SUPPLEMENTARY INFORMATION

EFFECT OF PROMOTER ARCHITECTURE ON THE CELL-TO-CELL VARIABILITY IN GENE EXPRESSION

Alvaro Sanchez¹, Hernan G. Garcia², Daniel Jones², Rob Phillips^{3,4}, Jane Kondev⁵

¹Graduate Program in Biophysics and Structural Biology and ⁵Department of Physics, Brandeis University, Waltham, MA.

²Department of Physics, ³Department of Bioengineering and ⁴Department of Applied Physics, California Institute of Technology, Pasadena, CA.

The moments of the mRNA probability distribution

We start by considering the same mechanism as in the text (see figure 1), in which the promoter switches between one active and one inactive state. There are only two stochastic variables in the model: the number of mRNA transcripts per cell (m), and the state of the promoter which reflects which transcription factors are bound where. The promoter state is always a discrete and finite stochastic variable (s) (for an example, see figure 1a). The example in figure 1a illustrates the simplest model of transcriptional activation by a transcription factor.

When the activator is bound to the promoter (state 1) mRNA is synthesized at rate r_1 . When the activator is not bound (state 2) mRNA is synthesized at a lower rate r_2 . The promoter switches stochastically from state 1 to state 2 with rate k_A^{off} , and from state 2 to state 1 with rate k_A^{on} . Each mRNA molecule is degraded with rate γ .

The time evolution for the joint probability of having the promoter in states 1 or 2, with m mRNAs in the cell (which we write as p(1,m) and p(2,m), respectively), is given by a master equation, which we can build by listing all possible reactions that lead to a change in cellular state, either by changing m or by changing s (figure 1b). The master equation takes the form:

$$\frac{d}{dt}p(1,m) = -k_A^{off}p(1,m) + k_A^{on}p(2,m) - r_1p(1,m) - \gamma mp(1,m) + r_1p(1,m-1) + \gamma(m+1)p(1,m+1),$$

$$\frac{d}{dt}p(2,m) = k_A^{off}p(1,m) - k_A^{on}p(2,m) - r_2p(2,m) - \gamma mp(2,m) + r_2p(2,m-1) + \gamma(m+1)p(2,m+1) .$$
(1)

Inspecting this system of equations, we notice that by defining the vector:

$$\vec{p}(m) = \begin{pmatrix} p(1,m) \\ p(2,m) \end{pmatrix},\tag{2}$$

and the matrices

$$\hat{K} = \begin{bmatrix} -k_A^{off} & k_A^{on} \\ k_A^{off} & -k_A^{on} \end{bmatrix} ; \quad \hat{R} = \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix} ; \quad \hat{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

we can rewrite the system of equations (1) in matrix form.

$$\frac{d}{dt}\vec{p}(m) = \left[\hat{K} - \hat{R} - m\gamma\hat{I}\right]\vec{p}(m) + \hat{R}\vec{p}(m-1) + (m+1)\gamma\hat{I}\vec{p}(m+1) \quad .$$
(3)

This approach can be generalized to any mechanism of transcriptional regulation at the promoter level. The only difference between the mechanisms rests on the particular dimensionality and form of the three matrices defined above. Examples of those matrices for all of the architectures and mechanisms investigated on this paper are given in Table S1 in Text S1. In steady state, the left hand side of equation (4) is equal to 0:

$$0 = \left[\hat{K} - \hat{R} - m\gamma \hat{I}\right]\vec{p}(m) + \hat{R}\vec{p}(m-1) + (m+1)\gamma \hat{I}\vec{p}(m+1) \quad .$$
(4)

In order to find the first two moments of the steady state mRNA probability distribution, we follow the same strategy as in references [1,2]: we multiply both sides of equation (5) by m and m^2 respectively, and then sum over all values of m, from 0 to ∞ . We start from the first moment of the mRNA distribution, which requires us to multiply equation (5) by m and then sum:

$$\sum_{m=0}^{\infty} m\left(\left[\hat{K} - \hat{R} - m\gamma\hat{I}\right]\vec{p}(m) + \hat{R}\vec{p}(m-1) + (m+1)\gamma\hat{I}\vec{p}(m+1)\right) = \sum_{m=0}^{\infty} m\hat{K}\vec{p}(m) - \sum_{m=0}^{\infty} m^{2}\gamma\hat{I}\vec{p}(m) - \sum_{m=0}^{\infty} m^{2}\gamma\hat{I}\vec{p}(m) + \sum_{m=0}^{\infty} m\hat{R}\vec{p}(m-1) + \sum_{m=0}^{\infty} m(m+1)\gamma\hat{I}\vec{p}(m+1) \quad .$$
(5)

Since none of the three matrices \hat{K} , \hat{R} and \hat{I} are functions of *m*, they can be taken out of the sums, and we find:

$$0 = \hat{K} \sum_{m=0}^{\infty} m \, \vec{p}(m) - \gamma \hat{I} \sum_{m=0}^{\infty} m^2 \, \vec{p}(m) - \hat{R} \sum_{m=0}^{\infty} m \, \vec{p}(m) + \hat{R} \sum_{m=0}^{\infty} m \, \vec{p}(m-1) + \gamma \hat{I} \sum_{m=0}^{\infty} m \, (m+1) \, \vec{p}(m+1)$$
(6)

