M. J. Dunlop, R. S. Cox, J. H. Levine, R. M. Murray, M. B. Elowitz
Regulatory Activity Revealed by Dynamic Correlations in Gene Expression Noise

Supplementary Figure

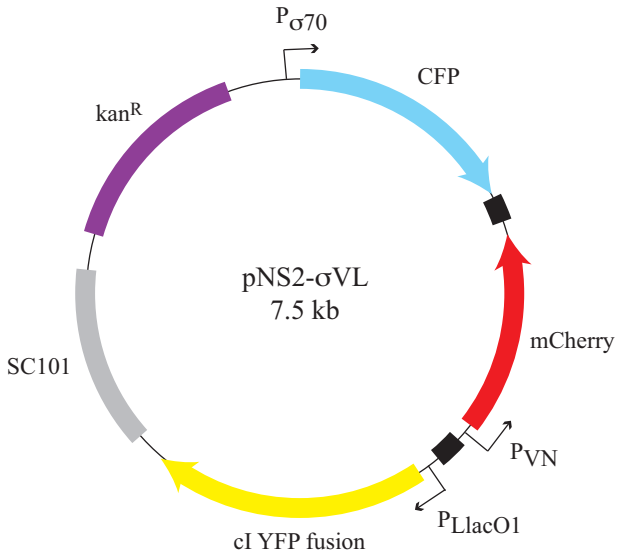

Figure S1: Map of plasmid pNS2-σVL. This plasmid contains the synthetic circuit shown in Fig. 3A of the text. Features include two fluorescent protein genes (*cfp* and *mCherry*), the *cI-yfp* gene fusion, a low-copy SC101 origin of replication, kanamycin resistance gene, and regulating promoters as indicated. In particular, P_{VN} represents the O_R2^* variant of the λ P_R promoter.

Supplementary Note

Analytic Solutions for Cross Correlation Functions Due to Noise

In this section we develop an analytic method for calculating arbitrary cross correlation functions.

We consider the following system of equations

$$\dot{A} = \alpha_A + E + I_A - \beta A \quad (1)$$

$$\dot{B} = \frac{\alpha_B}{1 + (A/K)^n} + E + I_B - \beta B \quad (2)$$

$$\dot{C} = \alpha_C + E + I_C - \beta C, \quad (3)$$

where E represents extrinsic noise and I_i is intrinsic noise for $i = \{A,\ B,\ C\}$. The noise sources have zero mean so the equilibrium point of the deterministic system can be calculated as $A_{eq} = \alpha_A/\beta$, $B_{eq} = \frac{\alpha_B}{\beta(1+(\alpha_A/(\beta K))^n)}$, and $C_{eq} = \alpha_C/\beta$. We expect perturbations due to noise to be small, so it is valid to linearize the system about the equilibrium point. Defining $a = A - A_{eq}$, $b = B - B_{eq}$, and $c = C - C_{eq}$ we obtain the following set of linear dynamics

$$\dot{a} = E + I_A - \beta a \quad (4)$$

$$\dot{b} = ga + E + I_B - \beta b \quad (5)$$

$$\dot{c} = E + I_C - \beta c \quad (6)$$

where g is the *local sensitivity*

$$g = -\frac{\alpha_B\ n\ (\frac{\alpha_A}{\beta K})^{n-1}}{K(1 + (\frac{\alpha_A}{\beta K})^n)^2}.$$

If we assume that the mean value of A is in the center of the Hill function so $A_{eq} = K$, as is the case with our simulated system, then this simplifies to $g = -\frac{\alpha_B n}{4K}$. This analysis is still possible regardless of the steady state value of A, but certain regimes (the middle of the Hill function, saturated edges of the nonlinearity) are better approximated by linear models. Note that the constant g is the only place that information about the nonlinearity enters the equations.

