The effect of feedback on the fidelity of information transmission of time-varying signals

Wiet Hendrik de Ronde,[1,*] Filipe Tostevin,[1] and Pieter-Rein ten Wolde[1]

[1] *FOM Institute for Atomic and Molecular Physics, Science Park 104, 1098 XG, Amsterdam*

(Dated: February 12, 2010)

Living cells are continually exposed to environmental signals that vary in time. These signals are detected and processed by biochemical networks, which are often highly stochastic. To understand how cells cope with a fluctuating environment, we therefore have to understand how reliably biochemical networks can transmit time-varying signals. To this end, we must understand both the noise characteristics and the amplification properties of networks. In this manuscript, we use information theory to study how reliably signalling cascades employing autoregulation and feedback can transmit time-varying signals. We calculate the frequency-dependence of the gain-to-noise ratio, which reflects how reliably a network transmits signals at different frequencies. We find that the gain-to-noise ratio may differ qualitatively from the power spectrum of the output, showing that the latter does not directly reflect signaling performance. Moreover, we find that auto-activation and auto-repression increase and decrease the gain-to-noise ratio for all of frequencies, respectively. Positive feedback specifically enhances information transmission at low frequencies, while negative feedback increases signal fidelity at high frequencies. Our analysis not only elucidates the role of autoregulation and feedback in naturally-occurring biological networks, but also reveals design principles that can be used for the reliable transmission of time-varying signals in synthetic gene circuits.

I. INTRODUCTION

Living cells constantly have to respond and adapt to a changing environment. In some cases, such as in response to a changing sugar concentration [1], a cell may wish to integrate out rapid variations and only respond to slow variations of the environmental signal, while in other cases, such as osmo adaptation [2] or bacterial chemotaxis [3], the cell needs to do the opposite — respond to rapid but not slow variations (adaptation). Indeed, to understand how cells cope with a fluctuating environment, we have to understand how cells transduce time-varying signals. Cells detect, process, and transduce signals via biochemical networks, which are the information processing devices of life. However, experiments in recent years have demonstrated that biochemical networks are often highly stochastic [4, 5]. This raises the question how reliably biochemical networks can transmit time-varying signals in the presence of noise.

Interestingly, biochemical networks exploit commonly recurring architectures [6, 7], such as autoregulation, cascades, and feedback, to process signals. These network motifs often implement signal amplification in order to raise the level of the input signal relative to the noise. Amplification can be characterised by the *gain*, the fold-change in the signal amplitude. However, it is important to recognise that such amplification can not only increase the levels of the desired signal, but can also amplify the noise itself. Therefore, to understand the possibilities and limitations of different network motifs for enhancing the fidelity of signal transduction, we need to understand how both the signal and the noise are propagated through these motifs. Specifically, information theory indicates that the reliability of signal transmission is determined by the ratio of the gain of the network to the total noise in the output signal — the gain-to-noise ratio. Moreover, to assess how reliably signals of different temporal characteristics are transduced, we have to understand the frequency dependence of the gain and the noise. Importantly, we expect that different network architectures will affect the frequency-dependence of the gain and the noise differently, which means that we have to study both these quantities. In this manuscript, we study the frequency-dependence of the gain-to-noise ratio for simple cascades, and for cascades employing autoregulation and feedback. This allows us to elucidate how autoregulation and feedback can shape the frequency range over which signals can be transduced reliably.

Information theory provides a formalism for quantifying the reliability of information transmission in the presence of noise [8]. A natural measure for the fidelity of signal transmission from an input signal S to an output signal X (the network response) is the mutual information between S and X, which is defined as

$$\begin{aligned} I(S,X) &= H(S) - H(S|X) \\ &= -\int dS\, p(S) \log\left[p(S)\right] - \\ &\left(-\int dX\, p(X) \int dS\, p(S|X) \log\left[p(S|X)\right]\right). \end{aligned} \quad (1)$$

Here, $p(S)$ and $p(X)$ are the probability distributions of possible input and output signals respectively, and $p(S|X)$ is the conditional probability of S once X is specified. The mutual information quantifies the reduction in entropy of (or uncertainty about) the signal after one ob-

*deronde@amolf.nl

tains knowledge of the network response, averaged over all possible responses. In other words, $I(S,X)$ is how much we learn (on average) about S by measuring X. For a deterministic system, every S leads to a unique X (we assume no degeneracy). Measuring X thus precisely specifies S, such that the uncertainty in S after a measurement of X is $H(S|X) = 0$ and $I(S,X) = H(S)$. However, in the presence of noise in the network each input S will lead to a distribution of possible outputs X. As a result, an observed X can correspond to multiple S values and $I(S,X) \leq H(S)$. For completely uncorrelated S and X, $I(S,X) = 0$. By construction, the mutual information is symmetric, such that $I(S,X) = I(X,S)$.

Recently, the mutual information has been used to study the reliability of information transmission in biochemical networks [9–12]. However, these studies considered only the steady-state response of a network to a distribution of *constant* input signals, which do not change on the timescale of the network response. Yet, in many biological systems, it cannot be assumed that the input signal is constant on the timescale of the network response.

Indeed, in many systems the message is encoded in the *temporal dynamics* of the input signal. A well-known example is bacterial chemotaxis, where the concentration of the intracellular messenger protein depends not on the steady-state ligand concentration, but rather on the change of this concentration in the recent past [13] — the response of the network thus depends on the history of the input signal. Moreover, the extracellular signal may be encoded in the temporal dynamics of the intracellular signal transduction pathway. An interesting example is provided by the rat PC-12 system: while stimulation with a neuronal growth factor gives rise to a sustained response of the Raf-Mek-Erk pathway, stimulation by an epidermal growth factor gives rise to a transient response of this pathway [14]. In all these cases, the message is encoded not in the concentration of some chemical species at a given moment in time, but rather in its concentration as a function of time. This means that to understand how reliably the network can transmit information, we need to know how accurately an input signal as a function of time — the input *trajectory* $s(t)$ — can be mapped onto an output trajectory $x(t)$. We thus need to understand the mutual information between the two trajectories, $I(s(t), x(t))$.

The ability of a biochemical network to transduce a time-varying input signal depends on the correlation time of the input signal and the architecture and response dynamics of the network. An instructive example is provided by the chemotaxis network of the bacterium *Escherichia coli*. This network employs integral negative feedback [15], as a result of which the intracellular messenger protein can adapt to a constant extracellular ligand concentration. This means that the signalling network cannot respond to changes in ligand concentration that occur on time scales longer than the adaptation time. At the other end of the frequency spectrum, changes in the messenger protein that occur on time scales shorter than the motor switching time will be integrated out; indeed, the network cannot respond reliably to rapidly varying input signals [16]. The architecture and the response dynamics of the processing network thus determines the frequency range over which signals can be transduced reliably.

Recently, we have applied information theory to biochemical networks and studied the mutual information between in- and output trajectories, $I(s(t), x(t))$ [16]. Here, we apply this framework to study the propagation of time-varying signals through a number of network motifs—cascades, autoregulation, and feedback. It is known that for *constant* signals (or, to be more precise, signals that do not vary on the time scale of the network response time), the mutual information decreases as a function of cascade length [11]. The same also holds true for time-varying signals. Indeed, the data-processing inequality states that in a cascade with n nodes, the information about the input encoded in the signal at node $i+1$ cannot be greater than the information at node i. Once lost, information about the input cannot be recovered later in the cascade. Simply increasing the length of a signalling cascade therefore can never increase the transmitted information. Conversely, maximising the total transmitted information cannot be the driving force behind the evolution of such cascades.

Cascades, however, often employ autoregulation and feedback, which can be used to shape the response of the network to signals of different frequencies. Importantly, autoregulation and feedback affect not only the frequency-dependent gain, which describes how strongly an input signal at a particular frequency is amplified in the absence of any biochemical noise, but also the frequency-dependence of the noise. While the frequency-dependence of the gain [17–19] and the noise [20] have been studied separately, the frequency-dependence of their ratio, the gain-to-noise ratio, has not. However, it is the gain-to-noise ratio which determines how reliably an input signal at a particular frequency can be transmitted [16]. In fact, as we will show, autoregulation and feedback affect the frequency-dependence of the gain and the noise differently, which means that it is essential to study these quantities together.

In this manuscript, we study the frequency-dependent gain-to-noise ratio using a Gaussian model. In the next section, we describe this model, and how we can use it to compute the frequency-dependent gain-to-noise ratio and the information transmission rate, which is given by the integral of this ratio over all frequencies [16]. In section *Results* we discuss the frequency-dependent gain-to-noise ratio of simple cascades, and cascades employing feedback and autoregulation. Our results highlight the idea that the output power spectrum is not a direct measure for the information content of the output signal—the output power spectrum can differ *qualitatively* from the spectrum of the gain-to-noise ratio. We also show (Fig. 7) that positive regulation tends to increase the

gain-to-noise ratio, while negative regulation tends to decrease it. Moreover, we show that the frequency spectra of motifs with negative feedback can exhibit windows in which the gain-to-noise ratio is increased; these motifs can thus act as band-pass filters for information transmission. Finally, we discuss some of the implications of our findings and the limitations of our analysis.

II. METHODS

We consider information transmission through a biochemical network from an input signal $s(t)$ to an output signal $x(t)$. The dynamics of the network can be described mathematically by a set of coupled Langevin equations [21] for the signal, response and an arbitrary number of intermediate components v_i, in vector form $\boldsymbol{v}$. In using the Langevin representation we assume that the copy number of each component is large such that the discrete number of molecules can be approximated by a continuous concentration.

$$\frac{ds}{dt} = f_s^+(s) + f_s^-(s) + \Gamma(t) \tag{2a}$$

$$\frac{d\boldsymbol{v}}{dt} = \boldsymbol{f}_v^+(s,\boldsymbol{v},x) - \boldsymbol{f}_v^-(s,\boldsymbol{v},x) + \boldsymbol{\eta}_v(t) \tag{2b}$$

$$\frac{dx}{dt} = f_x^+(s,\boldsymbol{v},x) - f_x^-(s,\boldsymbol{v},x) + \eta_x(t). \tag{2c}$$

Here, f_i^+ and f_i^- contain all the reactions involving the production and degradation of component i, respectively. f_s^+ and f_s^- only depend on s, so that we restrict our analysis to networks that do not feed back onto s itself. In these cases, the gain-to-noise ratio is independent of the input signal [22], as discussed in more detail below Eqn. 10. $\Gamma(t)$ is a stochastic driving process that serves to define the ensemble of possible input signals. The various noise sources η_i are taken to be independent and Gaussian-distributed [22–24], such that $\langle \eta_i(t)\,\eta_j(t')\rangle = \left\langle |\eta_i|^2 \right\rangle \delta_{ij}\delta(t-t')$. Here, we note that the assumption of independent noise sources is only made to simplify the analysis. (Anti)-correlations between noise sources can affect noise propagation [22], and can be included by a straightforward extension of the present discussion. Furthermore, we assume that $\left\langle |\eta_i|^2 \right\rangle = \langle f_i^+ \rangle + \langle f_i^- \rangle = 2\langle f_i^+ \rangle$ and $\left\langle |\Gamma|^2 \right\rangle = 2\langle f_s^+ \rangle$ [25], the sum of the production and degradation terms.

We introduce the vector $\boldsymbol{y} = (s;\boldsymbol{v};x)$ and $\boldsymbol{\eta} = (\Gamma, \boldsymbol{\eta}_v, \eta_x)$ and assume the network has a steady state $\langle \boldsymbol{y} \rangle$. Linearizing around this steady state we obtain

$$\frac{d\tilde{\boldsymbol{y}}}{dt} = \mathbf{J}|_{\boldsymbol{y}=\langle \boldsymbol{y} \rangle}\tilde{\boldsymbol{y}} + \boldsymbol{\eta}. \tag{3}$$

Here $\tilde{y}_i = y_i - \langle y_i \rangle$ is the deviation of the concentration of component i from its steady-state value, $\langle y_i \rangle$, and $\mathbf{J}$ is the Jacobian evaluated at the steady state [26]. J_{ij} describes the response of the component i to small changes in component j, while keeping all other components at their steady-state levels. The diagonal element $J_{ii} = -\tau_i^{-1}$ is the relaxation time or dissipative time scale of component i; it describes the time scale on which component i relaxes back to its steady-state value after a perturbation. After linearization, the architecture of the network is encoded in the structure of the Jacobian matrix (see Fig. 1): the diagonal terms correspond to autoregulation, the lower triangular part to downstream (feedforward) regulation and the upper triangular part to upstream (feedback) regulation. Since we restrict ourselves to systems without feedback from the network to the signal itself, we require that all elements on the first row of $\mathbf{J}$ are zero but for J_{ss}.

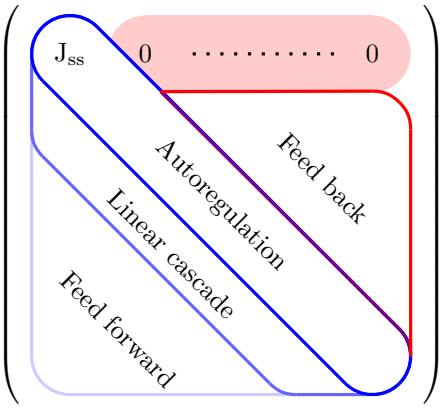


FIG. 1. The Jacobian matrix. The entries of the Jacobian matrix encode the structure of the reaction network.

We take as our input signal the variations $\tilde{s}$. A linear system does not change the frequency of the transmitted signal, but only the amplitude and the phase. Since Eqn. (3) is linear in $\tilde{\boldsymbol{y}}$, we can calculate exactly the power spectra of the network components [25],

$$\mathbf{P} = [i\omega\mathbf{I} - \mathbf{J}]^{-1}\,\boldsymbol{\Xi}\,\left[-i\omega\mathbf{I} - \mathbf{J}^T\right]^{-1}, \tag{4}$$

where $P_{ij}(\omega) = \left\langle \tilde{Y}_i(\omega)\tilde{Y}_j(-\omega) \right\rangle$ is the (cross-)power spectrum of $\tilde{y}_i$ and $\tilde{y}_j$, $\tilde{Y}_i(\omega)$ is the Fourier transform of $\tilde{y}_i(t)$, $\mathbf{I}$ is the identity matrix, and $\boldsymbol{\Xi}$ is the noise matrix with entries $\Xi_{ij} = \langle \eta_i(\omega)\eta_j(-\omega)\rangle = \left\langle |\eta_i|^2 \right\rangle \delta_{ij}$. The power spectrum is a commonly used tool to study time-varying signals, and describes how the total power of a signal is distributed over different frequencies. Power at low frequencies is related to slow variations of the signal, while power at high frequencies corresponds to rapid fluctuations. The integral of the power spectrum over all frequencies equals the total variance of the signal.

The information transmission rate for time-varying sig-

nals is [27, 28]

$$\lim_{T\to\infty}\frac{I\left(s\left(t\right),x\left(t\right)\right)}{T}=R\left(s\left(t\right),x\left(t\right)\right)$$
$$=-\frac{1}{2\pi}\int_0^\infty d\omega\ln\left[1-\Phi_{sx}(\omega)\right], \quad (5)$$

where T is the length of the trajectory and $\Phi_{sx}(\omega)$ is the coherence function, defined as

$$\Phi_{sx}(\omega)=\frac{\left|P_{sx}(\omega)\right|^2}{P_{ss}(\omega)P_{xx}(\omega)}. \quad (6)$$

$\Phi_{sx}(\omega)$ is a measure of the average correlation between the in- and output signals in the frequency domain. For completely independent in- and output signals, $\Phi_{sx}(\omega)=0$, while for a noiseless system $\Phi_{sx}(\omega)=1$.