It will be convenient in what follows to define the following vectors of partial moments of the mRNA probability distribution:

$$\vec{m}_{(0)} = \sum_{m=0}^{\infty} m^0 \vec{p}(m) = \begin{pmatrix} \sum_{m=0}^{\infty} m^0 p(1,m) \\ \sum_{m=0}^{\infty} m^0 p(2,m) \end{pmatrix} = \begin{pmatrix} p(1) \\ p(2) \end{pmatrix},$$

$$\vec{m}_{(1)} = \sum_{m=0}^{\infty} m \vec{p}(m) = \begin{pmatrix} \sum_{m=0}^{\infty} m p(1,m) \\ \sum_{m=0}^{\infty} m p(2,m) \end{pmatrix},$$

$$\vec{m}_{(2)} = \sum_{m=0}^{\infty} m^2 \vec{p}(m) = \begin{pmatrix} \sum_{m=0}^{\infty} m^2 p(1,m) \\ \sum_{m=0}^{\infty} m^2 p(2,m) \end{pmatrix}.$$
(7)

The usefulness of these vectors of partial moments of the mRNA distribution lies in the fact that they are related to the moments of the probability distribution. For instance, the mean mRNA is given by

$$\langle m \rangle = \sum_{s=1}^{2} \sum_{m=0}^{\infty} m \ p(s,m) = \sum_{m=0}^{\infty} m \ p(1,m) + \sum_{m=0}^{\infty} m \ p(2,m)$$
 (8)

If we define, again for convenience, the vector $\vec{u} = (1,1)$, we find that the mean of the mRNA distribution is related to the vectors of partial moments by $\langle m \rangle = \vec{u} \ \vec{m}_{(1)}$. Following this example, it is also straightforward to prove that the second moment of the mRNA distribution is given by: $\langle m^2 \rangle = \vec{u} \ \vec{m}_{(2)}$.

Given these definitions, we return to equation (7) which we can now write as:

$$\hat{K} \ \vec{m}_{(1)} - \gamma \hat{I} \ \vec{m}_{(2)} - \hat{R} \ \vec{m}_{(1)} + \hat{R} \sum_{m=0}^{\infty} m \ \vec{p}(m-1) + \gamma \hat{I} \sum_{m=0}^{\infty} m \ (m+1) \ \vec{p}(m+1) = 0 \quad .$$
(9)

We can re-arrange terms in the last two sums so that we write them as operations on the vectors of partial moments of the probability distributions. For instance, by making the change of variables: $m \rightarrow m+1$, and taking into account the fact that the number of mRNA molecules inside the cell can never fall below 0 (so that $\vec{p}(-1) = 0$), we find:

$$\sum_{m=0}^{\infty} m \, \vec{p}(m-1) = \sum_{m=0}^{\infty} (m+1) \, \vec{p}(m) = \vec{m}_{(1)} + \vec{m}_{(0)} \quad .$$
⁽¹⁰⁾

Similarly, by making the change of variables $m+1 \rightarrow m$, the last sum takes the simpler form:

$$\sum_{m=0}^{\infty} m(m+1) \vec{p}(m+1) = \sum_{m=0}^{\infty} m(m-1) \vec{p}(m) = \vec{m}_{(2)} - \vec{m}_{(1)} \quad .$$
(11)

Entering these results into equation (10), we finally find:

$$\hat{K} \ \vec{m}_{(1)} - \gamma \hat{I} \ \vec{m}_{(2)} - \hat{R} \ \vec{m}_{(1)} + \hat{R} \Big(\vec{m}_{(1)} + \vec{m}_{(0)} \Big) + \gamma \hat{I} \Big(\vec{m}_{(2)} - \vec{m}_{(1)} \Big) = \hat{K} \ \vec{m}_{(1)} - \gamma \hat{I} \ \vec{m}_{(1)} + \hat{R} \ \vec{m}_{(0)} \ .$$
(12)

The vector of partial moments $\vec{m}_{(1)}$ is therefore the solution to the matrix equation:

$$\left(\hat{K} - \gamma \hat{I}\right) \vec{m}_{(1)} + \hat{R} \vec{m}_{(0)} = 0 \quad .$$
(13)

The final step is to multiply both sides of equation (14) by the vector $\vec{u} = (1,1)$. Because of how it was constructed (i.e. p(1,m)s loss is p(2,m)s gain during transitions between promoter states), the matrix \hat{K} has the property that the sum of the elements of any one of its columns is always 0. Therefore, we find that $\vec{u} \ \hat{K} = 0$. The matrix \hat{R} is diagonal, so if we multiply matrix \hat{R} on the left by vector \vec{u} , we get a vector that is equal to the list of diagonal elements of matrix \hat{R} . Thus, we define the vector $\vec{r} = (\hat{R}_{11}, \hat{R}_{22}) = (r_1, r_2)$, as the vector for which it is true that $\vec{u} \ \hat{R} = \vec{r}$.

