

Feed-forward loops and diamond motifs lead to tunable transmission of information in the frequency domain

W. H. de Ronde,^{*} F. Tostevin, and P. R. ten Wolde*FOM Institute AMOLF, Science Park 104, 1098 XG Amsterdam, Netherlands*

(Received 12 March 2012; revised manuscript received 23 July 2012; published 13 August 2012)

Using a Gaussian model, we study the transmission of time-varying biochemical signals through feed-forward motifs and diamond motifs. To this end, we compute the frequency dependence of the gain, the noise, as well as their ratio, the gain-to-noise ratio, which measures how reliably a network transmits signals at different frequencies. We find that both coherent and incoherent feed-forward motifs can either act as low-pass or high-pass filters for information: The frequency dependence of the gain-to-noise ratio increases or decreases with increasing frequency, respectively. Our analysis of diamond motifs reveals that cooperative activation of the output component can increase the gain-to-noise ratio. This means that from the perspective of information transmission, it can be beneficial to split the input signal in two and recombine the two propagated signals at the output. Cooperative activation can be implemented via the formation of homo- or heteromultimers that then bind and activate the output component or via the binding of individual molecules of the intermediate species to the output component.

DOI: [10.1103/PhysRevE.86.021913](https://doi.org/10.1103/PhysRevE.86.021913)

PACS number(s): 87.18.Mp, 87.18.Tt, 87.18.Vf

I. INTRODUCTION

Cells live in a highly dynamic environment. While in some cases cells may wish to ignore rapid fluctuations and only respond to persistent changes, in other cases they may have to do the opposite. For example, in the case of chemotaxis or osmoadaptation, cells respond to changes in the stimulus but are insensitive to the absolute level of the stimulus. In contrast, in response to a changing sugar concentration, cells respond to the absolute steady-state sugar level, but may wish to integrate out rapid fluctuations of the sugar level. In general, to understand how cells cope with a changing environment, we have to understand how cells transduce time-varying signals. Moreover, given the observation that the biochemical networks which process the signals are stochastic in nature, we have to understand how reliably biochemical networks can process time-varying signals in the presence of noise.

Cells use recurring network motifs to specifically respond to temporal characteristics of the input signal. Negative feedback or incoherent feed-forward loops may be used to only respond to rapid variations and not to slow changes in the environment [1], while coherent feed-forward loops can be used to filter out transient fluctuations in the input and only respond to persistent changes in the environment [2]. To understand how specifically and reliably these motifs can respond to inputs with distinct temporal dynamics, we have to understand how they amplify input signals as a function of their frequency [3], which is characterized by the frequency-dependent gain. Moreover, we have to understand how they propagate biochemical noise as a function of frequency [3]. Indeed, information theory [4] tells us that the fidelity by which a signal of a given frequency is transmitted is determined by the gain-to-noise ratio at that frequency [3]. We have recently shown for motifs with different types of feedback regulation that different network architectures affect the frequency dependence of the gain and the noise differently [5], which means that both of these

quantities have to be studied together in order to understand how reliably a network transmits time-varying signals.

In this paper, we use a Gaussian model to study the frequency dependence of the gain, noise, and gain-to-noise ratio [3,5] of feed-forward loops and diamond motifs (Fig. 1). Both are common motifs in biochemical networks [6,7]. Feed-forward motifs have been shown to regulate many different cellular processes, and, indeed, they exhibit very rich dynamics. Feed-forward motifs can act as sign-sensitive circuits [8], perform adaptation [1], provide fold-change detection [9], or attenuate extrinsic noise [10].

While the mean-field response [1,6–10] and the noise characteristics [11,12] of feed-forward loops and diamond motifs have been well characterized, also in the context of information theory [13], how reliably they propagate time-varying signals has not been studied. This is of specific interest since the data-processing inequality dictates that information transmission decreases with the length of the cascade. If information transmission is the only constraint, the shortest cascade is the most reliable solution. However, as discussed above, cells may wish to respond to specific frequencies in the input signal and they may have to do so reliably. This is precisely what feed-forward loops and diamond motifs can achieve, in contrast to simple cascades.

We find that coherent feed-forward motifs and diamond motifs typically act as low-pass filters for information: They transmit input signals of low frequency more reliably than input signals of high frequency. In contrast, incoherent feed-forward motifs tend to act as high-pass filters for information. These results are not surprising: While for the coherent motifs the gain is large at low frequencies, for incoherent motifs the gain is large at high frequencies. We also show that, in contrast to the intuitive expectation, a coherent feed-forward loop can also act as a high-pass filter, while an incoherent feed-forward loop can also act as a low-pass filter for information.

Our results also reveal that diamond motifs can have a higher gain-to-noise ratio over all frequencies than simple two-level cascades, when the total cost of making the proteins is the same in the two networks under comparison. This means that from the perspective of information transmission,

^{*}deronde@amolf.nl

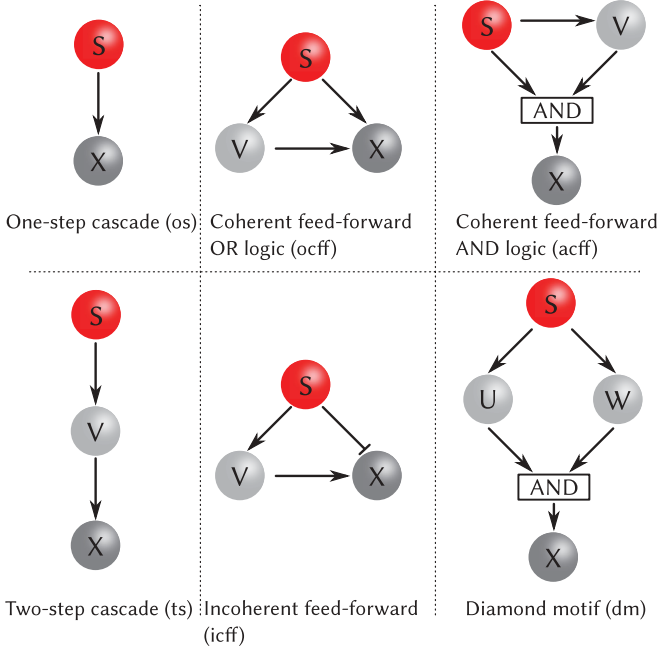


FIG. 1. (Color online) An overview of the different motifs studied in this paper, including their abbreviations as used in the main text.

it is beneficial to split the signal in two and combine the two transmitted signals again at the output. This could be considered as a form of coincidence detection. Interestingly, a diamond motif is not necessary: The same effect can also be achieved via cooperative activation of the output via an intermediate component. For example, the input may stimulate the formation of a homodimer or a homomultimer, which then activates the output; alternatively, the input activates a messenger, for example, a transcription factor, which then activates the output, the gene promoter, in a cooperative fashion. Our analysis suggests that the gain-to-noise ratio increases with the level of cooperativity.

In the next section we briefly introduce the mathematical background, and in the results section we describe the information transmission characteristics of the different feed-forward motifs. Finally, we discuss some of the implications of this work and its limitations.

II. METHODS

In this section, we briefly discuss the mathematical background of our analysis. A more in depth analysis is presented in Refs. [3,5,14]. The biochemical network consists of the components S , V_i (intermediate(s)), and X . Here S is the input and X is the output. We model the time evolution of these components using a set of coupled Langevin equations [15], which can be nonlinear, for example,

$$\frac{ds}{dt} = f_s(s) - \mu_s s + \Gamma(t), \quad (1)$$

$$\frac{dv_i}{dt} = f_{v_i}(s, \mathbf{v}) - \mu_{v_i} v_i + \eta_{v_i}(t), \quad (2)$$

$$\frac{dx}{dt} = f_x(s, \mathbf{v}, x) - \mu_x x + \eta_x(t). \quad (3)$$

For simplicity, we assume linear degradation of each component. The various noise sources (η_{v_i} , η_x) in Eqs. (2) and (3) are assumed to be independent and Gaussian distributed [16–18], which holds for birth-death reactions. These reactions are common in genetic networks and can be the result of coarse graining more complex networks [17]. As a result, we can study the mutual information for the linearized system [19–21]. We take the noise strength $\langle |\eta_{v_i}|^2 \rangle, \langle |\eta_x|^2 \rangle$ as the sum of the average number of production and degradation events per unit amount of time for component v_i, x [22–24]. We assume the noise source $\Gamma(t)$ to be a Gaussian white noise. It generates an ensemble of input trajectories $s(t)$ with Gaussian statistics. The “forces” $f_{s, v_i, x}(s, \mathbf{v}, x)$ model all the reactions involving the production events of s , v_i , and x .