The power spectrum of the output signal, $P_{xx}(\omega)$, can be decomposed as

$$P_{xx}(\omega)\equiv\Sigma(\omega)+N(\omega), \quad (7)$$
$$\equiv g^2(\omega)P_{ss}(\omega)+N(\omega). \quad (8)$$

Here, $\Sigma(\omega)\equiv g^2\left(\omega\right)P_{ss}(\omega)$ is the transmitted signal, $g^2(\omega)\equiv\left|P_{sx}(\omega)\right|^2/P_{ss}^2(\omega)$ is the frequency-dependent gain, $P_{ss}(\omega)$ is the power spectrum of the input signal and $N(\omega)$ is the frequency-dependent noise. With these definitions, the coherence function, Eq. 6, can be recast as

$$\Phi_{sx}(\omega)=\frac{\Sigma(\omega)}{N(\omega)+\Sigma(\omega)}, \quad (9)$$

and the mutual information rate can be rewritten as [16]

$$R\left(s\left(t\right),x\left(t\right)\right)=\frac{1}{2\pi}\int_0^\infty d\omega\ln\left[1+\frac{g^2(\omega)}{N(\omega)}P_{ss}(\omega)\right]. \quad (10)$$

We see that the information transmission rate depends on the power spectrum of the input signal, $P_{ss}(\omega)$, and on the gain-to-noise ratio $g^2(\omega)/N(\omega)$.

As discussed in Ref. [22], in a biological system the reaction that detects the input signal can, depending on the nature of the detection reaction, introduce significant correlations between the variations in the input signal and the intrinsic noise of the reactions that constitute the processing network. These correlations are a consequence of the molecular character of the components and are thus unique to biochemical networks. If the detection reaction does not introduce correlations, then Eqn. 8 is the spectral-addition rule [22]. The noise $N(\omega)$ is then the intrinsic noise of the processing network and also $g^2(\omega)$ only depends on properties of the processing network. On the other hand, if the detection reaction does introduce correlations, then the output power spectrum $P_{xx}(\omega)$ can be written in the form of Eqn. 8, but then $N(\omega)$ and $g^2(\omega)$ depend not only on characteristics of the processing network, but also on the statistics of the input signal; conversely, the variations of the input will also be affected by the noise in the processing network[22]. In what follows below, we assume for simplicity that the spectral-addition rule holds, which means that the gain, noise and gain-to-noise ratio are independent of the input signal, and that the input does not need to be specified.

Applying the linearization procedure outlined above may, in general, qualitatively change the dynamics of the network being considered. However, previous studies [9, 29] have shown that the Linear Noise Approximation provides an accurate description of many systems if the average copy numbers are of order 10 molecules or more. For the networks considered in this paper we also compared the power spectra calculated in the linear approximation with the results of stochastic simulations performed with Gillespie's algorithm [30], and again found good agreement when protein copy numbers are large (see S1). We therefore expect that the linear analysis presented in this paper provides an accurate description of the signalling characteristics of these networks.

III. RESULTS

First we study a simple cascade, where "simple" means that we consider a cascade where each component only regulates the activity of the next component in the cascade; a "simple" cascade is thus a cascade without autoregulation, feedback or feedfoward. We analyse this network in detail such that it can serve as an instructive example of the method described above. In addition, we will highlight general features of the results which recur in more complex networks. We then discuss network motifs including autoregulation and negative feedback loops, which are commonly observed in biochemical networks.

To understand the effects of autoregulation and feedback we will compare information transmission in these motifs to a corresponding simple cascade with the same number of components but without the additional regulation. In order to perform such a comparison of different motifs on an equal footing we constrain the average production rate of every component such that these are equal in the networks under comparison. We argue that from a biological perspective the rate of protein production is a more significant constraint on network design than average protein copy number, since the latter only depends on the *ratio* of the synthesis and degradation rate, while it is the absolute synthesis and degradation rate that determines the cost of having a protein at a particular copy number. This constraint also enforces that the noise strength at each level of the cascade $\left\langle\left|\eta_i\right|^2\right\rangle=2\left\langle f_i^+\right\rangle$ is the same in the motifs being compared. When comparing two systems with many parameters, equalising production rates is not a sufficient constraint to uniquely specify all parameter values. To reduce this potential parameter space we will (unless otherwise stated) hold constant as many of the network parameters as possible. For brevity we will only discuss networks in which all reg-

ulation occurs via the production reactions, with linear degradation of each component. However, our results are qualitatively unchanged if we instead consider regulation via protein degradation.

We characterise information transmission through these motifs in terms of the gain, noise and gain-to-noise ratio. Since we assume that the spectral-addition rule holds [22], these quantities are intrinsic, signal-independent properties of the network. We also wish to highlight differences between the information transmission characteristics of the network, as determined by the gain-to-noise ratio, and the output power spectrum $P_{xx}(\omega)$, since this is commonly discussed in studies of signal transmission. Since $P_{xx}(\omega)$ depends not only on the processing network but also on the input signal (see Eq. 8), we must therefore specify $P_{ss}(\omega)$; for this purpose we assume, for convenience, that the input signal $s(t)$ is generated via a Poisson birth-death process as in Eqn. 11a (*The simple cascade*).

A. The simple cascade

Initially we study a simple cascade with a single intermediate component. Extension of the cascade with more intermediate components is straightforward. The appropriate reaction scheme is

$$\frac{ds}{dt} = k_s - \mu_s s + \Gamma(t) \tag{11a}$$

$$\frac{dv}{dt} = k_v s - \mu_v v + \eta_v(t) \tag{11b}$$

$$\frac{dx}{dt} = k_x v - \mu_x x + \eta_x(t). \tag{11c}$$

We reiterate that we assume that there are no cross correlations in the noise; $\langle \eta_\alpha(t)\eta_\beta(t')\rangle = \delta_{\alpha\beta}\delta(t-t')$ and $\langle \Gamma(t)\,\eta_\alpha(t')\rangle = 0$. This means that the reactions are of the type $s \to s+v$ and $v \to v+x$, and not $s \to v$ and $v \to x$, respectively; put differently, the firing of a reaction does not consume a molecule of the reactant, and hence does not affect the fluctuations of the up-stream component [22]. In the *Discussion* section, we will briefly address some of the limitations of this assumption.

Fourier transformation gives

$$\tilde{X}(\omega) = \overbrace{\frac{k_x k_v \tilde{S}}{(i\omega+\mu_x)(i\omega+\mu_v)}}^{\text{signal}} + \underbrace{\frac{k_x \eta_v(\omega)}{(i\omega+\mu_x)(i\omega+\mu_v)} + \frac{\eta_x(\omega)}{i\omega+\mu_x}}_{\text{noise}}. \tag{12}$$

As indicated, we can identify the components of the output which are due to the input $\tilde{S}$ ("signal") and components which are due to intrinsic noise in the network. We obtain for the power spectrum of x,

$$P_{xx}(\omega) = \left\langle \tilde{X}\tilde{X}^* \right\rangle = \overbrace{\frac{k_x^2}{(\omega^2+\mu_x^2)}\frac{k_v^2}{(\omega^2+\mu_v^2)}}^{g^2(\omega)} \overbrace{\frac{2k_s}{(\omega^2+\mu_s^2)}}^{P_{ss}(\omega)} + \underbrace{\underbrace{\frac{k_x^2}{(\omega^2+\mu_x^2)}\frac{2k_v\langle s\rangle}{(\omega^2+\mu_v^2)}}_{N_{v\to x}(\omega)} + \underbrace{\frac{2k_x\langle v\rangle}{(\omega^2+\mu_x^2)}}_{N_x(\omega)}}_{N(\omega)} \tag{13}$$

Fig. 2A shows the output power spectrum of this network $P_{xx}(\omega)$ (red), as well as its decomposition into the noise $N(\omega)$ (green) and the transmitted signal $\Sigma(\omega) = g^2(\omega)P_{ss}(\omega)$ (black solid) (see also Eqn. 8). Simple cascades are characterised by a number of "knee" frequencies (vertical dashed), corresponding to the characteristic relaxation rates of the different components of the network (in this case μ_s, μ_v and μ_x). These knee frequencies are the inverse of the response times of the components — e.g. $\mu_v = \tau_v^{-1}$.

In order for the processing network to track variations in the input s on a time scale ω^{-1}, the network should be able to respond on this time scale. If any component of the processing network has a longer response time, this variation in s will be filtered. This filtering can be observed in the transmitted signal $\Sigma(\omega)$, where at frequencies above the first knee frequency, $\Sigma(\omega)$ scales with ω^{-2} and for every consecutive knee frequency, $\Sigma(\omega)$ decays with an additional factor ω^{-2} (Fig. 2A). In effect each level of the cascade acts as a low-pass filter, because the incoming signal is averaged over the protein response time. Mathematically, the transmitted signal $\Sigma(\omega)$ can be factored into the input signal $P_{ss}(\omega)$ (black dashed), and the total gain $g^2(\omega)$ (Fig. 2B, black), which is independent of the input signal (because we assume that the network does not feed back onto s). Moreover, the total gain of the network is the product of the gain of each cascade step: $g^2(\omega) = g^2_{s\to v}(\omega)g^2_{v\to x}(\omega)$; decaying as ω^{-4} for $\omega \gg \mu_v, \mu_x$ (Fig. 2B). Consequently, the transmitted signal $\Sigma(\omega)$ decays as ω^{-6} for $\omega \gg \mu_s, \mu_v, \mu_x$.

Since we assume that there are no cross-correlations between the different noise terms, the total noise $N(\omega)$ (green line in Figs. 2A and 2B) is given by the noise-addition rule [22, 31], which means that $N(\omega)$ is simply given by the sum of two independent contributions, $N_{v\to x}(\omega)$ (Fig. 2B, green dotted) and $N_x(\omega)$ (green dashed) (see Eqn. 13). Here, $N_x(\omega)$ is the noise in the concentration of x that arises from the intrinsic stochasticity in the production and decay events of x; $N_x(\omega)$ would be the total variance in the concentration of x if v, the input for x, would not vary over time. However, the upstream component v does vary in time, not only because it is driven by variations in the input s, but also because it fluctuates spontaneously due to the noise in its synthesis and decay events. This noise is propagated to x. Its contribution to the total noise power of x is $N_{v\to x}(\omega)$, which is given by the noise in v, $N_v(\omega)$,

multiplied by how much this noise is amplified at the level of x, given by $g^2_{v\to x}(\omega)$: $N_{v\to x}(\omega) = g^2_{v\to x}(\omega)N_v(\omega)$, where $g^2_{v\to x} = k_x^2/(\omega^2+\mu_x^2)$. The "extrinsic" contribution to the noise in x, $N_{v\to x}$, decays as ω^{-4} since the noise in v, decaying as ω^{-2}, is filtered by the finite lifetime of the protein x. The "intrinsic" contribution, $N_x(\omega)$, decays as ω^{-2}, meaning that for $\omega \gg \mu_v, \mu_x$, $N(\omega) \approx N_x(\omega)$. Hence, while the transmitted signal $\Sigma(\omega)$ decays as ω^{-6} for $\omega \gg \mu_s, \mu_v, \mu_x$, the noise $N(\omega)$ decays as ω^{-2} (Fig. 2B, green solid). As a result, for frequencies $\omega \gg \mu_s, \mu_v, \mu_x$, the transmitted signal $\Sigma(\omega)$ is completely obscured by the noise and the output $P_{xx}(\omega)$ is simply given by the noise $N(\omega)$ (Fig. 2A).

Finally, the gain-to-noise ratio (Fig. 2B, red) is

$$\frac{g^2(\omega)}{N(\omega)} = \frac{k_v k_x \mu_v}{2\langle s\rangle\left[\omega^2+\mu_v^2+\mu_v k_x\right]}. \tag{14}$$

This expression shows that the simple cascade effectively acts as a low-pass filter for information, meaning that it cannot reliably respond to signals that vary (much) faster than a characteristic cut-off frequency $\omega_c^2 = \mu_v(\mu_v + k_x)$. We note that the gain-to-noise ratio is independent of μ_x, since both the gain and the noise have the same functional dependence on μ_x. This is a general feature of the biochemical networks we will study: degradation of the output species occurs independently of the upstream components, and therefore provides no additional information about the input [16].

B. Autoregulation

In this section we consider direct feedback of a component onto its own production, as indicated in Fig. 3A and 4A. Autoregulation is one of the most common forms of regulation in signalling networks. It is well known that negative autoregulation speeds up the response time of components, which can also change the response time of the complete signalling cascade [32]. Positive autoregulation slows down the response time and can lead to bistability [32, 33].

Autoregulation alters only the diagonal entries of the Jacobian matrix (Fig. 1). This means that the characteristic timescale for dissipation of small fluctuations — the response time — changes, which is as expected. For the steady state of the system to be stable we require that the diagonal of the Jacobian has only negative terms. Thus autoregulation cannot qualitatively change the form of the output power spectrum $P_{xx}(\omega)$. In fact, once linearized, the dynamics of a network with autoregulation is equivalent to that of a simple cascade with a different degradation rate. In terms of information transmission, however, this is not always true, as we shall see below.

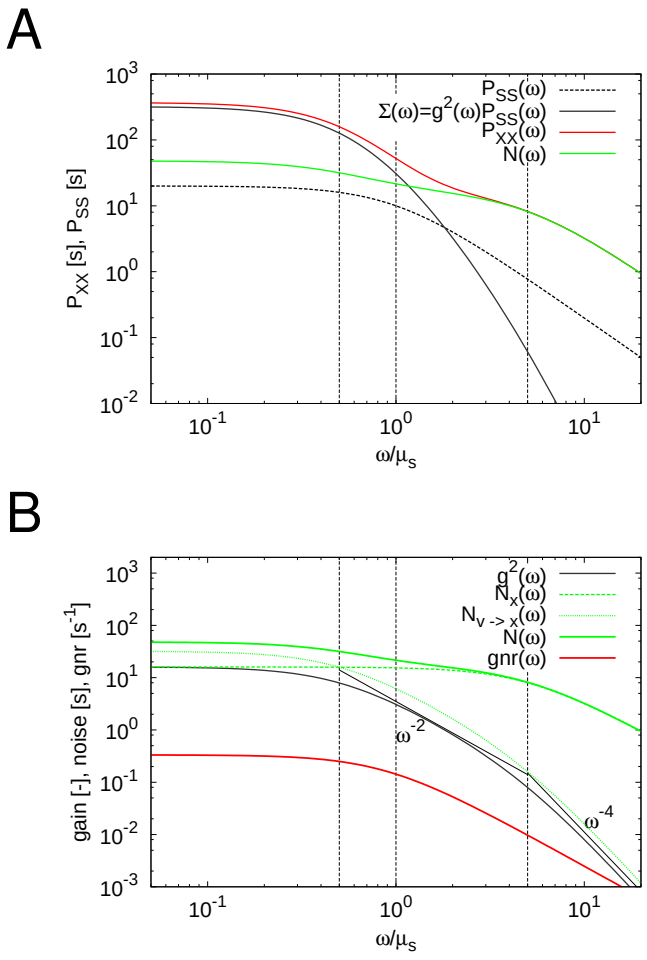


FIG. 2. Typical power spectra for a linear cascade. **A**: The power spectra of x and s, $P_{xx}(\omega)$ and $P_{ss}(\omega)$, together with the signal $\Sigma(\omega)$ and noise $N(\omega)$ components of the output, for the two-step cascade shown in Eqn. 11. **B**: The frequency-dependent gain $g^2(\omega)$, noise $N(\omega)$ and gain-to-noise ratio (GNR). Thin green lines indicate the two noise contributions, $N_{v\to x}(\omega)$ (dotted) and $N_x(\omega)$ (dashed). Parameters: $k_s = 10$, $k_v = 10$, $k_x = 1$, $\mu_v = 0.5$ and $\mu_x = 5$. Vertical lines indicate the degradation rates of the three components.

1. *Autoregulation at the response x does not affect information transmission*

We first consider autoregulation by the network output x on its own production, as depicted in Fig. 3A. For this motif the relaxation time of x is given by $\tau_x = -J_{xx}^{-1} = \left[\mu_x - \left\langle \frac{\partial}{\partial x} f(x) s\right\rangle\right]^{-1}$, where $f(x)$ describes the effect of the feedback of x onto its own production (see Eqn. 15 in Fig. 3B) For negative regulation $|J_{xx}| > \mu_x$, while for positive regulation $|J_{xx}| < \mu_x$. Negative (positive) regulation therefore reduces (increases) the response time of x to changes in s, compared to the equivalent simple

cascade network for which $f(x)$ =constant. In the output power spectrum $P_{xx}(\omega)$ this change in timescale appears as a shift in the knee frequency corresponding to τ_x^{-1}. A corresponding change can also be seen in both the gain and noise (see Eqn. 16 in Fig. 3C).