Finally, the identity matrix is $\hat{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Therefore, multiplying \hat{I} on the left by the vector \vec{u} leads us to: $\vec{u} \cdot \hat{I} = \vec{u}$. Therefore, when we multiply equation (14) by the vector \vec{u} we find:

leads us to:
$$u = u$$
. Therefore, when we multiply equation (14) by the vector u we find:

$$0 = \vec{u} \ \hat{K} \ \vec{m}_{(1)} - \vec{u} \ \gamma \hat{I} \ \vec{m}_{(1)} + \vec{u} \ \hat{R} \ \vec{m}_{(0)} = -\gamma \vec{u} \ \vec{m}_{(1)} + \vec{r} \ \vec{m}_{(0)} \ .$$
(14)

Knowing that the mean of the mRNA distribution is related to the vector of partial moments by: $\langle m \rangle = \vec{u} \ \vec{m}_{(1)}$, we find that:

$$\left\langle m\right\rangle = \frac{\vec{r} \ \vec{m}_{(0)}}{\gamma} \quad . \tag{15}$$

Note that, by definition,

$$\vec{m}_{(0)} = \sum_{m=0}^{\infty} m^0 \vec{p}(m) = \begin{pmatrix} \sum_{m=0}^{\infty} m^0 p(1,m) \\ \sum_{m=0}^{\infty} m^0 p(2,m) \end{pmatrix} = \begin{pmatrix} p(1) \\ p(2) \end{pmatrix} .$$
(16)

In other words, the first element of vector $\vec{m}_{(0)}$ is the steady state probability to find the promoter in state 1, and the second element is the steady state probability to find the promoter in state 2. This vector is straightforward to obtain by summing equation (5) over all *m*, and it is the solution of $\hat{K} \vec{m}_{(0)} = 0$, normalized so that p(1) + p(2) = 1.

In order to find the second moment, we just multiply equation (5) by m^2 and sum over all *m* from 0 to ∞ . As a result of this manipulation, we find:

$$\sum_{m=0}^{\infty} m^{2} \left(\left[\hat{K} - \hat{R} - m\gamma \hat{I} \right] \vec{p}(m) + \hat{R} \, \vec{p}(m-1) + (m+1)\gamma \hat{I} \, \vec{p}(m+1) \right) = \sum_{m=0}^{\infty} m^{2} \, \hat{K} \, \vec{p}(m) - \sum_{m=0}^{\infty} m^{3} \, \gamma \hat{I} \, \vec{p}(m) - \sum_{m=0}^{\infty} m^{3} \, \gamma \hat{I} \, \vec{p}(m) + \sum_{m=0}^{\infty} m^{2} \, \hat{R} \, \vec{p}(m-1) + \sum_{m=0}^{\infty} m^{2} \, (m+1) \, \gamma \hat{I} \, \vec{p}(m+1) =$$

$$\hat{K} \, \vec{m}_{(2)} - \gamma \hat{I} \, \vec{m}_{(3)} - \hat{R} \, \vec{m}_{(2)} + \sum_{m=0}^{\infty} m^{2} \, \hat{R} \, \vec{p}(m-1) + \sum_{m=0}^{\infty} m^{2} \, (m+1) \, \gamma \hat{I} \, \vec{p}(m+1) \, .$$
(17)

The last two terms of the right hand side of equation (18) can be simplified by writing the two sums in terms of the vectors of partial moments. In order to do that, we must make the same changes of variables that we invoked above when dealing with the mean. First, the change of variables $m \rightarrow m+1$ allows us to rewrite the first sum as:

$$\sum_{m=0}^{\infty} m^2 \vec{p}(m-1) = \sum_{m=0}^{\infty} (m+1)^2 \vec{p}(m) = \vec{m}_{(2)} + 2\vec{m}_{(1)} + \vec{m}_{(0)} \quad .$$
(18)

Finally, the change of variables $m+1 \rightarrow m$, allows us to re-write the last sum as:

$$\sum_{m=0}^{\infty} m^2 (m+1) \vec{p}(m+1) = \sum_{m=0}^{\infty} m(m-1)^2 \vec{p}(m) = \vec{m}_{(3)} - 2\vec{m}_{(2)} + \vec{m}_{(1)} \quad .$$
(19)

Entering these last two sums in equation (18), we find:

$$\hat{K} \ \vec{m}_{(2)} - \gamma \hat{I} \ \vec{m}_{(3)} - \hat{R} \ \vec{m}_{(2)} + \hat{R} \Big(\vec{m}_{(2)} + 2\vec{m}_{(1)} + \vec{m}_{(0)} \Big) + \gamma \hat{I} \Big(\vec{m}_{(3)} - 2\vec{m}_{(2)} + \vec{m}_{(1)} \Big) =
\hat{K} \ \vec{m}_{(2)} + \hat{R} \Big(2\vec{m}_{(1)} + \vec{m}_{(0)} \Big) + \gamma \hat{I} \Big(-2\vec{m}_{(2)} + \vec{m}_{(1)} \Big) = 0 \quad .$$
(20)