We assume that the network has a steady state and linearize about this steady state, so that we get a dynamical equation for the fluctuations of each component, $\tilde{v}_i(t) = v_i(t) - \langle v_i \rangle$, and similarly for the input s and output x . In the linearized form, the relation between the components i, j is described by the coupling coefficient J . These are the Jacobian elements of the deterministic equations ($\langle \eta(t) \rangle = 0$), and we assume that, for the calculation of the Jacobian, the correlation between two components can be linearized (e.g., $\langle SX \rangle = \langle S \rangle \langle X \rangle$), for example,

$$J_{v_i v_i} = \frac{\partial \frac{d\langle v_i \rangle}{dt}}{\partial \langle v_i \rangle} = -\mu_{v_i}, \quad J_{x v_i} = \frac{\partial \frac{d\langle x \rangle}{dt}}{\partial \langle v_i \rangle}, \quad (4)$$

where the first equation holds if f_{v_i} is independent of v_i itself. This holds for the motifs considered in this paper since there is no feedback by any component on its own expressions. Effects of feedback have been considered previously in Ref. [5]. We take as the input signal the variations of $s(t)$ around its mean $\langle s \rangle$, $\tilde{s}(t)$, and as the output the variations $\tilde{x}(t)$ of $x(t)$ around its mean $\langle x \rangle$. The mutual information rate between the in- and output trajectories, $s(t)$ and $x(t)$, respectively, is defined as [3]

$$R[\tilde{s}(t), \tilde{x}(t)] = \frac{1}{2\pi} \int_0^\infty d\omega \ln \left[1 + \frac{g^2(\omega)}{N(\omega)} P_{ss}(\omega) \right], \quad (5)$$

where P_{ss} is the power spectrum of the signal,

$$P_{ss}(\omega) = \langle \hat{s}(\omega) \hat{s}(-\omega) \rangle, \quad (6)$$

and $\hat{s}(\omega)$ the Fourier transform of $\tilde{s}(t)$. The gain $g^2(\omega)$ and noise $N(\omega)$ are defined through the power spectra

$$g^2(\omega) \equiv \frac{|P_{sx}(\omega)|^2}{P_{ss}^2(\omega)}, \quad (7)$$

$$N(\omega) \equiv P_{xx}(\omega) - g^2(\omega) P_{ss}(\omega). \quad (8)$$

These definitions are prescribed by using Eq. (5). A large gain-to-noise ratio $\Theta(\omega) = g^2(\omega)/N(\omega)$ leads to a high mutual information rate [Eq. (5)] and this implies reliable information transmission.

We have made a number of assumptions to obtain Eq. (5). First, we assume that the linearized system is an accurate representation of the nonlinear system. To ensure the accuracy of the linearization we use parameters such that the average copy number of any species is $\mathcal{O}(100)$, such that nonlinear effect because of low copy numbers are avoided [17,19]. Second, we assume that the variations $\tilde{s}(t)$ and $\tilde{x}(t)$ can be

described by a Gaussian joint-probability distribution. Third, and last, we assume the signal \mathbf{S} to be modular from the underlying network. Modularity of the signal with respect to the network indicates that no correlations exist between the variations in the input signal and the intrinsic noise of the reactions that constitute the processing network; it also implies that there is no feedback from the network onto the input signal. If signal modularity holds, then Eq. (8) is equal to the spectral-addition rule [17]. In this case, the gain-to-noise ratio $\Theta(\omega) = g^2(\omega)/N(\omega)$, does not depend on the input signal, but only on the information transmission characteristics of the processing network; it describes the ability of the network to reliably propagate input signals as a function of their frequency. As an additional simplification we assume that no (anti-)correlations between the different noise sources are present [5]. While these may quantitatively change the results presented below, they do not qualitatively change them. In the Appendix we discuss the effect of cross-correlations in more detail.

In the next section we describe the effect of the feed-forward motif on the information transmission through a biochemical network. We characterize the gain, noise, and gain-to-noise ratio (Θ), since these are intrinsic, signal-independent, properties of the network, when the spectral-addition rule holds [17]. We compare different motifs with simple linear one-step ($\mathbf{S} \rightarrow \mathbf{X}$) and linear two-step cascades ($\mathbf{S} \rightarrow \mathbf{V} \rightarrow \mathbf{X}$). We compare the different networks, unless specified otherwise, under the constraint that the average production rates of the respective components are the same in the networks under comparison.

Last, we comment on the differences between the gain-to-noise ratio, which describes information transmission and $P_{xx}(\omega)$, which describes power transmission.

III. RESULTS

A. Simple cascades

Simple cascades form the building blocks of the feed-forward motif. The feed-forward motif consists of two cascades, one in which the input \mathbf{S} directly regulates the output \mathbf{X} via a one-step cascade and one where \mathbf{S} indirectly regulates the output \mathbf{X} via a two-step cascade with an intermediate component \mathbf{V} . A diamond motif consists of two two-step cascades which start and end at the same component, \mathbf{S} and \mathbf{X} , respectively. Since we compare the behavior of these networks with simple cascades consisting of one or two steps, it is useful to briefly recall their main transmission characteristics. A more detailed discussion can be found in Ref. [5]. For readability we refer to the parameters of simple cascades with Roman symbols, while the parameters in feed-forward motifs are denoted by Greek symbols.

For a one-step cascade the gain is given by $k_x^2/(\omega^2 + m_x^2)$, where $k_x = J_{xs}$ is the coupling between s and x and m_x is the lifetime of \mathbf{X} , while the noise is given by $\langle |\eta_x|^2 \rangle / (\omega^2 + m_x^2)$. Consequently, Θ of a one-step cascade is constant for all frequencies (A5). For the two-step cascade the gain is $(k_v k_x)^2 / (\omega^2 + m_x^2)(\omega^2 + m_v^2)$, with two corner frequencies, m_v (the lifetime of \mathbf{V}), m_x , the noise is $k_x^2 \langle |\eta_v|^2 \rangle / [(\omega^2 + m_v^2)(\omega^2 + m_x^2)] + \langle |\eta_x|^2 \rangle / (\omega^2 + m_x^2)$, and as a result Θ decays with ω^{-2} for $\omega > (m_v + k_x \langle |\eta_v|^2 \rangle / \langle |\eta_x|^2 \rangle)^{1/2}$ (for more details see the Appendix).

B. The feed-forward motif

Two different types of feed-forward (ff) motifs exist. If the sign of the interaction of \mathbf{S} on \mathbf{X} and \mathbf{S} via \mathbf{V} on \mathbf{X} are of the same nature, both either active or inactive, the motif is referred to as *coherent* feed-forward (cff). If the signs of regulation are opposing, the motif is referred to as *incoherent* feed-forward (iff). For the coherent motif we can further differentiate with respect to the integration strategy at the reporter \mathbf{X} . If both \mathbf{S} and \mathbf{V} are required to produce \mathbf{X} , the node \mathbf{X} acts as an AND gate and we refer to the motif as AND-coherent feed-forward (acff), while if either \mathbf{S} or \mathbf{V} is sufficient to produce \mathbf{X} , we refer to the motif as OR-coherent feed-forward (ocff). The AND type is observed in the *ara* system of *E. coli* [8], while the OR type is present in the biosynthesis of the flagellar motor [25]. The motifs are shown in Figs. 2(a) (ocff), 3(a) (acff), and 4(a) (iff). In this section we study an iff motif for which the negative regulation is always at the response \mathbf{X} . As a result, a distinction between AND and OR regulation in the iff is not made. We discuss this assumption in more detail in the section discussing the iff. We now start by studying some general characteristics of the ff motif.