However, despite these changes in the response time, we find that the gain-to-noise ratio for an autoregulatory network (Eqn. 16c in Fig. 3B) is identical to the gain-to-noise ratio for a simple (two-component) cascade. The effect of changing J_{xx} on the noise and gain is identical (Eqn. 16a,b in Fig. 3B) and therefore cancels in the gain-to-noise ratio (as we also saw previously for the effect of μ_x in the simple cascade, Eqn. 14). The autoregulation by x of its own production alters the timing of production events. However, our constraint of equal average production means that the mean rate of this process in the two cascades is the same. Moreover, in the linearized regime the production of x is an identical Poissonian process in both simple and autoregulated cascades. Hence, to the extent that the system can be linearized, autoregulation at the output of a network does not affect information transmission. It is conceivable that non-linear effects cause autoregulation of the output component to affect information transmission, but a comparison of our analytical results discussed here with results of Gillespie simulations of the full system, suggest that the linearization approximation is surprisingly accurate (see also S1).

2. *Positive autoregulation within the cascade increases the gain-to-noise ratio*

In a cascade with autoregulation by an intermediate component the story is different (Fig. 4A and Eqn. 17 in Fig. 4B). First, we reiterate that since we compare the simple cascade and the cascade with autoregulation on the basis of equal average production and degradation rates, the noise strengths $\langle \eta_x^2 \rangle$ and $\langle \eta_v^2 \rangle$ are the same for both cascades. However, as noted above the effective relaxation timescale of component v, $\tau_v = -J_{vv}^{-1}$ (Eqn. 3), decreases with negative autoregulation and increases with positive autoregulation. This again leads to a reduction (increase) in both the gain (Fig. 4D, top left) and the noise (Fig. 4D, top right) of the network for negative (positive) autoregulation, as has been reported previously [20, 34]. However, unlike the case of autoregulation of the output x, the gain-to-noise ratio (Fig. 4D, bottom left) can change as a result.

Negative autoregulation (Fig. 4D, green) leads to a decrease in the response time compared to a simple cascade, corresponding to an increase in $|J_{vv}|$. This leads to a decrease in the gain of the autoregulated component $g^2_{s\to v}(\omega) = J^2_{vs}/(\omega^2 + J^2_{vv})$ at frequencies $\omega < |J_{vv}|$. Negative autoregulation therefore tends to suppress slowly varying signals relative to the simple cascade. Noise which is introduced upstream of or at the autoregulated component is filtered by the feedback-modified gain in exactly the same way as the signal, whereas noise introduced downstream of v is unaffected. Hence negative autoregulation reduces both the total gain of the network, which is the product of the individual reaction gains $g^2(\omega) = g^2_{s\to v}(\omega) g^2_{v\to x}(\omega)$, and the noise transmitted from v to x, $N_{v\to x}(\omega) = g^2_{v\to x} N_v(\omega)$, relative to the simple cascade. However, noise in the production and degradation of x is unchanged relative to the simple cascade. Since the total noise (Eqn. 19b in Fig. 4C) is the sum of independent noise contributions, $N(\omega) = N_x(\omega) + N_{v\to x}(\omega)$, the total noise decreases by a smaller factor than the gain, and the gain-to-noise ratio decreases compared with the simple cascade.

Conversely, positive autoregulation (Fig. 4D, red) increases the relaxation time of v, which increases $g^2_{s\to v}(\omega)$ at frequencies $\omega < |J_{vv}|$. We can therefore see that positive autoregulation amplifies slowly-varying signals. This leads to an increase in the network gain and the noise that is propagated from v to x. However, since the noise that is introduced at x is unchanged, positive autoregulation at v increases the gain-to-noise ratio compared to the simple cascade. Fig. 4D shows the comparison between a simple cascade and cascades with positive (red) and negative (green) autoregulation. Hornung and Barkai previously studied transmission of a *constant* signal with additive noise through a deterministic (noiseless) network [35], and found that positive autoregulation can increase the signal-to-noise ratio. Our results for time-varying signals with intrinsic network noise parallel their results.

Given a network with autoregulation, our constraint of equal production of each network component does not define a unique "equivalent" simple cascade. That is, different parameter combinations can be chosen for a simple cascade which satisfy the production constraint. The results in the preceding discussion correspond to one such parameter choice. Specifically, we choose the production rate of v in the simple cascade (Eqn. 11) to be $k_v = \langle f(v) \rangle$, while taking the same value for μ_v in both networks. A consequence of this choice is that the relaxation time τ_v changes between the two cascades, as discussed above. One can equally well construct a simple cascade for which the diagonal entries of the Jacobian, $J_{\alpha\alpha}$, are equal to those of the autoregulated cascade, so as to hold constant the relaxation time of each component between the two cascades. This is achieved by setting the spontaneous degradation rate for v in Eqn. 11 to be $\mu_v^{\text{new}} = \mu_v - \left\langle \frac{\partial}{\partial v} f(v) s \right\rangle$. By choosing this new rate, the average protein number $\langle v \rangle$ changes in the simple cascade, and as a result also the average production of x. To restore equal production of x we thus also require a rescaling of the kinetic production rate $k_x^{\text{new}} = k_x \mu_v^{\text{new}} / \mu_v$ in the simple cascade (Eqn. 11). Thus, in this comparison, the diagonal entries of the Jacobian matrices of the autoregulated and simple cascade are the same, while the off-diagonal entry $J_{xv} = k_x$ differs between the two.

Compared to a cascade with positive autoregulation, this new kinetic production rate in the simple cascade is smaller ($k_x^{\text{new}} < k_x$). The reduction in J_{xv} leads to a

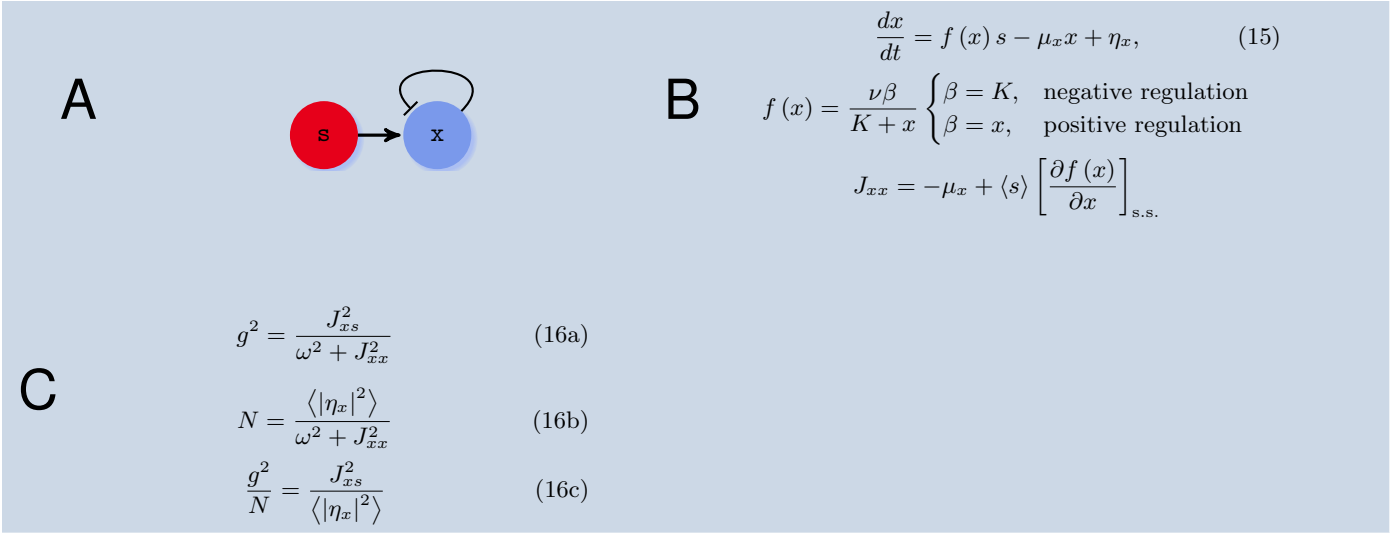


FIG. 3. Autoregulation of the output component. **A**: Schematic representation of the negative autoregulation motif, where s is the input signal and x the output signal, which negatively regulates its own production. **B**: The Langevin equations of the network. **C**: The characteristic equations for the gain, noise and gain-to-noise ratio (see also section *Autoregulation*).

uniform decrease in $g^2_{v\to x}(\omega)$ at all frequencies. As described above, this affects the signal and also the propagated noise $N_{v\to x}(\omega)$ equally, but not the intrinsic noise at x, $N_x(\omega)$. Thus, compared to a cascade with positive autoregulation, the gain-to-noise ratio is reduced at all frequencies in the simple cascade, as can be seen in Fig. 4E (black dotted). Interestingly, the decrease in the gain-to-noise ratio is most pronounced at high frequencies. This is because the propagated noise $N_{v\to x}(\omega)$ only has a significant contribution at frequencies $\omega < \mu_v^{\text{new}}$; at higher frequencies the total noise is dominated by $N_x(\omega)$, as discussed in section: *The simple cascade*. Thus at these higher frequencies, the gain is reduced relative to the positively-autoregulated cascade, but the noise is not, and so the change in the gain-to-noise is largest. For networks with negative autoregulation, the converse applies: the gain-to-noise ratio is higher in the simple cascade at all frequencies, but by a larger factor for $\omega < \mu_v^{\text{new}}$. Hence, the effect of positive or negative autoregulation is qualitatively the same in both parameterizations.

More generally, even if we relax the production constraint on each component, and instead require only the total production in the two cascades to be the same (i.e., $\langle f_v^+\rangle + \langle f_x^+\rangle$ = constant), we see the similar qualitative behaviour for the gain-to-noise ratio (see Eqn. S1-26-S1-28). Positively-autoregulated cascades have a larger gain-to-noise ratio than a simple cascade of the same length, while for a cascade with negative autoregulation the gain-to-noise ratio is smaller. For longer cascades drawing such general conclusions is more difficult. However, if the majority of parameters are kept the same between the simple and autoregulated cascades, as in the cases discussed in detail above, then we again find that positive autoregulation increases and negative autoregulation decreases the gain-to-noise ratio. Furthermore, given a specific simple cascade one can always add positive autoregulation to the network in such a way as to achieve a larger gain-to-noise ratio while maintaining the same total production cost.

We have here considered only autoregulation via the production of the intermediate v. However, for autoregulation via the degradation of v we observe similar results for the gain-to-noise ratio: if v suppresses its own degradation, the decrease in the effective turn-over rate leads to a reduction of the noise strength $N_{v\to x}(\omega)$, increasing the gain-to-noise ratio; when v enhances its own degradation rate the transmitted noise is increased, reducing signalling fidelity.

C. Feedback

Feedback, both positive and negative, corresponds to the upper-triangular part in the Jacobian of the linearized system (see Fig 1). It is known that negative feedback allows for adaptation as, for example, in the *E. coli* chemotaxis pathway [3, 13, 36]. Feedback can also shift noise to higher frequencies [20]. We will again consider separately the two cases of feedback by the output x onto an upstream component and feedback by an intermediate component onto a component higher up the cascade.

1. Feedback from x does not affect information transmission

For negative feedback from x to v (Fig. 5A and Eqn. 22 in Fig. 5B), the power spectrum of the response $P_{xx}(\omega)$ (Fig. 5D, red solid) can have a resonance peak while none is present in the input signal (black dotted). Surprisingly, this peak does not correspond to an increase in information transmission capabilities at the peak frequency (ω_{peak}), since no peak is present in the gain-to-noise ratio (Fig. 5D, red dashed). For positive feedback, no peak is present in either $P_{xx}(\omega)$ or the gain-to-noise ratio.

For a system with negative feedback from x to v the gain and noise both show a peak, but these can occur at different frequencies. We consider first the frequency dependence of the gain. At low frequencies the negative feedback leads to destructive interference at v between the input signal $\tilde{S}(\omega)$ and the signal that is fed back, $\tilde{X}(\omega)$. On the other hand, at high frequencies these two signals are exactly out of phase, and hence the interference becomes constructive (since the feedback combines the two signals negatively). However, at frequencies $\omega \gg \mu_v, \mu_x$ the *amplitude* of the fed-back signal decreases, due to averaging over the lifetimes of v and x; hence, even though the two signals interfere constructively, the significance of this interference decreases. Together, these three effects lead to a maximum in the gain. This maximum occurs at

$$\omega_{\text{res}}^2 = -\frac{1}{2}\left[\mu_x^2 + \mu_v^2 + 2J_{vx}J_{xv}\right], \tag{20}$$

which depends on the relaxation rates μ_x, μ_v and the coupling (feedback) loop between v and x, $J_{vx}J_{xv}$. This timescale corresponds to the imaginary part of the eigenvalues of the Jacobian (see Eqn. S1-39).

The frequency of the peak in the noise depends on the relative strengths of the two noise sources, η_v and η_x. The two noise terms are propagated differently through the network, because η_x originates at the regulator of the feedback loop, while η_v originates at the regulated component. We consider two limiting cases. If the total noise $N(\omega)$ (Eqn. 23b in Fig. 5C) is dominated by the transmitted noise, $N_{v\to x}(\omega)$, both the signal $\Sigma(\omega)$ and the dominant source of noise originate upstream of the feedback loop. Effectively, therefore, the feedback affects both the gain and noise of the network similarly. As a result the peak frequencies of both the noise and the gain are the same. On the other hand, when the total noise is dominated by $N_x(\omega)$, which is located downstream of the regulated component v, the feedback loop affects the signal and noise differently. As a result, the noise that is fed back has a different frequency profile than the signal, such that the peaks in the gain and the noise occur at different frequencies (Fig. 5D, black circles).

One might therefore expect that when $N_x(\omega) \gg N_{v\to x}(\omega)$ a peak in the gain-to-noise ratio is possible. However, an inspection of the expressions for the gain, Eqn. 23a, and the noise, Eqn. 23b (both in Fig. 5C), shows that they have the same denominator, such that the gain-to-noise ratio is a monotonically decreasing function of frequency (Eqn. 23c in Fig. 5C). The effect of the negative feedback is cancelled. Ultimately, this is due to the fact that the noise in the output x goes back into the feedback loop, such that the peaks in the gain and the noise cannot be controlled separately; in the next section, we show how this can be done. Furthermore, we note that the gain-to-noise ratio is again identical to a simple three-component cascade, as we also saw in the case of autoregulation of x. We conclude that feedback from x onto the cascade also has no effect on information transmission through the network.

This network (Eqn.23 in Fig. 5B) also highlights the idea that the power spectrum of the output $P_{xx}(\omega)$ may not be indicative of the information that is transmitted at different frequencies. We see in Fig. 5D that due to the negative feedback $P_{xx}(\omega)$ can have a peak at non-zero frequencies, even if none is present in the input signal. However, this peak does not correspond to the frequency at which the signal is transmitted most reliably. Instead, we can see that the peak is simply due to resonant amplification of both the signal and the noise at the characteristic frequency of the negative feedback loop.

It has been suggested [19] that a system where a negative feedback loop acts *on* the response component can have a large peak in the gain, such that signals on specific timescales can be selected for. If we take in Fig. 5A not x but v to be the output of the network, we obtain

$$\frac{g^2(\omega)}{N(\omega)} = \frac{J_{vs}^2\left(\omega^2 + \mu_x^2\right)}{J_{vx}^2\left\langle|\eta_x|^2\right\rangle + \left(\omega^2 + \mu_x^2\right)\left\langle|\eta_v|^2\right\rangle} \tag{21}$$

We observe that the gain-to-noise ratio is a monotonically increasing function of frequency and does not show a peak at any specific frequencies. Furthermore we note that as $\omega \to \infty$ the gain-to-noise ratio becomes equal to the gain-to-noise ratio for the one-step simple cascade ($J_{vs}/2\langle s\rangle$), since for large ω the noise from the downstream component is averaged out. Thus this network motif has a higher gain-to-noise at all frequencies than the cascade with x as the output. However, the information transmitted at low frequencies is less than if x were not present. Following the information processing inequality, the amount of information about s which is encoded in the dynamics of v is always larger than the corresponding information in x. By feeding back x to v we thus do not add more information to the signal, but essentially add an extra source of noise to the pathway from s to v. The strength of this noise is highest at frequencies $\omega < \mu_x$, and hence the effect of the feedback is to obscure the signal at these frequencies. As a result this motif acts as a high-pass filter for information.

