downstream gene negatively regulated	29
S16	Considerations about the experimental measurement of transfer functions	30

S1 Dependence of Hill function on external signal

The derivation is analogous as in [1], but for the sake of clarity we present it here in more detail. First, we will consider the simplest case of a single TF binding site on the operator ($n = 1$).

*Corresponding author: Anna Ochab-Marcinek, ochab@ichf.edu.pl

†Institute of Physical Chemistry, Polish Academy of Sciences, Kasprzaka 44/52, 01-224 Warsaw, Poland

‡Currently affiliated at: Broad Institute of MIT and Harvard, 415 Main Street, Cambridge MA 02142, USA

We assume that $R = R_a + R_i$ is the sum of active and inactive TFs. k_{on}^a, k_{off}^a are the rates of binding and unbinding of the active TF to the operator. k_{on}^i, k_{off}^i are the rates of binding and unbinding of the inactive TF. Then, the regulation part of the transfer function reads:

$$H(R) = \frac{1}{1 + \frac{k_{on}^a}{k_{off}^a} R_a + \frac{k_{on}^i}{k_{off}^i} R_i} \quad (S1)$$

$$= \frac{1}{1 + R \left[\frac{k_{on}^a}{k_{off}^a} f_a + \frac{k_{on}^i}{k_{off}^i} (1 - f_a) \right]}.$$

Under the assumption of negligible binding of inactive TFs,

$$H(R) = \frac{1}{1 + R \left[\frac{k_{on}^a}{k_{off}^a} f_a \right]} = \frac{1}{1 + R/K}. \quad (S2)$$

In the case of cooperative binding of TFs to 2 binding sites ($n = 2$), we have:

$$H(R) = \left[1 + \frac{k_{on,1}^a}{k_{off,1}^a} \frac{k_{on,2}^a}{k_{off,2}^a} R_a^2 + \frac{k_{on,1}^i}{k_{off,1}^i} \frac{k_{on,2}^i}{k_{off,2}^i} R_i^2 \right. \quad (S3)$$

$$+ \left(\frac{k_{on,1}^i}{k_{off,1}^i} \frac{\tilde{k}_{on,2}^a}{\tilde{k}_{off,2}^a} + \frac{k_{on,1}^a}{k_{off,1}^a} \frac{\tilde{k}_{on,2}^i}{\tilde{k}_{off,2}^i} \right) R_i R_a$$

$$\left. + \frac{k_{on,1}^a}{k_{off,1}^a} R_a + \frac{k_{on,1}^i}{k_{off,1}^i} R_i \right]^{-1}$$

Here, we simplistically assume that each of the two binding sites on the operator has the same affinity to TF unless the other binding site is already occupied. Therefore, in our notation, the subscript “1” denotes binding/unbinding of TF to the free operator and the subscript “2” denotes binding of TF to the already occupied operator. We also broadly assume that there may be a difference in the binding/unbinding rates depending on whether the other binding site is occupied by an active or inactive TF (see Table S1). Thus, the term $\frac{k_{on,1}^i}{k_{off,1}^i} \frac{\tilde{k}_{on,2}^a}{\tilde{k}_{off,2}^a}$ describes the reaction $R^a + R^i O \rightleftharpoons R^a R^i O$, where O is the operator. The term $\frac{k_{on,1}^a}{k_{off,1}^a} \frac{\tilde{k}_{on,2}^i}{\tilde{k}_{off,2}^i}$ describes the reaction $R^i + R^a O \rightleftharpoons R^a R^i O$.

Assuming strong cooperativity of R^a binding and weak binding of R^i even in the presence of R^a already bound, we obtain

$$H(R) = \frac{1}{1 + R^2 \left[\frac{k_{on,1}^a}{k_{off,1}^a} \frac{k_{on,2}^a}{k_{off,2}^a} f_a^2 \right]} = \frac{1}{1 + (R/K)^2}, \quad (S4)$$

and the derivation is analogous for other values of n .

TF binding/unbinding to the free operator	
Active TF	$R^a + O \xrightleftharpoons[k_{off,1}^a]{k_{on,1}^a} R^a O$
Inactive TF	$R^i + O \xrightleftharpoons[k_{off,1}^i]{k_{on,1}^i} R^i O$
TF binding/unbinding to the already occupied operator	
Active TF	$R^a + R^a O \xrightleftharpoons[k_{off,2}^a]{k_{on,2}^a} R_2^a O$
Active TF	$R^a + R^i O \xrightleftharpoons[k_{off,2}^a]{\tilde{k}_{on,2}^a} R^a R^i O$
Inactive TF	$R^i + R^i O \xrightleftharpoons[k_{off,2}^i]{k_{on,2}^i} R_2^i O$
Inactive TF	$R^i + R^a O \xrightleftharpoons[\tilde{k}_{off,2}^i]{\tilde{k}_{on,2}^i} R^a R^i O$

Table S1: Reactions of TF binding/unbinding to the operator assumed for derivation of our model, here for $n = 2$. For the sake of clarity, the reactions also include inactive TFs. Further on, we will assume that binding of inactive TFs is negligible ($k_{on,1}^i = k_{on,2}^i = \tilde{k}_{on,2}^i = 0$), see Table S2.

The assumptions of strong cooperativity and negligible binding of inactive TFs allows us to define K by

$$K^{-1} = \left(\frac{k_{on,1}^a \dots k_{on,n}^a}{k_{off,1}^a \dots k_{off,n}^a} \right)^{1/n} f_a, \quad (S5)$$

where the term in parentheses is the ratio of the on-rate and off-rate constants of (active) TF binding to the n sites on the operator, and f_a is the fraction of active TFs. And thus, for simplicity, we will use K as a measure of the strength of the signal that activates or deactivates the TFs (e.g., dependent on the concentration of effector molecules or TF phosphorylation), analogously as in [1]. We assume that the concentration of signal molecules is much greater than TF concentration – that is, the number of free signal molecules is well approximated by the total number of signal molecules, so that a given signal level makes a fraction f_a of TFs active independent of the TF level in the cell.

Gene 1 is autoregulated and TF binding to its operator is described by Hill kinetics with the transfer function $h_1(R)$, with the signal parameter K_1 , cooperativity n , and leakage $\epsilon_1 = k_{ml1}/k_{m1}$:

$$h_1(R) = H_1(R)(1 - \epsilon_1) + \epsilon_1, \quad \text{where} \quad H_1(R) = \frac{1}{1 + (R/K_1)^n} \quad (S6)$$

Gene 2 encoding for the target protein P is also regulated by the transcription factor R, whose binding to the operator of that gene is described by Hill kinetics with the transfer function $h_2(R)$ with the corresponding parameters K_2 , m , and $\epsilon_2 = k_{ml2}/k_{m2}$:

$$h_2(R) = H_2(R)(1 - \epsilon_2) + \epsilon_2, \quad \text{where} \quad H_2(R) = \frac{1}{1 + (R/K_2)^m}. \quad (S7)$$

Activation/deactivation of TF	
	$R^i \xrightleftharpoons[k_-]{k_+} R^a$
TF binding/unbinding	
to the operator of Gene 1	$R^a + O^r \xrightleftharpoons[k_{off,1}^r]{k_{on,1}^r} R^a O^r$
	$R^a + R^a O^r \xrightleftharpoons[k_{off,2}^r]{k_{on,2}^r} R_2^a O^r$
to the operator of Gene 2	$R^a + O^t \xrightleftharpoons[k_{off,1}^t]{k_{on,1}^t} R^a O^t$
	$R^a + R^a O^t \xrightleftharpoons[k_{off,2}^t]{k_{on,2}^t} R_2^a O^t$
Transcription	
from Gene 1, leakage	$O^r \xrightarrow{k_{ml1}} O^r + M^r$
autoactivated	$R^a O^r \xrightarrow{k_{m1}} R^a O^r + M^r$
	$R_2^a O^r \xrightarrow{k_{m1}} R_2^a O^r + M^r$
In the case of activation of Gene 2	
from Gene 2, leakage	$O^t \xrightarrow{k_{ml2}} O^t + M^t$
activated	$R^a O^t \xrightarrow{k_{m2}} R^a O^t + M^t$
	$R_2^a O^t \xrightarrow{k_{m2}} R_2^a O^t + M^t$
In the case of repression of Gene 2	
from Gene 2, unrepressed transcription	$O^t \xrightarrow{k_{m2}} O^t + M^t$
leakage	$R^a O^t \xrightarrow{k_{ml2}} R^a O^t + M^t$
	$R_2^a O^t \xrightarrow{k_{ml2}} R_2^a O^t + M^t$
Translation	
from Gene 1	$M^r \xrightarrow{k_{p1}} M^r + R^i$
from Gene 2	$M^t \xrightarrow{k_{p2}} M^t + P$
Degradation	
	$M^r \xrightarrow{k_{dm1}} \emptyset$
	$M^t \xrightarrow{k_{dm2}} \emptyset$
	$R^i \xrightarrow{k_{dp1}} \emptyset$
	$R^a \xrightarrow{k_{dp1}} \emptyset$
	$P \xrightarrow{k_{dp2}} \emptyset$

Table S2: Reactions assumed in our simulations. Note that binding of inactive TFs is assumed to be negligible (cf. Table S1).

the parameters in the Hill functions, K_1 for the Gene 1, and K_2 for the Gene 2, are both inversely proportional to the signal-dependent fraction of active TFs (Eq. S5), under the assumptions of our model, i.e. strong cooperativity and negligible binding of inactive TFs. K_1 is therefore proportional to K_2 at any signal level, so the signal strength can be measured by any of these parameters.

S2 Intrinsic noise of the target gene

The distribution of target proteins numbers in cell population is the superposition of the distribution of the transcription rates $q(h_2)$ and the intrinsic noise $g(P; k_{m2}h_2)$ corresponding to

a given transcription rate:

$$p_2(P) = \int_{\epsilon_2}^1 q(h_2)g(P; k_{m2}h_2)dh_2. \quad (\text{S8})$$

Under the assumption that target mRNA degrades faster than target proteins [2], g can be a negative binomial distribution, as used in [3], or a gamma distribution as its continuous limit [4, 5]:

$$g(P, k_{m2}h_2) = \gamma(P; ah_2, b) = \frac{1}{b^{ah_2}\Gamma(ah_2)} P^{ah_2-1} \exp\left(\frac{-P}{b}\right), \quad (\text{S9})$$

where $a = k_{m2}/k_{dp2}$ is the mean frequency of bursts at fully active promoter ($h_2 = 1$) and $b = k_{p2}/k_{dm2}$ is the mean burst size.

For $g(P; k_{m2}h_2)$ being a gamma distribution, it can be shown [3] that noise measured by the square of coefficient of variation, $\eta_P = \sigma_P^2/\langle P \rangle^2$ (σ_P^2 being the variance of the protein number distribution), is an additive quantity containing the contribution from the intrinsic noise as well as from the regulatory noise, for any mixing function $q(h_2)$,

$$\eta_P = \eta_{reg} + \eta_{int} = \frac{\sigma_{h_2}^2}{\langle h_2 \rangle^2} + \frac{1}{a\langle h_2 \rangle}, \quad (\text{S10})$$

where $\sigma_{h_2}^2$ is the variance of $q(h_2)$. Intrinsic noise of the target gene decreases as a increases, and for large a the distribution of target protein numbers $p_2(P)$ can be approximated by the rescaled function $\frac{1}{ab}q\left(\frac{P}{ab}\right)$ [3].

The method presented in this paper is, therefore, valid when the intrinsic noise of the target gene, quantified by $1/a$, is sufficiently small, such that the transcription rate distribution itself gives a significant information about the shape of the target gene expression.

S3 Extrema of the transcription rate distribution

S3.1 2-gene cascade in which the upstream gene is not regulated

The derivation has been described in detail in ref. [3] (SI therein), but, for the sake of clarity, we briefly remind it here, with the modification that the transfer function contains the term responsible for transcriptional leakage. A non-regulated Gene 1 produces TFs whose levels in cell population are gamma-distributed,

$$p_0(R) = \gamma(R; \alpha, \beta) = \frac{1}{\beta^\alpha \Gamma(\alpha)} R^{\alpha-1} \exp\left(\frac{-R}{\beta}\right), \quad (\text{S11})$$

with $\alpha = k_{m1}/k_{dp1}$ and $\beta = k_{p1}/k_{dm1}$. The input distribution $p_0(R)$ transferred through the nonlinear filter $h_2(R)$ produces an output distribution of transcription rates of the Gene 2 given by the relation

$$q_0(h_2) = p_0(R(h_2)) \left| \frac{dR(h_2)}{dh_2} \right|, \quad (\text{S12})$$

where $R(h_2)$ is the inverse function of $h_2(R)$,

$$R(h_2) = K_2 \left(\frac{h_2 - 1}{\epsilon_2 - h_2} \right)^{\frac{1}{m}}. \quad (\text{S13})$$

Thus, according to the relation (S12), the distribution of transcription rates of the target gene reads

$$q_0(h_2; \alpha, \beta, K_2) = \frac{1}{|m|} \frac{K_2^\alpha (1 - \epsilon_2)}{\beta^\alpha \Gamma(\alpha) (1 - h_2) (h_2 - \epsilon_2)} \left(\frac{1 - h_2}{h_2 - \epsilon_2} \right)^{\alpha/m} \exp \left[- \left(\frac{1 - h_2}{h_2 - \epsilon_2} \right)^{1/m} \frac{K_2}{\beta} \right]. \quad (\text{S14})$$

The formula is almost the same as in [3] except that the transfer function $h_2(R)$ contains the leakage term ϵ_2 .