1. General characteristics

We first study the gain. The gain for the ff is

$$g^2(\omega) = \frac{\overbrace{g_{s \rightarrow x}^2}^{J_{xs}^2}}{\omega^2 + J_{xx}^2} + \frac{\overbrace{(J_{vs} J_{xv})^2}^{J_{vs} J_{xv}}}{(\omega^2 + J_{xx}^2)(\omega^2 + J_{vv}^2)} - \frac{\overbrace{2J_{xs} J_{vs} J_{xv} J_{vv}}^{\text{coherence}}}{(\omega^2 + J_{xx}^2)(\omega^2 + J_{vv}^2)}. \quad (9)$$

The first term is the gain due to the direct regulation of \mathbf{X} by \mathbf{S} ($g_{s \rightarrow x}^2$), the second term is due to the pathway $\mathbf{S} \rightarrow \mathbf{V} \rightarrow \mathbf{X}$ ($g_{s \rightarrow v \rightarrow x}^2$), and the third term is a term due to the interaction between the two pathways.

The first term $g_{s \rightarrow x}^2$ is the gain of a one-step cascade in which the input s regulates the output x . It depends on the coupling constant J_{xs} and the lifetime of the protein \mathbf{X} , $\mu_x^{-1} = -J_{xx}^{-1}$. If the lifetime μ_x^{-1} of \mathbf{X} is longer than the time scale ω^{-1} on which the input signal varies, $\omega \gg \mu_x$, then variations in the input $s(t)$ are filtered out by the slow response of \mathbf{X} .

The second term $g_{s \rightarrow v \rightarrow x}^2$ is the gain of a two-step cascade in which the input s regulates the output x via an intermediate v . It is seen that the gain of the two-step cascade is the product of the gain in each cascade step. This gives rise to two corner frequencies in $g_{s \rightarrow v \rightarrow x}^2$, one at $J_{vv} = -\mu_v$ and another at $J_{xx} = -\mu_x$, where μ_v and μ_x are the degradation rates of proteins \mathbf{V} and \mathbf{X} , respectively. The gain is large for $\omega \ll \mu_v, \mu_x$, since in this frequency regime both \mathbf{V} and \mathbf{X} can respond rapidly on the time scale of the signal variations, ω^{-1} , while for frequencies $\omega \gg \mu_v, \mu_x$, the gain decreases strongly, scaling as ω^{-4} , because in this regime the input variations are filtered by the finite lifetime of both \mathbf{V} and \mathbf{X} . Note that a lower degradation rate of the proteins increases the gain at low frequencies, but also reduces the corner frequencies beyond which the gain rapidly drops. This is a generic trade-off between the bandwidth of information transmission

(the frequency range, bounded by the corner frequency, with large Θ) and the magnitude of the gain in the band.

The third term describes the coherence of the interaction between the signal transmitted via the direct pathway $\mathbf{S} \rightarrow \mathbf{X}$ and the signal transmitted via the indirect pathway $\mathbf{S} \rightarrow \mathbf{V} \rightarrow \mathbf{X}$. Both signals originate from the source signal $s(t)$, which means that their variations are correlated. If $\text{sgn}(J_{xs}) = \text{sgn}(J_{vs}J_{xv})$ the third term is positive (since by construction $J_{vv} = -\mu_v$) and we have a coherent interaction. Such a coherent interaction is present in cff networks and leads to an increase in the gain. If $\text{sgn}(J_{xs}) \neq \text{sgn}(J_{vs}J_{xv})$, as in the iff motifs, the coherence term is negative and thus the gain $g^2(\omega)$ is decreased.

The coherence term in the gain is a function of the phase of each pathway, and to be more precise the phase difference, which is the difference between the phase of the signal at \mathbf{X} that comes directly from \mathbf{S} , $\phi_{s \rightarrow x}$ and that which comes indirectly via \mathbf{V} , $\phi_{s \rightarrow v \rightarrow x}$. The signal oscillations are described by its amplitude and phase, and both properties are propagated through the system to the final response \mathbf{X} . Without a phase difference, the two incoming paths have a constructive interference, while for an increase in the phase difference the interference becomes destructive, reducing the effect of the coherent term. Therefore, it is instructive to compare the phase of the direct pathway $\phi_{s \rightarrow x}$ and the indirect pathway ($\phi_{s \rightarrow v \rightarrow x}$) at x

$$\phi_{s \rightarrow x} = \phi_s, \quad (10)$$

$$\phi_{s \rightarrow v \rightarrow x}(\omega) = \phi_s - \arctan \left[\frac{\omega}{\mu_v} \right]. \quad (11)$$

For $\omega = 0$, both signals are in phase. As the frequency increases, the phase of the signal that is transmitted via the indirect pathway decreases with respect to that which is transmitted via the direct pathway. At the corner frequency $-\mu_v = J_{vv}$ the phase difference is $\phi_s - \phi_{s \rightarrow v \rightarrow x}(J_{vv}) = \pi/4$ and for $\omega \rightarrow \infty$, the signals are even more out of phase, $\phi_s - \phi_{s \rightarrow v \rightarrow x} = \pi/2$. Combining both the phase and amplitude information, we see that the coherence between the two signals decreases with increasing ω for two related reasons. First, there is a decrease in the coherence term for large ω due to the time averaging over the fast signal fluctuations resulting from the limited response time of \mathbf{X} and \mathbf{V} . Second, the coherence decreases because the phase difference increases (see the Appendix).

Before we consider the noise, it is useful to briefly consider the three terms of the gain together. While the gain of the direct pathway scales for $\omega \gg \mu_x$ as ω^{-2} , the gain of the indirect pathway and the coherence term scale for $\omega \gg \mu_v, \mu_x$ as ω^{-4} . This means that for frequencies $\omega \gg \mu_v$ the gain of the direct pathway dominates.

The noise in the linearized ff motif is

$$N(\omega) = \frac{\overbrace{J_{xv}^2 \langle |\eta_v|^2 \rangle}^{N_{v \rightarrow x}(\omega)}}{(\omega^2 + J_{vv}^2)(\omega^2 + J_{xx}^2)} + \frac{\overbrace{\langle |\eta_x|^2 \rangle}^{N_x(\omega)}}{(\omega^2 + J_{xx}^2)} \quad (12)$$

$$= g_{v \rightarrow x}^2(\omega) N_v(\omega) + \frac{\langle |\eta_x|^2 \rangle}{(\omega^2 + J_{xx}^2)}. \quad (13)$$

It is seen that the total noise in x is independent of the sign of regulation of either pathway, since all terms are positive. The expression also reveals that the noise is the sum of two noise

sources. One is the intrinsic noise arising from the stochastic production and decay of \mathbf{X} , given by $N_x(\omega) = \langle |\eta_x|^2 \rangle / (\omega^2 + J_{xx}^2)$. The other is the extrinsic noise coming from the stochastic production and decay of \mathbf{V} , which is given by the intrinsic noise of v , $N_v(\omega) = \langle |\eta_v|^2 \rangle / (\omega^2 + J_{vv}^2)$, multiplied by a frequency-dependent gain $g_{v \rightarrow x}^2 = J_{xv}^2 / (\omega^2 + J_{xx}^2)$ which reflects how the noise from v is amplified by J_{xv} and integrated by x as a result of its finite lifetime. While the intrinsic noise in x , $N_x(\omega)$, scales as ω^{-2} for $\omega \gg \mu_x = -J_{xx}$, the extrinsic noise $g_{v \rightarrow x}^2(\omega) N_v(\omega)$ scales as ω^{-4} for $\omega \gg \mu_v, \mu_x$. Indeed, for $\omega \gg \mu_v$, the noise that originates from v in the indirect pathway becomes negligible. In Eq. (12) we present a general expression for the noise without the effect of cross-correlations ($\langle \eta_v \eta_x \rangle = 0$). In the Appendix we comment on the effect of cross-correlations.

Finally, we obtain Θ for this three component motif:

$$\begin{aligned} \Theta(\omega) &= \frac{g^2(\omega)}{N(\omega)} \\ &= \frac{(\omega^2 + J_{vv}^2) J_{xs}^2 + (J_{vs} J_{xv})^2 - 2 J_{xs} J_{vs} J_{xv} J_{vv}}{J_{xv}^2 \langle |\eta_v|^2 \rangle + (\omega^2 + J_{vv}^2) \langle |\eta_x|^2 \rangle}. \end{aligned} \quad (14)$$

While the noise is independent of the sign of regulation in the pathways, the gain, and hence the gain-to-noise ratio, depends on the total sign of regulation of each pathway (either $\mathbf{S} \rightarrow \mathbf{X}$ or $\mathbf{S} \rightarrow \mathbf{V} \rightarrow \mathbf{X}$). This indicates that it is not important which of the reactions in a specific pathway acts negatively; only the overall effect of the pathway is important for information transmission.