2. *Negative feedback within a cascade can lead to a peak in the gain-to-noise ratio*

In section *Autoregulation* we saw that the gain-to-noise ratio is sensitive to the precise position of autoregulation in a cascade. In this section we therefore study a cascade where the feedback is not from x to v, but between two intermediate components w and v (see Fig. 6A and Eqn. 28 in Fig. 6B). This also corresponds to taking the output of the previous feedback cascade (Fig. 5A) as the input to another downstream process.

Expressions for the gain, noise and gain-to-noise ratio are given in Fig. 6C. For positive feedback the gain, noise and gain-to-noise ratio are once again monotonically decreasing with increasing frequency. However, we find that for a network with strong negative feedback (Hill coefficient $n > 1$, see Eqn. S1-54), the gain-to-noise ratio can have a maximum as a function of frequency at

$$\omega^2_{\text{peak}} = -\frac{1}{2}\left(\overbrace{J^2_{xw}\frac{\left\langle|\eta_w|^2\right\rangle}{\left\langle|\eta_x|^2\right\rangle}}^{\text{noise}} + \overbrace{\mu_v^2+\mu_w^2+2J_{vw}J_{wv}}^{\text{resonance}\left(\omega^2_{res}\right)}\right)$$
$$= -\frac{1}{2}\left(J^2_{xw}\frac{\left\langle|\eta_w|^2\right\rangle}{\left\langle|\eta_x|^2\right\rangle}+\mu_v^2+\mu_w^2-2\frac{n\mu_w\mu_v\langle w\rangle^n}{K^n+\langle w\rangle^n}\right). \tag{25}$$

This peak frequency depends on the characteristic resonance frequency of the feedback loop, ω_{res}, which is determined by the interactions between v and w: μ_v, μ_w, J_{vw} and J_{wv}. It is additionally dependent on the relative strengths of the noise introduced into the network at w and at x.

We can understand the appearance of this peak as follows. For a network with negative feedback, $g^2(\omega)$ (Fig. 6D, bottom right) has a maximum as a function of frequency at ω_{res}, the characteristic resonance frequency of the feedback loop. Input signals at this frequency are amplified by the constructive interference between the signal transmitted to v from s and the signal which is fed back from w to v. We note that the resonance frequency has the same form as Eqn. 20, and depends only on the interactions between v and w. The behaviour of the noise power spectrum (Fig. 6D, top right) is more complex. We consider two limiting cases in which different noise terms dominate. When the total noise is dominated by noise introduced at v or w, the noise is processed through the feedback loop together with the signal. As discussed in the previous section, $N(\omega)$ therefore shows a peak at a similar frequency to the gain (black line). These two peaks cancel, and hence the gain-to-noise ratio (Fig. 6D, top left, black line) is monotonically decreasing with frequency. On the other hand, when the total noise is dominated by $N_x(\omega)$ (top right, red line) the noise in the network is not affected by the feedback loop. Hence no peak is found in the noise power spectrum. In this limit, the peak in the gain-to-noise ratio corresponds to the peak in the gain at ω_{res} (top left, red line).

From these arguments we see that the peak in the gain-to-noise ratio becomes more pronounced as the relative contribution of $N_x(\omega)$ to the total noise increases. Additionally, increasing the strength of the negative feedback by reducing K or increasing n leads to a more pronounced peak. However, this increase in the *relative* peak height comes at the expense of a reduction in the value of the gain-to-noise ratio at all frequencies.

How does the gain-to-noise ratio of the network with feedback compare to the corresponding (four-component) simple cascade? We examine the ratio of the gain-to-noise for the network with feedback to the gain-to-noise of the simple cascade,

$$G_{\text{fb}}(\omega) = \left[\frac{g^2(\omega)}{N(\omega)}\right]_{\text{fb}} \Big/ \left[\frac{g^2(\omega)}{N(\omega)}\right]_{\text{simple}}, \tag{26}$$

and find that (Fig. 6E,F)

$$G_{\text{pos}}(\omega) > 1 \text{ if } \omega^2 < \mu_v\mu_w\left[1-\frac{n}{2}\frac{K^n}{K^n+\langle w\rangle^n}\right], \tag{27a}$$

$$G_{\text{neg}}(\omega) > 1 \text{ if } \omega^2 > \mu_v\mu_w\left[1+\frac{n}{2}\frac{\langle w\rangle^n}{K^n+\langle w\rangle^n}\right] \tag{27b}$$

Interestingly, for both types of feedback there is a range of frequencies over which the gain-to-noise ratio increases relative to the simple cascade. This contrasts to the results of section *Autoregulation*, where we found that autoregulation affected the gain-to-noise ratio in the same way at all frequencies.

This difference can again be understood in terms of the interference of the two signals arriving at v. As described above (and in Eqn.S1-43-S1-44), at low frequencies the signal propagated from s to v and the feedback signal from w to v are in phase, while at high frequencies the two signals are exactly out of phase. Hence for a positive feedback loop (Fig. 6E,F; red line) the signals combine constructively at low frequencies, increasing the gain, but destructively at high frequencies, decreasing the gain. Recall that, since we are comparing networks with equal production, the noise strengths $\left\langle|\eta_v|^2\right\rangle$, $\left\langle|\eta_w|^2\right\rangle$ and $\left\langle|\eta_x|^2\right\rangle$ are equal in the regulated and simple cascades. In an analogous way to the autoregulation discussed in section *Autoregulation*, the presence of feedback between w and v affects both the signal and noise introduced upstream of x, but not noise introduced at x. Hence, at low frequencies positive feedback amplifies the signal and the noise introduced at the levels of v and w, but not noise introduced at x. Hence at low frequencies the gain-to-noise ratio increases relative to the simple cascade. At high frequencies, however, positive feedback reduces the gain and the noise upstream of x, but not the intrinsic noise $N_x(\omega)$; consequently, the gain-to-noise ratio is reduced compared to the simple cascade. Conversely, a network with negative feed-

back (Fig. 6E,F;green line) reduces the gain at low frequencies, reducing the gain-to-noise ratio. However, at high frequencies, the feedback amplifies the signal but not $N_x(\omega)$, leading to an increase in the gain-to-noise ratio.

From these results we conclude that if a cell is only concerned with low frequency input signals, it is beneficial in terms of information transmission to add positive feedback within the signalling cascade. If the system wishes to respond specifically to high-frequency signals, negative feedback can be used to increase the fidelity of transmission for these signals. Additionally for a strong negative feedback ($n \gg 1$ or $K \ll \langle w \rangle$, see Eqn. S1-61) the gain-to-noise can have a peak in the regime where signalling is more reliable than for a simple cascade, allowing the cell to focus on signals in a particular frequency band. We note that the negative feedback motifs considered here do not lead to perfect adaptation to constant input signals, which is characterised by $g^2(\omega = 0) = 0$ and is necessary for true band-pass behaviour. Perfect adaptation requires that the feedback to be implemented via a buffer node or side branch [37]. An example of this network architecture is the *E. coli* chemotaxis pathway [15], for which the gain-to-noise ratio does indeed indicate a band-pass for information [16].

IV. DISCUSSION

In this paper we have analysed information transmission through a number of network motifs, namely cascades, autoregulation and feedback. One of the most important conclusions of our analysis is that to understand how reliably biochemical networks can transmit time-varying signals, we have to study the frequency-dependent gain-to-noise ratio [16]. In particular, the power spectrum of the output signal may not be a good measure for how biochemical networks transduce time-varying input signals. The power spectrum of the output signal depends on the power spectrum of the input signal, the frequency-dependent gain, and the frequency-dependent noise. Only the latter two quantities are intrinsic properties of the network, provided that the network detects the input via biochemical reactions that do not affect the statistics of the input signal [22]. Moreover, we have seen that the power spectrum of the output signal may differ qualitatively from that of the frequency-dependent gain-to-noise ratio. A striking example is provided by the network with negative feedback from the output component, which shows a peak in the output signal (see Fig. 5D): while one might be tempted to conclude that input signals at this frequency are transduced most reliably, our analysis shows that this peak in the output spectrum is simply the result of resonant amplification of both the input signal and the noise in the network.

Our analysis leads us to draw the following conclusions on the effect of autoregulation and feedback on the transmission of time-varying signals: 1) autoregulation of the output component does not affect the gain-to-noise ratio, and hence does not affect information transmission (Fig. 3C); 2) positive autoregulation of an intermediate component increases the gain-to-noise ratio over all frequencies, while negative autoregulation tends to decrease it over all frequencies (Fig. 4D); 3) negative feedback from the output component onto an upstream component may lead to a peak in the power spectrum of the output, and those of the gain and the noise; yet, even though the peaks of gain and the noise can be at different frequencies, negative feedback from the output component onto an upstream component can *not* lead to a peak in the spectrum of the gain-to-noise ratio (Fig. 5D); 4) positive feedback between upstream components enhances the gain-to-noise ratio at low frequencies, while negative feedback increases the gain-to-noise ratio at high frequencies (Fig. 6F). Further, we note that it is possible to achieve a peak in the gain-to-noise ratio via negative feedback between components that are upstream of the output component (Fig. 6D); however, this comes at the expense of a reduction in the gain-to-noise ratio for all frequencies. We also note here that stronger band-pass filtering of information can be obtained with networks employing integral feedback in a side branch[16], as found in the networks of osmo adaptation [2] or bacterial chemotaxis [15]. Alternatively, band-pass filters for information transmission can be obtained via feed*forward* loops, which we will discuss in a forthcoming publication.

Taken together these results reveal the following design principles for the use of feedback and autoregulation in signal transduction cascades (see the schematic drawing Fig. 7). Firstly, feedback and autoregulation can improve information transmission, but only if they occur upstream of the dominant source of noise in the cascade. Feedback or autoregulation downstream of the dominant noise source affects the gain and the noise similarly. Secondly, if signals over the full frequency range have to be transmitted reliably, positive autoregulation is advantageous, while if the cell is concerned only with low- or high-frequency signals, then positive or negative feedback can be employed.

The approach employed here has a number of limitations. Firstly, we have used the linear-noise approximation, and the power spectra calculated using this approximation may deviate from those of the full non-linear system. We argue that this effect does not significantly affect our results, since we find excellent agreement between the power spectra calculated analytically using the linear-noise approximation and those obtained from stochastic simulations of the full system (see S1). The second potential source of inaccuracy is the use of Eqns. 5 and 10, which are exactly only for linear Gaussian systems. However, the information rate calculated in this approximation provides a lower bound on the information transmission rate of the full system [38].

Another limitation of our analysis is that to reduce the complexity of the problem, we have assumed that the net-

works obey the spectral-addition rule [22], meaning that reactants are not consumed during reaction events. However, irreversible modifications of a substrate molecule are common in biochemical networks, and reactions of this type can significantly change the correlations between different network components. For instance, in a cascade of the type $X_0 \to X_1 \to \dots X_{n-1} \to X_n$, where in each reaction step the reactant is consumed, correlations of the form $\langle \eta_i \eta_{i+1} \rangle = -k \langle X_i \rangle$ appear between different noise terms. As a result, for this cascade the covariance between different components $\langle x_i x_{j \neq i} \rangle = 0$ [22, 39], and hence the mutual information between *instantaneous* levels of components X_i and $X_{j \neq i}$ is zero [16]. This may suggest that these cascades cannot effectively transmit information. Yet, the analysis of [16] indicates that this motif can, in fact, reliably transmit time-varying signals. It would therefore be of interest to study the effect of cross-correlations in the noise on the information transmission in the motifs studied here. We leave this for future work.

Lastly, how could our predictions be tested experimentally? It is increasingly being recognised that stimulating biochemical networks with time-varying signals provides a wealth of information on the dynamics of these networks [2, 3, 40–43]. These experiments can also be used to study the reliability by which biochemical networks can transmit time-varying signals. By measuring not only the power spectra of the in- and output signals, $P_{ss}(\omega)$ and $P_{xx}(\omega)$, but also their cross-power spectrum $P_{sx}(\omega)$, one can obtain the frequency-dependent gain $g^2(\omega) \equiv |P_{sx}(\omega)|^2 / P_{ss}(\omega)^2$ and the frequency-dependent noise $N(\omega)$ (see Eqn. 8), and hence the gain-to-noise ratio. Stimulating synthetic gene circuits or existing signal transduction pathways and gene regulation networks of known architecture with time-varying signals, for example using microfluidic devices, would make it possible to test our predictions on the effect of feedback and autoregulation on information transmission.

V. ACKNOWLEDGMENTS

We thank Martin Frimmer and Nils Becker for a careful reading of this manuscript. This work is part of the research program of the "Stichting voor Fundamenteel Onderzoek der Materie (FOM)", which is financially supported by the "Nederlandse organisatie voor Wetenschappelijk Onderzoek (NWO)".

[1] M.R. Bennett, W.L. Pang, N.A. Ostroff, B.L. Baumgartner, S. Nayak, L.S. Tsimring, and J. Hasty. Metabolic gene regulation in a dynamically changing environment. *Nature*, 454:1119–1122, 2008.

[2] J.T. Mettetal, D. Muzzey, C. Gomez-Uribe, and A. van Oudenaarden. The Frequency Dependence of Osmo-Adaptation in *Saccharomyces cerevisiae*. *Science*, 319:482–484, 2008.

[3] Y. Tu, T.S. Shimizu, and H.C. Berg. Modeling the chemotactic response of *Escherichia coli* to time-varying stimuli. *Proc. Natl Acad. Sci. USA*, 105:14855–14860, 2008.

[4] M.B. Elowitz, A.J. Levine, E.D. Siggia, and Peter S. Swain. Stochastic gene expression in a single cell. *Science*, 297:1183–1186, 2002.

[5] P.S. Swain, M.B. Elowitz, and E.D. Siggia. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc. Natl Acad. Sci. USA*, 99:12795–12800, 2002.

[6] A. Ma'ayan, S.L. Jenkins, S. Neves, A. Hasseldine, E. Grace, B. Dubin-Thaler, N.J. Eungdamrong, G. Weng, P.T. Ram, J.J. Rice, A. Kershenbaum, G.A. Stolovitzky, R.D. Blitzer, and R. Iyengar. Formation of Regulatory Patterns During Signal Propagation in a Mammalian Cellular Network. *Science*, 309:1078–1083, 2005.

[7] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon. Network Motifs: Simple Building Blocks of Complex Networks. *Science*, 298:824–827, 2002.

[8] C.E. Shannon. A Mathematical Theory of Communication. *Bell Sys. Tech. J.*, 27:379–423, 623–656, 1948.

[9] E. Ziv, I. Nemenman, and C.H. Wiggins. Optimal signal processing in small stochastic biochemical networks. *PLoS ONE*, 2:e1077, 2007.

[10] G. Tkačik, C.G. Callan Jr., and W. Bialek. Information capacity of genetic regulatory elements. *Phys. Rev. E*, 78:011910, 2008.

[11] A.M. Walczak, A. Mugler, and C.H. Wiggins. A stochastic spectral analysis of transcriptional regulatory cascades. *Proc. Natl Acad. Sci. USA*, 106:6529–6534, 2009.

[12] P. Mehta, S. Goyal, T. Long, B.L. Bassler, and N.S. Wingreen. Information processing and signal integration in bacterial quorum sensing. *Mol. Syst. Biol.*, 5:325, 2009.

[13] J.E. Segall, S.M. Block, and H.C. Berg. Temporal comparisons in bacterial chemotaxis. *Proc. Natl Acad. Sci. USA*, 83:8987–8991, 1983.

[14] C.J. Marshall. Specificity of receptor tyrosine kinase signaling: Transient versus sustained extracellular signal-regulated kinase activation. *Cell*, 80:179–185, 1995.

[15] T.-M. Yi, Y. Huang, M.I. Simon, and J. Doyle. Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proc. Natl Acad. Sci. USA*, 97:4649–4653, 2000.

[16] F. Tostevin and P.R. ten Wolde. Mutual information between input and output trajectories of biochemical networks. *Phys. Rev. Lett.*, 102:218101, 2009.

[17] M. Samoilov, A. Arkin, and J. Ross. Signal processing by simple chemical systems. *J. Phys. Chem. A*, 106:10205–10221, 2002.

[18] B.P. Ingalls. A frequency domain approach to sensitivity analysis of biochemical systems. *J. Phys. Chem. B*, 108:1143–1152, 2004.

[19] J. Locasale. Signal duration and the time scale dependence of signal integration in biochemical pathways. *BMC Syst. Biol.*, 2:108, 2008.