S3.2 2-gene cascade in which the upstream gene is self-regulated

When the Gene 1 is self-regulated, the TF number distribution is given by [1]

$$p_1(R; \alpha, \beta) = AR^{\alpha-1} \exp \left(\frac{-R}{\beta} \right) H_1(R)^{\alpha(1-\epsilon_1)/n}, \quad (\text{S15})$$

where α and β are defined as above, and $\epsilon_1 = k_{ml1}/k_m$ is the leakage of the Gene 1. A is a normalization constant, whose explicit form was, to the best of our knowledge, not known to date. We have found that for a given n it can be presented using standard special functions. In Section S5, we have shown example formulas for A , for $n = \pm 1, \pm 2$. The distribution of transcription rates $q(h_2)$ obtained from the relation (S12) with the input distribution $p_1(R; \alpha, \beta)$ is the following:

$$q(h_2) = \frac{A}{|m|} \frac{K_2^\alpha (1 - \epsilon_2) e^{-\frac{K_2}{\beta} \left(\frac{1-h_2}{h_2-\epsilon_2} \right)^{\frac{1}{m}}}}{(1 - h_2) (h_2 - \epsilon_2)} \left(\frac{1 - h_2}{h_2 - \epsilon_2} \right)^{\frac{\alpha}{m}} \left(1 + \left(\frac{1 - h_2}{h_2 - \epsilon_2} \right)^{\frac{n}{m}} \left(\frac{K_2}{K_1} \right)^n \right)^{-\frac{\alpha(1-\epsilon_1)}{n}}. \quad (\text{S16})$$

However, the explicit knowledge of that formula is not needed if we just want to know the number and positions of its minima and maxima. We note that (S15) belongs to the same class as the gamma distribution (S11), such that

$$\frac{dp_1(R)}{dR} = p_1(R)F(R), \quad (\text{S17})$$

where $F(R) = \frac{\alpha-1}{R} - \frac{1}{\beta}$ for a non-regulated gene, and $F(R) = \frac{\alpha h_1(R)-1}{R} - \frac{1}{\beta}$ for an autoregulated gene. This property allows us to derive the analogous geometric construction as in [3]. We search for the extrema of the transcription rate distribution $q(h_2)$:

$$\frac{d}{dh_2} q(h_2) = 0. \quad (\text{S18})$$

Because

$$\frac{dp_1(R)}{dR} = p_1(R) \left[\frac{\alpha h_1(R) - 1}{R} - \frac{1}{\beta} \right], \quad (\text{S19})$$

the Eq. (S18) reads:

$$p_1(R) \left[\left(\frac{\alpha h_1(R) - 1}{R} - \frac{1}{\beta} \right) \left(\frac{dR(h_2)}{dh_2} \right)^2 + \frac{d^2R(h_2)}{dh_2^2} \right] = 0. \quad (\text{S20})$$

We note that for $\alpha < 1$, $p_1(R) = 0$ when $R = \infty$. For $\alpha > 1$, $p_1(R) = 0$ when $R = 0$ and $R = \infty$. We need to calculate the conditions for the term in the square bracket to be equal to zero:

$$\left(\frac{\alpha h_1(R) - 1}{R} - \frac{1}{\beta} \right) \left(\frac{dR(h_2)}{dh_2} \right)^2 + \frac{d^2R(h_2)}{dh_2^2} = 0. \quad (\text{S21})$$

It turns out that

$$\left(\frac{dR(h_2)}{dh_2} \right)^2 = \frac{1}{R(h_2)^2} f_1 \quad (\text{S22})$$

and

$$\frac{d^2R(h_2)}{dh_2^2} = \frac{1}{R(h_2)^3} f_2, \quad (\text{S23})$$

where

$$f_1 = \left(\frac{h_2 - 1}{\epsilon_2 - h_2} \right)^{\frac{4}{m}} \frac{K_2^4 (\epsilon_2 - 1)^2}{(\epsilon_2 - h_2)^2 m^2 (h_2 - 1)^2} \quad (\text{S24})$$

and

$$f_2 = f_1 \frac{m - 2h_2m + 1 + (m - 1)\epsilon_2}{1 - \epsilon_2}. \quad (\text{S25})$$

Substituting (S22), (S23), and (S24) into (S21), we obtain a relation that can be transformed by rearrangement of terms into the form:

$$h_2(R) = \left[-\frac{1}{2\beta m} R + \frac{\alpha h_1(R)}{2m} + \frac{1}{2} \right] (1 - \epsilon_2) + \epsilon_2 \quad (\text{S26})$$

The points of intersection of the transfer function $h_2(R)$ and the curve given by the right-hand side of the equation (S26), projected onto the vertical axis, define the positions of the minima and maxima of $q(h_2)$. If the number of intersections is even, then one more maximum is at $h_2 = \epsilon_2$ for positively regulated Gene 2 or at $h_2 = 1$ for negatively regulated Gene 2. Since the terms with ϵ_2 cancel out on both sides of Eq. S26, the construction can be presented in a simpler way as a Hill function intersecting a rescaled curve:

$$H_2(R) = -\frac{1}{2\beta m} R + \frac{\alpha h_1(R)}{2m} + \frac{1}{2} \equiv L(R). \quad (\text{S27})$$

S4 Extrema of the TF number distribution

From [1] we know that the minima and maxima of the protein number distribution $p_1(R)$ produced by the autoregulated Gene 1 are also given by a geometric construction, which shows that bimodal distribution of TF numbers is possible only for a positively autoregulated Gene 1 [1]. Note that the positions of the maxima and minima are in this case projected onto the horizontal axis R :

$$h_1(R) = \frac{1}{\alpha\beta}R + \frac{1}{\alpha} \quad (\text{S28})$$

(Additionally, if for $R > 0$ there are two intersections of $h_1(R)$ and the straight line on the right-hand side of Eq. (S28), then one more maximum is at $R = 0$.) After multiplication of both sides of Eq. (S28) by $\alpha/(2m)$, addition of $1/2$ to both sides and rearrangement of terms, the rescaled equation can be presented in a graphical form on the same plot as the geometric construction (S27) for the downstream gene and it takes the form:

$$\frac{m+1}{2m} = L(R), \quad (\text{S29})$$

where the curve $L(R)$ is given by Eq. (S27). The number of the extrema of $p_1(R)$ is given by the number of intersections of the horizontal straight line given by the left-hand side of the equation and the curve $L(R)$ defined by the right-hand side of Eq. (S27). Again, if for $R > 0$ there are two intersections, then one more maximum is at $R = 0$.

Steady states of the corresponding deterministic model are given by the equation in which the noise term $1/\alpha$ is absent:

$$h_1(R) = \frac{1}{\alpha\beta}R, \quad (\text{S30})$$

which leads to the equivalent of the Eq. S29:

$$\frac{1}{2} = L(R). \quad (\text{S31})$$

S5 Normalization constant for the protein number distribution of an autoregulated gene

A is a normalization constant of the distribution $p_1(R)$ of TF numbers. The explicit form of A was, to the best of our knowledge, not known to date. We have found that for a given n it can be presented using standard special functions. A is, in other words, the 0-th moment of $p_1(R)$.

$$\mu_0 \equiv \int_0^\infty p_1(R)dR, \quad (\text{S32})$$

of the following (not normalized) probability distribution

$$p_1(R) = R^{a-1} e^{-R/b} \left[1 + \left(\frac{R}{K} \right)^n \right]^{-\frac{a(1-\epsilon)}{n}}. \quad (\text{S33})$$

Interestingly, not only μ_0 , but also higher moments of $p_1(R)$ (S33) may be expressed with the help of known special functions. In this way, one may obtain analytical formulas for the moments as a function of the model parameters,

$$\mu_m = \mu_m(a, b, K, \epsilon, n), \quad m = 0, 1, 2 \quad (\text{S34})$$

We assume that a, b, K , and ϵ are real and non-negative, and that n is integer. We consider here $n = \pm 1, \pm 2$. The explicit form of μ_m (S34) can be obtained with the help of *Mathematica* symbolic algebra package.

S5.1 n=1

$$\begin{aligned} \mu_0(a, b, K, \epsilon, 1) &= \pi K^a \csc(\pi a \epsilon) \left(\frac{K}{b} \right)^{-a \epsilon} {}_1\tilde{F}_1 \left(a - a \epsilon; 1 - a \epsilon; \frac{K}{b} \right) \\ &- \pi K^a \csc(\pi a \epsilon) \frac{\Gamma(a) {}_1\tilde{F}_1 \left(a; a \epsilon + 1; \frac{K}{b} \right)}{\Gamma(a - a \epsilon)}. \end{aligned} \quad (\text{S35})$$

In above, $\csc(\pi a \epsilon) = 1/\sin(\pi a \epsilon)$, $\Gamma(z)$ is a Gamma function, whereas ${}_1\tilde{F}_1(\alpha; \beta; z)$ is the regularized confluent hypergeometric function

$${}_1\tilde{F}_1(\alpha; \beta; z) = \frac{{}_1F_1(\alpha; \beta; z)}{\Gamma(\beta)}, \quad (\text{S36})$$

where by ${}_1F_1(\alpha; \beta; z)$ we denote the confluent hypergeometric function.

S5.2 n=2

$$\begin{aligned} \mu_0 &= \mu_0(a, b, K, \epsilon, 2) = \frac{K^{(a+1)}}{2\Gamma\left(\frac{1}{2}(a - a\epsilon)\right)} \left\{ \right. \\ &+ \frac{2}{K} \left(\frac{b}{K} \right)^{a\epsilon} \Gamma(a\epsilon) \Gamma\left(\frac{1}{2}(a - a\epsilon)\right) {}_1F_2 \left(\frac{a}{2} - \frac{a\epsilon}{2}; \frac{1}{2} - \frac{a\epsilon}{2}, 1 - \frac{a\epsilon}{2}; -\frac{K^2}{4b^2} \right) \\ &+ \frac{1}{K} \Gamma\left(\frac{a}{2}\right) \Gamma\left(-\frac{a\epsilon}{2}\right) {}_1F_2 \left(\frac{a}{2}; \frac{1}{2}, \frac{a\epsilon}{2} + 1; -\frac{K^2}{4b^2} \right) \\ &\left. - \frac{1}{b} \Gamma\left(\frac{a+1}{2}\right) \Gamma\left(-\frac{a\epsilon}{2} - \frac{1}{2}\right) {}_1F_2 \left(\frac{a}{2} + \frac{1}{2}; \frac{3}{2}, \frac{a\epsilon}{2} + \frac{3}{2}; -\frac{K^2}{4b^2} \right) \right\}. \end{aligned} \quad (\text{S37})$$

In above, ${}_1F_2(\alpha; \beta_1, \beta_2; z)$ is the generalized hypergeometric function.

S5.3 n=-1

$$\begin{aligned} \mu_0(a, b, K, \epsilon, -1) &= \frac{K^a \Gamma(-a) \Gamma(a\epsilon) {}_1F_1\left(a\epsilon; a+1; \frac{K}{b}\right)}{\Gamma(a(\epsilon-1))} \\ &+ \left(\frac{1}{b}\right)^{-a} \Gamma(a) {}_1F_1\left(a(\epsilon-1); 1-a; \frac{K}{b}\right), \end{aligned} \quad (\text{S38})$$

${}_1F_1(\alpha; \beta; z)$ is the confluent hypergeometric function.

S5.4 n=-2

$$\begin{aligned} \mu_0 &= \mu_0(a, b, K, \epsilon, -2) = \frac{\left(\frac{1}{b}\right)^{1-a} K^{a+1}}{2\Gamma\left(\frac{1}{2}(a-a\epsilon)\right)} \left\{ \right. \\ &+ \frac{1}{K} \left(\frac{1}{b}\right)^{a-1} \Gamma\left(-\frac{a}{2}\right) \Gamma\left(\frac{a\epsilon}{2}\right) {}_1F_2\left(\frac{a\epsilon}{2}; \frac{1}{2}, \frac{a}{2}+1; -\frac{K^2}{4b^2}\right) \\ &- \left(\frac{1}{b}\right)^a \Gamma\left(-\frac{a}{2}-\frac{1}{2}\right) \Gamma\left(\frac{1}{2}(a\epsilon+1)\right) {}_1F_2\left(\frac{a\epsilon}{2}+\frac{1}{2}; \frac{3}{2}, \frac{a}{2}+\frac{3}{2}; -\frac{K^2}{4b^2}\right) \\ &+ \left. \frac{2b}{K^{a+1}} \Gamma(a) \Gamma\left(\frac{1}{2}a(\epsilon-1)\right) {}_1F_2\left(\frac{a\epsilon}{2}-\frac{a}{2}; \frac{1}{2}-\frac{a}{2}, 1-\frac{a}{2}; -\frac{K^2}{4b^2}\right) \right\}. \end{aligned} \quad (\text{S39})$$

In above, ${}_1F_2(\alpha; \beta_1, \beta_2; z)$ denotes the generalized hypergeometric function.

S6 Simulations

We carried out the simulations based on the Gillespie algorithm [6], using the custom program Mesokin, available at <http://pepe.ichf.edu.pl/tabaka/mesokin.tar.gz>. Since the trajectories generated by the simulation are ergodic, we used for simplicity single long trajectories ($t = 2 \times 10^9$) to create histograms. In Tables S3-S8 we present the values of parameters used in our simulations to generate the data shown in Figs. 5, S1-S19. Input files for the simulations are attached as a supplementary file.

S7 Time scales and parameters of the system

The typical abundance of TFs in *E. coli* is between several tens and several hundreds of molecules per cell [7], whereas the levels of proteins, in general, are typically higher, with top 17% of highly abundant proteins, whose numbers are greater than 2050 copies per cell [8]. In our simulations, the mean TF number at full activation, $\alpha\beta = 125$, the mean target protein number at full activation, $ab = 1250$.

In order to fulfil the assumption of Hill kinetics, the promoter fluctuations driven by TF binding/unbinding should be sufficiently fast. On the other hand, to keep the model realistic, we

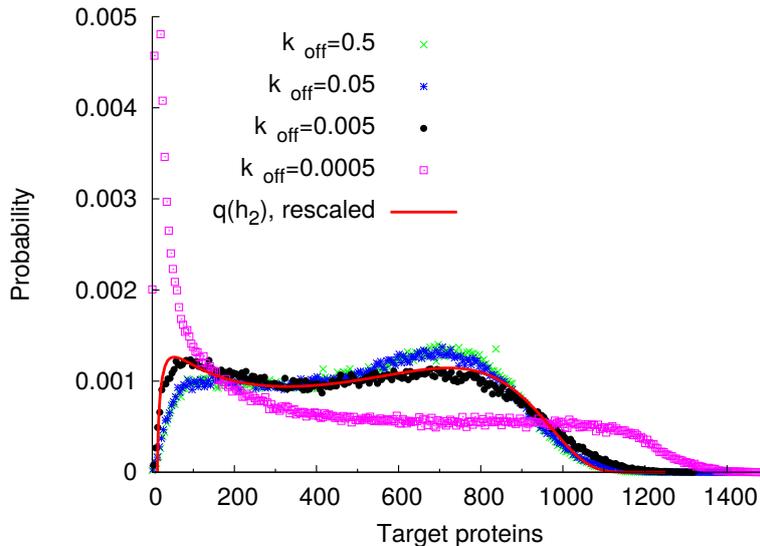


Figure S1: Test of Hill kinetics of TF binding/unbinding. The rate of promoter fluctuations is here varied by varying the values of $k_{off} \equiv k_{off,1}^r = k_{off,2}^r = k_{off,1}^t = k_{off,2}^t$ and $k_{on,1}^r = 10^{-3} k_{on,2}^r = k_{on,1}^t = 10^{-3} k_{on,2}^t$ such that $K_1 = K_2 = 70$. The values corresponding to $k_{off} = 0.005$ (black points) have been chosen for simulations described in the main text. The values of other parameters are shown in Table S4.

have chosen the promoter fluctuations to be as slow as possible without exceeding the range of validity of our model. In Fig. S1, we tested the protein number distributions corresponding to different values of the $k_{i,off}$ constant ($k_{i,on}$ being fixed). For the simulations described in the main text, $k_{i,off} = 0.005$ has been chosen.