We are now in a position to study in more detail the ff motif. As discussed above, for $\omega \gg \mu_v = -J_{vv}$, the gain of the direct pathway dominates the total gain, because the finite lifetime of \mathbf{V} averages out the variations in the signal that are transmitted via the indirect pathway. Also the noise that originates at v in the indirect pathways becomes negligible in the total noise in x . Therefore, in this frequency regime, the direct pathway is dominant and Θ becomes that of a one-step cascade, which means that it approaches a constant value, independent of frequency. For smaller frequencies, the behavior of the ff motif depends on the relation between J_{xs} , J_{vs} , J_{xv} , and $J_{vv} = -\mu_v$, which are the coupling constants of the direct and indirect pathway and the degradation rate of \mathbf{V} , respectively [Eq. (4)]. When $|J_{xs}| \ll |J_{vs} J_{xv} / J_{vv}|$, then for $\omega \ll \mu_v$, the signal is transmitted more strongly via the indirect pathway than via the direct pathway, in which case the ff motif resembles a two-step cascade [Figs. 2(c), 3(c), and 4(c), thick red dotted line]; clearly, in the limit that J_{xs} reaches zero, the ff motif becomes a two-step cascade for all frequencies, in which case Θ [Eq. (A10)] is constant up to a cut-off frequency $\omega_c^2 = \mu_v(\mu_v + J_{xv})$, falling off as ω^{-2} for frequencies much larger than that [5]. When on the other hand $|J_{xs}| \gg |J_{vs} J_{xv} / J_{vv}|$, then the direct pathway dominates the gain for all frequencies [Figs. 2(c), 3(c), and 4(c), thick red solid line]. Indeed, in terms of the gain, the ff network effectively becomes a one-step cascade. However, the noise via the indirect pathway still contributes to the total noise and therefore this pathway effectively acts as a noise source. This has interesting consequences for the gain-to-noise ratio, as we describe in the next paragraph, since this allows any ff

motif to function as either a high-pass or a low-pass filter for information.

The gain-to-noise ratio [Eq. (14)] of a ff motif varies monotonically with frequency, but it can either be a decreasing or increasing function of ω . Indeed, both a coherent and an iff motif can act as either a high-pass or a low-pass filter for information. We can determine whether Θ increases or decreases monotonically with frequency, by comparing Θ at $\omega \rightarrow \infty$ to that at $\omega = 0$ [see Eq. (14)]. The Θ is monotonically increasing, meaning that the network acts as a high-pass filter for information, if

$$\langle |\eta_v|^2 \rangle > \langle |\eta_x|^2 \rangle \left(\frac{J_{vs}^2}{J_{xs}^2} + \frac{2\mu_v J_{vs}}{J_{xs} J_{xv}} \right). \quad (15)$$

In the Appendix we give a more detailed explanation for each of the terms in Eq. (15). The frequency of the inflection point is

$$\omega_{ip} = \sqrt{\frac{J_{xv}^2 \langle |\eta_v|^2 \rangle + J_{vv}^2 \langle |\eta_x|^2 \rangle}{3 \langle |\eta_x|^2 \rangle}}. \quad (16)$$

In the next two sections, we discuss these conditions in more detail for the coherent and iff motifs, respectively.

The data-processing inequality states that information that is lost cannot be recovered. Consequently, increasing the length of a cascade reduces information transmission. For this reason, for equal total production rate of the components within a cascade, Θ of a one-step cascade is always larger than that of a ff motif. The gain by itself can be larger in a ff motif than in a one-step cascade [Fig. 3(c), dashed and solid red

lines]. However, the intermediate component V in the indirect pathway introduces an additional noise source, which is not present in the one-step cascade. The coherence term in the gain is not large enough to compensate for this increase in the noise.

We have now specified some general characteristics. In the following we study the coherent and incoherent motif separately.

2. The coherent feed-forward motif

We will first compare the cff motif, both the OR and the AND types, with two-step cascades. Next we compare the two motifs with each other. First, however, we start with a number of observations which apply to both the acff and the ocff. Unless specified otherwise, we assume equal degradation rates for the respective components in the respective motifs and cascades.

The steady-state gain is completely determined by the average copy number X ($=\partial \langle x \rangle / \partial \langle s \rangle$). As a result, if the production of x in a cff equals the production of x in a two-step simple cascade, the gain at zero frequency ($\omega = 0$) is equal. Next, in the linear-noise approximation assumed here, the intrinsic noise arising from the production and degradation events of X , $N_x(\omega)$, is equal for the cff and the two-step cascade if the production rates of x are equal [see Eq. (12)], since we assume throughout that the degradation rates are the same. If the production of v is equal as well, the noise $N_v(\omega)$ is equal in the cff and the two-step cascade. However, the transmitted noise $N_{v \rightarrow x}(\omega)$, and thus the total noise $N(\omega)$ can be different, as discussed in more detail below.

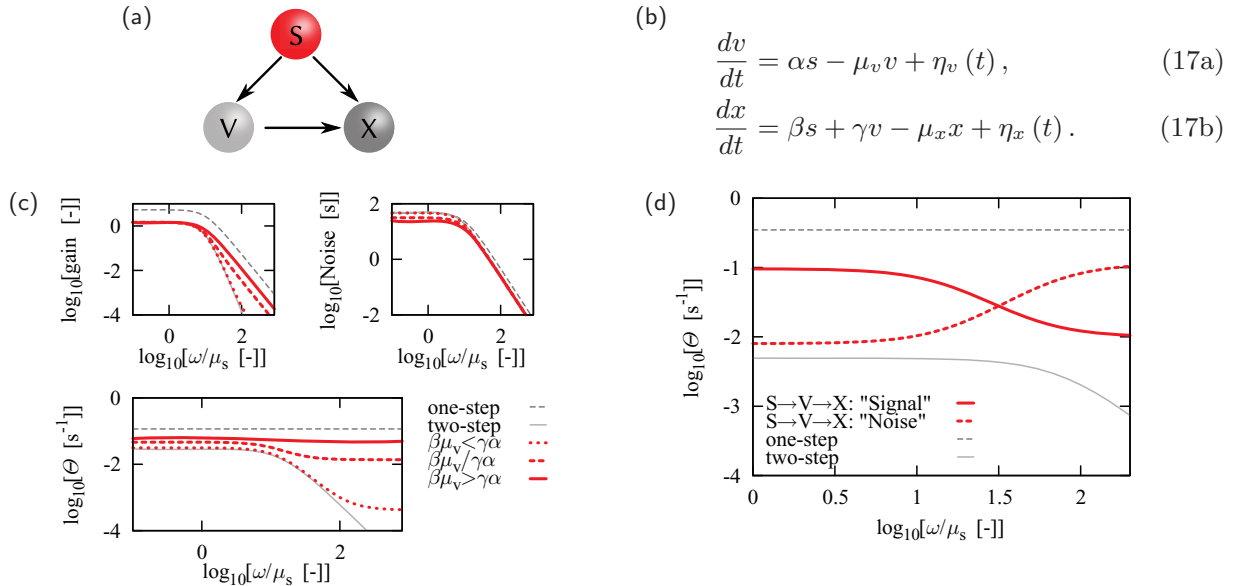


FIG. 2. (Color online) The ocff motif. (a) In the ocff, either S or V is required to produce X . (b) The Langevin equations for the ocff. (c) For equal production in the total cascade Θ is shown for different weightings of the two pathways. If the pathway $S \rightarrow X$ dominates ($\beta\mu_v > \gamma\alpha$) (solid red line), the ocff is similar to a one-step cascade (thin gray dashed line) although with smaller gain. If the pathway $S \rightarrow V \rightarrow X$ dominates (thick red dotted solid line), two-step cascade (thin gray solid) behavior is obtained. However, for large ω the ocff is similar to the one-step cascade (scaling as ω^0), because the signal and noise fluctuations through v are averaged out due to the finite response time of V . Parameters: $k_s = 10$, $\mu_v = 10$, $\mu_x = 10$, $k_x^1 = 23.1$, $k_v^2 = 11$, $k_x^2 = 11$, $\alpha = 11$, $\gamma = 10.1, 5.8, 1.1$, and $\beta = 1, 5.8, 10.1$ for, respectively, thick red solid, thick red dashed, thick red dotted solid lines. β sets the production of x in the cascade, and μ_s sets the time scale. (d) The gain-to-noise ratio for a cff motif can have both high-pass and low-pass characteristics. For high-pass characteristics, the indirect pathway functions as a “noise” source at low frequencies, while at high frequencies the noise is filtered out due to the finite response time of V . Parameters: $k_s = 100$, $\mu_v = 5$, $\mu_x = 10$, $k_x^1 = 70$, $k_v^2 = 1$, $k_x^2 = 690$, $\alpha = 20, 0.3$, $\gamma = 20, 1000$, $\beta = 10, 40$, respectively solid, dashed.