[20] M.L. Simpson, C.D. Cox, and G.S. Sayler. Frequency domain anlysis of noise in autoregulated gene circuits. *Proc. Natl Acad. Sci. USA*, 100:4551–4556, 2003.

[21] D.T. Gillespie. The chemical Langevin equation. *J. Chem. Phys.*, 113:297–306, 2000.

[22] S. Tănase-Nicola, P.B. Warren, and P.R. ten Wolde. Signal detection, modularity, and the correlation between extrinsic and intrinsic noise in biochemical networks. *Phys. Rev. Lett.*, 97:068102, 2006.

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

[24] M. Scott, T. Hwa, and B. Ingalls. Deterministic characterization of stochastic genetic circuits. *Proc. Natl Acad. Sci. USA*, 104:7402–7407, 2007.

[25] P.B. Warren, S. Tănase-Nicola, and P.R. ten Wolde. Exact results for noise power spectra in linear biochemical reaction networks. *J. Chem. Phys.*, 125:144904, 2006.

[26] (We note that with the chosen expression for the size of the random events $\langle \eta_i \eta_j \rangle$, the linearized Langevin equations lead to the Fluctuation-Dissipation theorem or Linear Noise Approximation [31, 44]).

[27] A. Borst and F.E. Theunissen. Information theory and neural coding. *Nat. Neurosci.*, 2:947–957, 1999.

[28] T. Munakata and M. Kamiyabu. Stochastic resonance in the FitzHugh-Nagumo model from a dynamic mutual information point of view. *Eur. Phys. J. B*, 53:239–243, 2006.

[29] F.J. Bruggeman, N. Blüthgen, and H.V. Westerhoff. Noise management by molecular networks. *PLoS Comput. Biol.*, 5:e1000506, 2009.

[30] D.T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.*, 22:403–434, 1976.

[31] J. Paulsson. Models of stochastic gene expression. *Phys. Life Rev.*, 2:157–175, 2005.

[32] R. Heinrich, B.G. Neel, and T.A. Rapoport. Mathematical models of protein kinase signal transduction. *Mol. Cell*, 9:957–970, 2002.

[33] U. Alon. *An introduction to systems biology.* Chapman & Hall/CRC, 2007.

[34] M. Thattai and A. van Oudenaarden. Intrinsic noise in gene regulatory networks. *Proc. Natl Acad. Sci. USA*, 98:8614–8619, 2001.

[35] G. Hornung and N. Barkai. Noise propagation and signaling sensitivity in biological networks: A role for positive feedback. *PLoS Comput. Biol.*, 4:0055–0061, 2008.

[36] R.M. Macnab and D.E. Koshland Jr. The Gradient-Sensing Mechanism in Bacterial Chemotaxis. *Proc. Natl Acad. Sci. USA*, 69:2509–2512, 1972.

[37] W. Ma, A. Trusina, H. El-Samad, W.A. Lim, and C. Tang. Defining network topologies that can achieve biochemical adaptation. *Cell*, 138:760–773, 2009.

[38] P.P. Mitra and J.B. Stark. Nonlinear limits to the information capacity of optical fibre communications. *Nature*, 411:1027–1030, 2001.

[39] E. Levine and T. Hwa. Stochastic fluctuations in metabolic pathways. *Proc. Natl Acad. Sci. USA*, 104:9224–9229, 2007.

[40] O. Lipan and W.H. Wong. The use of oscillatory signals in the study of genetic networks. *Proc. Natl Acad. Sci. USA*, 102:7063–7068, 2005.

[41] D.W. Austin, M.S. Allen, J.M. McCollum, R.D. Dar, J.R. Wilgus, G.S. Sayler, N.F. Samatova, C.D. Cox, and M.L. Simpson. Gene network shaping of inherent noise spectra. *Nature*, 439:608–611, 2006.

[42] C. Tan, F. Reza, and L. You. Noise-limited frequency signal transmission in gene circuits. *Biophys. J.*, 93:3753–3761, 2007.

[43] A. Cournac and J.-A. Sepulchre. Simple molecular networks that respond optimally to time-periodic stimulation. *BMC Syst. Biol.*, 3:29, 2009.

[44] N.G. van Kampen. *Stochastic processes in physics and chemistry.* North-Holland Publishing Company, 1981.

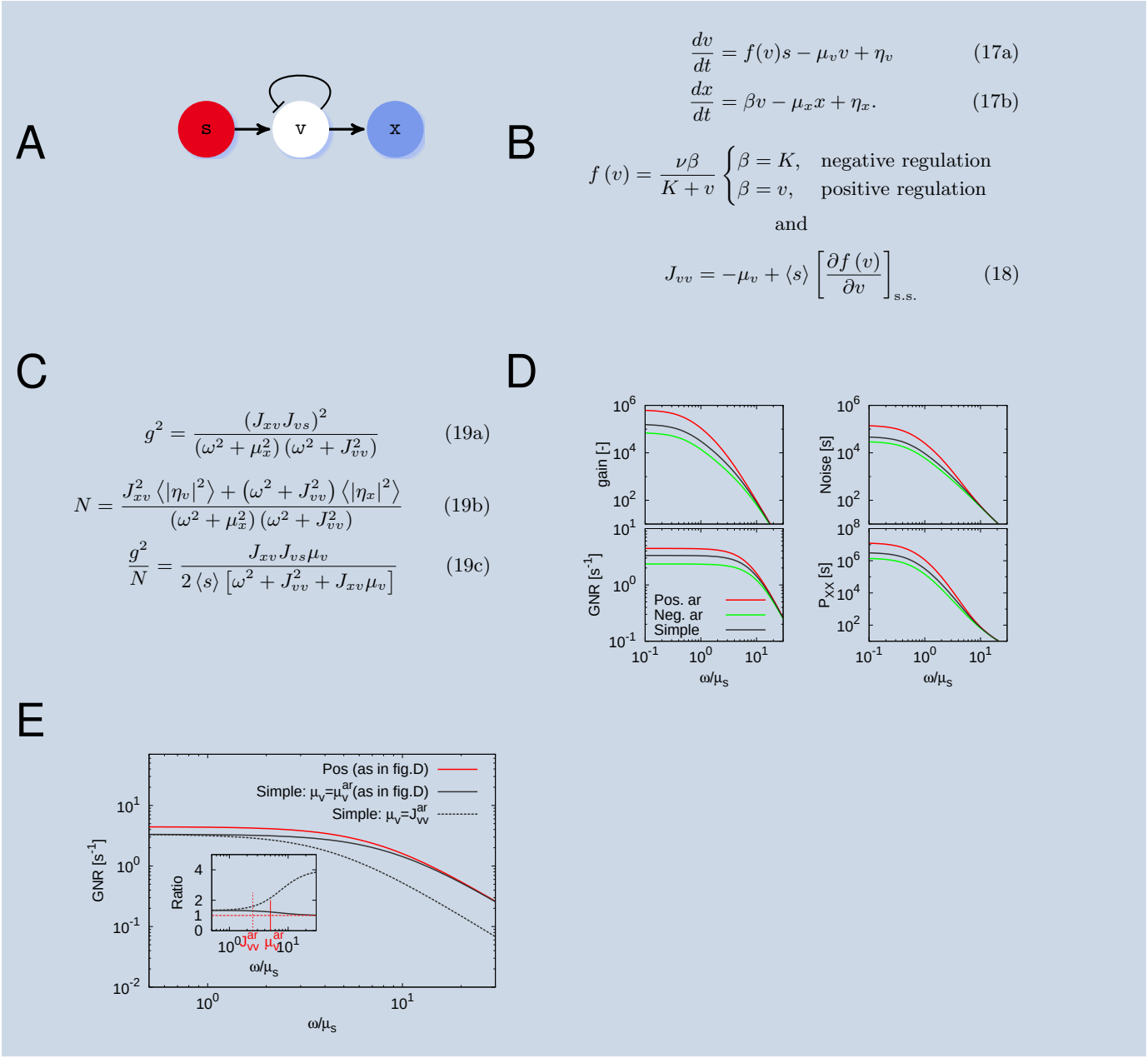
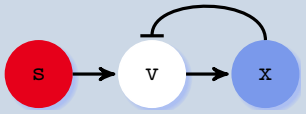


FIG. 4. A two-step cascade with autoregulation of the intermediate component. **A**: Cartoon of the negative autoregulation motif, where the intermediate component v negatively regulates its own production. **B**: The Langevin equations describing the network. **C**: The characteristic equations for the gain $g^2(\omega)$, noise $N(\omega)$ and gain-to-noise ratio. **D**: The gain, noise, gain-to-noise ratio (GNR) and output power spectrum $P_{xx}(\omega)$ plotted as a function of frequency for three different cascades: simple (black), positive autoregulation (red) and negative autoregulation (green). Negative autoregulation reduces the gain, noise and gain-to-noise ratio. For positive autoregulation the opposite holds. Positive autoregulation has a smaller knee frequency in the gain-to-noise ratio than negative autoregulation (see also section *Autoregulation*). Parameters: $k_s = 10$, $k_v = 100$, $k_x = 10$, $\mu_v = 5$, $\mu_x = 0.5$, $K = \langle v\rangle$ and $\nu^a = 200$, $\nu^r = 200$. **E**: The gain-to-noise ratio for a cascade with postive autoregulation (red line) and two simple cascades (black). Solid: as in **D**, the degradation rate of the simple cascade equals that of the cascade with positive autoregulation, $\mu_v = \mu_v^{ar}$. Hence $|J_{vv}|$ is smaller in the cascade with auto-activation, and the gain-to-noise ratio is larger at low frequencies. Dotted: In the simple cascade we take $\mu_v = J_{vv}^{\text{ar}}$, and instead increase the production rate J_{xv}. This decreases the gain-to-noise ratio of the simple cascade with respect to the autoregulated cascade over the full frequency spectrum. Inset: the ratio of the gain-to-noise ratio of the cascade with positive autoregulation to that of the simple cascade; solid: $\mu_v = \mu_v^{ar}$, dotted $\mu_v = J_{vv}^{ar}$. The dashed red vertical line indicates J_{vv}^{ar}, the vertical solid red line μ_v^{ar}, which shows the shift in frequency dependence. Parameters: $k_s = 10$, $k_v = 100$, $k_x = 10$, $\mu_v = 5$, $\mu_x = 0.5$, $K = \langle v\rangle$ and $\nu = 200$.

A



B

$$\frac{dv}{dt} = f(x)\, s - \mu_v v + \eta_v \tag{22a}$$

$$\frac{dx}{dt} = \beta v - \mu_x x + \eta_x, \tag{22b}$$

where

$$f(x) = \frac{\nu C^n}{K^n + x^n} \begin{cases} C = x, & \text{positive feedback} \\ C = K, & \text{negative feedback} \end{cases}$$

C

$$g^2 = \frac{(J_{vs} J_{xv})^2}{(\omega^2 + \mu_v^2)(\omega^2 + \mu_x^2) + F(\omega)} \tag{23a}$$

$$N = \frac{\beta^2 \langle |\eta_v|^2 \rangle + (\omega^2 + \mu_v^2) \langle |\eta_x|^2 \rangle}{(\omega^2 + \mu_v^2)(\omega^2 + \mu_x^2) + F(\omega)} \tag{23b}$$

$$\frac{g^2}{N} = \frac{(J_{vs} J_{xv})^2}{\beta^2 \langle |\eta_v|^2 \rangle + (\omega^2 + \mu_v^2) \langle |\eta_x|^2 \rangle}, \tag{23c}$$

$$P_{xx}(\omega) = \frac{J_{vs}^2 J_{xv}^2 P_{ss}(\omega) + J_{xv}^2 \langle |\eta_v|^2 \rangle + (\omega^2 + \mu_v^2) \langle |\eta_x|^2 \rangle}{(\omega^2 + \mu_v^2)(\omega^2 + \mu_x^2) + F(\omega)} \tag{23d}$$

with

$$F(\omega) = J_{vx} J_{xv} \left[J_{vx} J_{xv} + 2 (\omega^2 - \mu_x \mu_v) \right] \tag{24}$$

D

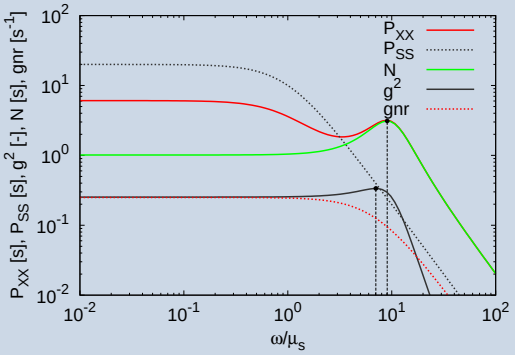


FIG. 5. Feedback from the output signal x to an upstream component, discussed in section *Feedback* . **A**: Cartoon of the negative feedback motif, where the output signal x negatively regulates v. **B**: The Langevin equations describing the network. **C**: The characteristic equations: gain $g^2(\omega)$, noise $N(\omega)$ and gain-to-noise ratio $g^2(\omega)/N(\omega)$. **D**: Power spectra of the output, $P_{xx}(\omega)$, input, $P_{ss}(\omega)$, gain $g^2(\omega)$ and noise $N(\omega)$. $P_{xx}(\omega)$, $g^2(\omega)$ and $N(\omega)$ all exhibit a peak due to the negative feedback, while the gain-to-noise ratio is monotonically decreasing. The black dots indicate the peaks of the gain and noise, which occur at different frequencies (see section *Feedback*). Parameters: $k_s = 10$, $K = 0.2 \langle x \rangle$, $\nu = 1260$, $k_x = 5$, $\mu_v = 5$, $\mu_x = 5$ and $n = 3$.