The cross correlation theorem states that cross correlation in the time domain is equal to multiplication in the frequency domain

$$R_{f,g}(\tau) = F^{-1}[\tilde{f}^*\tilde{g}],$$

where * denotes the complex conjugate and $f(t)$ and $g(t)$ are the two signals. However, we average over many cross correlation functions, so we need to calculate the expected value of the cross correlation function over many realizations of the noise

$$E\{R_{f,g}(\tau)\} = E\{F^{-1}[\tilde{f}^*\tilde{g}]\}. \quad (7)$$

Taking the Fourier transform of Eqns. (4)–(6) we find

$$\tilde{a} = \frac{1}{\beta + i\omega}(\tilde{E} + \tilde{I}_A) \quad (8)$$

$$\tilde{b} = \frac{1}{\beta + i\omega}(\tilde{E} + \tilde{I}_B + g\tilde{a}) \quad (9)$$

$$\tilde{c} = \frac{1}{\beta + i\omega}(\tilde{E} + \tilde{I}_C) \quad (10)$$

$$\tilde{E} = \frac{\theta}{\beta + i\omega}\tilde{\eta}_E \quad (11)$$

$$\tilde{I}_i = \frac{\lambda_i}{\kappa + i\omega}\tilde{\eta}_i. \quad (12)$$

Below, we calculate cross correlation expressions for two cases: two independent genes (A and C in Fig. 2 of the main text) and a simple regulatory link with repression (A and B). The first, and simpler, case is worked through in detail, while the results of the second case are summarized.

Unregulated Case

We substitute Eqns. (8)–(12) into Eqn. (7), dropping tildes to simplify notation

$$\begin{aligned} E\{R_{a,c}(\tau)\} &= E\{F^{-1}[\frac{1}{\beta - i\omega}(\frac{\theta}{\beta - i\omega}\eta_E^* + \frac{\lambda_A}{\kappa - i\omega}\eta_a^*)\frac{1}{\beta + i\omega}(\frac{\theta}{\beta + i\omega}\eta_E + \frac{\lambda_C}{\kappa + i\omega}\eta_c)]\} \\ &= E\{F^{-1}[\frac{1}{\beta^2 + \omega^2}(\frac{\theta^2}{\beta^2 + \omega^2}\eta_E^*\eta_E + \frac{\theta\lambda_C}{(\beta - i\omega)(\kappa + i\omega)}\eta_E^*\eta_c \\ &\quad + \frac{\theta\lambda_A}{(\beta + i\omega)(\kappa - i\omega)}\eta_a^*\eta_E + \frac{\lambda_A\lambda_C}{\beta^2 + \omega^2}\eta_a^*\eta_c)]\}. \end{aligned}$$

Because the Fourier transform is a linear operation we can analyze each of the four terms individually. We use two features of white noise to simplify analysis. First, white noise has a flat power spectral density $\eta_i^*(\omega)\eta_i(\omega) = W_i$ and second $E\{\eta_i(t)\eta_j(t)\} = 0$ for $i \neq j$. Thus,

$$\begin{aligned} E\{\eta_i(t)\eta_j(t)\}_{i\neq j} &= E\{F^{-1}[F[\eta_i(t)\eta_j(t)]]\} \\ &= E\{F^{-1}[\eta_i^*(\omega)\eta_j(\omega)]\} \\ &= 0. \end{aligned}$$

Therefore if we have a deterministic function $G(\omega)$

$$\begin{aligned} E\{F^{-1}[G(\omega)\eta_i^*(\omega)\eta_j(\omega)]]\}_{i\neq j} &= E\{\frac{1}{\sqrt{2\pi}}F^{-1}[G(\omega)] \star F^{-1}[\eta_i^*(\omega)\eta_j(\omega)]\} \\ &= \frac{1}{\sqrt{2\pi}}F^{-1}[G(\omega)] \star E\{F^{-1}[\eta_i^*(\omega)\eta_j(\omega)]\} \\ &= 0, \end{aligned}$$

where $\star$ represents convolution. Due to these white noise properties the last three terms in the cross correlation expression become zero and the remaining term simplifies to

$$E\{R_{a,c}(\tau)\} = F^{-1}[\frac{1}{\beta^2+\omega^2}\frac{\theta^2}{\beta^2+\omega^2}\ W_E].$$

Applying the inverse Fourier transform we find

$$E\{R_{a,c}(\tau)\} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{1}{\beta^2+\omega^2}\frac{\theta^2}{\beta^2+\omega^2}\ W_E\ e^{-i\omega\tau}d\omega.$$