With respect to Θ , we observe that for large ω , the cff always has a larger Θ than the two-step cascade, while for small ω , no general results can be presented. Comparing to the one-step cascade, Θ of the cff is always smaller than that of a one-step cascade if the total production in both networks is equal. This is because η_v acts as an additional noise source that corrupts the signal [Figs. 2(c) and 3(c), bottom].

We discuss the ocff where the two pathways combine according to OR logic. We first compare the ocff with a two-step cascade on the footing of equal production costs for each of the components separately. For $\omega > 0$, the gain of the ocff is always larger than the gain of the two-step cascade (Table II): For $0 < \omega < \mu_v, \mu_x$ the gain of the cff is boosted by the coherent interaction between the signals propagated via the direct and indirect pathways, while for $\omega > \mu_v$ the signal is attenuated in the two-step cascade by the finite lifetime of V, whereas it can still be propagated in the ocff via the direct pathway. The noise $N(\omega)$ in the ocff motif is smaller than that in the two-step cascade for all ω . The intrinsic noise $N_x(\omega)$ is equal in both networks. However, the extrinsic noise coming from v , $g_{s \rightarrow v}^2(\omega)N_v(\omega)$ [see Eq. (12)] is not. Since the intrinsic noise in v , $N_v(\omega)$, is the same in both networks because the production of v is equal, the difference must lie in how this noise is propagated to x , which is given by $g_{s \rightarrow v}^2 = J_{xv}^2/(\omega^2 + J_{xx}^2)$; since J_{xx} is the same, this means that the coupling between V and X in the ocff motif, J_{xv}^{cff} , must be less than that in the two-step cascade, J_{xv}^{ts} . Indeed, in the ocff the production of x depends on both s and v , while in the two-step cascade it only depends on v . Together with the constraint that the production rates of X are equal, this indeed implies that $J_{xv}^{\text{cff}} < J_{xv}^{\text{ts}}$: $J_{xv}^{\text{ts}}\langle v \rangle = J_{sv}^{\text{cff}}\langle s \rangle + J_{xv}^{\text{cff}}\langle v \rangle$, showing that $J_{xv}^{\text{cff}} < J_{xv}^{\text{ts}}$. In the ocff network, the noise in x is thus smaller because less noise is propagated from v because of the smaller amplification of the transmitted signal between v and x . The higher gain and the lower noise means that the gain-to-noise ratio of the ocff motif is higher than that of a two-step cascade, as shown in Fig. 2(c). The ocff is thus able to signal more reliably than the corresponding two-step cascade.

As discussed above [Eq. (15)] the cff can act as either a low-pass or as a high-pass filter. It can be shown for the ocff motif that for equal production of the species separately in the network, the plateau value of Θ in the high-frequency regime is higher for the high-pass filter than for the low-pass filter. Therefore, if the goal is to transmit signals reliably at high frequencies, then a suitable parameter set can be chosen that yields a high-pass filter with a high Θ at high frequencies, irrespective of Θ behavior at low frequencies [Fig. 2(d), thick red dashed line].

For different constraints on the production the results are summarized in the Appendix in Tables I and II.

In the next paragraphs we study the acff motif. The acff motif combines the two paths at X according to AND logic and is described by Eq. (20). The linearized system has an identical structure to the ocff, and therefore the results are qualitatively similar to those of the ocff. However, quantitatively, the results can be different.

First we compare the acff with the corresponding two-step cascade. For equal production of v and x individually the gain of the acff [Eq. (A54)] is always larger than the gain

of the two-step cascade [Table III]. This again is due to the coherent interaction between the direct and indirect pathways at low frequencies, $\omega \ll \mu_v, \mu_x$, while for high frequencies, $\omega \gg \omega_v, \omega_x$, the gain of a two-step cascade falls off more rapidly with frequency than the gain of the direct pathway. Interestingly, the mathematical dependence of the total noise $N(\omega)$ on the network parameters (J_{xv}, μ_v, \dots) of the acff is the same as that of the two-step cascade. Since $N_v(\omega)$ and $N_x(\omega)$ are equal in both networks, this implies that the noise propagated from v to x ($N_{v \rightarrow x}(\omega)$) is equal:

$$N_{v \rightarrow x, \text{AND}} = \frac{\beta^2 \langle s \rangle^2 \langle |\eta_v|^2 \rangle}{(\omega^2 + \mu_x^2)(\omega^2 + \mu_v^2)}, \quad (18)$$

$$N_{v \rightarrow x, \text{ts}} = \frac{k_x^2 \langle |\eta_v|^2 \rangle}{(\omega^2 + \mu_x^2)(\omega^2 + \mu_v^2)}. \quad (19)$$

The total rate of production of x by v in the two-step cascade is $k_x \langle v \rangle$ [Eq. (A7)], while, in the linear-noise approximation used here, the average total rate of production of x by s and v in the acff motif is $\beta \langle s \rangle \langle v \rangle$ [Eq. (A53)]; the latter means that the coupling of x to v is given by $J_{xv} = \beta \langle s \rangle$. For equal rate of production of x in the two-step cascade and the acff— $k_x \langle v \rangle = \beta \langle s \rangle \langle v \rangle$ —the coupling strengths $J_{xv} = k_x = \beta \langle s \rangle$ in the two networks are the same, meaning that the extrinsic noise coming from v , $N_{v \rightarrow x}$ is indeed the same. Since the gain is larger, but the noise is the same, Θ of the acff is always larger than that of a two-step cascade.

The acff cannot transmit information better at small frequencies than a one-step cascade, given equal total production. An interesting observation is, however, that for equal production of x —thus, at the cost of producing more V—the acff can perform better than the one-step cascade [Eq. (A57)]. This is in contrast to the ocff, which also under this constraint has a smaller Θ .

Again, we can choose different constraints for the production. We have summarized these results in the Appendix in Tables III and IV.

Finally, we compare the two motifs ocff and acff for equal separate production of V and X. The acff has a larger gain, because it has a steeper response function at equal production. Not only the gain, but also the intrinsic noise ($N_{v \rightarrow x}$) is larger in the acff. These terms have an opposite effect on Θ . For $\omega \rightarrow \infty$ the acff has a larger Θ , since the noise fluctuations from v are averaged out and only the gain remains important. For small ω , however, no general conclusion for the Θ can be drawn, but a few observations can be made. An increase in the coupling J_{xs} of $S \rightarrow X$ (increasing β) leads to a larger Θ for the ocff than for the acff. In this case, the direct coupling between S and X dominates in the ocff, whereas this connection is not directly present in the acff due to the AND-logic integration strategy. The ocff “averages” the effects of the direct and indirect pathways, while for the acff gate the pathways are more connected [Fig. 3(d)]. Indeed, one necessary requirement for the ocff to have a larger Θ than the acff at low frequencies is that the direct pathway dominates the output. Another necessary requirement for this is that Θ of the acff has to have high-pass characteristics [Eqs. (A73) and (A74)].