In Fig. S2, we compare the time scales of the simulated processes for $K_1 = K_2 = 70$, $n = m = -2$ (see also Table S5 for the parameter values used in the simulation for this case). Consistently with typical biological time scales, target mRNA fluctuations are faster than target protein fluctuations. For our model to be valid, the total number of TFs fluctuates more slowly than the on-off fluctuations of the downstream promoter, such that the promoter experiences an approximately constant level of TFs, compared to the time scale of its own transcription. Interestingly enough, the levels of active TFs fluctuate much faster, constituting the fastest time scale in the system, but nevertheless, the simulation results presented in this paper show a very good agreement with the predictions of the theoretical model. Thus, the assumption of Hill kinetics as the fastest time scale in the system has been slightly relaxed without detriment to the validity of the model.

The parameters of the downstream gene in our simulations fulfil the assumption under which the non-regulated gene produces a gamma distribution of proteins [2, 4, 5]: The protein lifetime is longer than the mRNA lifetime. We checked the expression of such a non-regulated gene, at three different transcription levels (Fig. S3), and indeed, the distributions of the protein number

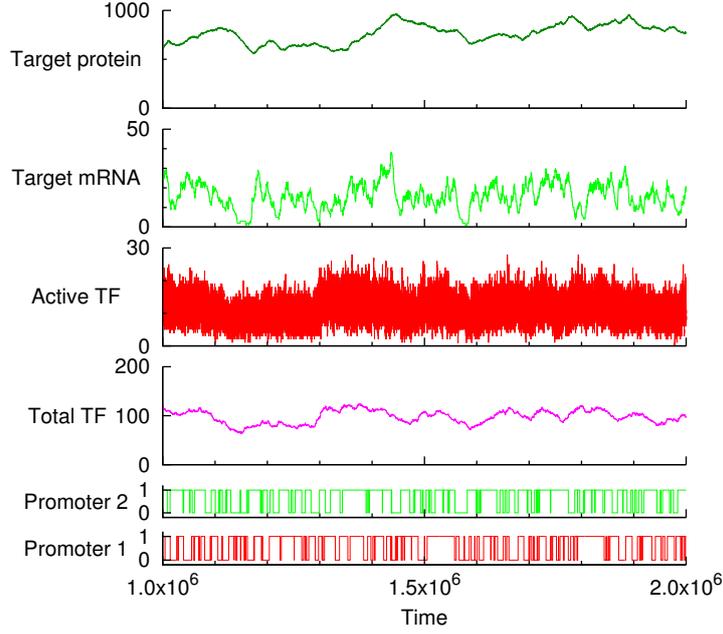


Figure S2: Comparison of the time scales of the simulated biochemical reactions. Parameter values are shown in Table S5.

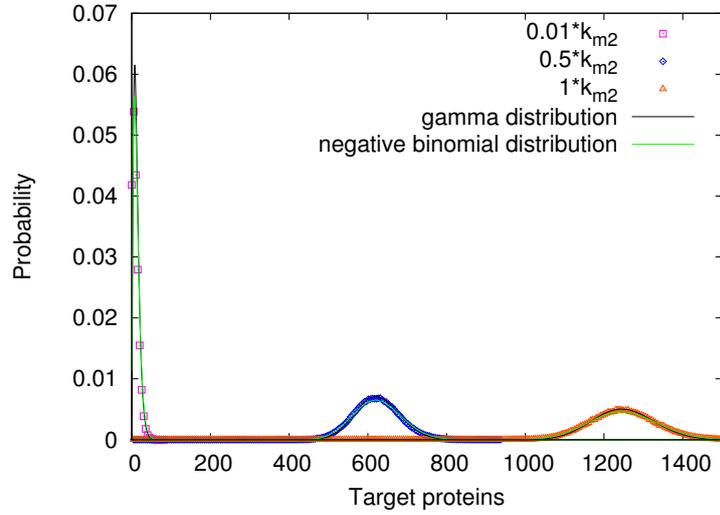


Figure S3: Protein number distributions of the unregulated downstream gene are well approximated by the gamma or negative binomial distributions. The three distributions shown correspond to the following transcription rates: a) leaky level: $0.01k_{m2} = 2.5 \times 10^{-5}$, b) 1/2 of the maximum level: $0.5k_{m2} = 1.25 \times 10^{-3}$, c) maximum level: $k_{m2} = 2.5 \times 10^{-3}$. The values of other parameters are shown in Table S6.

P are well modelled by gamma distributions $\gamma(P; a, b)$ or negative binomial distributions

$$g(P; a, b) = \frac{\Gamma(a + P)}{\Gamma(P + 1)\Gamma(a)} \left(\frac{b}{1 + b}\right)^P \left(1 - \frac{b}{1 + b}\right)^a. \quad (\text{S40})$$

S8 Response of the cascade in absence of self-regulation

In order to compare the behaviour of the system with a non-regulated and self-regulated Gene 1, we simulated three different levels of constitutive transcription of the non-regulated Gene 1, corresponding to the mean burst frequencies: $\alpha\epsilon_1$, $\alpha/2$, and α . The case of the lowest expression level, $\alpha\epsilon_1$, mimics the removal of self-regulation (i.e., the rate of binding of TF to its own operator $k_{on} = 0$), however, we also test other expression levels because the removal of feedback may perhaps change the properties of the promoter so that it could influence the rate of transcription from its gene. In all the cases discussed below, the Gene 1, as non-regulated, does not respond to signal and the number of TFs in cells is gamma-distributed (Eq. S11). At the same time, the downstream genes have graded responses, distribution of their transcription rates being given by $q_0(h_2) = q_0(h_2; \alpha, \beta, K_2)$ (Eq. S14).

The response of the cascade with a removed feedback to an increasing signal is shown in Figs. S4–S6. Note that the parameter values used for our case study have been chosen such that the cascade with no feedback has a graded response. The sensitivity of the response corresponds to different ranges of signal levels, depending on the promoter-TF affinity (early or late responses). (In general, a cascade without feedback can also have a binary response [3].) The distributions of the total number of TFs are constant at any signal level because the signal does not affect TF production, it only defines the active TF fraction.

In Fig. S4, the non-regulated Gene 1 has a very weak constitutive transcription corresponding to $\alpha\epsilon_1$, i.e. to that of the self-regulated Gene 1 at its fully inactive, leaky state. In this case, the high-affinity promoter responds sensitively. This is also depicted by the intersections of the transfer function H_2 and the corresponding dashed lines, as the limiting case of the geometric construction in Fig. 4 in the main manuscript. Fig. S5 shows the behaviour of the gene cascade with the non-regulated Gene 1 set at an intermediate level corresponding to $\alpha/2$. Here, the intermediate-affinity promoter responds sensitively. In Fig. S6 we present the response of the cascade in the case where the transcription of the non-regulated Gene 1 is set at the maximal level corresponding to α , i.e. to that of the self-regulated Gene 1 at its fully active state. In this case, the low-affinity promoter responds sensitively. Again, the maxima of these distributions are shown as the limiting cases of the construction in Fig. 4 in the main manuscript, by the intersections of H_2 and the corresponding dashed lines.

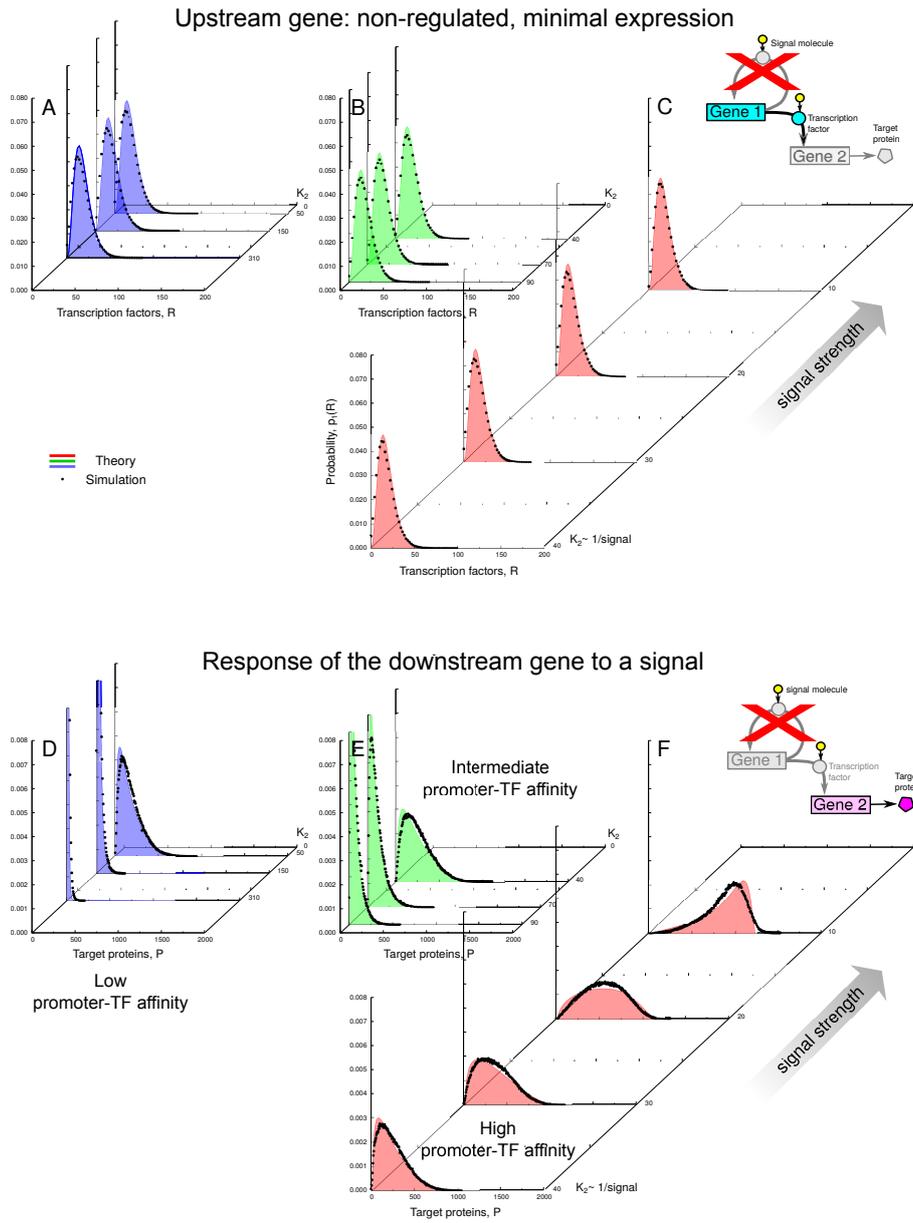


Figure S4: Gene 1, as non-regulated, does not respond to signal and the downstream genes have then graded responses. Gene 1 has a very weak constitutive transcription corresponding to mean burst frequency $\alpha\epsilon_1$. Parameters: $n = -2$, $m = -2$, $\alpha = 25$, $\beta = 5$, $a = 250$, $b = 5$, $\epsilon_1 = 0.15$, $\epsilon_2 = 0.01$. Black dots: simulated protein number distributions. Filled curves: theoretical distributions, $p_0(R)$ in (A,B,C), and $q_0(h_2/[ab]; \alpha\epsilon_1, \beta)/[ab]$ in (D,E,F). The values of other parameters are shown in Table S8.

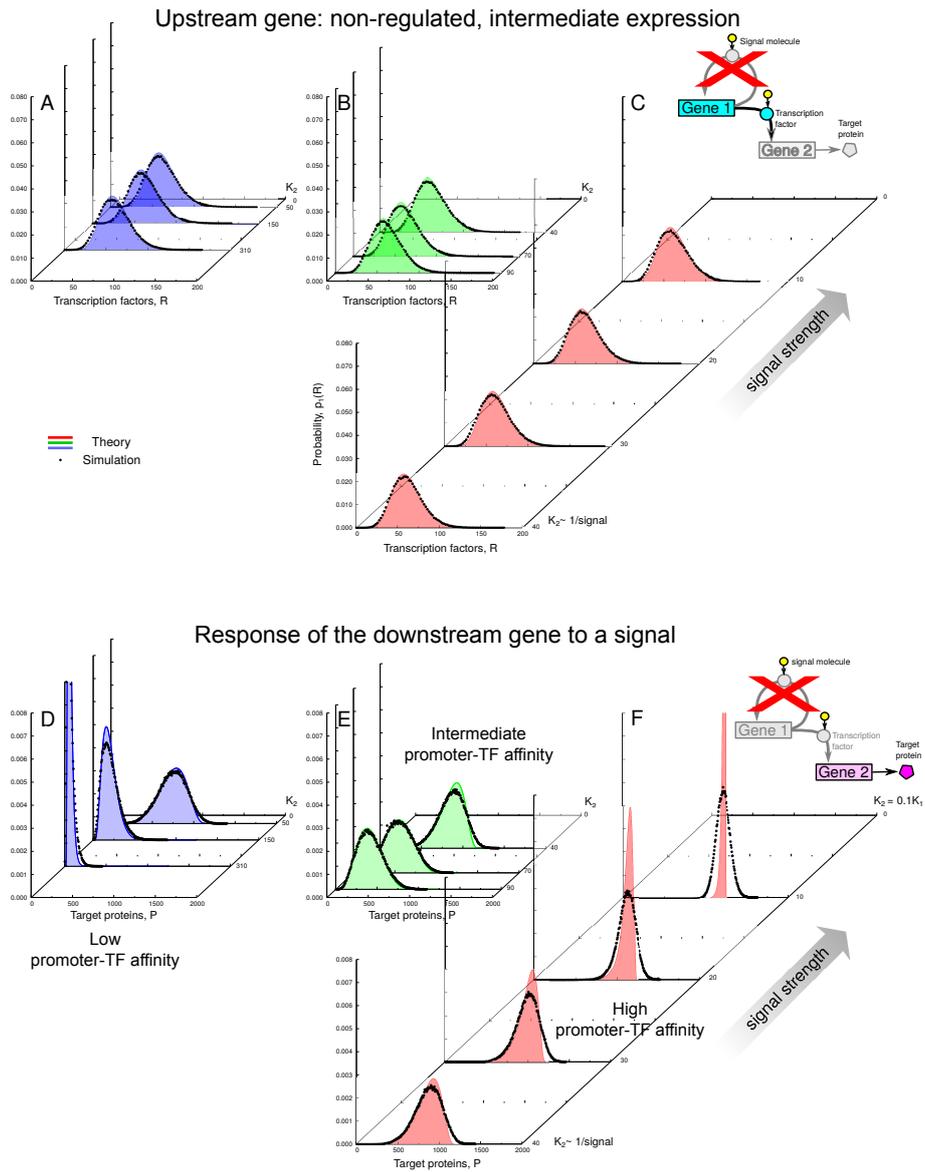


Figure S5: Gene 1, as non-regulated, does not respond to signal and the downstream genes have then graded responses. Gene 1 has an intermediate constitutive transcription corresponding to mean burst frequency $\alpha/2$. Parameters: $n = -2$, $m = -2$, $\alpha = 25$, $\beta = 5$, $a = 250$, $b = 5$, $\epsilon_1 = 0.15$, $\epsilon_2 = 0.01$. Black dots: simulated protein number distributions. Filled curves: theoretical distributions, $p_0(R)$ in (A,B,C), and $q_0(h_2/[ab]; \alpha/2, \beta)/[ab]$ in (D,E,F). The values of other parameters are shown in Table S8.