This integral can be solved using Cauchy's Residue theorem. Specifically, we need to consider two cases: $\tau < 0$ and $\tau \geq 0$. In the first case we can apply Jordan's lemma if we use a contour that encircles the upper half plane (Fig. T1a). Using Cauchy's Residue theorem we find

$$\lim_{R\to\infty}\int_{C_R} f(z)dz + \int_{-R}^{R} f(z)dz = 2\pi i\sum \text{Res}.$$

By Jordan's lemma the contour at infinity, C_R, becomes zero and we are left with the integral we want to evaluate. In our integral we have a second-order pole at $z = i\beta$ and a second-order pole at $z = -i\beta$. Since we are closing the contour in the upper half plane we need to evaluate the residue at $z = i\beta$:

$$\begin{aligned}\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{1}{\beta^2+\omega^2}\frac{\theta^2}{\beta^2+\omega^2}\ W_E\ e^{-i\omega\tau}d\omega &= \frac{1}{2\pi}2\pi i Res(i\beta)\\ &= \frac{\theta^2 W_E}{4\beta^3}e^{\beta\tau}(1-\beta\tau).\end{aligned}$$

For $\tau \geq 0$ we can use Jordan's lemma if we choose a contour that encircles the lower half plane (Fig. T1b). Since the direction of encirclement is now clockwise, the residue theorem has an additional negative sign:

$$\begin{aligned}\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{1}{\beta^2+\omega^2}\frac{\theta^2}{\beta^2+\omega^2}\ W_E\ e^{-i\omega\tau}d\omega &= \frac{1}{2\pi}(-2\pi i Res(-i\beta))\\ &= \frac{\theta^2 W_E}{4\beta^3}e^{-\beta\tau}(1+\beta\tau).\end{aligned}$$

Combining these two results we find

$$E\{R_{a,c}(\tau)\} = \frac{\theta^2 W_E}{4\beta^3}e^{-\beta|\tau|}(1+\beta|\tau|).$$

Note that this expression for the cross correlation used a and c, the mean subtracted versions of A and C. These expressions are consistent with those calculated directly from the nonlinear simulations because we used the mean subtracted version of the cross correlation function.

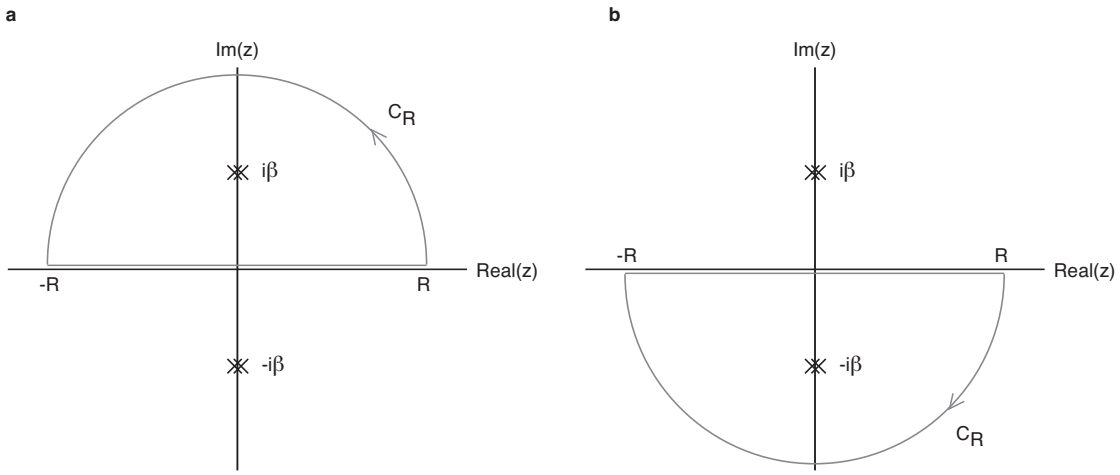

Figure T1: Contours used for evaluating Cauchy's Residue theorem. Pole locations shown are for the unregulated cross correlation expression.