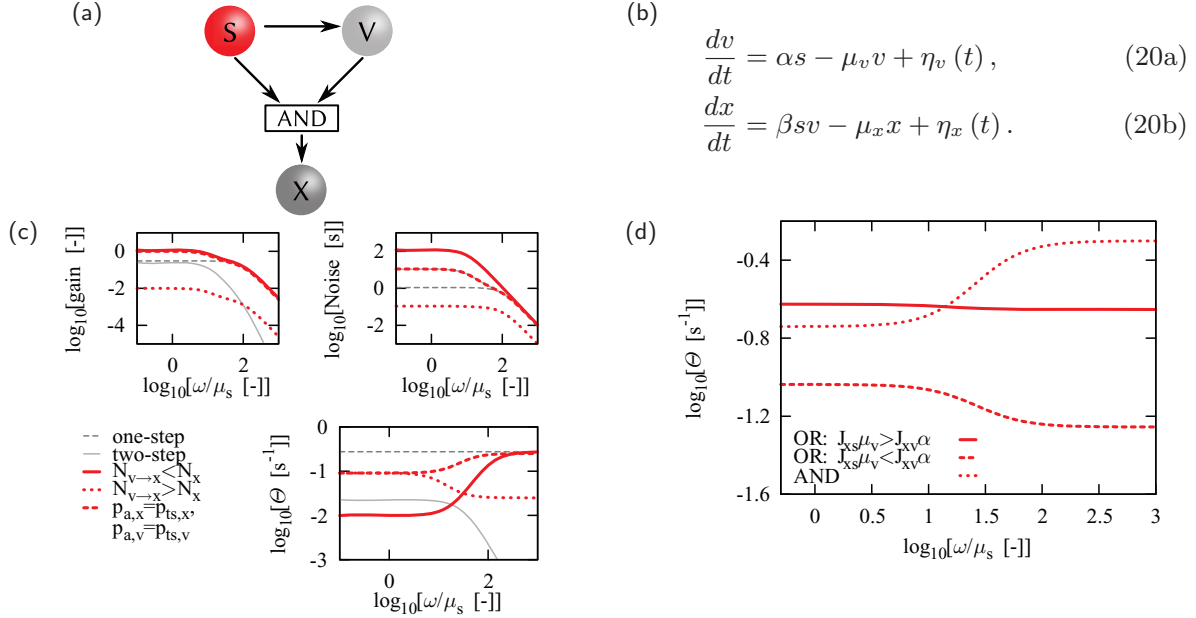


FIG. 3. (Color online) The acff loop. (a) In the acff, both S and V are required to produce X . (b) The Langevin equations for the acff. (c) Θ of the acff for different weighings (changing α) of the two pathways for equal total production and equal degradation rates. First, for large ω the direct pathway is dominant and both gain and noise [Eqs. (A54) and (A55)] scale with ω^{-2} , leading to a constant Θ , which is larger than that of a two-step cascade. If α is small (thick red solid), the production of v , p_v is small. To compensate for the small production of v , the production of x should be large, leading to a large β . Note that $\alpha\beta \neq 1$. The gain $g_{v \rightarrow x}^2$ scales with $\beta^2\alpha^2$, while the noise $N_{v \rightarrow x} \propto \beta^2\alpha$, and $N_x \propto \beta\alpha$. For small ω , therefore, Θ scales as α , while for $\omega \rightarrow \infty$, $N_{v \rightarrow x}$ is averaged out and Θ scales as $\alpha\beta^0$. The dependence of the gain and the noise on μ_v is slightly different. Therefore, a small bandwidth exists for which the gain decreases more slowly than the noise for increasing ω (thick red solid). For large α the opposite reasoning holds. The thick red dashed line shows the acff with equal production of v and x individually, as the two-step cascade. Both the gain, noise, and Θ are larger. It is interesting to note that for small frequencies and equal total production, the acff can have a smaller gain-to-noise ratio than the simple two-step cascade (thick red solid and thin gray solid). This is the case if $N_{v \rightarrow x}(\omega)$ becomes dominant. Parameters: $k_s = 100$, $\mu_v = 10$, $\mu_x = 100$, $k_x^1 = 55$, $k_x^2 = 100$, $k_v^2 = 5$, $\alpha = 0.5, 5, 50$, respectively thick red solid line, thick red dashed line, thick red dotted solid line. β such that the total production is equal, respectively $\beta = 10.9, 1, 0.01$. μ_s sets the time scale. (d) The acff (red dotted line) for large ω always has a larger Θ than the ocff (thick red solid line and thick red dashed line). However, for small ω , depending on kinetic rates, Θ of the ocff can be larger than the acff (thick red solid line). For this to occur, in the ocff the direct pathway $S \rightarrow X$ should dominate the indirect $S \rightarrow V \rightarrow X$ pathway. Then, for small ω the ocff acts as a one-step cascade. For the acff this is not the case, because the coupling between s and v is more complex. Parameters: $k_s = 100$, $\mu_v = 10$, $\mu_x = 100$, $J_{vs}^{\text{AND}} = J_{vs}^{\text{OR}} = 10$, $\beta = 1$, $J_{xv}^{\text{OR}} = 16, 67$ and $J_{xs}^{\text{OR}} = 83, 33$ respectively, solid and dashed lines. μ_s sets the time scale.

3. The incoherent feed-forward motif

The last class of ff loops that we consider is the iff motif [Fig. 4(a)]. We discuss one type of iff motif, one in which negative regulation is implemented in the direct pathway. It is known that the precise topology can influence the noise behavior [26,27]; however, the qualitative behavior of Θ of the different possible architectures is similar to that of the networks studied here. For a more detailed discussion, we refer to the Appendix.

An inspection of the gain in Eq. (9) reveals that the coherence term is negative in both types of iff motifs, since either J_{xs} or the product $J_{xv}J_{vs}$ is negative, respectively. This leads to a reduction of the gain on time scales that are smaller than the response times of V and X ($\omega < \min\{\mu_v, \mu_x\}$).

For $\omega < \mu_v$, the two pathways are exactly out of phase, that is, the phase difference is $-\pi$, while as ω increases, the phase difference reduces to $-\pi/2$ (10).

We model the incoherent regulation using a repressive Hill function [Eq. (23b)], which is a commonly used coarse-grained description of protein interactions, gene, or enzyme regulation.

If the direct pathway is repressive, the repression strength depends on the ratio $\langle s \rangle / K$, with K being the value of s at which X is reduced to half its maximal value. In the limit that $K \gg \langle s \rangle$, the repression is very weak, and the influence of the direct pathway is negligible. The iff then effectively reduces to a two-step cascade. In the opposite limit, repression is very strong and the iff becomes adaptive [1,28]. This means that $\langle x \rangle$ does not depend on $\langle s \rangle$. Indeed,

$$\langle x \rangle = \frac{v}{\mu_x} \frac{K(v)}{K + \langle s \rangle} = \frac{v\alpha}{\mu_x \mu_v} \frac{K \langle s \rangle}{K + \langle s \rangle} \approx \frac{K v \alpha}{\mu_x \mu_v}. \quad (21)$$

In terms of the frequency response, adaptation to constant signals corresponds to a zero gain at zero frequency:

$$\begin{aligned} \lim_{K \ll \langle s \rangle} g^2(\omega = 0) &= \frac{(J_{vs}J_{xv} - J_{vv}J_{xs})^2}{\mu_v^2 \mu_x^2} \\ &= \frac{K^2}{(K + \langle s \rangle)^2} \frac{(\alpha v)^2}{\mu_v^2 \mu_x^2} \left[1 - \frac{\langle s \rangle}{K + \langle s \rangle} \right]^2 \approx 0. \end{aligned} \quad (22)$$