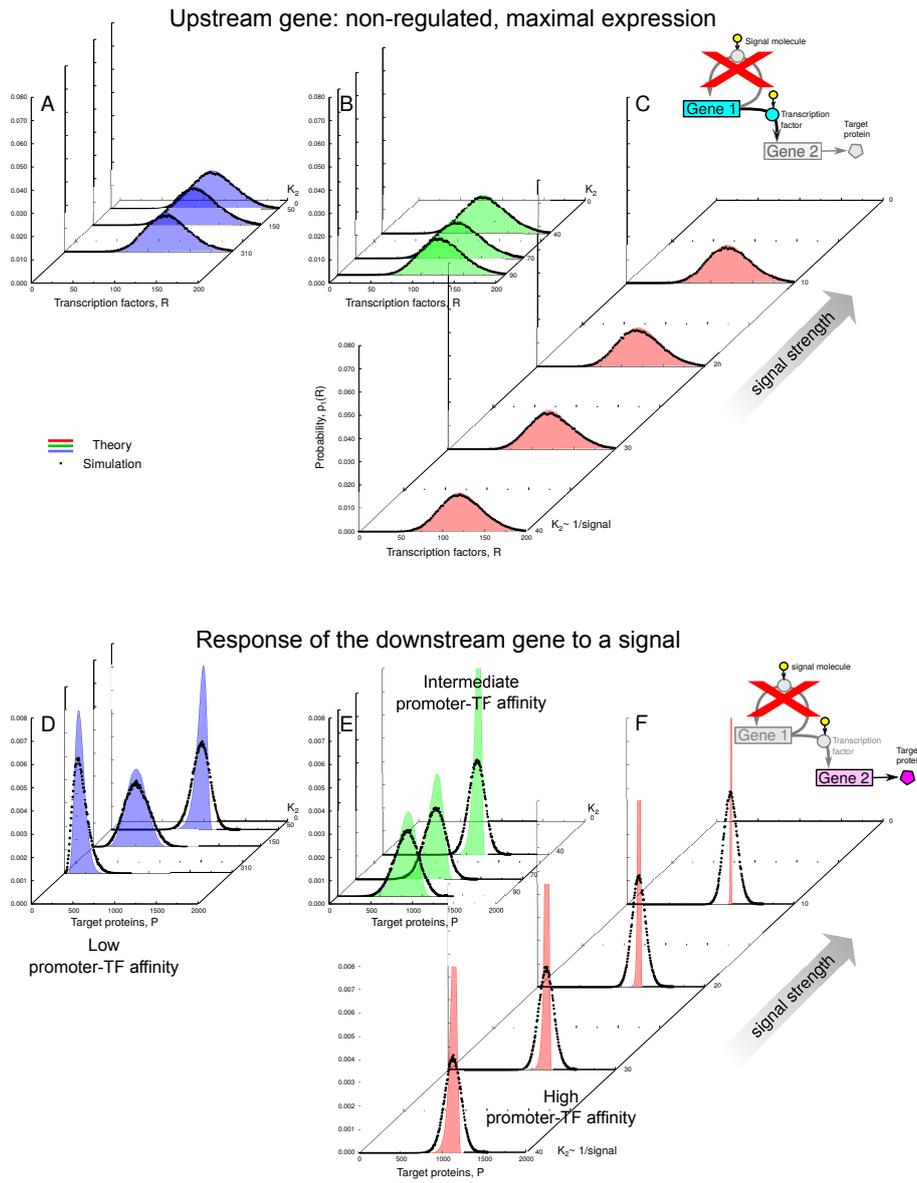


Figure S6: Gene 1, as non-regulated, does not respond to signal and the downstream genes have then graded responses. Gene 1 has a maximal constitutive transcription corresponding to mean burst frequency α . Parameters: $n = -2$, $m = -2$, $\alpha = 25$, $\beta = 5$, $a = 250$, $b = 5$, $\epsilon_1 = 0.15$, $\epsilon_2 = 0.01$. Black dots: simulated protein number distributions. Filled curves: theoretical distributions, $p_0(R)$ in (A,B,C), and $q_0(h_2/[ab]; \alpha, \beta)/[ab]$ in (D,E,F). The values of other parameters are shown in Table S8.

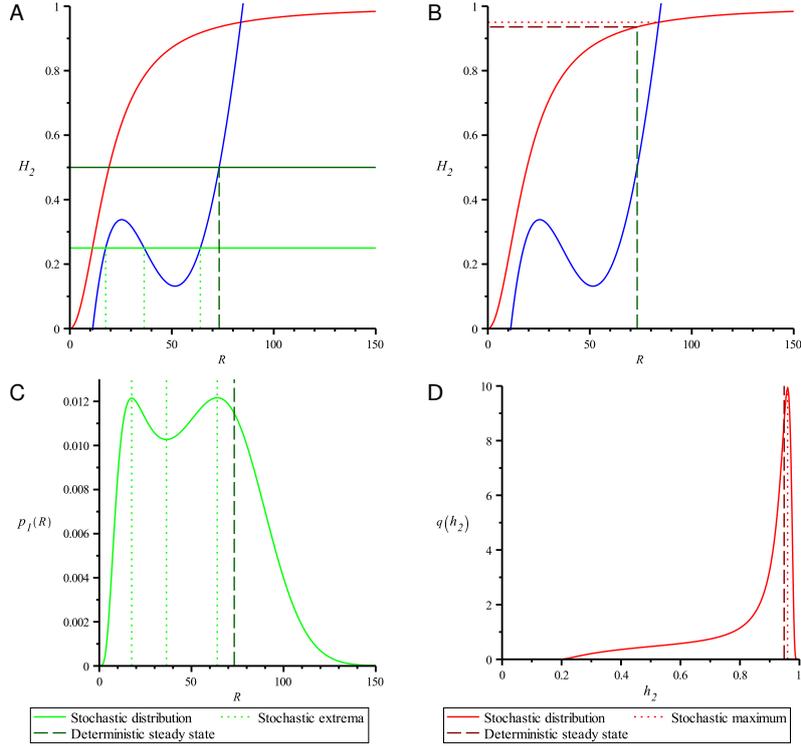


Figure S7: A, B: Geometric construction as in Fig.2B1. Here, it is shown in two panels for the sake of clarity. C, D: Distributions as in Fig.2B1, compared with the deterministic steady states. Dotted lines: positions of the distribution extrema. Dashed lines: positions of deterministic steady states. In A, the intersections of $L(R)$ (blue) and the $\frac{m+1}{2m}$ line (green) indicate the extrema of the distribution of TF levels. The intersection of $L(R)$ with the $1/2$ line (dark green) indicates the deterministic steady state (Eq. S31). In B, the intersections of $L(R)$ and $H_2(R)$ (red) indicate the extrema of the transcription rate distribution.

S9 Comparison between deterministic and stochastic model for Figs. 2B1 and 2B2 in main text

In this section, we take a closer look at the relationship between bimodality and bistability in the stochastic and deterministic versions of our model, using the examples shown in Figs. 2B1 and 2B2 in the main text. We show that the ranges of stochastic bimodality and deterministic bistability of TF levels do not overlap in Fig. 2B1, and, although they are close, the behaviour of the system differs in these ranges: When the studied example system from Fig. 2B1 has a bimodal input and a unimodal output, its deterministic counterpart has a monostable input and a monostable output. If that same system had a bistable input and a bistable output, its stochastic version would have a unimodal input and a unimodal output. In Fig.2B2, a unimodal stochastic input gives rise to a bimodal stochastic output, but bistable input and output are impossible in the deterministic version of this model in any range of signal intensity.

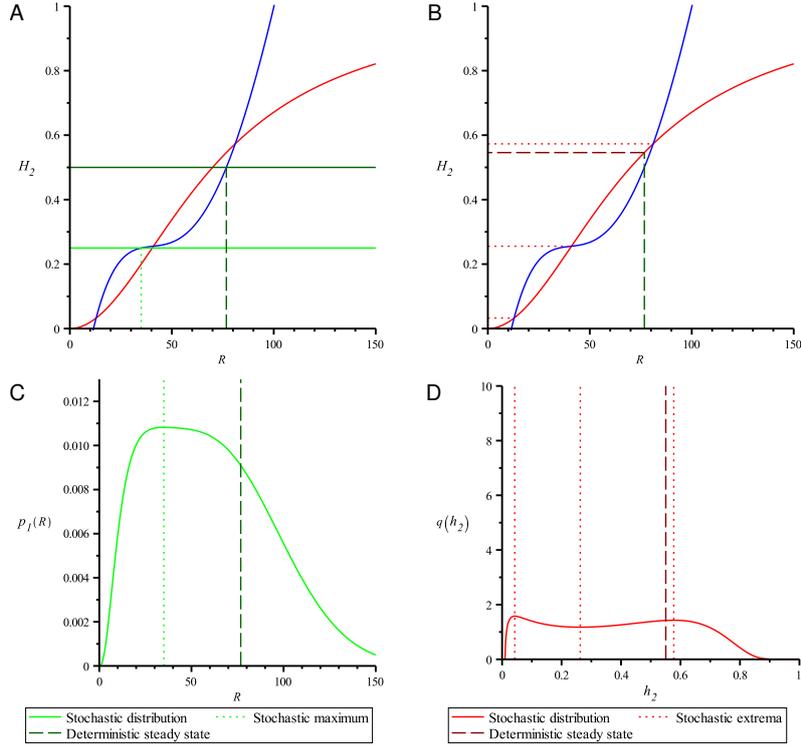


Figure S8: A, B: Geometric construction as in Fig. 2B2. Here, it is shown in two panels for the sake of clarity. C, D: Distributions as in Fig. 2B2, compared with the deterministic steady states. Dotted lines: positions of the distribution extrema. Dashed lines: positions of deterministic steady states. In A, the intersection of $L(R)$ (blue) and the $(m+1)/(2m)$ line (green) indicates the extrema of the distribution of TF levels. The intersection of $L(R)$ with the $1/2$ line (dark green) indicates the deterministic steady state (Eq. S31).

In Figs. S7 and S8 we show what the Figs. 2B1 and 2B2 would look like if the model were deterministic.

For the stochastic system as in Fig. 2B1, bimodal input gives rise to a unimodal output, whereas in the deterministic model the input is monostable and, consequently, so is the output (Fig. S7). Note that the deterministic input differs much from its stochastic counterpart, but the outputs are similar in the deterministic and stochastic model. This is because the downstream gene regulation is overdriven, i. e., it responds to the input in the region where the transfer function h_2 is almost saturated.

For the stochastic system as in Fig. 2B2, a unimodal input gives rise to a bimodal output, but again, the deterministic input is monostable and it maps into a monostable output (Fig. S8). Here, the deterministic and stochastic versions differ considerably for both input and output. It is easy to see that bistable input and output are impossible in this system in any range of signal intensity.

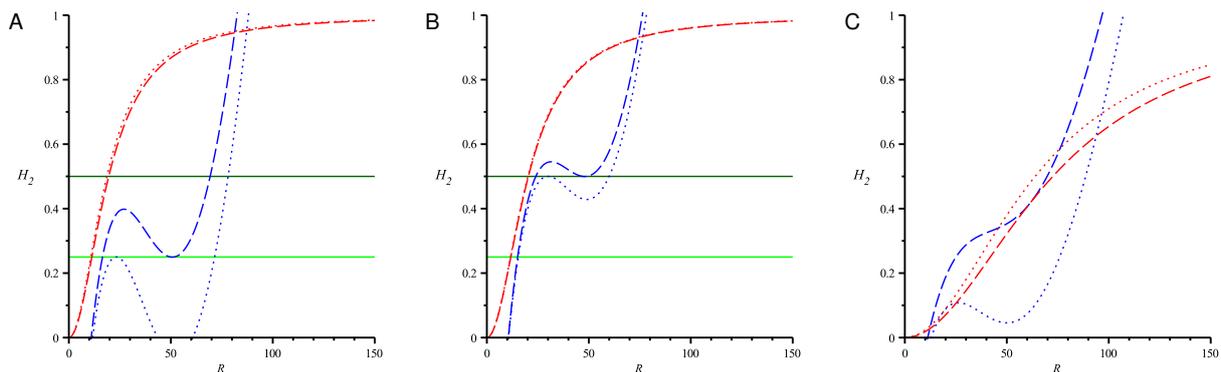


Figure S9: Geometric constructions showing the ranges of bimodality/bistability in Fig. 2B. We vary the signal parameters K_1 and K_2 and the other parameters are same as in Fig. 2B1 and B2. A: Stochastic bimodality range for the distribution of TF levels in 2B1: $K_1 \approx 62.2 \dots 65.2$ and $K_2 \approx 18.7 \dots 19.6$. B: Bistability range for TFs for the deterministic version of the model in 2B1: $K_1 \approx 67.1 \dots 67.9$ and $K_2 \approx 20.1 \dots 20.4$. C: Stochastic bimodality range for the distribution of transcription rates in Fig.2B2: $K_1 = K_2 \approx 63.9 \dots 72.6$. Dotted lines: upper limit; dashed lines: lower limit. Red: $H_2(R)$; blue: $L(R)$; green: $(m+1)/(2m)$ level; dark green: $1/2$ level.