Regulated Case

Analysis of the cross correlation function due to repression is similar. After simplification using the white noise properties discussed in the previous section we find

$$\begin{aligned} E\{R_{a,b}(\tau)\} &= F^{-1}[\frac{\theta^2\ W_E}{(\beta+i\omega)^2(\beta-i\omega)^2}] \\ &+ F^{-1}[\frac{g\ \theta^2\ W_E}{(\beta+i\omega)^3(\beta-i\omega)^2}] \\ &+ F^{-1}[\frac{g\ \lambda_A^2\ W_A}{(\beta+i\omega)^2(\beta-i\omega)(\kappa+i\omega)(\kappa-i\omega)}]. \end{aligned}$$

These three terms have a convenient interpretation: The first term is the artificial correlation due to extrinsic noise, the second term is extrinsic noise that has propagated through the link, the third term is intrinsic noise that has propagated through the link. Cauchy's Residue theorem and Jordan's lemma are applied to find

$$R_{a,b}(\tau) = \begin{cases} \frac{\theta^2 W_E}{4\beta^3} e^{\beta\tau}(1-\beta\tau) \\ +\frac{g\theta^2 W_E}{16\beta^4} e^{\beta\tau}(3-4\beta\tau+2\beta^2\tau^2) \\ +\lambda_A^2 g W_A (e^{\beta\tau}\frac{\kappa^2(1-2\beta\tau)-\beta^2(5-2\beta\tau)}{4\beta^2(\beta^2-\kappa^2)^2} + e^{\kappa\tau}\frac{1}{2\kappa(\beta-\kappa)^2(\beta+\kappa)}) & \tau < 0 \\ \\ \frac{\theta^2 W_E}{4\beta^3} e^{-\beta\tau}(1+\beta\tau) \\ +\frac{g\theta^2 W_E}{16\beta^4} e^{-\beta\tau}(3+2\beta\tau) \\ +\lambda_A^2 g W_A (\frac{e^{-\beta\tau}}{4\beta^2(\kappa^2-\beta^2)} + \frac{e^{-\kappa\tau}}{2\kappa(\beta+\kappa)^2(\beta-\kappa)}) & \tau \geq 0. \end{cases}$$

Summary

The cross correlation relations are summarized as

$$
\begin{aligned}
E\{R_{A,C}(\tau)\} &= N_{A,C}\frac{\theta^2 W_E}{4\beta^3}e^{-\beta|\tau|}(1+\beta|\tau|) \\
E\{R_{A,B}(\tau)\} &= \begin{cases} N_{A,B}(\frac{\theta^2\ W_E}{16\beta^4}e^{\beta\tau}(2g\beta^2\tau^2-4\beta(g+\beta)\tau+3g+4\beta) \\ +\lambda_A^2 g W_A(e^{\beta\tau}\frac{\kappa^2(1-2\beta\tau)-\beta^2(5-2\beta\tau)}{4\beta^2(\kappa^2-\beta^2)^2}+e^{\kappa\tau}\frac{1}{2\kappa(\beta-\kappa)^2(\beta+\kappa)})) & \tau<0 \\ \\ N_{A,B}(\frac{\theta^2\ W_E}{16\beta^4}e^{-\beta\tau}(2\beta(g+2\beta)\tau+3g+4\beta) \\ +\lambda_A^2 g W_A(e^{-\beta\tau}\frac{1}{4\beta^2(\kappa^2-\beta^2)}+e^{-\kappa\tau}\frac{1}{2\kappa(\beta+\kappa)^2(\beta-\kappa)})) & \tau\geq 0 \end{cases}
\end{aligned}
$$

where the normalization factors are

$$
\begin{aligned}
N_{A,B} &= \frac{1}{\sqrt{R_{A,A}(0)R_{B,B}(0)}} \\
N_{A,C} &= \frac{1}{\sqrt{R_{A,A}(0)R_{C,C}(0)}} \\
R_{A,A}(0) &= \frac{\theta^2 W_E}{4\beta^3}+\frac{\lambda_A^2 W_A}{2\beta\kappa(\kappa+\beta)} \\
R_{B,B}(0) &= \frac{\theta^2(3g^2+6g\beta+4\beta^2)\ W_E}{16\beta^5}+\frac{\lambda_A^2 g^2(\kappa+2\beta)\ W_A}{4\kappa\beta^3(\kappa+\beta)^2}+\frac{\lambda_B^2\ W_B}{2\kappa\beta(\kappa+\beta)} \\
R_{C,C}(0) &= \frac{\theta^2 W_E}{4\beta^3}+\frac{\lambda_C^2 W_C}{2\beta\kappa(\kappa+\beta)}.
\end{aligned}
$$