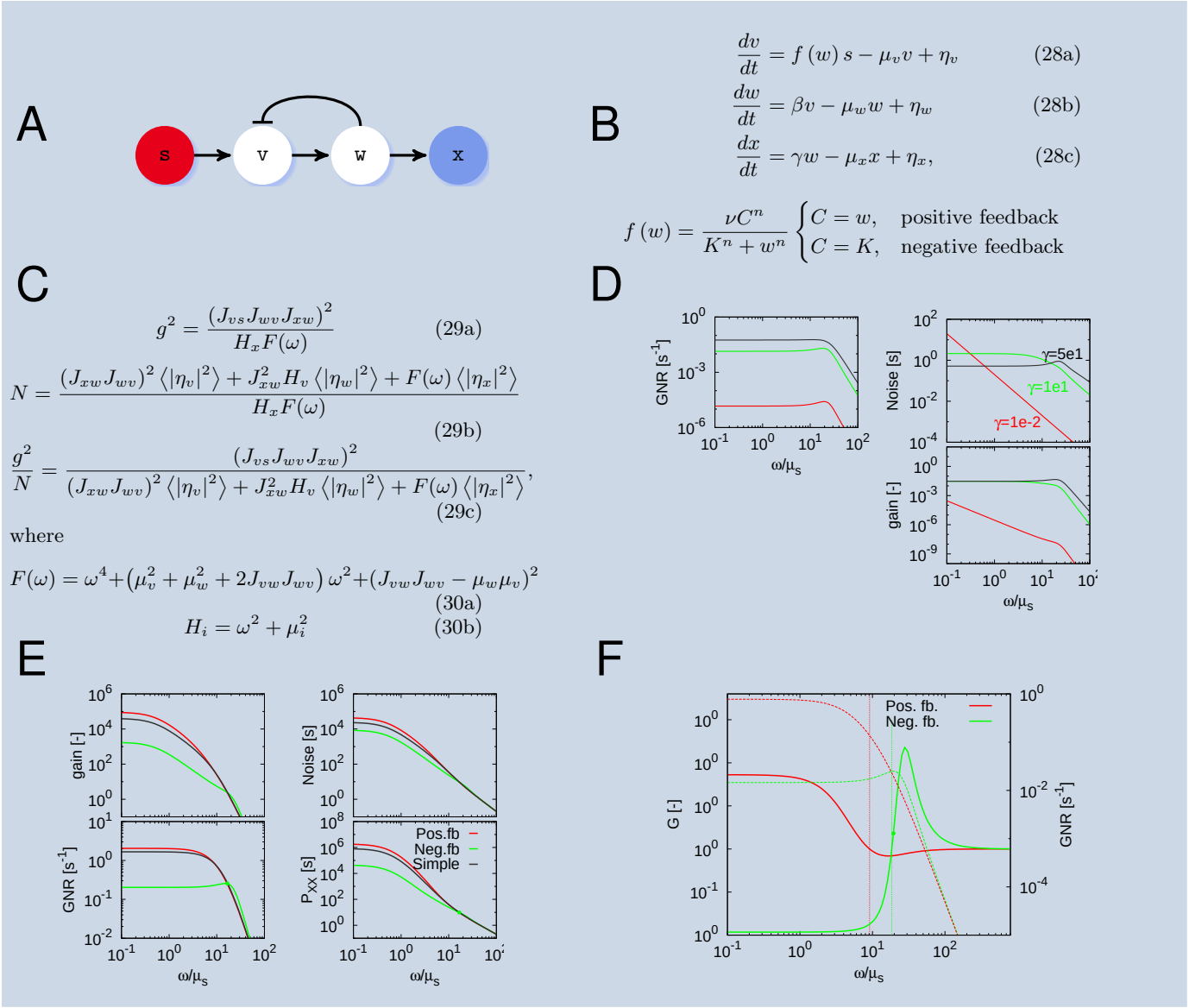


FIG. 6. A three-step cascade with feedback from an intermediate component, discussed in section *Feedback* . **A**: Cartoon of a negative feedback motif, where s is the signal and x the response, and w negatively regulates v. **B**: The Langevin equations of this motif. **C**: The characteristic equations: gain $g^2(\omega)$, noise $N(\omega)$ and gain-to-noise ratio $g^2(\omega)/N(\omega)$. **D**: The effect of changing the strength of the intrinsic noise in x, $N_x(\omega)$, on the spectra of the gain, noise, and gain-to-noise ratio of a cascade with negative feedback. $N_x(\omega)$ is varied by changing $\gamma(= J_{xw})$ and μ_x, in such a way that $\langle x\rangle$ remains constant. Lines show: black, $\gamma = 50$; green, $\gamma = 10$; red, $\gamma = 0.01$. Decreasing γ and μ_x in this way increases the relative contribution of $N_x(\omega)$ to the total noise. We see that as γ is reduced the gain and noise decrease at frequencies $\omega < \mu_x$, but the noise increases at lower frequencies. The gain-to-noise ratio decreases at all frequencies. However, the peak in the gain-to-noise ratio becomes more pronounced. Parameters: $k_s = 10$, $\beta = 10$, $\nu = 330$, $K = 0.5\langle w\rangle$, $n = 5$, $\mu_v = 10$, $\mu_w = 10$. **E**: The spectra of the gain, noise, gain-to-noise ratio, and the output power, $P_{xx}(\omega)$. For small ω, positive feedback (red line) enhances the gain, noise, and gain-to-noise ratio, while negative feedback (green line) decreases these. For higher frequencies, negative feedback increases the gain, enhancing the gain-to-noise ratio. With negative feedback a peak in the gain-to-noise ratio is present (denoted by the green dot), while none is present in the output power spectrum $P_{xx}(\omega)$. Parameters: $k_s = 10$, $\mu_w = 10$, $\mu_v = 10$, $\mu_x = 0.5$, $\beta = 10$, $\gamma = 10$. For positive feedback: $K = 0.5\langle w\rangle$, $n = 1$ and $\nu = 150$. For negative feedback: $K = 0.5\langle w\rangle$, $n = 4$ and $\nu = 1700$. **F**: Solid lines show $G_{\rm fb}(\omega)$ (left axis), the gain-to-noise ratio for networks with positive (red) or negative (green) feedback divided by that of the corresponding simple cascade. Relative to the simple cascade, positive feedback increases the gain-to-noise ratio at low frequencies, while negative feedback increases the gain-to-noise ratio at high frequencies. Vertical lines indicate the frequencies at which $G_{\rm fb}(\omega) = 1$ (Eqn. 27). Dashed lines show the gain-to-noise ratios for the positive and negative feedback motifs (right axis). Parameters: $k_s = 10$, $\mu_w = 10$, $\mu_v = 10$, $\mu_x = 1$, $\beta = 10$, $\gamma = 1$ and $K = 0.5\langle w\rangle$. For positive feedback: $n = 1$ and $\nu = 150$. For negative feedback $n = 5$ and $\nu = 3300$.

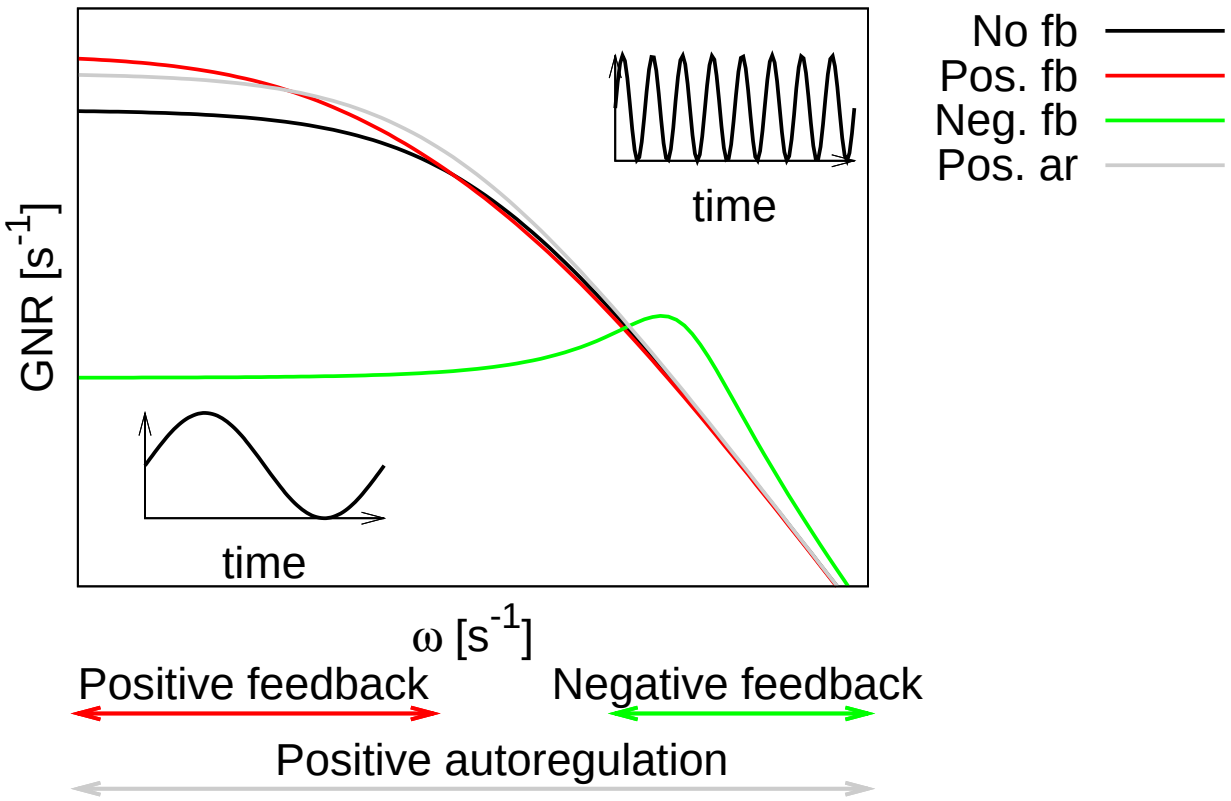


FIG. 7. A schematic drawing of the main conclusions of this paper. The frequency of the variations of the input is shown on the x-axis. For three different motifs the gain-to-noise ratio is shown. The arrows indicate the specific frequency regime for which each motif performs better with respect to a simple cascade.

S1-1. S1

All cascades have the following simple (linear) birth-death process for the signal

$$\frac{ds}{dt} = k_s - \mu_s s + \Gamma(t) \tag{S1-1}$$

S1-2. GILLESPIE SIMULATIONS

The linearization used in the derivation can change the characteristics of the frequency response. A linearized system does not change the frequency of the transmitted signal. However, this may not be the case for a non-linear system. To study this, we performed Gillespie simulations of the full system. The positive and negative regulation in our networks arises from Hill-like interactions between components. In the Gillespie simulation we calculated the propensities for every reaction with identical expressions. For example, in the network with negative feedback from w to v, we model reactions like Eqn. 1a in 6B as

$$S \xrightarrow{r} S + V \tag{S1-2}$$

where r is

$$r = \frac{\nu K^n s}{K^n + w^n}. \tag{S1-3}$$

In these equations the actual copy number w is used, and not $\langle w \rangle$, as in the linearized expressions (Eqns. 2a,2b in Fig. 6C).

The power spectra are calculated using 2^{11} (2048) exponentially distributed frequencies from $\omega = 10^{-3}$ to $\omega = 10^3$ and averaged over 2^4 neighbouring frequencies to obtain a single data point. In total we have 2^7 data-points. The length of the simulation is 10^6 seconds, or a maximum of 10^9 events. For every run 50 blocks are averaged.

The positive feedback loops considered here display bistability. For the positive feedback loops a constant low level production is added to drive the system to the stable state with high copy numbers, instead of the stable state where the copy number equals zero. For the positively autoregulated component this is described by that

$$\frac{dv}{dt} = -\mu_v v + \eta_v + \begin{cases} \frac{\nu v s}{K+v} \text{ if } v \neq 0 \\ \frac{1}{1000} \text{ if } v = 0 \end{cases} \tag{S1-4}$$

Linearizing this we find that the fluctuations follow

$$\frac{d\tilde{v}}{dt} = -\mu_v \tilde{v} + \eta_v - \frac{\nu K \langle s \rangle}{K + \langle v \rangle}\tilde{v} + \frac{\nu K}{K + \langle v \rangle}\tilde{s} \tag{S1-5}$$

which is equivalent to the linearization of Eqn. 1 in Fig. 4A. The addition of the basal expression therefore drives the system to a specific steady state, but does not change the dynamic behaviour around this steady state.
For positive feedback within the cascade, the motif is described by

$$\frac{dw}{dt} = a + k_w v - \mu_w w + \eta_w. \tag{S1-6}$$

Taking different values for $a = 0.1, 1, 10$ does not lead to qualitatively different answers (see Fig. S1-4). Again, the basal production changes the steady state, but not the dynamical behaviour around the steady state.

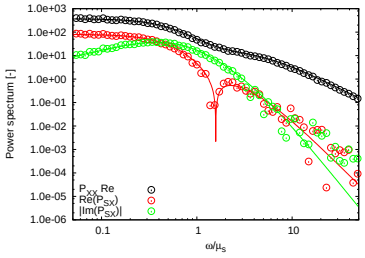


Figure S1-1. The results (circles) of the Gillespie simulations for the linear cascade (Eqn. 11), together with the results of the linear noise approximation (solid lines) as employed in the main text. Kinetic rates as in figure 2.

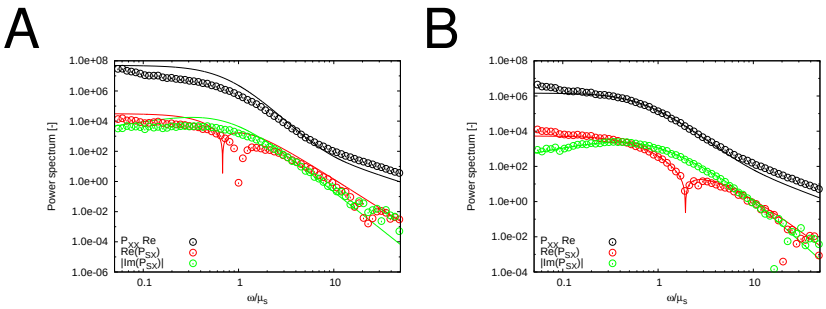


Figure S1-2. **A**) The results (circles) of the Gillespie simulations for the network with positive autoregulation of v (Eqn.17 in Fig. 4B) (kinetic rates as in Fig. 4C with positive autoregulation), together with the results of the linear noise approximation (solid lines) as employed in the main text. To drive the system to the non-zero steady state, basal production of v is present (Eqn. S1-4). The steady state of the full non-linear Gillespie simulation is slightly different from the steady state derived from the mathematical expressions for s, v and x. This causes the slight difference between the results of the linearization and the simulations. **B**) The results (circles) of the Gillespie simulations for a network with negative autoregulation on v (Eqn.17 in Fig. 4B). Together with the results of the linear noise approximation (solid lines) as employed in the main text. Kinetic rates as in Fig. 4D with negative autoregulation.

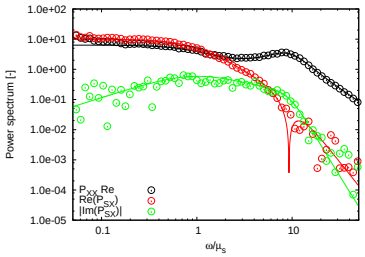


Figure S1-3. The results (circles) of the Gillespie simulations for a network with negative feedback from x to v (Eqn.22 in Fig. 5B), together with the results of the linear noise approximation (solid lines) as employed in the main text. Kinetic rates as in Fig. 5D.

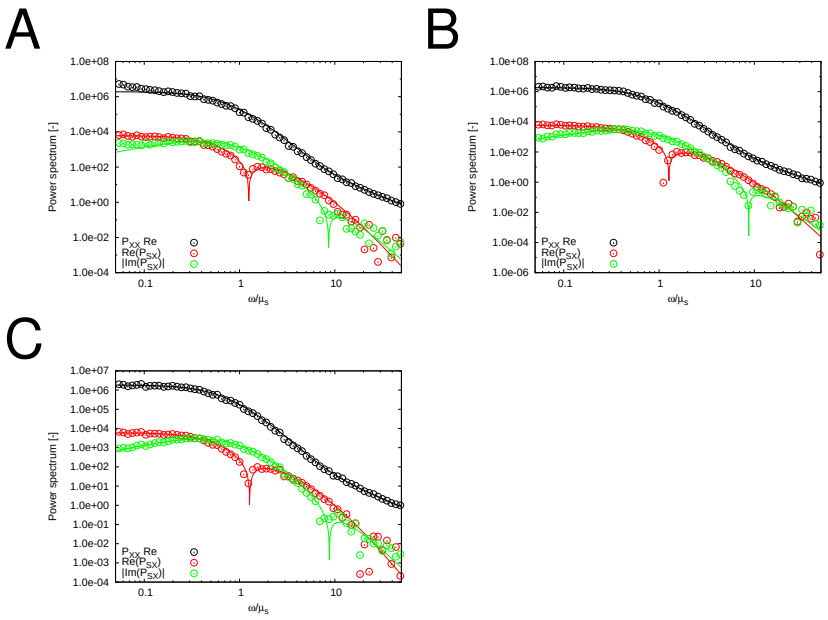


Figure S1-4. The results (circles) of the Gillespie simulations for a network with positive feedback from w to v (Eqn.28 in Fig. 6B), together with the results of the linear noise approximation (solid lines) as employed in the main text. Kinetic rates as in Fig. 6D with positive feedback. To drive the system to the non-zero steady state, basal production of w is present (Eqn. S1-6), **A**) with $a = 0.1$, **B**) with $a = 1$, **C**) with $a = 10$

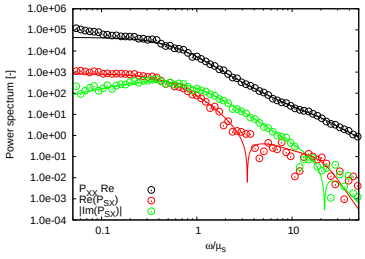


Figure S1-5. The results (circles) of the Gillespie simulations for a network with negative feedback from w to v (Eqn.28 in Fig. 6B), together with the results of the linear noise approximation (solid lines) as employed in the main text. Kinetic rates as in Fig. 6D with negative feedback.