In Fig. S9 we compare the ranges of the signal parameter values in which bimodality or bistability occurs in the stochastic or deterministic versions of the model. These ranges do not overlap. In Fig.2B1, the range of stochastic bimodality of the distribution of TF levels was $K_1 \approx 62.2 \dots 65.2$ and $K_2 \approx 18.7 \dots 19.6$ (Fig. S9A; note that K_1 is proportional to K_2). The corresponding range of deterministic bistability is different, although quite close: $K_1 \approx 67.1 \dots 67.9$ and $K_2 \approx 20.1 \dots 20.4$ (Fig. S9B). For completeness, in Fig. S9C, we show the range of stochastic bimodality of the distribution of transcription rates for Fig.2B2: for parameter values used in that figure, the range was $K_1 = K_2 \approx 63.9 \dots 72.6$.

Using the above information, we checked what the Fig.2B1 would look like if the signal parameters were in the range of deterministic bistability. In that range, there is no stochastic bimodality. We have chosen the values of signal parameters $K_1 = 67.4$ and $K_2 = 20.22$. Fig. S10 demonstrates that in the range of deterministic bistability the behaviour of the system is different than it was in the range of stochastic bimodality: The relationship between deterministic steady states and stochastic extrema is different than it was in Fig. S9A,C. Here, two stable steady states of TF levels will lead to two stable steady states of transcription rates according to the deterministic model. The corresponding stochastic model predicts, on the other hand, a unimodal input and a unimodal output.

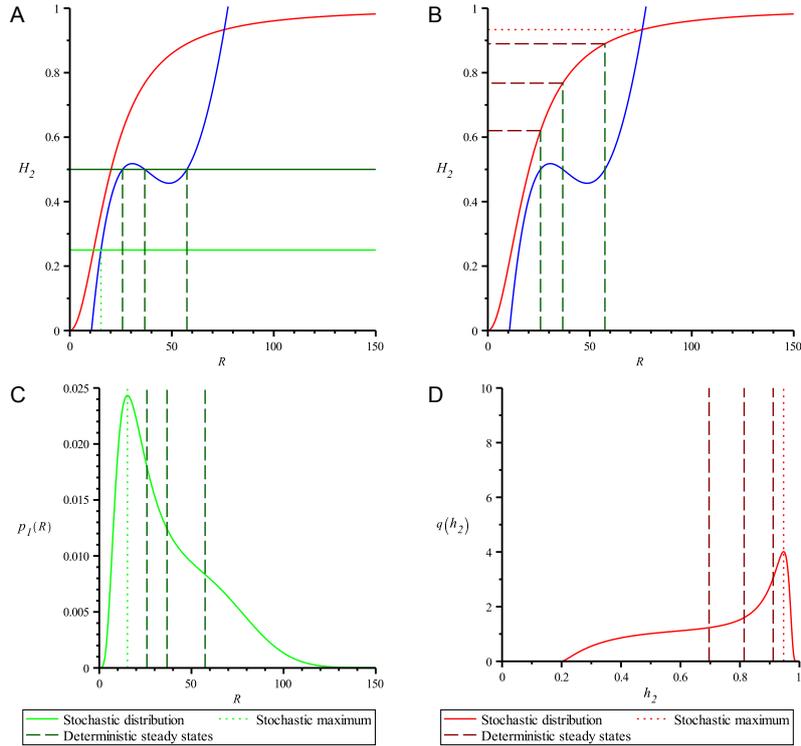


Figure S10: A, B: Geometric construction as in Fig.2B1, but with $K_1 = 67.4$ and $K_2 = 20.22$, i. e., in the range of bistability of the deterministic version of the model. Here, the construction is shown in two panels for the sake of clarity. C, D: Distributions of TF levels (B) and transcription rates (C), compared with the deterministic steady states. Dotted lines: positions of the distribution extrema. Dashed lines: positions of deterministic steady states. In A, the intersection of $L(R)$ (blue) and the $(m+1)/(2m)$ line (green) indicates the maximum of the distribution of TF levels. The intersection of $L(R)$ with the $1/2$ line (dark green) indicates the deterministic steady states (Eq. S31).

S10 Dependence of distribution shape on leakage

In the examples shown in Fig.2B in the main text, we assumed a significant amount of promoter leakage, at least in the regulatory promoter ($\epsilon_1 = 0.1$, $\epsilon_2 = 0.2$ in Fig.2B1, $\epsilon_1 = 0.15$, $\epsilon_2 = 0.01$ in Fig.2B1). The reason for this choice was twofold: 1. Considerable basal transcription is common in wild-type genes [9, 10], see also [1]. 2. Our method of geometric construction suggests that a higher leakage allows for a richer spectrum of behaviours of the gene cascade with a positively self-regulated upstream gene. Namely, the “bimodal input to unimodal output” or “unimodal input to bimodal output” behaviours are obtained more easily when the regulatory promoter is more leaky. We show this on the examples below:

In Fig. S11 we plot the geometric construction, input and output distributions for a system with the same parameters as in Fig.2B1, except for the leakage, which is much lower: $\epsilon_1 = 0.02$, $\epsilon_2 = 0.02$. Here, the leakiness of the regulatory promoter is so low that it is more difficult to

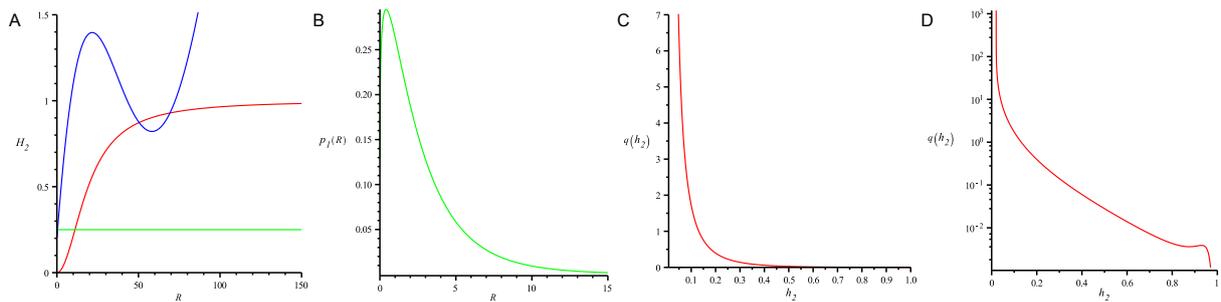


Figure S11: Behaviour of the system with the same parameters as in Fig.2B1 in the main text, except for the leakage, which is much lower: $\epsilon_1 = 0.02$, $\epsilon_2 = 0.02$. The “bimodal input to unimodal output” behaviour is not present here: The input is unimodal and the output is visually unimodal as well (although it is mathematically bimodal). A: Geometric construction. Blue: $L(R)$; red: $H_2(R)$; green: $(m + 1)/(2m)$. B: Distribution of TF levels. C: Distribution of transcription rates. D: Distribution of transcription rates, log scale, to show an extremely small bimodality.

obtain the situation when the expression of the upstream gene is bimodal, and the expression of the downstream gene is unimodal. This is because it is difficult to obtain the $L(R)$ curve that intersects the $(m + 1)/(2m)$ line twice, and, at the same time, to place the $L(R)$ curve below the $H_2(R)$ curve in such a way that $L(R)$ and $H_2(R)$ have just one intersection. In this example, at low leakage in both genes, we get a unimodal input and a practically unimodal output, although it is mathematically bimodal.

It is, however, still possible to observe a behaviour at low leakage that is very close to the “bimodal input to unimodal output” behaviour (Fig. S12). At a certain choice of parameters ($n = m - 2$, $\alpha = 150$, $\beta = 2$, $K_1 = 130$, $K_2 = 1$, $\epsilon_1 = 0.02$, $\epsilon_2 = 0.02$), we can get a strongly bimodal input leading to a practically unimodal output (although here the output distribution is mathematically bimodal).

In Fig. S13 we plot the geometric construction, and the input and output distributions for a system with same parameters as in Fig.2B2, where the “unimodal input to bimodal output” behaviour was shown, but we changed the leakage to $\epsilon_1 = 0.02$ and $\epsilon_2 = 0.02$, as above. Here, at a lower leakiness of the regulatory promoter, the $L(R)$ curve is moved upwards, so that it cannot intersect $H_2(R)$ twice. As a result, we obtain a unimodal input and a unimodal output.

Again, it is possible to find a set of parameters for which the gene system shows the “unimodal input to bimodal output” behaviour at low leakage in the regulatory promoter. At $n = m = -2$, $\alpha = 25$, $\beta = 5$, $K_1 = 70$, $K_2 = 1$, $\epsilon_1 = 0.02$, $\epsilon_2 = 0.02$, we have a unimodal input and a strongly bimodal output (Fig. S14).

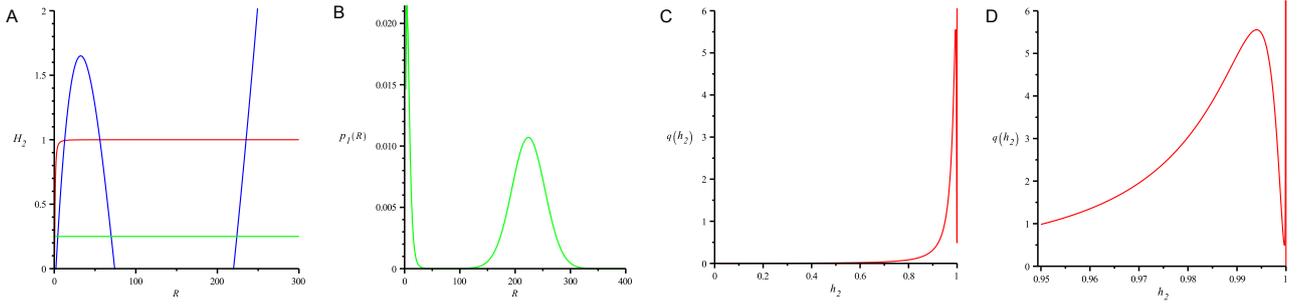


Figure S12: Behaviour at low leakage that is very close to the “bimodal input to unimodal output” behaviour. Parameters: $n = m - 2$, $\alpha = 150$, $\beta = 2$, $K_1 = 130$, $K_2 = 1$, $\epsilon_1 = 0.02$, $\epsilon_2 = 0.02$. A: Geometric construction. Blue: $L(R)$; red: $H_2(R)$; green: $(m + 1)/(2m)$. B: Distribution of TF levels. C: Distribution of transcription rates. D: Distribution of transcription rates, zoomed, to show an extremely small bimodality.

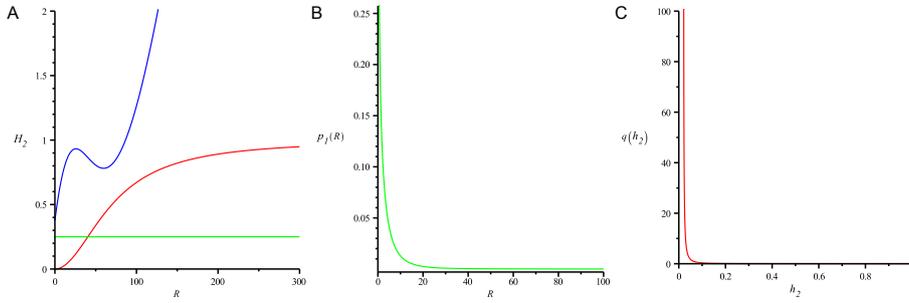


Figure S13: Behaviour of the system with the same parameters as in Fig.2B2 in the main text, except for the leakage, which is much lower: $\epsilon_1 = 0.02$, $\epsilon_2 = 0.02$. The “unimodal input to bimodal output” behaviour is not present here: The input and output are unimodal. A: Geometric construction. Blue: $L(R)$; red: $H_2(R)$; green: $(m + 1)/(2m)$. B: Distribution of TF levels. C: Distribution of transcription rates.

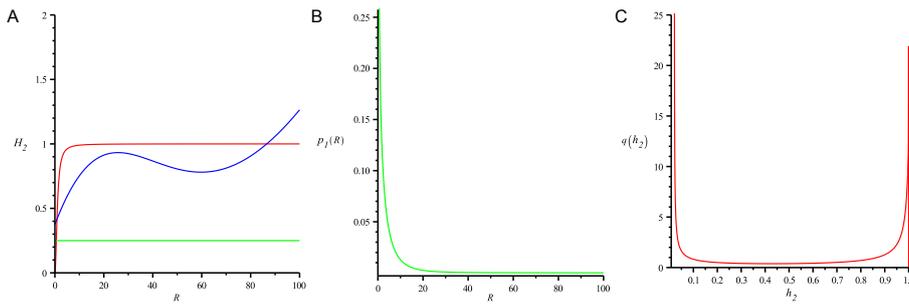


Figure S14: “Unimodal input to bimodal output” behaviour at low leakage. Parameters: $n = m = -2$, $\alpha = 25$, $\beta = 5$, $K_1 = 70$, $K_2 = 1$, $\epsilon_1 = 0.02$, $\epsilon_2 = 0.02$. A: Geometric construction. Blue: $L(R)$; red: $H_2(R)$; green: $(m + 1)/(2m)$. B: Distribution of TF levels. C: Distribution of transcription rates.

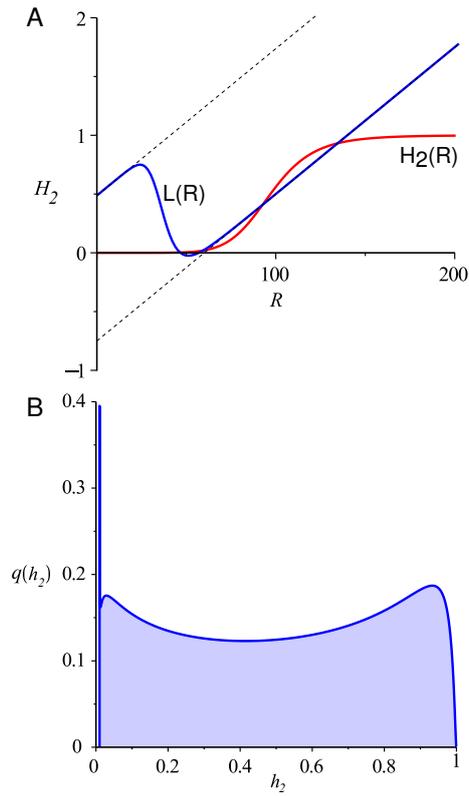


Figure S15: Example of a trimodal distribution $q(h_2)$, parameters: $n = -8$, $m = -8$, $\alpha = 20$, $\beta = 5$, $\epsilon_1 = 0.01$, $\epsilon_2 = 0.01$, $K_1 = 37$, $K_2 = 97$. (A) Geometric construction. (B) Distribution of transcription rates, $q(h_2)$.