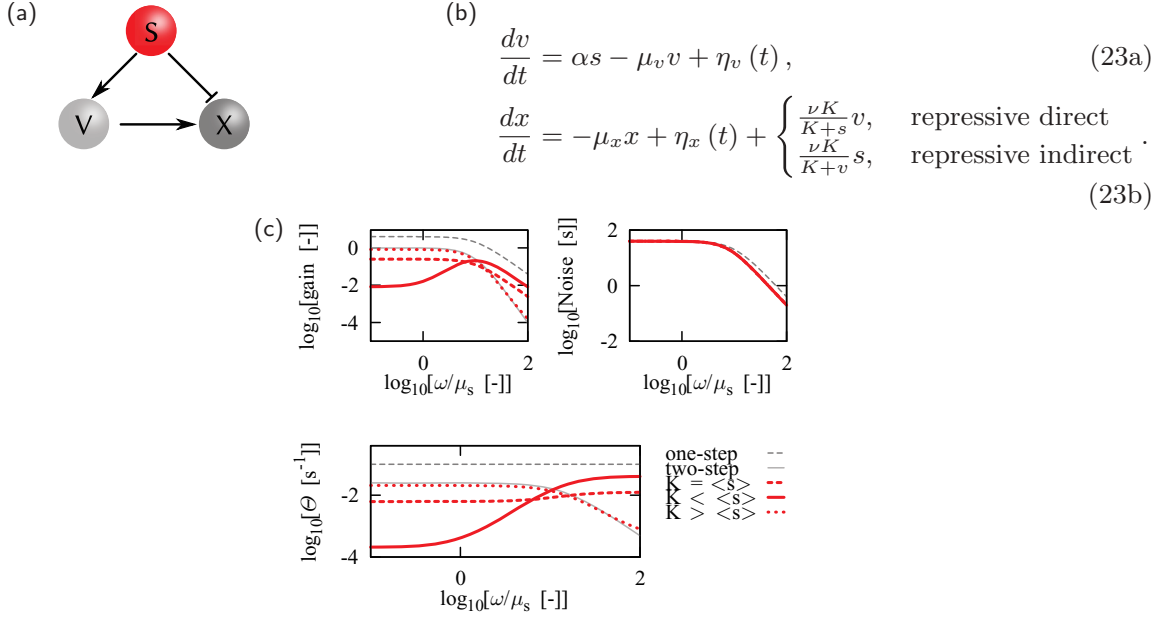


FIG. 4. (Color online) The iff motif. (a) The incoherent motif. (Left) The pathway $S \rightarrow V \rightarrow X$ is positive, while the pathway $S \rightarrow X$ is negative. (b) The Langevin description of the network. (c) The coherence terms acts destructively on small time scales, reducing the gain, but not the noise, so that Θ is reduced. The gain-to-noise ratio of the incoherent motif with a repressive direct link $S \rightarrow X$ has different characteristics for three different values of K (23). If $K > \langle s \rangle$ (thick red dotted line) the negative regulation by S diminishes and the iff motif resembles a two-step cascade. For large ω , the indirect pathway is averaged out; at these frequencies one-step characteristics are observed. If $K < \langle s \rangle$ (red solid line), the iff becomes purely adaptive, leading to $g^2 = 0$ for $\omega = 0$. Note that the lines in the noise plot are overlapping. Thin gray dashed line, one-step cascade; thin gray solid line, two-step cascade. Parameters: $k_s = 100$, $\alpha = 1$, $\mu_v = 1$, $\mu_x = 1$ and $\nu = 110$, 20, 11 for $K/\langle s \rangle = 0.1, 1, 10$. μ_s sets the time scale.

For large frequencies, only the direct pathway, which in this example is repressive, transmits information. For information transmission it is not important whether the pathway acts negatively or positively on X . The variations in s still affect the variations in x , but with an opposing sign. For the motif where the repression occurs in the indirect pathway, similar conclusions hold, but now the motif functions as a one-step cascade in the case that repression is weak.

An iff motif with a strong negative interaction in one of the two pathways acts as a high-pass filter for information. This is because of the destructive interference of the two pathways at small frequencies, $\omega \ll \mu_v, \mu_x$. For higher frequencies, the gain increases because the phase difference between the two pathways decreases and also because the indirect pathway becomes less important as the finite lifetime of V increasingly averages out the variations in S . The gain therefore has high-pass characteristics. Since the noise [Eq. (12)] is not affected by the destructive interference and has low-pass characteristics, Θ is high-pass [Fig. 4(c), red line].

For the incoherent motif, a low-pass Θ is observed only if the negative regulation is small (e.g., $K \gg \langle s \rangle$) and the direct pathway is negligible compared to the indirect pathway. Then the motif for small ω resembles a two-step cascade [Fig. 4(c), thick red dotted line], which indeed exhibits a low-pass Θ . For large ω the direct pathway will dominate, which means that in contrast to a two-step cascade, Θ reaches a constant as a function of frequency for $\omega \gg \mu_v, \mu_x$.

Finally, we compare, for completeness, the iff motif to a two-step cascade. For equal production of v and x separately,

the iff has a lower Θ for small ω than the two-step cascade (Table VI). This is because of the destructive interference between the direct and indirect pathway in the iff, which reduces the gain at low frequencies.

C. Multimerization

In this section we examine multimerization of intermediate signaling components [29]. In this motif, an intermediate component v is activated by the input signal s , which then cooperatively activates the output component x [Fig. 5(a)]. The intermediate components could form a protein complex that then binds and activates the output, but it need not be: The intermediate components could also bind the output component, which could be a gene promoter or an enzyme, separately but cooperatively, thereby activating it. This is a common motif in gene regulation and also enzyme regulation.

This system is described by

$$\frac{dv}{dt} = \alpha_v s(t) - \mu_v v(t) + \eta_v(t), \quad (24)$$

$$\frac{dx}{dt} = \gamma_v \frac{v^n(t)}{K^n + v^n(t)} - \mu_x x(t) + \eta_x(t) \quad (25)$$

$$\approx \gamma_n v^n(t) - \mu_x x(t) + \eta_x(t). \quad (26)$$

Here, n is a measure for the cooperativity, the number of V molecules that are required to activate X . We assume that the concentration v is very low, $K \gg v$, in which case Eq. (25) reduces to Eq. (26), with $\gamma_n \simeq \gamma_v/K^n$. For this network, the concentration of X and the coupling between v and x , J_{xv} ,