As we did before, we can transform this equation into an equation for the moments of the mRNA distribution by multiplying both sides of this equation on the left by the vector \vec{u} . Performing these operations, we find:

$$\vec{u} \ \hat{K} \ \vec{m}_{(2)} + \vec{u} \ \hat{R} \Big(2\vec{m}_{(1)} + \vec{m}_{(0)} \Big) + \vec{u} \ \gamma \hat{I} \Big(-2\vec{m}_{(2)} + \vec{m}_{(1)} \Big) = \vec{r} \Big(2\vec{m}_{(1)} + \vec{m}_{(0)} \Big) + \gamma \vec{u} \ \Big(-2\vec{m}_{(2)} + \vec{m}_{(1)} \Big) = 2 \ \vec{r} \ \vec{m}_{(1)} + \vec{r} \ \vec{m}_{(0)} - 2 \ \gamma \Big\langle m^2 \Big\rangle + \gamma \Big\langle m \Big\rangle = 0 \ .$$

$$(21)$$

Therefore, the second moment of the mRNA distribution in steady state is given by:

$$\left\langle m^{2}\right\rangle = \frac{\vec{r} \ \vec{m}_{(1)}}{\gamma} + \frac{\vec{r} \ \vec{m}_{(0)} + \gamma \left\langle m\right\rangle}{2\gamma} \quad .$$

$$(22)$$

Using the fact that the first moment is given by:

$$\left\langle m\right\rangle = \frac{\vec{r} \ \vec{m}_{(0)}}{\gamma} \quad . \tag{23}$$

We can further simplify the second moment as:

$$\left\langle m^2 \right\rangle = \left\langle m \right\rangle + \frac{\vec{r} \cdot \vec{m}_{(1)}}{\gamma} \quad .$$
 (24)

Therefore, the normalized variance can be written as:

$$\eta^{2} = \frac{\left\langle m^{2} \right\rangle - \left\langle m \right\rangle^{2}}{\left\langle m \right\rangle^{2}} = \frac{1}{\left\langle m \right\rangle} + \frac{1}{\left\langle m \right\rangle^{2}} \left(\frac{\vec{r} \ \vec{m}_{(1)}}{\gamma} - \left\langle m \right\rangle^{2} \right) . \tag{25}$$

The moments of the protein probability distribution

We can use the same method to compute the normalized variance of the protein distribution. We will start from a promoter that is constitutively active, and then extend our analysis to a promoter that switches between two or more active and inactive states. We assume that each transcription event leads to the production of multiple proteins (a "burst"). The number of proteins produced per mRNA (which we denote as β) obeys a geometric distribution [3,4,5] with an average burst

size b. Therefore, the probability for β is given by: $h(\beta) = \frac{b^{\beta}}{(1+b)^{\beta+1}}$. We assume that proteins

are also degraded with a constant rate per molecule of δ . In order to write down the master equation for this process, we have to consider all the possible ways in which the cell can enter or leave a state with *n* proteins during a small increment of time dt. If we assume that mRNA lifetime is much shorter than protein lifetime (an approximation that is realistic in many experimental systems –see refs [4,5,6]), then all of the proteins may be assumed to be made simultaneously. Therefore, we need to consider the possibility that the cell will jump from a state with *n* proteins to a state with $n + \beta$, for all possible values of β . The probability that the cell will leave a state with *n* proteins, by making a transition to a state with $n + \beta$ proteins is equal to the product of the probability that the cell is in a state with *n* proteins (p(n)), the probability that the cell will make a transcript during dt (rdt), and the probability that the transcript makes β proteins before it is degraded $(h(\beta))$. Thus, the total probability per unit time to abandon a state with *n* proteins is given by $r h(\beta) p(n)$. Since β can in principle take any integer value, the total probability to abandon the state with *n* proteins by the occurrence of a protein burst is given by the sum of $r h(\beta) p(n)$ over all possible values of β . This term will be given by:

 $\sum_{\beta=1}^{\infty} r h(\beta) p(n) = r p(n) \sum_{\beta=1}^{\infty} h(\beta)$ Also, we need to consider that the cell may enter a state with *n* proteins from any state with less than *n* proteins. The probability per unit time that the cell enters a state with n proteins, from a state with $n - \beta$ proteins is given by: $rh(\beta) p(n - \beta)$. Therefore, following the same logic as we did before, the net probability per unit time that the cell enters a state with *n* proteins is $\sum_{\beta=1}^{n} r h(\beta) p(n - \beta)$. With these considerations, the master equation for a constitutive promoter is given by:

$$\frac{d}{dt}p(n) = -\sum_{\beta=1}^{\infty} rh(\beta)p(n) + \sum_{\beta=1}^{n} rh(\beta)p(n-\beta) - \delta np(n) + \delta(n+1)p(n+1) \quad .$$
(26)

As discussed above, the first sum can be further simplified to:

$$\sum_{\beta=1}^{\infty} rh(\beta) p(n) = r \ p(n) \sum_{\beta=1}^{\infty} h(\beta) = r \ p(n) \sum_{\beta=1}^{\infty} \frac{b^{\beta}}{(1+b)^{\beta+1}} = r\left(\frac{b}{1+b}\right) p(n) \ .$$
(27)

As a result, the master equation takes the form:

$$\frac{d}{dt}p(n) = -r\left(\frac{b}{1+b}\right)p(n) + \sum_{\beta=1}^{n} rh(\beta)p(n-\beta) - \delta np(n) + \delta(n+1)p(n+1).$$
(28)