S11 Example of a trimodal transcription rate distribution

In Fig. S15 we show an example of an exotic, trimodal distribution $q(h_2)$, for parameters: $n = -8$, $m = -8$, $\alpha = 20$, $\beta = 5$, $\epsilon_1 = 0.01$, $\epsilon_2 = 0.01$, $K_1 = 37$, $K_2 = 97$.

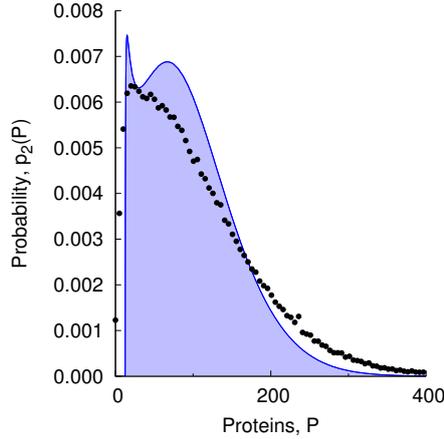


Figure S16: Zoomed view of Fig. 5D from the main manuscript: The slight bimodality predicted by the model for $K_1 = 62$ and $K_2 = 5K_1$ is blurred by intrinsic noise. The values of other parameters are shown in Table S3.

S12 Zoomed view of Fig. 5D from the main manuscript

In Fig. S16 we present a zoomed view of Fig. 5D from the main manuscript. It shows that the slight bimodality predicted by the model for $K_1 = 62$ and $K_2 = 5K_1$ is blurred by intrinsic noise.

S13 Coefficients of variation and Fano factors of the distributions

In Fig. S17 we compare the coefficient of variation (CV) and Fano factor (FF) of the distributions shown in Fig. 5 in the main manuscript (parameters: $n = m = -2$, $\alpha = 25$, $\beta = 5$, $\epsilon_1 = 0.15$, $\epsilon_2 = 0.01$, $K_2 = 5K_1$, $K_1, 0.1K_1$). CV is the ratio of standard deviation to mean of the distribution. FF is the ratio of variance to mean of the distribution. CV (Fig. S17A), as a dimensionless quantity, is more suitable for such comparison. In particular, FF (Fig. S17B) for distributions of transcription rates $q(h_2)$ is much lower than FF for the distribution of protein numbers $p_1(R)$ because the values of the relative transcription rate h_2 are less than 1, whereas the TF numbers are much greater than 1.

Interestingly enough, the values of CV and FF are not very informative as to the apparent width or the bimodality of a distribution, when we look at the whole possible range of distributions driven by the external signal, i.e., the TF numbers from $\alpha\beta\epsilon_1$ to $\alpha\beta$, or the relative transcription rates of the downstream genes from ϵ_2 to 1. In Figs. S17C-F we show distributions for arbitrarily chosen values of parameters K_1 , such that their shapes are *visually* the widest possible for a given set of the remaining parameters. (To date, there is no widely accepted universal statistical measure of bimodality, and for this reason, we decided to make the arbitrary choice, just for the purpose of visualisation.) The corresponding values of FF and CV

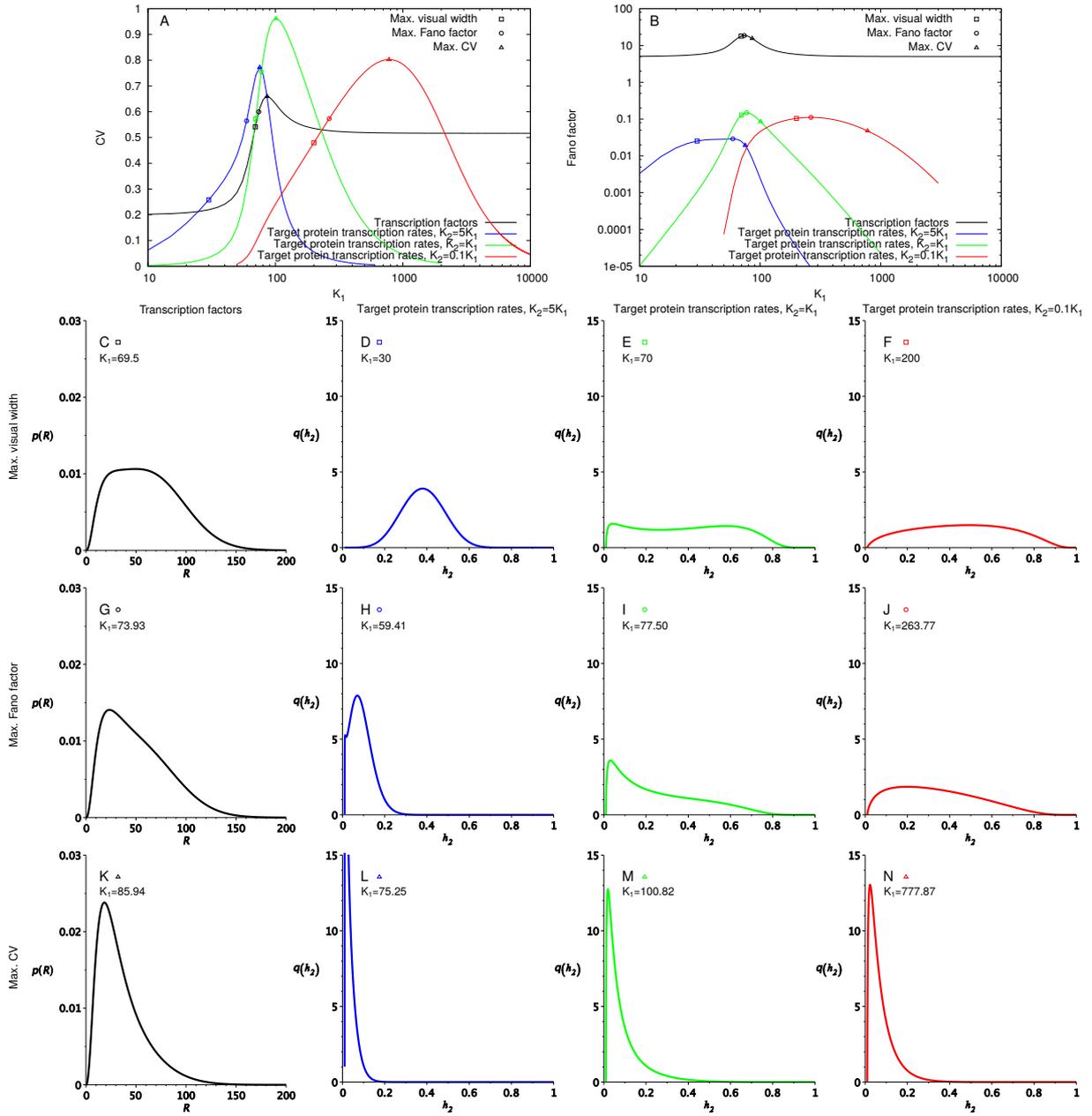


Figure S17: Coefficient of variation (CV) and Fano factor (FF) of the distributions shown in Fig. 5 in the main manuscript (parameters: $n = m = -2$, $\alpha = 25$, $\beta = 5 \epsilon_1 = 0.15$, $\epsilon_2 = 0.01$, $K_2 = 5K_1, K_1, 0.1K_1$). A and B: CV and FF of the distribution of the number of TFs, $p_1(R)$ (black), compared with the CV and FF of the distributions of transcription rates of the downstream gene, $q(h_2; K_1, K_2)$ (blue, green, red). Squares denote the values of CV and FF for arbitrarily chosen values of parameters K_1 , such that the distributions are *visually* the widest possible for a given set of the remaining parameters. Circles denote the values of CV and FF for the maximal FF. Triangles denote the values of CV and FF for the maximal CV. C-F: Corresponding visually widest distributions. G-J: Distributions having the maximal FF. K-N: Distributions having the maximal CV.

for these distributions are marked by squares in Figs. S17A and B. It turns out that these are not maximal values of FF or CV. The shapes of the distributions corresponding to maximal FF (Fig. S17G-J) and maximal CV (Fig. S17K-M) appear narrower and usually are not bimodal.

It can, however, be noticed that out of the three downstream genes whose promoter affinities to TF are lower, same or higher than in the upstream gene, the highest CV and FF is reached by the gene whose promoter is equally sensitive to TF as the regulatory gene's promoter. In other words, the overlap of sensitivity regions of their transfer functions leads to the increased noise.

S14 Precision of response of the target gene to external signal

The relative distribution width, being a measure of (im)precision of gene's response to external signal, is given by the formula

$$W(K_1) = \frac{\sigma_{K_1}}{|\mu_\infty - \mu_0|}, \quad (\text{S41})$$

where σ_{K_1} is the standard deviation of a distribution at the signal level measured by K_1 , μ_∞ is the mean of the distribution at $K_1 = \infty$, and μ_0 is the mean of the distribution at $K_1 = 0$. Numerical calculation of means and standard deviations of the TF number distribution $p_1(R)$ (Eq. S15) and transcription rate distribution of the target gene $q(h_2)$ (Eq. S16) can be easily performed (we used Maple software). However, direct calculation of means and standard deviations of the target protein distribution $p_2(P)$ in the form of intrinsic noise representation (Eq. S8) is more computationally expensive, as it involves integration of $q(h_2)$ with the distribution of intrinsic noise $g(P, h_2)$ of the target gene. To overcome this problem, we note that when the intrinsic noise of the target gene is gamma-distributed (Eq. S9), we can use the intrinsic noise representation (Eq. S8) to separate the moments of $q(h_2)$ and $g(P, h_2)$:

$$\mu_{p_2}(P, K_1) = ab \mu_q(h_2, K_1), \quad (\text{S42})$$

$$\sigma_{p_2}^2(P, K_1) = ab^2 \mu_q(h_2, K_1) + a^2 b^2 \sigma_q^2(h_2, K_1). \quad (\text{S43})$$

In the above equations, $\mu_{p_2}(P, K_1)$ and $\sigma_{p_2}^2(P, K_1)$ are the mean and variance of the distribution $p_2(P, K_1)$ at a given value of the signal parameter K_1 , whereas $\mu_q(h_2, K_1)$ and $\sigma_q^2(h_2, K_1)$ are the mean and variance of the distribution $q(h_2, K_1)$, which can be calculated numerically without unnecessary computational cost.

The mean of $g(P)$ is ab and variance is ab^2 , a being the mean burst frequency of the downstream gene at its maximal activity, and b being the mean burst size of the target gene. Then the

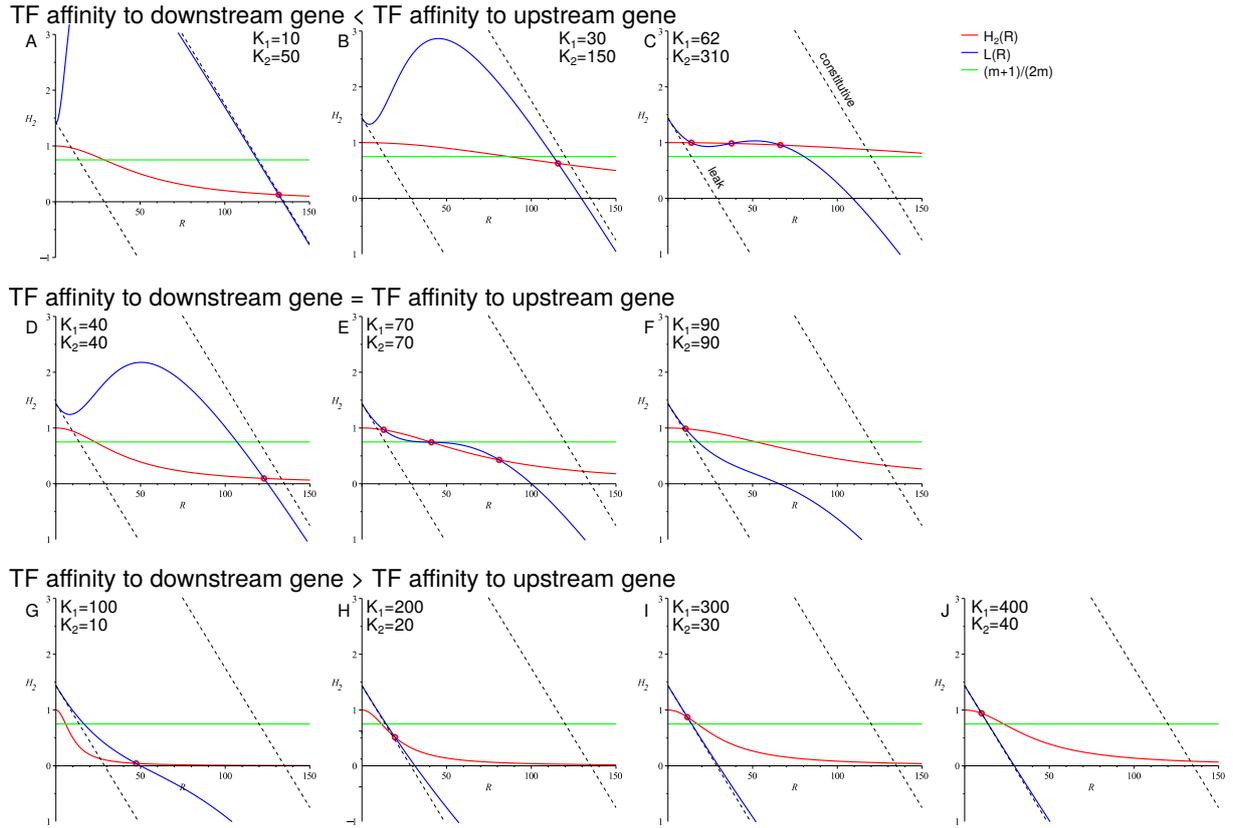


Figure S18: Geometric construction for a cascade where the upstream gene is positively autoregulated and the downstream gene is negatively regulated. The construction is a shifted mirror image of that for both genes regulated positively (Fig. 5 in the main text). Parameters: $n = -2$, $m = 2$, $\alpha = 25$, $\beta = 5$, $a = 250$, $b = 5$, $\epsilon_1 = 0.15$, $\epsilon_2 = 0.01$. Coloured dots are manually added guides to the eye that mark the intersections of the curves.

relative width of the distribution of target proteins is

$$\begin{aligned}
 W(K_1) &= \frac{\sqrt{ab^2 \mu_q(h_2, K_1) + a^2 b^2 \sigma_q^2(h_2, K_1)}}{ab(1 - \epsilon_2)}, \\
 &= \frac{1}{1 - \epsilon_2} \sqrt{\frac{\mu_q(h_2, K_1)}{a} + \sigma_q^2(h_2, K_1)}.
 \end{aligned} \tag{S44}$$

where ϵ_2 is the leakage of the target gene.