S1-3. SIMPLE CASCADE

The one step simple cascade is described by

$$\frac{dx}{dt} = k_x s - m_x x + \eta_x(t) \tag{S1-7}$$

with the following characteristic equations

$$g^2(\omega) = \frac{k_x^2}{\omega^2 + m_x^2} \tag{S1-8a}$$

$$N(\omega) = \frac{\left\langle |\eta_x|^2 \right\rangle}{\omega^2 + m_x^2} \tag{S1-8b}$$

$$\frac{g^2}{N} = \frac{k_x^2}{\left\langle |\eta_x|^2 \right\rangle} \tag{S1-8c}$$

Here, $\left\langle |\eta_x|^2 \right\rangle = k_x \langle s \rangle + m_x \langle x \rangle = 2k_x \langle s \rangle$. The three component simple cascade is described by (compare Eqn. 11)

$$\frac{dv}{dt} = k_v s - m_v v + \eta_v(t) \tag{S1-9a}$$

$$\frac{dx}{dt} = k_x v - m_x x + \eta_x(t) \tag{S1-9b}$$

with the following characteristic equations

$$g^2(\omega) = \frac{k_v^2 k_x^2}{(\omega^2 + m_x^2)(\omega^2 + m_v^2)} \tag{S1-10a}$$

$$N(\omega) = \frac{k_x^2 \left\langle |\eta_v|^2 \right\rangle + (\omega^2 + m_v^2) \left\langle |\eta_x|^2 \right\rangle}{(\omega^2 + m_x^2)(\omega^2 + m_v^2)} \tag{S1-10b}$$

$$\frac{g^2}{N} = \frac{k_v^2 k_x^2}{k_x^2 \left\langle |\eta_v|^2 \right\rangle + (\omega^2 + m_v^2) \left\langle |\eta_x|^2 \right\rangle} \tag{S1-10c}$$

The simple cascade is used as a reference. For the kinetic rates of the simple cascade we use roman symbol (k and m). For the kinetic rates of the cascades with feedback regulation we use greek symbols.

S1-4. AUTOREGULATION

A. Autoregulation by x

An elementary network for autoregulation by x onto itself is

$$\frac{dx}{dt} = f(x)\, s - \mu_x x + \eta_x(t), \text{where} \tag{S1-11a}$$

$$f(x)\, s = \frac{\nu\kappa}{K + x} s \begin{cases} \kappa = K, & \text{negative regulation} \\ \kappa = x, & \text{positive regulation} \end{cases} \tag{S1-11b}$$

The gain, noise and gain-to-noise for this network are

$$g^2(\omega) = \frac{J_{xs}^2}{\omega^2 + J_{xx}^2} = \frac{1}{\omega^2 + \left(\mu_x - \frac{\partial f(\langle x\rangle)}{\partial \langle x\rangle}\langle s\rangle\right)^2} f(\langle x\rangle)^2 \tag{S1-12a}$$

$$N(\omega) = \frac{\left\langle |\eta_x|^2 \right\rangle}{\omega^2 + J_{xx}^2} \tag{S1-12b}$$

$$\frac{g^2}{N} = \frac{J_{xs}^2}{\left\langle |\eta_x|^2 \right\rangle} = \frac{\left(\frac{\partial f(\langle s\rangle, \langle x\rangle)}{\partial \langle s\rangle}\right)^2}{\left\langle |\eta_x|^2 \right\rangle} \tag{S1-12c}$$

For equal average production, as the simple three component cascade, (production rate k_x), we chose

$$\langle f(x)\, s\rangle \equiv f(\langle x\rangle)\langle s\rangle = k_x \langle s\rangle \tag{S1-13}$$

where the first equation expresses the fact that we assume that the average rates can be expressed by the rates at the deterministic steady state, thus ignoring fluctuations. Thus

$$k_x = J_{xs} = \frac{\partial f(\langle s\rangle, \langle x\rangle)}{\partial \langle s\rangle} = \frac{\nu\kappa}{(K + \langle x\rangle)} \tag{S1-14}$$

and $\left\langle |\eta_x|^2 \right\rangle = 2k_x \langle s\rangle$. Expressed in terms of the kinetic rates of the simple cascade, the autoregulated cascade has the following form

$$GNR = \frac{k_x^2}{2k_x \langle s\rangle} = \frac{k_x}{2\langle s\rangle} \tag{S1-15}$$

which is identical to Eqn. S1-8c. The power spectrum of x for the autoregulated cascade is

$$P_{XX}(\omega) = \frac{J_{xs}^2 \left\langle |\Gamma|^2 \right\rangle + (\omega^2 + \mu_s^2)\left\langle |\eta_x|^2 \right\rangle}{(\omega^2 + \mu_s^2)\left(\omega^2 + \left(\mu_x - \frac{\partial f(\langle s\rangle, \langle x\rangle)}{\partial \langle x\rangle}\right)^2\right)} \tag{S1-16}$$

Following a rescaling of the kinetic degradation rate μ_x, such that $\mu_x^{\text{new}} = \mu_x - J_{xx}$, we observe that the power spectrum of the simple cascade and the autoregulated cascade agree. This is because the noise term η_x depends on the mean rate of the production and degradation events. In steady state the average number of production events equals the average number of degradation events. Since by the rescaling the production is not changed, the noise η_x is constant. The change in $\mu_x \to \mu_x^{\text{new}}$ will lead to a new steady state value $\langle x\rangle$, but not to a different number of degradation events.

B. Autoregulation by v

For autoregulation of one of the intermediate components the network is

$$\frac{dv}{dt} = f(v)\, s - \mu_v v + \eta_v(t) \tag{S1-17a}$$

$$\frac{dx}{dt} = \beta v - \mu_x x + \eta_x(t), \text{where} \tag{S1-17b}$$

$$f(v)\, s = \frac{\nu\kappa}{K + v} s \begin{cases} \kappa = K, & \text{negative regulation} \\ \kappa = v, & \text{positive regulation} \end{cases} \tag{S1-17c}$$

The gain and noise for this network are

$$g^2(\omega) = \frac{(J_{xv}J_{vs})^2}{(\omega^2+\mu_x^2)(\omega^2+J_{vv}^2)} \tag{S1-18a}$$

$$N(\omega) = \frac{\beta^2\left\langle|\eta_v|^2\right\rangle + (\omega^2+J_{vv}^2)\left\langle|\eta_x|^2\right\rangle}{(\omega^2+\mu_x^2)(\omega^2+J_{vv}^2)} \tag{S1-18b}$$

where $J_{vs} = f(\langle v\rangle)$. We equalize the production for v between the autoregulated cascade and the three component simple cascade (with rates k_v and k_x) to obtain

$$k_v = J_{vs} = \frac{\nu\kappa}{(K+\langle v\rangle)} \tag{S1-19}$$

and the gain-to-noise ratio for the autoregulated cascade expressed in terms of the kinetic rates of the simple cascade (thus using k_v, k_x and m_v, m_x where applicable) is

$$\frac{g^2}{N} = \frac{(J_{vs}J_{xv})^2}{\beta^2\left\langle|\eta_v|^2\right\rangle + (\omega^2+J_{vv}^2)\left\langle|\eta_x|^2\right\rangle} = \frac{k_v^2\beta^2}{\beta^2\left\langle|\eta_v|^2\right\rangle + \left(\omega^2+\left(\mu_v^2-\frac{\partial f(\langle s\rangle,\langle v\rangle))}{\partial\langle v\rangle}\right)^2\right)\left\langle|\eta_x|^2\right\rangle}. \tag{S1-20}$$

We keep all kinetic rates equal in the autoregulated and simple cascade that do not influence the constraint condition (Eqn. S1-19) (thus $\mu_v = m_v$ and $\beta = k_x$). We then obtain

$$\frac{g^2}{N} = \frac{k_v^2k_x^2}{k_x^2\left\langle|\eta_v|^2\right\rangle + \left(\omega^2+\left(m_v^2-\frac{\partial f(\langle s\rangle,\langle v\rangle))}{\partial\langle v\rangle}\right)^2\right)\left\langle|\eta_x|^2\right\rangle}. \tag{S1-21}$$

We note that for positive autoregulation $|J_{vv}| < \mu_v$ while for negative autoregulation $|J_{vv}| > \mu_v$. Thus the *GNR* is larger for the positively autoregulated than the three component cascade, especially for $\omega < J_{vv}$. For the negatively autoregulated cascade the opposite holds.

The constraint does not lead to a unique relation between autoregulated and non-autoregulated cascade. An alternative choice would be a simple three component cascade for which the degradation rate μ_v is equivalent to the "effective" degradation rate in the autoregulated cascade. Thus $m_v = J_{vv}^{ar}$. The production of x is then

$$\overbrace{\beta\frac{k_v}{\mu_v}}^{\text{autoregulated}} = \overbrace{k_x\frac{k_v}{m_v}}^{\text{three component cascade}} \tag{S1-22}$$

Equalizing this leads to

$$\beta\frac{k_v}{\mu_v} = k_x\frac{k_v}{m_v} \tag{S1-23}$$

$$\rightarrow k_x = \beta\frac{m_v}{\mu_v}, \tag{S1-24}$$

which leads to

$$\frac{g^2}{N} = \frac{\left(\beta\frac{\mu_v^{casc}}{\mu_v}k_v\right)^2}{\left(\beta\frac{m_v}{\mu_v}\right)^2\left\langle|\eta_v|^2\right\rangle + \left(\omega^2+(m_v)^2\right)\left\langle|\eta_x|^2\right\rangle} = \frac{(\beta k_v)^2}{\beta^2\left\langle|\eta_v|^2\right\rangle + \left(\frac{\mu_v}{m_v}\right)^2\left(\omega^2+(m_v)^2\right)\left\langle|\eta_x|^2\right\rangle} \tag{S1-25}$$

for the *GNR* of the three component simple cascade. Since for positive feedback $m_v < \mu_v$, the *GNR* of the positively autoregulated cascade is larger than that of the simple cascade, especially if $\omega \gg m_v$ or $\omega^2 \gg \beta^2\left\langle|\eta_v|^2\right\rangle$.

If we allow for even more differences between the kinetic rates, but require equal production, we obtain the following equations (we still assume the signal to be identical in both cases)

$$\mu_v = Cm_v \text{ and } \beta = Ck_x, \tag{S1-26}$$

where C is an arbitrary constant. We note that the mean level of v differs between the autoregulated and the simple cascade

$$\langle v\rangle^{ar} = \frac{1}{C}\langle v\rangle^{\text{simple}} \tag{S1-27}$$

As a result we derive for the gain-to-noise ratio for the regulated cascade (using Eqns. S1-20, S1-26 and S1-27)

$$\begin{aligned}\frac{g^2}{N} &= \frac{\left(Ck_xk_v\right)^2}{\left(Ck_x\right)^2\left\langle|\eta_v|^2\right\rangle + \left(\omega^2 + \left(C\mu_v - \frac{\partial f(\langle s\rangle,\langle v\rangle)}{\partial\langle v\rangle}\right)^2\right)\left\langle|\eta_x|^2\right\rangle} \\ &= \frac{\left(k_xk_v\right)^2}{\left(k_x\right)^2\left\langle|\eta_v|^2\right\rangle + \left(\frac{\omega^2}{C^2} + \left(\mu_v - \frac{1}{C}\frac{\partial f(\langle s\rangle,\langle v\rangle)}{\partial\langle v\rangle}\right)^2\right)\left\langle|\eta_x|^2\right\rangle}\end{aligned} \tag{S1-28}$$

For small ω the conclusions on positive and negative feedback are still valid, but for $\omega\to\infty$ the ratio of the *GNR* for positive feedback and a three component cascade is a function of C. Similar arguments can be made about comparing negative and positive feedback for $\omega\to\infty$, where again the ratio of the gain-to-noise ratio's depends on C.

S1-5. FEEDBACK

A. Feedback from x to v

An elementary system with feedback from x to v is

$$\frac{dv}{dt} = f(x)\,s - \mu_v v + \eta_v(t) \tag{S1-29a}$$

$$\frac{dx}{dt} = \beta v - \mu_x x + \eta_x(t), \tag{S1-29b}$$

where

$$f(x)\,s = \frac{\nu\kappa^n s}{K^n + x^n}\begin{cases}\kappa = K, & \text{negative feedback}\\ \kappa = x, & \text{positive feedback}\end{cases} \tag{S1-30}$$

For the gain, noise and $P_{XX}(\omega)$ we obtain

$$g^2(\omega) = \frac{\left(J_{vs}\beta\right)^2}{\left(\omega^2+\mu_v^2\right)\left(\omega^2+\mu_x^2\right) + J_{vx}\beta\left[J_{vx}\beta + 2\left(\omega^2 - \mu_x\mu_v\right)\right]} \tag{S1-31a}$$

$$N(\omega) = \frac{\beta^2\left\langle|\eta_v|^2\right\rangle + \left(\omega^2+\mu_v^2\right)\left\langle|\eta_x|^2\right\rangle}{\left(\omega^2+\mu_v^2\right)\left(\omega^2+\mu_x^2\right) + J_{vx}\beta\left[J_{vx}\beta + 2\left(\omega^2 - \mu_v\mu_x\right)\right]} \tag{S1-31b}$$

$$P_{XX}(\omega) = \frac{\left(J_{vs}\beta\right)^2 P_{SS}(\omega) + \beta^2\left\langle|\eta_v|^2\right\rangle + \left(\omega^2+\mu_v^2\right)\left\langle|\eta_x|^2\right\rangle}{\left(\omega^2+\mu_v^2\right)\left(\omega^2+\mu_x^2\right) + J_{vx}\beta\left[J_{vx}\beta + 2\left(\omega^2 - \mu_v\mu_x\right)\right]}, \tag{S1-31c}$$

where

$$J_{vx} = \frac{\partial\langle f(x)\,s\rangle}{\partial\langle x\rangle} = \frac{\partial\dfrac{\nu\langle\kappa\rangle^n\langle s\rangle}{K^n + \langle x\rangle^n}}{\partial\langle x\rangle} \tag{S1-32}$$

We note that the *GNR* is independent of J_{vx}. The peak in P_{XX} only exist if $J_{vx} < 0$, since

$$P_{XX} = \frac{(J_{vs}\beta)^2 P_{SS}}{(\omega^2+\mu_v^2)(\omega^2+\mu_x^2)+J_{vx}\beta\left[J_{vx}\beta+2(\omega^2-\mu_v\mu_x)\right]}+ \\ + \frac{\beta^2\left\langle|\eta_v|^2\right\rangle}{(\omega^2+\mu_v^2)(\omega^2+\mu_x^2)+J_{vx}\beta\left[J_{vx}\beta+2(\omega^2-\mu_v\mu_x)\right]}+ \\ + \frac{\mu_v^2\left\langle|\eta_x|^2\right\rangle}{(\omega^2+\mu_v^2)(\omega^2+\mu_x^2)+J_{vx}\beta\left[J_{vx}\beta+2(\omega^2-\mu_v\mu_x)\right]}+ \\ + \frac{\omega^2\left\langle|\eta_x|^2\right\rangle}{(\omega^2+\mu_v^2)(\omega^2+\mu_x^2)+J_{vx}\beta\left[J_{vx}\beta+2(\omega^2-\mu_v\mu_x)\right]} \quad \text{(S1-33)}$$

which are four monotonic decreasing functions of ω for $J_{vx} > 0$. So only for negative feedback a peak can exist in the power spectrum, gain and noise (since the same argument applies to gain and noise).
The frequency of the maximum of the gain can easily be obtained, since it coincides with the minimum of the denominator D

$$D = \left(\omega^2+\mu_v^2\right)\left(\omega^2+\mu_x^2\right)+J_{vx}\beta\left[J_{vx}\beta+2\left(\omega^2-\mu_x\mu_v\right)\right]. \quad \text{(S1-34)}$$

This frequency, where the gain has a maximum, is

$$\omega_{\text{res}}^2 = -\frac{1}{2}\left[\mu_v^2+\mu_x^2+2J_{xv}\beta\right], \quad \text{(S1-35)}$$

such that we require $\mu_v^2+\mu_x^2+2J_{xv}\beta < 0$. As a check we note that $D > 0$ for ω_{res} so divergence is not possible. The maximum frequency for the noise is not the minimum of D, due to the ω-dependence in the numerator. If $\beta^2\left\langle|\eta_v|^2\right\rangle \gg \left\langle|\eta_x|^2\right\rangle$, the ω-dependence in the noise is less strong, and the frequency of the peak of the noise shifts to the frequency of the peak in the gain. Although a peak in P_{XX} can be derived analytically ($\frac{dP_{XX}}{d\omega}$ is 4th order in ω^2), it is not insightful. We note that P_{XX} is the sum of the noise (N) and the signal (Σ), such that if one of these two dominates in P_{XX} the peak is likely to coalesce with the peak of the dominating term. We also note that the signal Σ depends on μ_s, so the peak in P_{XX} is not likely to coincide exactly with the peak in the gain, since the gain is independent of μ_s.
Compared with a three component cascade (rates k_v, k_x) , requiring equal production, we note that

$$k_v = \frac{\nu\kappa^n}{K^n+\langle x\rangle^n} \quad \text{(S1-36)}$$

and the three component cascade has an identical *GNR* as the cascade with regulation.