In steady state, the right hand side of equation (29) is equal to 0, and we have:

$$0 = -r \left(\frac{b}{1+b}\right) p(n) + \sum_{\beta=1}^{n} rh(\beta) p(n-\beta) - \delta np(n) + \delta(n+1) p(n+1).$$
⁽²⁹⁾

The first two moments of the steady state protein distribution p(n) can be obtained, in exactly the same way we used to find out the moments of the mRNA distribution in the previous section: by multiplying both sides of equation (30) by n and n^2 respectively, and then summing over all

n. Before we do that, it is useful to evaluate the sums $\sum_{n=0}^{\infty} n^2 \sum_{\beta=1}^{n} rh(\beta) p(n-\beta)$ and

 $\sum_{n=0}^{\infty} n \sum_{\beta=1}^{n} rh(\beta) p(n-\beta)$. We can find the general term of the first sum by expanding the series:

$$\sum_{n=0}^{\infty} n^{2} \sum_{\beta=1}^{n} h(\beta) p(n-\beta) = 1^{2} (h(1)p(0)) + 2^{2} (h(1)p(1) + h(2)p(0)) + 3^{2} (h(1)p(2) + h(2)p(1) + h(3)p(0)) + \dots =$$

$$\left(1^{2} h(1) + 2^{2} h(2) + 3^{2} h(3) \dots\right) p(0) + \left(2^{2} h(1) + 3^{2} h(2) + 4^{2} h(3) \dots\right) p(1) + \left(3^{2} h(1) + 4^{2} h(2) + 5^{2} h(3) \dots\right) p(2) + \dots =$$

$$\sum_{n=0}^{\infty} p(n) \left(\sum_{\beta=1}^{\infty} h(\beta) (n+\beta)^{2}\right) = \sum_{n=0}^{\infty} \left(b + 2b^{2} + 2bn + \frac{b}{1+b}n^{2}\right) p(n) .$$
(30)

We can do the same for the second sum, and we find:

$$\sum_{n=0}^{\infty} n \sum_{\beta=1}^{n} h(\beta) p(n-\beta) = 1(h(1)p(0)) + 2(h(1)p(1) + h(2)p(0)) + 3(h(1)p(2) + h(2)p(1) + h(3)p(0)) + \dots =$$

$$(h(1) + 2h(2) + 3h(3)\dots) p(0) + (2h(1) + 3h(2) + 4h(3)\dots) p(1) + (3h(1) + 4h(2) + 5h(3)\dots) p(2) + \dots =$$

$$\sum_{n=0}^{\infty} p(n) \left(\sum_{\beta=1}^{\infty} h(\beta)(n+\beta)\right) = \sum_{n=0}^{\infty} \left(b + \frac{b}{1+b}n\right) p(n) \quad .$$
(31)

Likewise, it will be necessary to recall from the first section of this supplement, that the sum $\sum_{n=0}^{\infty} n(n+1)p(n+1) \text{ can be computed by using the change of variables: } n+1 \rightarrow n, \text{ and we find:}$ $\sum_{n=0}^{\infty} n(n+1)p(n+1) = \sum_{n=0}^{\infty} n(n-1)p(n) . \qquad (32)$

With these results in hand, we can finally solve the first two moments of the protein distribution p(n). As explained above, we can find the first moment by multiplying both sides of equation (30) by *n* and then summing over all *n*. In order to find the second moment, we multiply both sides of equation (30) by n^2 and then sum over all *n*. For the first moment, we find:

$$0 = -r\left(\frac{b}{1+b}\right)\sum_{n=0}^{\infty} np(n) + r\sum_{n=0}^{\infty} n\sum_{\beta=1}^{n} h(\beta) p(n-\beta) - \delta \sum_{n=0}^{\infty} n^2 p(n) + \delta \sum_{n=0}^{\infty} n(n+1)p(n+1) =$$

$$= -r\left(\frac{b}{1+b}\right) \langle n \rangle + r\left(b + \frac{b}{1+b} \langle n \rangle\right) - \delta \langle n^2 \rangle + \delta \langle n^2 \rangle - \delta \langle n \rangle = rb - \delta \langle n \rangle.$$
(33)

Solving this equation, we find that the mean protein per cell is equal to:

$$\left\langle n\right\rangle = \frac{rb}{\delta}.\tag{34}$$

For the second moment, we find:

$$0 = -r\left(\frac{b}{1+b}\right)\sum_{n=0}^{\infty} n^2 p(n) + r\sum_{n=0}^{\infty} n^2 \sum_{\beta=1}^{n} h(\beta) p(n-\beta) - \delta \sum_{n=0}^{\infty} n^3 p(n) + \delta \sum_{n=0}^{\infty} n^2 (n+1) p(n+1) =$$
$$= -r\left(\frac{b}{1+b}\right) \langle n \rangle + r\left(b+2b^2+2b\langle n \rangle + \frac{b}{1+b}\langle n^2 \rangle\right) - \delta \langle n^3 \rangle + \delta \langle n^3 \rangle - 2\delta \langle n^2 \rangle + \delta \langle n \rangle = (35)$$
$$= r\left(b+2b^2+2b\langle n \rangle\right) - 2\delta \langle n^2 \rangle + \delta \langle n \rangle.$$