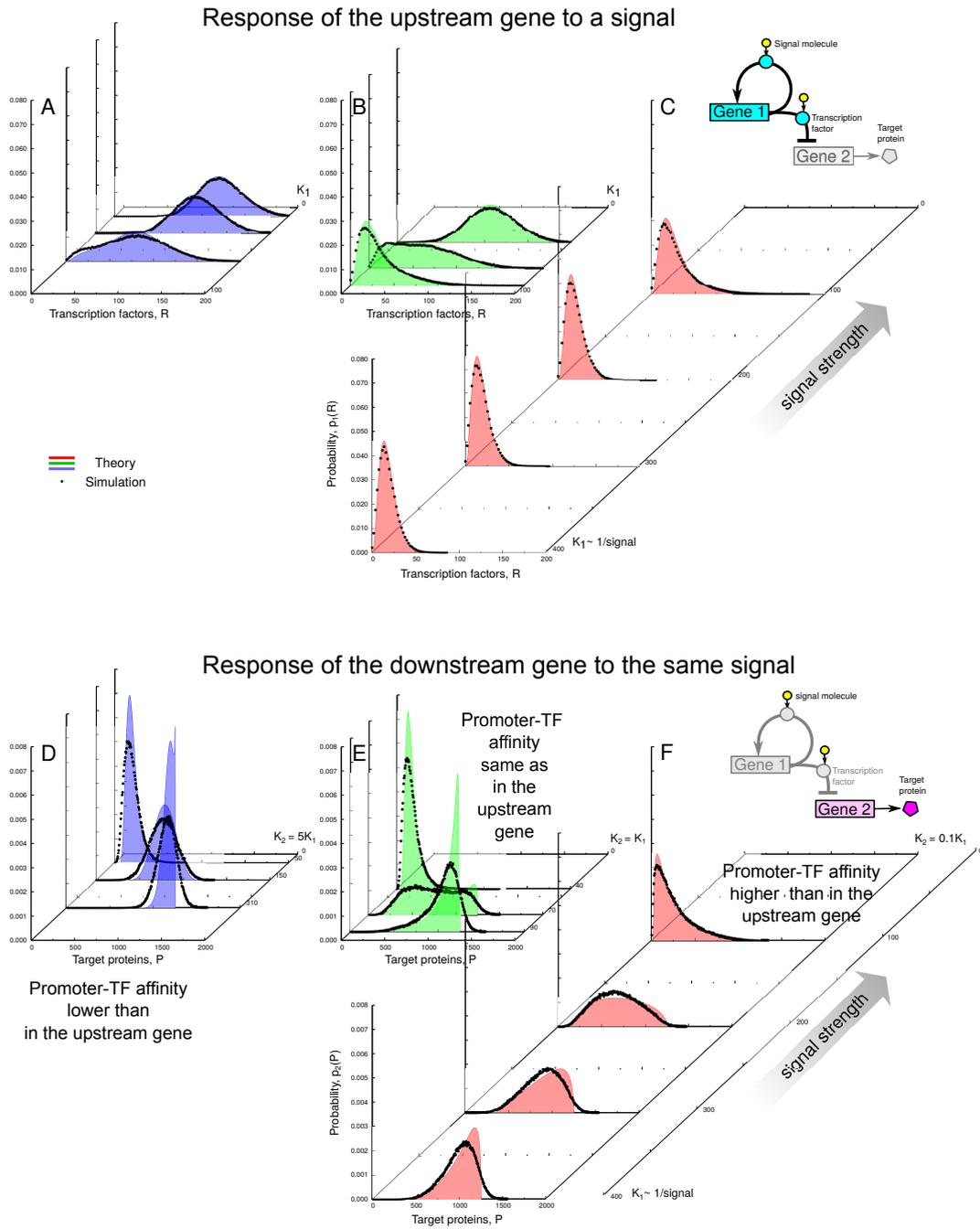


Figure S19: The response of the positively autoregulated upstream gene and the corresponding response of the negatively regulated downstream gene to varying signal. Parameters are the same as in Fig. S18. Black dots represent protein number distributions $p_2(P)$ obtained from simulation. Theoretical curves are $p_1(R)$ in the top panels, and the rescaled distributions of transcription rates, $\frac{1}{ab}q(\frac{P}{ab})$ in the bottom panels. The values of other parameters are shown in Table S3.

S15 Upstream gene positively autoregulated, downstream gene negatively regulated

In Fig. S18 we present the geometric construction for a cascade where the upstream gene is positively autoregulated ($n = -2$) and the downstream gene is negatively regulated ($m = 2$). The construction is a shifted mirror image of that for both genes regulated positively (Fig. 5 in the main text), where m was equal to -2 .

Fig. S19 shows the response of the positively autoregulated upstream gene and the corresponding response of a negatively regulated downstream gene to a varying signal. The theoretical curves $q(h_2)$ are mirror images of those for both genes regulated positively (Fig. 6 in the main text).

S16 Considerations about the experimental measurement of transfer functions

Below we use the following notation: In the original system, Gene 1 consists of the promoter ρ_R and the coding sequence of the protein R. Gene 2 consists of the promoter ρ_P and the protein P coding sequence. R is a TF for both Gene 1 and Gene 2 (Fig. S20A).

In experimental realisation of our system, one has to face two main difficulties in measurement of the transfer functions $h_1(R)$ and $h_2(R)$ of the two promoters: Firstly, Gene 1 is self-regulating. Therefore, one needs to disrupt the feedback loop in order to gain the control over the input of the ρ_R promoter [11]. Secondly, the input levels can be noisy even if one controls their mean. During transmission of the noisy input into a noisy output, the shape of the input distribution will be nonlinearly distorted [3]. It is therefore not certain whether the “mean input vs. mean output” curve will have the same shape as the transfer function $h_i(R)$ in the deterministic model. Below we test the parameter values used in Fig. 2 to see whether the “mean input vs. mean output” curves, produced by our stochastic model for Gene 1 and 2, well approximate the corresponding transfer functions.

S16.1 Opening of the feedback loop

We propose to apply a method similar to that described in [11], where the self-regulating gene is divided into two parts (Fig. S20):

1. In order to produce R in a controlled manner, one needs to place the R coding sequence under the control of a custom promoter ρ_A , regulated by a protein A.
2. In order to measure the response of ρ_R promoter to R, we need to place the reporter protein B coding sequence under the control of ρ_R .
3. The construct described in the point 1 can also be used to measure the response of ρ_P to R.

There are two ways of control of the mean level of R in the above system: By the level of A or by the level of the inducer that activates R.

S16.2 Mean level of R protein controlled by the level of A protein

In this version of the method, we vary A concentration in cells, to obtain varying mean levels of the transcription factor R, which in turn causes a varying response of the promoter under study. At the same time, we keep the inducer (i. e., the signal affecting the transfer function) at some fixed level such that a given fraction f_a of all R molecules are active. Importantly, we need to ensure that the levels of A are the same in all cells, i. e., that our control over A is not noisy. Otherwise, we would add another level of noise to the system.

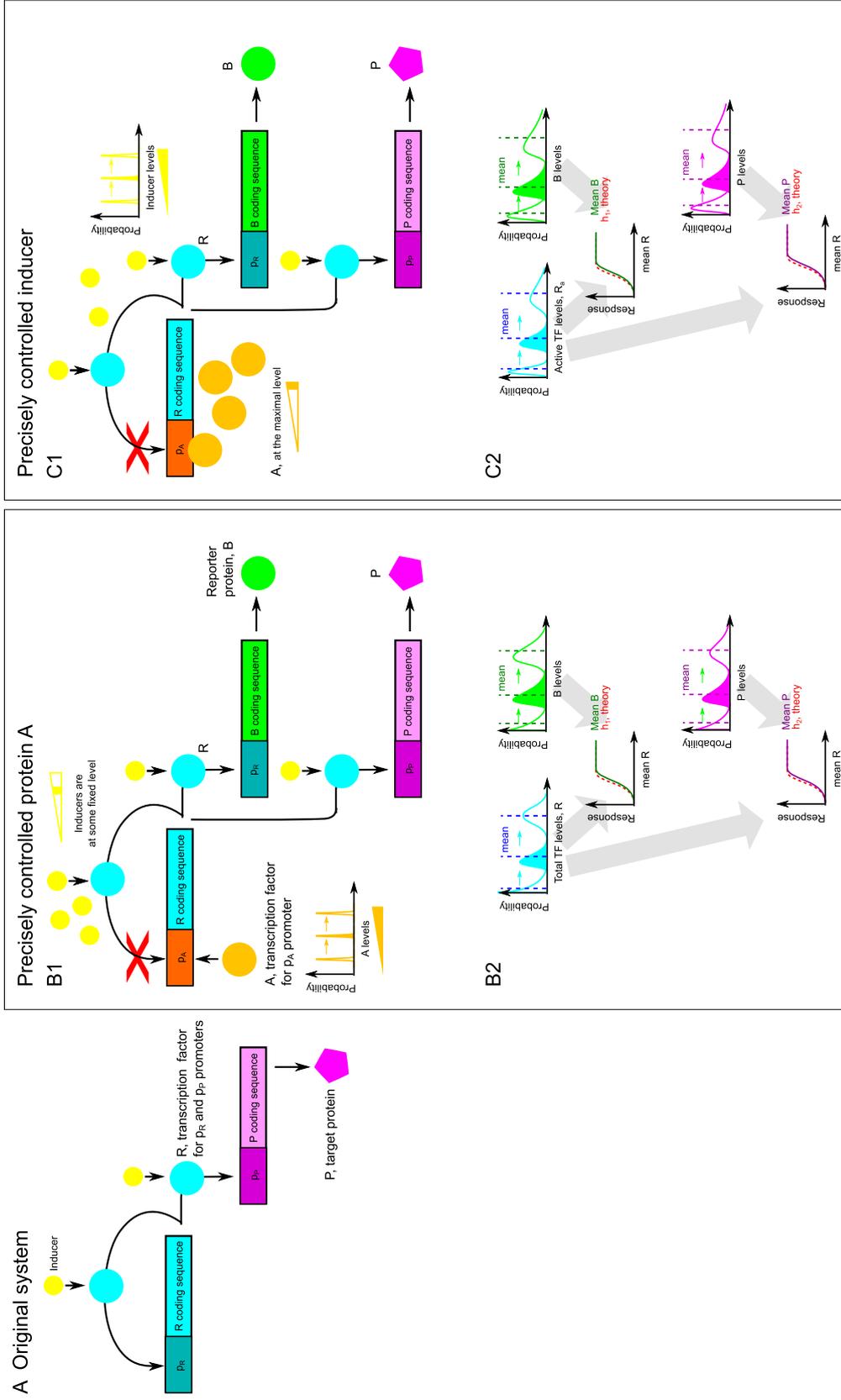


Figure S20: Two proposed methods of measurement of transfer functions in our system, based on [11]. A: The original gene cascade. B: Mean level of R protein controlled by the level of inducer. C: Mean level of R protein controlled by the level of inducer. B1, C1: Feedback opening [11] in order to gain control over the PR promoter. B2, C2: “mean input” vs. mean output” curves are expected to approximate the theoretical deterministic transfer functions. Note that in Fig. C2 we have the varying distribution of active TFs, R_a , because the distribution of total TFs, R , does not vary.

Depending on the A level, we get a varying input distributions of R proteins, given by the gamma distribution (Eq. S11). Since A is a transcription factor for \mathbf{p}_A , its level affects the rate of transcription from that promoter, resulting in variation in parameter $\alpha = \alpha(A)$:

$$p_0(R; \alpha(A), \beta) = \gamma(R; \alpha(A), \beta) \quad (\text{S45})$$

Note that the varying $\alpha(A)$ changes the shape of the distribution of R. $p_0(R; \alpha(A), \beta)$ affects the promoter \mathbf{p}_R and leads to a distribution $q_0(h_1; \alpha(A), \beta, K_1)$ of transcription rates from that promoter (Eq. S14) in cell population.

In reality, it is impossible to directly measure the response $q_0(h_1; \alpha(A), \beta, K_1)$ of \mathbf{p}_R promoter. instead, we need to measure the distribution of reporter proteins B:

$$p_B(B; A) = \int_{\epsilon_1}^1 q_0(h_1; \alpha(A), \beta, K_1) \gamma(B; ah_1, b) dh_1, \quad (\text{S46})$$

where B is the level of B protein, ϵ_1 is the leakage of the \mathbf{p}_R promoter, a is the mean burst frequency of B proteins, and b is their mean burst size.

However, the mean response $\langle B \rangle$ of the \mathbf{p}_R promoter scales linearly with the mean transcription rate $\langle h_1 \rangle$:

$$\langle B(A) \rangle = \int_0^\infty B p_B(B; A) dB \quad (\text{S47})$$

$$= \int_0^\infty B \int_{\epsilon_1}^1 q_0(h_1; \alpha(A), \beta, K_1) \gamma(B; ah_1, b) dh_1 dB \quad (\text{S48})$$

$$= \int_{\epsilon_1}^1 \left[\int_0^\infty B \gamma(B; ah_1, b) dB \right] q_0(h_1; \alpha(A), \beta, K_1) dh_1 \quad (\text{S49})$$

$$= ab \int_{\epsilon_1}^1 h_1 q_0(h_1; \alpha(A), \beta, K_1) dh_1 \quad (\text{S50})$$

$$= ab \langle h_1(A) \rangle. \quad (\text{S51})$$

And therefore, below we will safely use $\langle h_1 \rangle = \int_{\epsilon_1}^1 h_1 q_0(h_1) dh_1$ instead of $\langle B \rangle$, because, within our model, intrinsic noise of downstream promoters does not distort their *mean* response.