To compare these results to the nonlinear simulation we use the constants specified in Table T1, where $g=-\frac{\alpha_B n}{4K}$. W_E, W_i for $i=\{A,B,C\}$ are treated as binary variables and set to 0 or 1 to turn off and on extrinsic and intrinsic noise for comparison to the full nonlinear simulations shown in the text. We assume $W_A=W_B=W_C$ for simplicity. Fig. 2 in the text shows that the analytic solutions for cross correlations match the simulated nonlinear system extremely well.

We have calculated the cross correlation functions for two types of regulation. This method can be applied more generally to larger networks, provided perturbations due to noise are small enough that linearization is a valid approximation.

Sensitivity Analysis

Two prominent features of the cross correlation curve for repression are the location of the dip, τ_{reg}, and the magnitude of the dip, M (shown schematically in Fig. T2). We calculate how sensitive these features are to variations in the system parameters. These results indicate which parameters play a primary role in setting the features of the cross correlation function.

To find τ_{reg} for repression we take

$$
\frac{dR_{a,b}(\tau)}{d\tau}\Big|_{\tau=\tau_{reg}}=0.
$$

Parameter	Value
α_A	1.39 molecules/cell/min
α_B	4.5 molecules/cell/min
α_C	1.39 molecules/cell/min
β	0.0116 1/min
K	120 nM
n	1.7
κ	0.139 1/min
θ	0.0532 (molecules/cell)$^{1/2}$/min
λ_A	0.621 (molecules/cell)$^{1/2}$/min
λ_B	1.12 (molecules/cell)$^{1/2}$/min
λ_C	0.621 (molecules/cell)$^{1/2}$/min
T_{cc}	60 mins
T_{int}	5 mins

Table T1: Simulation Parameters

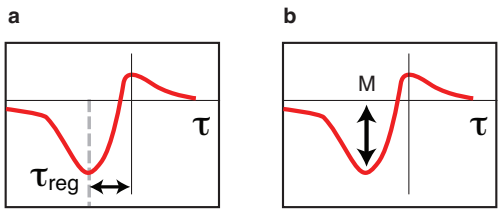

Figure T2: Schematic of cross correlation function features (a) τ_{reg} and (b) M.

In general it is not possible to find a closed form solution for τ_{reg}, however numerical root finding methods can be applied. The nominal parameter values for sensitivity analysis are those listed in Table T1. For a feature of the cross correlation function, y, we find the normalized sensitivity

$$S_i = \frac{y(p_i + \Delta p_i) - y(p_i - \Delta p_i)}{2\ \Delta p_i\ y(p_i)}.$$

For $\Delta p_i = 0.05 p_i$ the parameteric sensitivities are shown in Table T2. Large values indicate that the feature y is very sensitive to that parameter. The sign of the sensitivity indicates whether y will get larger ($S_i > 0$) or smaller ($S_i < 0$) as the parameter is increased.

Parameter	τ_{reg}	M
β	-108.5	2.39
g	18.9	-85.7
θ	11.3	5.5
κ	4.0	-12.4
λ_A	-0.9	2.5
λ_B	0.0	0.0
λ_C	0.0	0.0

Table T2: Normalized Sensitivities

τ_{reg} is most sensitive to the parameter β, which sets the time scale of both protein decay and extrinsic noise. As the cell cycle ($\log(2)/\beta$) gets longer, the location of the dip moves further away from zero. M is most sensitive to the local sensitivity g, which is negative for repression. As g becomes less negative, the repressor has less of an effect on its target and the dip gets smaller. In the extreme case when $g = 0$, which indicates an inactive or non-existent regulatory connection, the dip disappears.