B. Linear Stability Analysis and Control theory

We now shift gears and use some methods from linear stability analysis to study the biochemical network from a slightly different perspective. After linearizing, the solution to the linear differential equations for the perturbations is (ignoring the added noise)

$$\frac{d\tilde{\boldsymbol{y}}(t)}{dt} = \mathbf{J}\tilde{\boldsymbol{y}} \quad \text{(S1-37a)}$$

$$\tilde{\boldsymbol{y}}(t) = e^{\mathbf{J}t}\tilde{\boldsymbol{y}}(0) = \mathbf{R}e^{\lambda t}\mathbf{L}\tilde{\boldsymbol{y}}(0) \quad \text{(S1-37b)}$$

Where $\mathbf{J}$ is the Jacobian, with eigenvalues λ_i and right eigenvectors r_i. The exponential matrix ($e^{\mathbf{J}t}$) describes the time dependency, and can decomposed in a matrix with diagonal entries e^{λ_i}, $\mathbf{R}$ with the right eigenvectors (as columns) and $\mathbf{L}$ with the left eigenvectors (as rows). Alternatively, we could write down the solution in terms of the right eigenvectors

$$\tilde{\boldsymbol{y}}(t) = c_1 e^{\lambda_1 t}\boldsymbol{r_1} + \ldots + c_n e^{\lambda_n t}\boldsymbol{r_n} \quad \text{(S1-38)}$$

where $c_1 \dots c_n$ are weighing coefficients which are obtained by solving for the initial condition. In both expressions we note that the exponential exponent involves λ. If λ is complex, we can rewrite the exponent as

$$e^{\lambda t} = e^{(\Re_\lambda + i\Im_\lambda)t} = e^{\Re_\lambda t}\left(\cos\left(\Im_\lambda t\right) + i\sin\left(\Im_\lambda t\right)\right) \tag{S1-39}$$

and the fluctutations decay (if $\Re_\lambda < 0$) with characteristic frequency $\Im_\lambda$. For stability we require that $\mu_v\mu_x - J_{xv}J_{vx} > 0$.

Yet another different method is control theory, which we can use to describe our system. In control theory we describe a linear system using the convolution of a response function with the input to determine the output. In the fourier space this becomes multiplication, such that we have (again ignoring noise)

$$V(\omega) = H_1(\omega)\, S(\omega) + X_{fb}(\omega) \tag{S1-40a}$$

$$X_{fb}(\omega) = G_1(\omega)\, X(\omega) \tag{S1-40b}$$

$$X(\omega) = H_2(\omega)\, V(\omega) \tag{S1-40c}$$

so that the total response function between input and output is

$$X(\omega) = \frac{H_1(\omega)\, H_2(\omega)}{1 - H_2(\omega)\, G_1(\omega)} S(\omega) \tag{S1-41}$$

which is, if we take as transfer functions

$$H_1(\omega) = \frac{J_{vs}}{i\omega + \mu_v} \qquad H_2(\omega) = \frac{J_{xv}}{i\omega + \mu_x} \qquad G_1(\omega) = \frac{\gamma}{i\omega + \mu_v} \tag{S1-42}$$

equal to $g(\omega)$.

The phase of the gain, which identifies the phase shift between s and x is

$$\Delta\phi = \arctan\left(\frac{-\omega\left(\mu_v + \mu_x\right)}{-\omega^2 - J_{xv}\gamma + \mu_v\mu_x}\right) \tag{S1-43}$$

We can now define ω_ϕ

$$\omega_\phi = \sqrt{\mu_v\mu_x - J_{xv}\gamma} \tag{S1-44}$$

which defines the frequency for which the phase difference between x and s shifts by a factor π. Since x is also the fedback signal, this is the phase difference between the signals in the feedback loop. For negative feedback ($\gamma < 0$) $\Delta\phi$ moves from 0 to π for ω changing from 0 to ∞.

C. Feedback from w to v

For the regulated four component cascade, the network is

$$\frac{dv}{dt} = \frac{\nu\kappa^n s}{K^n + w^n} - \mu_v v + \eta_v(t) \tag{S1-45a}$$

$$\frac{dw}{dt} = \beta v - \mu_w w + \eta_w(t) \tag{S1-45b}$$

$$\frac{dx}{dt} = \gamma w - \mu_x x + \eta_x(t)\,, \tag{S1-45c}$$

with κ as before (e.g. Eqn. S1-30). We linearize and obtain

$$J_{vs} = \frac{\nu\kappa^n}{K^n + \langle w\rangle^n} \tag{S1-46a}$$

$$J_{vw} = \frac{\partial f\left(\langle w\rangle\right)}{\partial \langle w\rangle}\langle s\rangle = -\frac{\nu n \langle w\rangle^n K^n \langle s\rangle}{\langle w\rangle \left(K^n + \langle w\rangle^n\right)^2}, \tag{S1-46b}$$

where the equation S1-46b is for negative feedback (for positive feedback the sign would be positive). The gain and noise are

$$g^2(\omega) = \frac{J_{vs}^2 \beta^2 \gamma^2}{\left(\omega^2 + \mu_x^2\right) F(\omega)} \tag{S1-47a}$$

$$N(\omega) = \frac{\gamma^2\beta^2\left\langle|\eta_v|^2\right\rangle + \gamma^2\left(\omega^2+\mu_v^2\right)\left\langle|\eta_w|^2\right\rangle + F(\omega)\left\langle|\eta_x|^2\right\rangle}{\left(\omega^2+\mu_x^2\right)F(\omega)} \tag{S1-47b}$$

and

$$F(\omega) = \omega^4 + \left(\mu_v^2 + \mu_w^2 + 2J_{vw}\beta\right)\omega^2 + \left(J_{vw}\beta - \mu_w\mu_v\right)^2, \tag{S1-48}$$

where $F(\omega)$ is a function of the parameters in the feedback loop only. The *GNR* is described by $C/a(\omega)$, i.e. a constant divided by a function of ω. For this to have an extremum, the denominator should have an extremum. We differentiate and obtain

$$\omega_{\text{peak}}^2 = -\frac{1}{2}\left(\gamma^2\frac{\left\langle|\eta_w|^2\right\rangle}{\left\langle|\eta_x|^2\right\rangle} + \mu_v^2 + \mu_w^2 + 2J_{vw}\beta\right). \tag{S1-49}$$

Since this expression is negative, to have $\omega^2 > 0$ we require negative feedback. Explicitly writing $J_{wv}J_{vw}$, we have for the requirement that a peak exists

$$2J_{wv}J_{vw} = 2\frac{\nu n \langle w\rangle^n K^n \langle s\rangle}{\langle w\rangle\left(K^n + \langle w\rangle^n\right)^2}\beta > \left(\gamma^2\frac{\left\langle|\eta_w|^2\right\rangle}{\left\langle|\eta_x|^2\right\rangle} + \mu_v^2 + \mu_w^2\right) \tag{S1-50}$$

$$\langle w\rangle = \frac{\beta\langle v\rangle}{\mu_w} = \frac{\beta}{\mu_v\mu_w}\frac{\nu K^n\langle s\rangle}{K^n + \langle w\rangle^n} \tag{S1-51}$$

which gives n solutions for $\langle w\rangle$ (of which only one is real and positive). If we constrain the production rate of v and w to be constant - and we assume $\langle v\rangle = \frac{k_v\langle s\rangle}{m_v}$ - then we obtain

$$\frac{\nu K^n}{K^n + \langle w\rangle^n} = k_v$$

and the following expression for Eqn. S1-50

$$\langle w\rangle = \frac{\beta\langle v\rangle}{\mu_w} = \frac{k_v\beta\langle s\rangle}{\mu_w\mu_v}. \tag{S1-52}$$

we rewrite the coupling strength J_{vw}

$$J_{vw}^{\text{pos}} = \frac{\nu n \langle w\rangle^{n-1} K^n\langle s\rangle}{\left(K^n + \langle w\rangle^n\right)^2} = \frac{n\langle s\rangle k_v}{\langle w\rangle}\frac{K^n}{K^n + \langle w\rangle^n} = \frac{n\langle s\rangle k_v}{\langle w\rangle}\frac{\left(\mu_w\mu_v K\right)^n}{\left(\mu_w\mu_v K\right)^n + \left(k_v k_w\langle s\rangle\right)^n} \tag{S1-53a}$$

$$J_{vw}^{\text{neg}} = -\frac{\nu n \langle w\rangle^{n-1} K^n\langle s\rangle}{\left(K^n + \langle w\rangle^n\right)^2} = -\frac{n\langle s\rangle k_v}{\langle w\rangle}\frac{\langle w\rangle^n}{K^n + \langle w\rangle^n} = -\frac{n\langle s\rangle k_v}{\langle w\rangle}\frac{\left(k_v k_w\langle s\rangle\right)^n}{\left(\mu_w\mu_v K\right)^n + \left(k_v k_w\langle s\rangle\right)^n} \tag{S1-53b}$$

For $K \ll \langle w\rangle$ positive regulation is maximized and J_{vw}^{pos} is maximal, while negative regulation is greatly suppressed and $|J_{vw}^{\text{neg}}|$ is minimal. The limit $n \to \infty$ is more complicated. If $K < \langle w\rangle$, $J_{vw}^{\text{neg}} \to -\infty$, while $J_{vw}^{\text{pos}} \to \infty$ for $\langle w\rangle < K$. In the opposite scenario's the limits tend to zero. This is only valid if while changing n, $\langle w\rangle$ remains constant, which

is true due the constraint.
With Eqn. S1-53 we can study ω_{peak} in more detail and we obtain

$$\begin{aligned} 2\frac{\nu n \langle w\rangle^n K^n \langle s\rangle}{\langle w\rangle \left(K^n + \langle w\rangle^n\right)^2}\beta &> \left(\gamma\mu_w + \mu_v^2 + \mu_w^2\right) \\ \frac{2k_v\beta n \langle s\rangle}{\langle w\rangle}\frac{\langle w\rangle^n}{K^n + \langle w\rangle^n} &> \left(\gamma\mu_w + \mu_v^2 + \mu_w^2\right) \\ 2\frac{\langle w\rangle^n}{K^n + \langle w\rangle^n}\mu_v\mu_w n &> \left(\gamma\mu_w + \mu_v^2 + \mu_w^2\right) \end{aligned} \tag{S1-54}$$

which, interestingly, only has a solution for $n > 1$.
The power spectrum of x is

$$P_{XX}(\omega) = \frac{J_{vs}^2\beta^2\gamma^2 P_{SS} + \gamma^2\beta^2\left\langle|\eta_v|^2\right\rangle + \gamma^2\left(\omega^2 + \mu_v^2\right)\left\langle|\eta_w|^2\right\rangle + F(\omega)\left\langle|\eta_x|^2\right\rangle}{\left(\omega^2 + \mu_x^2\right)F(\omega)}, \tag{S1-55}$$

which depends on μ_s through P_{SS} and therefor will have a peak for a different ω than the *GNR*.
The *GNR* for the simple four component cascade is

$$\frac{g^2(\omega)}{N(\omega)} = \frac{\left(k_v k_w k_x\right)^2}{k_x^2\left[k_w^2\left\langle|\eta_v|^2\right\rangle + \left(\omega^2 + m_v^2\right)\left\langle|\eta_w|^2\right\rangle\right] + \left(\omega^2 + m_v^2\right)\left(\omega^2 + m_w^2\right)\left\langle|\eta_x|^2\right\rangle} = \frac{\left(k_v k_w k_x\right)^2}{D} \tag{S1-56}$$

where we chose ν such that

$$k_v = \frac{\nu\kappa^n}{K^n + \langle w\rangle^n} \tag{S1-57}$$

to obtain equal production. We then obtain for the ratio of the *GNR* of the feedback cascade and a simple cascade

$$G = \frac{GNR_{\text{fb}}}{GNR_{\text{simple}}} = \frac{D}{D + J_{vw}\beta\left[J_{vw}\beta + 2\left(\omega^2 - \mu_v\mu_w\right)\right]\left\langle|\eta_x|^2\right\rangle} \tag{S1-58}$$

So that the feedback is larger if $J_{vw}\beta\left[J_{vw}\beta + 2\left(\omega^2 - \mu_v\mu_w\right)\right]\left\langle|\eta_x|^2\right\rangle < 0$.
The result of this inequality is

$$G_{\text{pos}}(\omega) > 1 \text{ if } \omega^2 < \mu_v\mu_w\left(1 - \frac{n}{2}\frac{K^n}{K^n + \langle w\rangle^n}\right), \tag{S1-59}$$

$$G_{\text{neg}}(\omega) > 1 \text{ if } \omega^2 > \mu_v\mu_w\left(1 + \frac{n}{2}\frac{\langle w\rangle^n}{K^n + \langle w\rangle^n}\right). \tag{S1-60}$$

which are Eqns. 27 from the article. The peak for the negative feedback occurs at ω_{peak} (Eqn. S1-49). The negative feedback cascade is larger than the four component simple cascade if $\omega > \omega_{\text{switch}}$ (Eqns. 27). Thus if $\omega_{\text{peak}} > \omega_{\text{switch}}$ the *GNR* for the negative feedback at the peak is larger than the four component cascade

$$\begin{aligned} \omega_{\text{peak}}^2 &> \mu_v\mu_w\left[1 + \frac{n}{2}\frac{\left(k_v\beta\langle s\rangle\right)^n}{\left(\mu_v\mu_w K\right)^n + \left(k_v\beta\langle s\rangle\right)^n}\right] \\ \left(\gamma^2\frac{\left\langle|\eta_w|^2\right\rangle}{\left\langle|\eta_x|^2\right\rangle} + \mu_v^2 + \mu_w^2 + 2J_{vw}\beta\right) &< -\mu_v\mu_w\left[2 + n\frac{\left(k_v\beta\langle s\rangle\right)^n}{\left(\mu_v\mu_w K\right)^n + \left(k_v\beta\langle s\rangle\right)^n}\right] \\ n\mu_v\mu_w M &> \left(\mu_v + \mu_w\right)^2 + 2\gamma^2\frac{\left\langle|\eta_w|^2\right\rangle}{\left\langle|\eta_x|^2\right\rangle} \end{aligned} \tag{S1-61}$$

which is possible for large n and large $M = \frac{\langle w\rangle^n}{\langle w\rangle^n + K^n}$, which indicates that $K \ll \langle w\rangle$, in both cases representing a strong negative feedback.

S1-6. COMMENTS ON FIG. 6E

Here we list some additional explanation on Fig. 6E. In this figure, we keep the parameters μ_v, μ_w, ν, $\beta\,(=J_{wv})$ and K constant, since they dictate the feedback cycle (Eqn. 1 in Fig. 6B). We vary J_{xw} and μ_x, so that in this case, not the average production rate of x is constrained, but the average copy number $\langle x\rangle$.
To understand the dependence of the gain, noise and gain-to-noise ratio on $\gamma = J_{xw}$ and μ_x,we note that $g^2 \sim \gamma^2 g^2_{s\to w}/(\mu_x^2+\omega^2)$ and $N \sim \gamma^2/(\mu_x^2+\omega^2)N_v(\omega) + \gamma^2/(\mu_x^2+\omega^2)N_w(\omega) + N_x(\omega)$, where $N_v(\omega)$ and $N_w(\omega)$ are independent of γ and μ_x and $N_x(\omega) = 2\gamma\,\langle w\rangle\,/(\mu_x^2+\omega^2)$ (with $\langle w\rangle$ being independent of γ and μ_x).

For $\omega \ll \mu_x$, the contributions of v and w to $N(\omega)$ are proportional to γ^2/μ_x^2, while the contribution of x is given by $N_x(\omega) \propto \gamma/\mu_x^2$. Hence, for $\omega \ll \mu_x$, the contributions of v and w to the noise are constant, while the contribution of x decreases with increasing γ and μ_x, leading to a decrease of $N(\omega)$. Since the gain is constant in this regime, the gain-to-noise ratio increases with increasing γ and μ for $\omega \ll \mu_x$. For $\omega \gg \mu_x$, the gain, and the contributions of v and w to the noise increase with γ^2 while the contribution of x to the noise increases with γ, meaning that also in this regime the gain-to-noise ratio increases with γ and μ_x.