Solving this last equation, we find that the second moment of the protein distribution is equal to:

$$\left\langle n^{2}\right\rangle = \frac{r}{2\delta}b + \frac{r}{2\delta}2b^{2} + \frac{r}{2\delta}2b\left\langle n\right\rangle + \frac{\left\langle n\right\rangle}{2} = (1+b)\left\langle n\right\rangle + \left\langle n\right\rangle^{2}.$$
(36)

Therefore, the normalized variance of the protein distribution for a constitutive promoter takes the form:

$$\frac{Var(n)}{\langle n \rangle^2} = \frac{\langle n^2 \rangle - \langle n \rangle^2}{\langle n \rangle^2} = \frac{(1+b)\langle n \rangle + \langle n \rangle^2 - \langle n \rangle^2}{\langle n \rangle^2} = \frac{(1+b)}{\langle n \rangle}$$
(37)

If now we consider that the promoter can exist in two states, characterized by having different rates of transcription, then the cell's state is characterized not only by the number of proteins present, but also by the state of the promoter. Therefore, the master equation must consider two variables: one characterizing the state of the promoter (s), and one representing the number of proteins per cell (n). By analogy with the mRNA master equation, and the master equation for the protein distribution of a constitutive promoter, the two-state master equation for the protein distribution can be written as:

$$\frac{d}{dt}p(1,n) = -k_{A}^{on}p(1,n) + k_{A}^{off}p(2,n) - \sum_{\beta=1}^{\infty} r_{1}h(\beta)p(1,n) + \sum_{\beta=1}^{n} r_{1}h(\beta)p(1,n-\beta) - \delta np(1,n) + \delta(n+1)p(1,n+1),$$

$$\frac{d}{dt}p(2,n) = k_{A}^{on}p(1,n) - k_{A}^{off}p(2,n) - \sum_{\beta=1}^{\infty} r_{2}h(\beta)p(2,n) + \sum_{\beta=1}^{n} r_{2}h(\beta)p(2,n-\beta) - \delta np(2,n) + \delta(n+1)p(2,n+1) .$$
(38)

Just as we did in order to compute the moments of the mRNA distribution, we can define the vector $\vec{p}(n) = (p(1,n), p(2,n))$. By doing so, we will be able to re-write the master equation (39) as a matrix equation, that will be applicable to any promoter with any number of states. This matrix equation can be written in terms of exactly the same matrices we used for the mRNA probability distribution. We find:

$$\frac{d}{dt}\vec{p}(n) = \left[\hat{K} - \frac{b}{1+b}\hat{R} - n\delta\,\hat{I}\right]\vec{p}(n) + \hat{R}\sum_{\beta=1}^{n}h(\beta)\vec{p}(n-\beta) + (n+1)\delta\,\hat{I}\,\vec{p}(n+1) \quad .$$
(39)

In steady state, the left side of equation (40) is equal to 0, and the master equation has the form:

$$0 = \left[\hat{K} - \frac{b}{1+b}\hat{R} - n\delta \hat{I}\right]\vec{p}(n) + \hat{R}\sum_{\beta=1}^{n}h(\beta)\vec{p}(n-\beta) + (n+1)\delta \hat{I}\vec{p}(n+1) \quad .$$
(40)

Just as we did in order to calculate the moments of the mRNA distribution, it will be convenient to define the vectors of partial moments:

$$\vec{n}_{(0)} = \sum_{n=0}^{\infty} n^{0} \vec{p}(n) = \begin{pmatrix} \sum_{n=0}^{\infty} n^{0} p(1,n) \\ \sum_{n=0}^{\infty} n^{0} p(2,n) \end{pmatrix} = \begin{pmatrix} p(1) \\ p(2) \end{pmatrix},$$

$$\vec{n}_{(1)} = \sum_{n=0}^{\infty} n \vec{p}(n) = \begin{pmatrix} \sum_{n=0}^{\infty} n p(1,n) \\ \sum_{n=0}^{\infty} n p(2,n) \end{pmatrix},$$

$$\vec{n}_{(2)} = \sum_{n=0}^{\infty} n^{2} \vec{p}(n) = \begin{pmatrix} \sum_{n=0}^{\infty} n^{2} p(1,n) \\ \sum_{n=0}^{\infty} n^{2} p(2,n) \end{pmatrix}.$$
(41)

It is straightforward to see that the vector $\vec{n}_{(0)}$ is exactly identical to the vector $\vec{m}_{(0)}$. The next two vectors $\vec{n}_{(1)}$ and $\vec{n}_{(2)}$ can be obtained by multiplying equation (41) by *n* and *n*² respectively, and then summing over all n. We end up with the following two equations:

$$0 = \sum_{n=0}^{\infty} n \left[\hat{K} - \frac{b}{1+b} \hat{R} - n\delta \hat{I} \right] \vec{p}(n) + \hat{R} \sum_{n=0}^{\infty} n \sum_{\beta=1}^{n} h(\beta) \vec{p}(n-\beta) + \sum_{n=0}^{\infty} n(n+1)\delta \hat{I} \cdot \vec{p}(n+1) = \\ = \hat{K} \vec{n}_{(1)} - \frac{b}{1+b} \hat{R} \vec{n}_{(1)} - \delta \hat{I} \vec{n}_{(2)} + \delta \hat{I} \left(\vec{n}_{(2)} - \vec{n}_{(1)} \right) + \hat{R} \left(\frac{b}{1+b} \vec{n}_{(1)} + b \vec{n}_{(0)} \right) \\ = \left(\hat{K} - \delta \hat{I} \right) \vec{n}_{(1)} + b \hat{R} \vec{n}_{(0)},$$
(42)