Degenerate Cross Correlation Functions

Different regulatory networks can generate similar cross correlation functions. This property is referred to as the "degeneracy" of the cross correlation function, and can interfere with attempts to use the cross correlation function to uniquely identify unknown regulatory architectures. Here we explore how degeneracy appears and ways in which its limits may be partially circumvented.

Redundant Network Elements

The problem is shown clearly with the network diagram in Fig. T3. Here we assume that regulatory parameters are identical for all the links (except for the sign of regulation, repression vs. activation). Proteins B and C are redundant because they are controlled by the same input, A, and extrinsic noise affects them in the same way. Thus, in the extreme case where only extrinsic noise is present, $B(t)$ and

$C(t)$ will be identical. Consequently, the cross correlations of each with D are equal, $R_{B,D}(\tau) = R_{C,D}(\tau)$, even though there is no link between B and D.

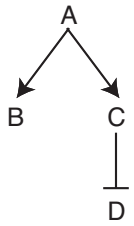

Figure T3: Network with redundant components. Cross correlation functions $R_{B,D}(\tau)$ and $R_{C,D}(\tau)$ may look similar.

Intrinsic noise helps to discriminate between $R_{B,D}(\tau)$ and $R_{C,D}(\tau)$ because intrinsic noise in C propagates to D causing additional time-lagged correlation. Intrinsic noise in B, on the other hand, does not propagate to affect D. Thus, one obtains an additional time-lagged component in $R_{C,D}(\tau)$ compared to $R_{B,D}(\tau)$.

More specifically, analytic solutions for the two cross correlation functions are summarized by

$$\begin{aligned} R_{B,D}(\tau) &= E\{F^{-1}[f^*_{Be} f_{De} + f^*_{BA} f_{DA}]\} \\ R_{C,D}(\tau) &= E\{F^{-1}[f^*_{Ce} f_{De} + f^*_{CA} f_{DA} + f^*_{CC} f_{DC}]\}, \end{aligned}$$

where $f_{i,j}$ is the Fourier transform of the differential equations describing the dynamics of protein j in response to noise from i, * is the complex conjugate, e is extrinsic noise, and E refers to the expected value. Assuming the network connections are identical for A-B and A-C we find the difference between the two cross correlations

$$R_{C,D}(\tau) - R_{B,D}(\tau) = E\{F^{-1}[f^*_{CC} f_{DC}]\}.$$

This term represents propagation of noise in C to affect D. It is non-zero due to intrinsic noise in C.

Fig. T4 compares $R_{B,D}(\tau)$ and $R_{C,D}(\tau)$. $R_{C,D}(\tau)$ is lower because intrinsic noise in C propagates through the repression link. The difference increases when intrinsic noise in C gets larger.

This simple example illustrates how redundant network elements can confound analysis. More complicated networks will have similar problems any time there are two or more network elements that are controlled by the same inputs. Intrinsic noise or other signals that affect individual genes can help to distinguish correlations due to regulation from correlations due to symmetric network elements.

This degeneracy arises due to the effective redundancy of two genes, B and C, which are regulated identically and have similar function. In practice, such redundancy may be unusual. For example, in the case of the galactose network considered in the text, GalR and GalS have similar function, and even similar binding sites. However, GalS and not GalR, is regulated by CRP. Thus, the two genes are not redundant.

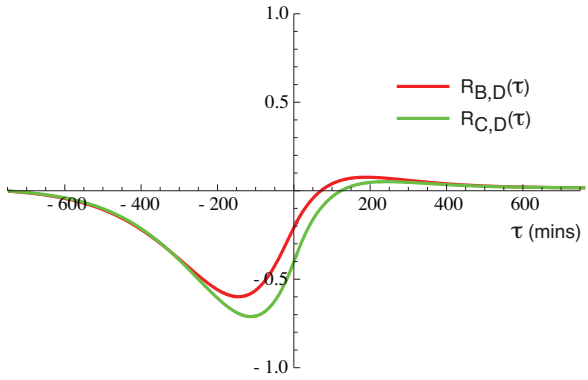

Figure T4: Cross correlation functions for network with redundant elements.