To a given mean input at some A level, $\langle R \rangle_A = \alpha(A)\beta$, we ascribe a given mean output $\langle h_1 \rangle_A$:

$$h_{stoch,1} : \langle R \rangle_A \rightarrow \langle h_1 \rangle_A. \quad (\text{S52})$$

We compare this stochastic transfer function $h_{stoch,1}(\langle R \rangle_A)$ of the \mathbf{p}_R promoter with the theoretical deterministic transfer function $h_1(\langle R \rangle_A) = h_1(\alpha(A)\beta)$, given by the Eq. S6. In the same way, we compare the stochastic transfer function $h_{stoch,2}(\langle R \rangle_A) : \langle R \rangle_A \rightarrow \langle h_2 \rangle_A$ of the \mathbf{p}_P

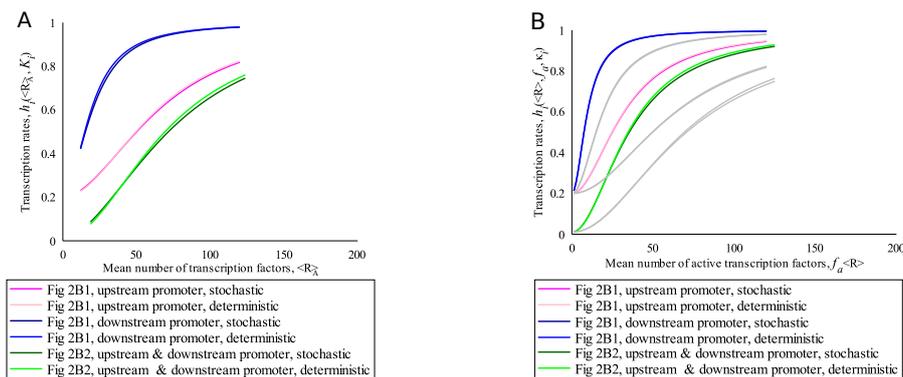


Figure S21: Very good agreement between the stochastic and deterministic transfer functions for our model with parameter values same as in Fig. 2 in the main text. A: Mean level of R protein controlled by the level of A protein assuming that $f_a=1$. B: Mean level of R protein controlled by the level of inducer, with assumption that $f_a = 1/2$ in Fig. 2. Assuming $f_a = 1$ we get the results equivalent to those in Fig. A (grey curves). The index i in h_i is equal to 1 or 2, for the upstream or downstream gene.

promoter with the corresponding deterministic transfer function $h_2(\langle R \rangle_A)$, given by the Eq. S7.

Using this method in experiment, one determines the transfer function $h_1(R)$ (or $h_2(R)$) for a given signal level (which corresponds to some fraction f_a of active TFs), using the cooperativity n (or m) and the K parameter as fitting parameters, while varying the (known) mean number of TFs, $\langle R \rangle_A$. Note that the second method, described below in subsection S16.3, requires different fitting parameters.

In Fig S21A we present the comparison between the stochastic and deterministic transfer functions for our model with parameter values same as in Fig. 2 in the main text. We assumed that the parameter α of the \mathbf{p}_A promoter is same as in Fig. 2, i. e., as in the original Gene 1 with \mathbf{p}_R promoter. Parameter β is also assumed to be the same as in the original Gene 1. Note that the range in which the curves can be measured is limited by the native minimal and maximal mean transcription frequencies, $\alpha\epsilon_1$ and α of the \mathbf{p}_A promoter. The curves obtained from the stochastic model (mimicking the experimental curves) are in a very good agreement with the deterministic transfer functions. One should keep in mind, however, that the differences between experimental data and theory may be greater if the real gene system contains more sources of noise than those covered by our model [11], or, perhaps, if the parameter values are in a different range.

S16.3 Mean level of R protein controlled by the level of inducer

In this version of the method, we vary the inducer concentration in cells, to obtain varying mean levels of active R. Here, we keep the promoter \mathbf{p}_A always maximally active by applying a high level of protein A. One can also use a highly active constitutive promoter as \mathbf{p}_A , without

the regulator A. As a result, the mean levels of R will be maximal, and we only control the fraction f_a of R which are active due to inducer binding. We need to ensure that the levels of inducer in all cells are the same, i. e., that we have a deterministic control over the inducer. The difficulty of this method lies in the fact that we need to be able to measure not only the levels of R but also, separately, the levels of its active fraction f_a .

Now the promoter p_A produces a distribution of R proteins, which *is not* varied:

$$p_0(R; \alpha, \beta) = \gamma(R; \alpha, \beta). \quad (\text{S53})$$

We only change the active fraction f_a of R proteins, so, differently than in the subsection [S16.2](#), the shape of the distribution of the active proteins, R_a , will always be the same, only rescaled by the factor f_a :

$$p_{0,a}(R_a) = \frac{1}{f_a} \gamma\left(\frac{R_a}{f_a}; \alpha, \beta\right). \quad (\text{S54})$$

It should be noted that varying f_a we vary the K_1 parameter in the transfer function h_1 . For $h_1(R, f_a) = H_1(R, f_a)(1 - \epsilon_1) + \epsilon_1$, we have

$$H_1(R, f_a) = \frac{1}{1 + \left(\frac{R}{K_1}\right)^n} = \frac{1}{1 + \left(\frac{f_a R}{\kappa_1}\right)^n}, \quad (\text{S55})$$

where

$$\kappa_1 = f_a K_1 = \left(\frac{k_{on,1}^a \dots k_{on,n}^a}{k_{off,1}^a \dots k_{off,n}^a}\right)^{-1/n}, \quad (\text{S56})$$

see Eq. [S5](#).

The varying fraction of active R proteins affects the promoter p_R and gives rise to a varying distribution $q_0(h_1; \alpha, \beta, \kappa_1/f_a)$ of transcription rates (Eq. [S14](#)). We calculate the mean of that distribution, $\langle h_1 \rangle_{f_a}$.

To a given mean input $\langle R_a \rangle_{f_a} = f_a \alpha \beta = f_a \langle R \rangle$ we ascribe a given mean output $\langle h_1 \rangle_{f_a}$, similarly as in Eq. [S52](#), in order to get the stochastic transfer function, $h_{stoch,1}(\langle R \rangle, f_a)$. We compare it to the deterministic transfer function $h_1(\langle R \rangle, f_a)$. Note that $\langle R \rangle$ is fixed in these functions and we only vary f_a . The analogous procedure should be used for measurement of the transfer function $h_2(R)$ of the p_P promoter.

In the present version of the method, the experimentalist should determine the transfer function by fitting the cooperativity n (or m) and the parameter κ_1 (or κ_2), knowing the total mean number of TFs, $\langle R \rangle$, and their (varying) active fraction f_a . Note that these are different fitting parameters than in the method described above in the subsection [S16.2](#). The original K_1 (K_2) parameter of the transfer function for a given active fraction f_a of TFs is obtained from the Eq. [S56](#).

Fig S21B shows the comparison between the stochastic and deterministic transfer functions for our model with parameter values same as in Fig. 2 in the main text. We assumed that the parameter α of the p_A promoter is same as in Fig. 2, i. e., as in the original Gene 1 with p_R promoter. Parameter β is also assumed to be the same as in the original Gene 1. Note that in order to parametrize the model as in the measurement method described in this section, we need to know what value of f_a was present in Fig. 2 (in an experiment, this value should be known from measurement). Here, for the purpose of illustration, we arbitrarily assumed that f_a was equal to $1/2$ (coloured curves). Note that assuming $f_a = 1$ we get the results equivalent to those in the subsection S16.2 (grey curves).

References

- [1] Ochab-Marcinek A, Tabaka M. Transcriptional leakage versus noise : A simple mechanism of conversion between binary and graded response in autoregulated genes. *Phys Rev E*. 2015;91:012704.
- [2] Shahrezaei V, Swain PS. Analytical distributions for stochastic gene expression. *Proc Natl Acad Sci USA*. 2008;105(45):17256–17261.
- [3] Ochab-Marcinek A, Tabaka M. Bimodal gene expression in noncooperative regulatory systems. *P Natl Acad Sci USA*. 2010;107(51):22096–22101.
- [4] Cai L, Friedman N, Xie XS. Stochastic protein expression in individual cells at the single molecule level. *Nature*. 2006;440(7082):358–362.
- [5] Friedman N, Cai L, Xie XS. Linking stochastic dynamics to population distribution: an analytical framework of gene expression. *Phys Rev Lett*. 2006;97(16):168302.
- [6] Gibson MA, Bruck J. Efficient exact stochastic simulation of chemical systems with many species and many channels. *J Phys Chem A*. 2000;104(9):1876–1889.
- [7] Ishihama A, Kori A, Koshio E, Yamada K, Maeda H, Shimada T, et al. Intracellular concentrations of 65 species of transcription factors with known regulatory functions in *Escherichia coli*. *J Bacteriol*. 2014;196(15):2718–2727.
- [8] Ishihama Y, Schmidt T, Rappsilber J, Mann M, Hartl FU, Kerner MJ, et al. Protein abundance profiling of the *Escherichia coli* cytosol. *BMC Genomics*. 2008;9:102.
- [9] Yanai I, Korbelt JO, Boue S, McWeeney SK, Bork P, Lercher MJ. Similar gene expression profiles do not imply similar tissue functions. *Trends in genetics : TIG*. 2006;22(3):132–8.
- [10] Ingolia NT, Murray AW. Positive-Feedback Loops as a Flexible Biological Module. *Current Biology*. 2007;17(8):668–677.
- [11] Hsu C, Jaquet V, Maleki F, Becskei A. Contribution of bistability and noise to cell fate transitions determined by feedback opening. *Journal of molecular biology*. 2016;428(20):4115–4128.

Figure	K_1	K_2	k_+	$k_{on,1}^t$	$k_{on,2}^t$
5 A,D; S19 A,D	10	50	0.03774	4e-6	4e-3
5 A,D; S19 A,D	30	150	0.003578	4e-6	4e-3
5 A,D; S16; S19 A,D	62	310	0.001461	4e-6	4e-3
5 B,E; S19 B,E	40	40	0.002463	2e-5	2e-2
5 B,E; S19 B,E	70	70	0.001273	2e-5	2e-2
5 B,E; S19 B,E	90	90	0.0009630	2e-5	2e-2
5 C,F; S19 C,F	100	10	0.0008584	2e-4	2e-1
5 C,F; S19 C,F	200	20	0.0004116	2e-4	2e-1
5 C,F; S19 C,F	300	30	0.0002707	2e-4	2e-1
5 C,F; S19 C,F	400	40	0.0002016	2e-4	2e-1

k_-	$k_{on,1}^r$	$k_{off,1}^r$	$k_{on,2}^r$	$k_{off,2}^r$	$k_{off,1}^t$	$k_{off,2}^t$
1e-2	2e-5	5e-3	2e-2	5e-3	5e-3	5e-3

k_{ml1}	k_{m1}	k_{ml2}	k_{m2}	k_{p1}	k_{p2}
3.75e-5	2.5e-4	2.5e-5	2.5e-3	5e-4	5e-4

k_{dm1}	k_{dm2}	k_{dp1}	k_{dp2}	Simulation time
1e-4	1e-4	1e-5	1e-5	2e9

Table S3: Parameters for Figs. 5, S16, S19

$k_{on,1}^r$	$k_{off,1}^r$	$k_{on,2}^r$	$k_{off,2}^r$	$k_{on,1}^t$	$k_{off,1}^t$	$k_{on,2}^t$	$k_{off,2}^t$
2e-3	5e-1	2	5e-1	2e-3	5e-1	2	5e-1
2e-4	5e-2	2e-1	5e-2	2e-4	5e-2	2e-1	5e-2
2e-5	5e-3	2e-2	5e-3	2e-5	5e-3	2e-2	5e-3
2e-6	5e-4	2e-3	5e-4	2e-6	5e-4	2e-3	5e-4

k_+	k_-	k_{ml1}	k_{m1}	k_{ml2}	k_{m2}
0.001273	1e-2	3.75e-5	2.5e-4	2.5e-5	2.5e-3

k_{p1}	k_{p2}	k_{dm1}	k_{dm2}	k_{dp1}	k_{dp2}
5e-4	5e-4	1e-4	1e-4	1e-5	1e-5

Simulation time
2e9

Table S4: Parameters for Fig. S1

$k_{on,1}^r$	$k_{off,1}^r$	$k_{on,2}^r$	$k_{off,2}^r$	$k_{on,1}^t$	$k_{off,1}^t$	$k_{on,2}^t$	$k_{off,2}^t$
2e-5	5e-3	2e-2	5e-3	2e-5	5e-3	2e-2	5e-3

k_+	k_-	k_{ml1}	k_{m1}	k_{ml2}	k_{m2}
0.001273	1e-2	3.75e-5	2.5e-4	2.5e-5	2.5e-3

k_{p1}	k_{p2}	k_{dm1}	k_{dm2}	k_{dp1}	k_{dp2}	Simulation time
5e-4	5e-4	1e-4	1e-4	1e-5	1e-5	2e6

Table S5: Parameters for Fig. S2, $K_1 = K_2 = 70$.

k_{m2}			
2.5e-5			
1.25e-3			
2.5e-3			
k_{p2}	k_{dm2}	k_{dp2}	Simulation time
5e-4	1e-4	1e-5	2e9

Table S6: Parameters for Fig. S3 (non-regulated gene).

k_{ml2}		k_{m2}		k_{dp2}			
2.5e-4		25e-3		1e-4			
12.5e-5		12.5e-3		5e-5			
2.5e-5		2.5e-3		1e-5			
12.5e-6		12.5e-4		5e-6			
2.5e-6		2.5e-4		1e-6			
$k_{on,1}^r$	$k_{off,1}^r$	$k_{on,2}^r$	$k_{off,2}^r$	$k_{on,1}^t$	$k_{off,1}^t$	$k_{on,2}^t$	$k_{off,2}^t$
2e-5	5e-3	2e-2	5e-3	2e-5	5e-3	2e-2	5e-3
k_+		k_-		k_{ml1}		k_{m1}	
0.001273		1e-2		3.75e-5		2.5e-4	
k_{p1}	k_{p2}	k_{dm1}	k_{dm2}	k_{dp1}	Simulation time		
5e-4	5e-4	1e-4	1e-4	1e-5	2e9		

Table S7: Parameters for Fig. 7, $K_1 = K_2 = 70$.

Figure	K_2	k_+	$k_{on,1}^t$	$k_{on,2}^t$
S4, S5, S6 A,D	50	0.03774	4e-6	4e-3
S4, S5, S6 A,D	150	0.003578	4e-6	4e-3
S4, S5, S6 A,D	310	0.001461	4e-6	4e-3
S4, S5, S6 B,E	40	0.002463	2e-5	2e-2
S4, S5, S6 B,E	70	0.001273	2e-5	2e-2
S4, S5, S6 B,E	90	0.0009630	2e-5	2e-2
S4, S5, S6 C,F	10	0.0008584	2e-4	2e-1
S4, S5, S6 C,F	20	0.0004116	2e-4	2e-1
S4, S5, S6 C,F	30	0.0002707	2e-4	2e-1
S4, S5, S6 C,F	40	0.0002016	2e-4	2e-1
Figure	k_{m1}			
S4	3.75e-5			
S5	12.5e-5			
S6	2.5e-4			

k_-	$k_{off,1}^t$	$k_{off,2}^t$		
1e-2	5e-3	5e-3		
k_{ml2}	k_{m2}	k_{p1}	k_{p2}	
2.5e-5	2.5e-3	5e-4	5e-4	
k_{dm1}	k_{dm2}	k_{dp1}	k_{dp2}	Simulation time
1e-4	1e-4	1e-5	1e-5	2e9

Table S8: Parameters for Figs. S4, S5, S6.