and

$$0 = \sum_{n=0}^{\infty} n^{2} \left[\hat{K} - \frac{b}{1+b} \hat{R} - n\delta \hat{I} \right] \vec{p}(n) + \hat{R} \sum_{n=0}^{\infty} n^{2} \sum_{\beta=1}^{n} h(\beta) \vec{p}(n-\beta) + \sum_{n=0}^{\infty} n^{2}(n+1)\delta \hat{I} \cdot \vec{p}(n+1) =$$

$$= \hat{K} \vec{n}_{(2)} - \frac{b}{1+b} \hat{R} \vec{n}_{(2)} - \delta \hat{I} \vec{n}_{(3)} + \delta \hat{I} \left(\vec{n}_{(3)} - 2\vec{n}_{(2)} + \vec{n}_{(1)} \right) + \hat{R} \left(\frac{b}{1+b} \vec{n}_{(2)} + 2b\vec{n}_{(1)} + b(1+2b)\vec{n}_{(0)} \right) =$$

$$= \hat{K} \vec{n}_{(2)} + \delta \hat{I} \left(-2\vec{n}_{(2)} + \vec{n}_{(1)} \right) + \hat{R} \left(2b\vec{n}_{(1)} + b(1+2b)\vec{n}_{(0)} \right) =$$

$$= \left(\hat{K} - 2\delta \hat{I} \right) \vec{n}_{(2)} + \left(\delta \hat{I} + 2b\hat{R} \right) \vec{n}_{(1)} + b(1+2b)\hat{R} \vec{n}_{(0)} \quad .$$
(43)

Now by multiplying the vector $\vec{u} = (1,1)$ on the left of equations (43) and (44), we find

$$0 = -\delta \langle n \rangle + b \ \vec{r} \ \vec{n}_{(0)} \,, \tag{44}$$

and:

$$0 = -2\delta \langle n^2 \rangle + \delta \langle n \rangle + 2b \ \vec{r} \ \vec{n}_{(1)} + b(1+2b)\vec{r} \ \vec{n}_{(0)} \quad .$$

$$\tag{45}$$

Thus, we find analytical equations for the first two moments of the protein distribution:

$$\left\langle n\right\rangle = \frac{b\ \vec{r}\ \vec{n}_{(0)}}{\delta},\tag{46}$$

$$\left\langle n^{2}\right\rangle = (1+b)\left\langle n\right\rangle + \frac{b\ \vec{r}\ \vec{n}_{(1)}}{\delta}.$$
(47)

Where $\vec{n}_{(1)}$ is the solution of equation (43):

$$0 = \left(\hat{K} - \delta \ \hat{I}\right) \ \vec{n}_{(1)} + b \ \hat{R} \ \vec{n}_{(0)}, \tag{48}$$

Armed with these equations, we can finally compute the stationary variance of the protein distribution:

$$\frac{Var(n)}{\langle n \rangle^2} = \frac{(1+b)\langle n \rangle + \frac{b \ \vec{r} \cdot \vec{n}_{(1)}}{\delta} - \langle n \rangle^2}{\langle n \rangle^2} = \frac{(1+b)}{\langle n \rangle} + \frac{1}{\langle n \rangle^2} \left(\frac{b \ \vec{r} \ \vec{n}_{(1)}}{\delta} - \langle n \rangle^2\right) \ . \tag{49}$$

Exploration of the space of parameter values

In order to test how some of the key qualitative and quantitative conclusions discussed in the main text depend on choice of rate constants that characterize the different architectures, we computed the Fano factor for a large set of parameter values drawn randomly from the space of possible values. The results of these calculations are shown in figures S2, S3, and S4.

REFERENCES

- 1. Sanchez A, Kondev J (2008) Transcriptional control of noise in gene expression. Proc Natl Acad Sci: 105,5081-5086.
- 2. Kepler TB, Elston TC (2001) Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations. Biophys J: 81,3116-3136.
- 3. Berg O (1978) A model for statistical fluctuations of protein numbers in a microbialpopulation. J Theor Biol: 71:587-603.
- 4. Cai L, Friedman N, Xie XS (2006) Stochastic protein expression in individual cells at the single molecule level. Nature: 440, 358-362.
- 5. Yu J, Xiao, J., Ren, X., Lao, K., S. Xie (2006) Probing gene expression in live cells one protein at a time. Science: (311)1600-1603.
- 6. Kennell D, Riezman, H. (1977) Transcription and translation initiation frequencies of the Escherichia coli lac operon. J Mol Biol 114: 1-21.