Parametric Degeneracies

A second class of degeneracies comes from uncertainty in parameters. Consider the two cascades in Fig. T5. If we measure $R_{A,C}(\tau)$ is it possible to determine that there is a middle element in the network or will the cross correlation function be indistinguishable from that of $R_{X,Y}(\tau)$ in certain cases? In this example extrinsic

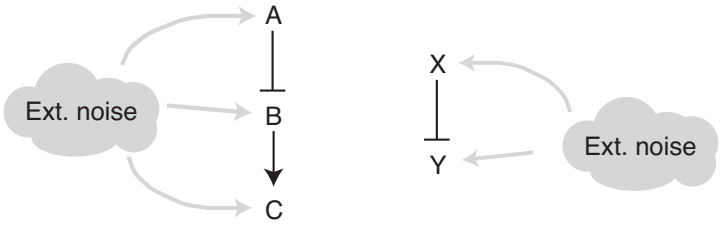

Figure T5: Example of a network showing parametric degeneracy. Gray arrows show how extrinsic noise affects the system.

noise is helpful in discriminating between the two cascades. In the two-step cascade, extrinsic noise affects A, B, and C, and thus enters into the cross correlation function in three ways, while in the one-step cascade it only enters twice.

Without extrinsic noise it is possible to adjust parameters to make the two cross correlation functions look nearly identical (Fig. T6a). For example, if B degrades quickly, but the net strength of the two networks are the same then the resulting cross correlation functions are very similar. When extrinsic noise is considered, the two cross correlations separate (Fig. T6b). Extrinsic noise increases overall positive correlation near $\tau = 0$ for both cascades, but the effect is more pronounced in the longer cascade. Additionally, extrinsic noise increases the the effective regulatory timescale of repression in the longer cascade, shown by the location of the dip at $\tau < 0$.

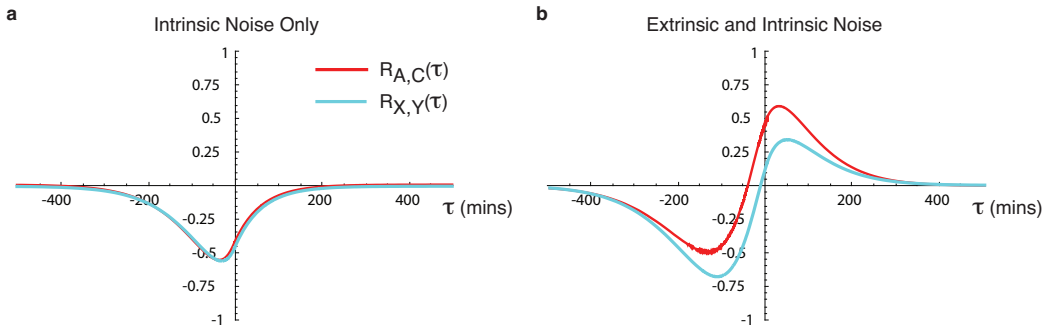

Figure T6: Cross correlation functions for two cascades with degenerate parameters. (a) No extrinsic noise, $W_E = 0$. (b) Extrinsic and intrinsic noise. Simulation parameters are $\beta_x = \beta_y = \beta_a = \beta_c = 1/60$ mins^{-1}, $\beta_b = 1/5$ mins^{-1}, $g_{xy} = g_{ab}g_{bc} = -0.01$, $W_i = 1$, $W_E = 0.064$.

In this example the intrinsic noise-only dynamics are described by

$$\begin{aligned} \dot{a} &= -\beta_a a + \eta_a \\ \dot{b} &= -\beta_b b + \eta_b + g_{ab} a \\ \dot{c} &= -\beta_c c + \eta_c + g_{bc} b \end{aligned}$$

$$\begin{aligned} \dot{x} &= -\beta_x x + \eta_x \\ \dot{y} &= -\beta_y y + \eta_y + g_{xy} x \end{aligned}$$

where intrinsic noise has been approximated by white noise and parameters are given in the figure caption.

Since biochemical parameters are often unknown or uncertain there are many possible situations where the shape of two cross correlation functions may look very similar, but knowledge of noise and network properties can help to clarify cross correlations, as described here.