SUPPLEMENTARY FIGURES



Figure 1. Cartoon depiction of the construction of kinetic rate matrices and vectors. (A) Cartoon representation of the kinetic rate matrix \hat{K} . The diagonal elements represent the net rate at which the promoter abandons each state. For instance, element $\{\hat{K}\}_{11}$ is the rate at which the promoter abandons state 1 due to stochastic association of the activator with the promoter: $\{\hat{K}\}_{11} = -k_A^{on}$, and element $\{\hat{K}\}_{22} = -k_A^{off}$ is the rate of dissociation of the activator from the promoter, abandoning state 2. The non-diagonal element $\{\hat{K}\}_{21} = k_A^{on}$ is the rate at which the promoter makes a transition from state 1 to state 2 (by dissociation association of one activator to the promoter), and the non-diagonal element $\{\hat{K}\}_{12} = k_A^{off}$ is the rate at which the promoter makes a transition from state 1 to state 2 (by dissociation association of one activator to the promoter), and the non-diagonal element $\{\hat{K}\}_{12} = k_A^{off}$ is the rate at which the promoter makes a transition from state 2 to state 1 (by dissociation of the activator). (B) The transcription rate matrix \hat{R} contains, in its diagonal elements, the net rate of transcription at each promoter state. Element $\{\hat{R}\}_{11} = r_1$ is the rate of transcription in promoter state 2. (C) The vector $\vec{r} = (r_1, r_2)$ contains the rates of transcription at states 1 and 2, and is identical to the diagonal of matrix \hat{R} .



Figure 2. Effect of parameter choice on Fano factor for independent and cooperative repression architectures. We sample the parameter space by randomly selecting 10,000 different values for the mean mRNA $\langle m \rangle$ (within 0.005 and 100), k_R^{off} (from $k_R^{off}/\gamma = 0.01$ to $k_R^{off}/\gamma = 100$), $\langle m \rangle_{max}$ (from 5 to 100), and Ω (from 0.001 to 1). The Fano factor is calculated for both independent and cooperative repression architectures, when the mean is the same for both. In the X axis we plot the Fano Factor for independent repression. In the Y axis we plot the Fano factor for cooperative repression. As is the case throughout the paper, we assume that we vary the mean by titrating the amount of repressor inside the cell. Each point in the figure corresponds to two architectures with the same mean. We find that cooperative binding always results in larger cell-to-cell variability than non-cooperative binding. The red solid line marks the region where the Fano factor is the same for both architectures.



Figure 3. Simple activation tends to be noisier than simple repression at low expression levels. We follow the same procedure as in figure S1, and sample the parameter space by randomly selecting 1,000 different values for the mean mRNA $\langle m \rangle$ (within 0.01 and 100), k_R^{off} and k_A^{off} (from 0.01 γ to 100 γ), $\langle m \rangle_{max}$, the enhancement factor f (from 10 to 100). For each one of these 10,000 sets of parameters, we compute the Fano factor for the simple activation and the simple repression architectures. We plot the ratio between the Fano factor for simple activation and repression as a function of the mean. We find that at low mRNA levels ($\langle m \rangle < 10$), the simple activation architecture is noisier than the simple repression architecture in over 99% of the sets of rates tested here. In contrast, at high mRNA levels, it is the other way around. In order for the comparison between both architectures to be meaningful, we have assumed that the repressor and the activator have the same affinity for their operators (even if we vary this affinity over 4 orders of magnitude). The red solid line marks the region where the Fano factor is the same for both architectures (and thus the ratio between the two is 1)



Figure 4. Effect of parameter choice on Fano factor for the repression by DNA looping architecture We sample the parameter space by randomly selecting 10,000 different values for the mean mRNA $\langle m \rangle$ (within 0.005 and 100), k_R^{off} (from $k_R^{off}/\gamma = 0.01$ to $k_R^{off}/\gamma = 100$), $\langle m \rangle_{max}$ (from 5 to 100), k_{loop} (from 0.01 γ to 100 γ), and the parameter characterizing the rate of dissociation in the presence of the auxiliary operator, relative to that in its absence (c). We first assume that c = 1 for all parameter sets (A), and then we randomly sample it within 1 and 10 (B). In the X axis we plot the Fano Factor for simple repression. In the Y axis we plot the Fano factor for repression by DNA looping. As is the case throughout the paper, we assume that we vary the mean by titrating the amount of repressor inside the cell. Each point in the figure corresponds to two architectures with the same mean. We find that whether DNA looping enhances or diminishes noise depends on the value of c. If c = 1, meaning that DNA looping does not affect the rate of dissociation of the repressor from the operator, the Fano factor for the DNA looping architecture is larger than the Fano factor for the simple repression architecture. On the other hand, if c > 1, DNA looping may decrease noise (as observed for ~40% of the parameters chosen). The red solid line marks the region where the Fano factor is the same for both architectures.



Table S1: Kinetic rate matrices for all mechanisms in the text. In the first column, we represent the kinetic mechanisms of gene regulation for all of the architectures considered in the text. In the second and third columns, we show the corresponding promoter kinetic transition rate matrices \hat{K} and the vector $\vec{r} = \vec{u} \hat{R}$ for all of the mechanisms.



Table S2: Fold-change in noise for different promoter architectures. The fold-change in promoter noise is shown as a function of the different kinetic parameters corresponding to each promoter architecture considered throughout the text. Refer to Table I for the definition and value of each rate.