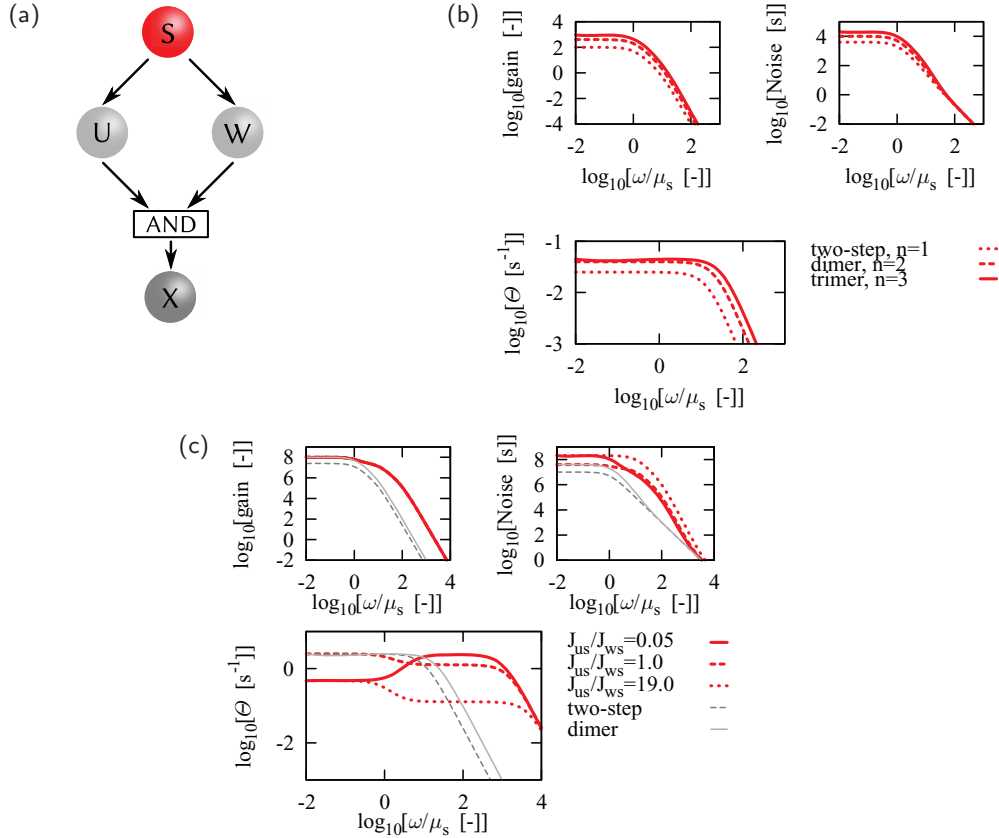


FIG. 5. (Color online) The diamond motif. (a) The diamond motif. The diamond motif combines two pathways ($S \rightarrow V \rightarrow X$ and $S \rightarrow W \rightarrow X$), originating from the same source at the response X . With homodimerization U and W correspond to two molecules of the species V . (b) The difference between the gain, noise, and Θ for a network with different levels of cooperativity (24). For increasing cooperativity (n increases), the gain and the noise both increase. The increase in the noise is smaller than in the gain, since N_x is unaffected by the cooperative interactions and therefore Θ increases. Next, also the knee frequency, and thus the bandwidth, increases, since Θ scales with ω^{-2} for $\omega^{-2} > n^2 \mu_v$. Parameters: $k_s = 100, k_v = 10, k_x = 10, \mu_v = 100, \mu_x = 1, \mu_s$ sets the time scale. (c) Θ for the diamond motif where U and W have different degradation rates $\mu_u < \mu_w$. If $J_{us} > J_{ws}$ Θ is low-pass (thick red dotted), since on long time scales ($t > \mu_u^{-1}$), the signal is transmitted while on shorter time scales most transmission is corrupted by the intrinsic noise. If $J_{us} < J_{ws}$ (thick red solid), the slow signal variations transmitted via w are corrupted by noise from u . For faster variations, the noise from u is averaged out and the signal can be transmitted with larger reliability. Note that the gain for all three parameter sets is equal, and the lines thus overlap. The bandpass characteristic for Θ is thus due to the different dependence of the noise on ω . Parameters: $k_s = 100, \mu_x = 10$, two-step cascade: $\mu_v = 10, k_v^{ts} = 500, k_x^{ts} = 100, \alpha$ to equalize production of v, γ_v to equalize production of x, β to equalize $\langle v^{dm} \rangle = \langle u^{dm} \rangle + \langle w^{dm} \rangle$ and $\alpha = 0.05k_v^{ts}, 0.5k_v^{ts}, 0.95k_v^{ts}$, respectively, thick red solid line, thick red dashed line, thick red dotted solid line. μ_s sets the time scale.

depends on the degree of cooperativity n . The concentration X is given by $\langle x \rangle = \gamma_n \langle v^n \rangle / \mu_x$. In the linear-noise approximation this concentration is given by $\langle x \rangle \simeq \gamma_n \langle v \rangle^n / \mu_x$, and the coupling $J_{xv,n} = n \gamma_n \langle v \rangle^{n-1}$, where the subscript n in $J_{xv,n}$ indicates that we consider the coupling between v and x when the degree of cooperativity is n . Since we compare the networks on the footing of equal productions costs and the degradation rates are kept constant, we find that the coupling constant $J_{xv,n}$ for a system in which n V molecules are required to activate X is related to the coupling constant $J_{xv,n=1}$ for a system in which only one V molecule is required to activate X , simply via $J_{xv,n} = n J_{xv,n=1}$.

For this motif, the gain is given by

$$g^2(\omega) = \frac{n^2 J_{xv,n=1}^2 J_{vs}^2}{(\omega^2 + \mu_x^2)(\omega^2 + \mu_v^2)}. \quad (27)$$

It is seen that the gain increases with the cooperativity n .

The noise for this motif is given by

$$N(\omega) = \frac{\overbrace{n^2 J_{xv,n=1}^2 \langle |\eta_v|^2 \rangle}^{N_{v \rightarrow x}(\omega)}}{(\omega^2 + \mu_v^2)(\omega^2 + \mu_x^2)} + \frac{\overbrace{\langle |\eta_x|^2 \rangle}^{N_x(\omega)}}{(\omega^2 + \mu_x^2)} \quad (28)$$

$$= g_{v \rightarrow x}^2(\omega) N_v(\omega) + \frac{\langle |\eta_x|^2 \rangle}{(\omega^2 + \mu_x^2)}. \quad (29)$$

Clearly, increasing n increases the extrinsic noise in x that originates in v , $N_{v \rightarrow x}(\omega) = g_{v \rightarrow x}^2(\omega) N_v(\omega)$ but not the intrinsic noise in x , $N_x(\omega)$. Increasing the coupling $J_{xv,n} = n J_{xv,n=1}$ does not change the intrinsic noise, but it does affect how fluctuations in v are amplified at the level of x .

Θ then reads

$$\Theta(\omega) = \frac{J_{vs}^2 J_{xv,n=1}^2}{J_{xv,n=1}^2 \langle |\eta_v|^2 \rangle + \frac{1}{n^2} (\omega^2 + \mu_v^2) \langle |\eta_x|^2 \rangle}. \quad (30)$$

From Eq. (30) it is clear that Θ increases with the cooperativity n . This is because while the overall gain $g_{s \rightarrow x}^2$ and the extrinsic noise $N_{v \rightarrow x}(\omega)$ both increase with n , the intrinsic noise $N_x(\omega)$ does not. Interestingly, not only the amplitude of Θ increases, but also the knee frequency and thus the bandwidth for reliable information transmission. Indeed, the knee frequency ω_c is set by $\omega_c^2 = \mu_v^2 + n^2 J_{xv,n=1}^2 \langle |\eta_v|^2 \rangle / \langle |\eta_x|^2 \rangle = \mu_v(\mu_v + n^2 \gamma_{n=1})$, showing that it increases with n [Fig. 5(b)].

To summarize, we observe that information transmission can be increased by cooperatively activating the output. This could either be achieved via homomultimerization of the intermediate component, or by separate binding of the intermediate molecules to the output component. We note that an increase in n increases the nonlinearity of our system and therefore the approximation might break down. However, for the parameters used here [Fig. 5(b)], numerical simulations of the nonlinear system agree very well with the linear theory (see the Appendix).

D. Diamond motif

The multimerization motif discussed above could be considered to be a special case of a diamond motif, in which the intermediate components are identical. Here, we consider the general scheme in which they are different: the diamond motif [30–34]. We compare the Θ of this motif to that discussed above. Moreover, we compare the performance of this motif to that of a two-step cascade; if Θ of the diamond motif is higher than that of a two-step cascade, then this indicates that from

the perspective of information transmission it is beneficial to split the signal after the input and recombine the signals downstream, at the output.

We consider a diamond motif in which the two intermediate components U and W can either form a homodimer U_2 or W_2 , respectively, or a heterodimer UW ; note that in this network the intermediate components effectively activate the output via AND logic. We compare the performance of this motif to that of a homodimer motif in which there is only one intermediate component V , which forms a homodimer V_2 ; this corresponds to the scenario discussed before with $n = 2$. The diamond motif is described by

$$\begin{aligned} \frac{du}{dt} &= \alpha s(t) - \mu_u u(t) + \eta_u(t), \\ \frac{dw}{dt} &= \beta s(t) - \mu_w w(t) + \eta_w(t), \\ \frac{dx}{dt} &= \gamma_u u^2(t) + \gamma_w w^2(t) + 2\gamma_{uw} u(t)w(t) \\ &\quad - \mu_x x(t) + \eta_x(t). \end{aligned} \quad (31)$$

The factor 2 is introduced so that the diamond motif reduces to the motif with only one intermediate component V when the properties of the components U and W are identical and equal to those of V . If $\gamma_u = \gamma_w = 0$, only heterodimer activation is possible. In general, the values of the γ 's might be different depending on the respective binding kinetics of the components U and W .

The gain of the diamond motif of Eq. (31) is

$$g^2(\omega) = \frac{\overbrace{(J_{xu} J_{us})^2}^{g_{s \rightarrow u \rightarrow x}^2}}{(\omega^2 + \mu_x^2)(\omega^2 + \mu_u^2)} + \frac{\overbrace{(J_{xw} J_{ws})^2}^{g_{s \rightarrow w \rightarrow x}^2}}{(\omega^2 + \mu_x^2)(\omega^2 + \mu_w^2)} + \frac{\overbrace{2J_{xu} J_{us} J_{xw} J_{ws} (\omega^2 + \mu_u \mu_w)}^{\text{coherence}}}{(\omega^2 + \mu_x^2)(\omega^2 + \mu_u^2)(\omega^2 + \mu_w^2)}. \quad (32)$$

Here, $J_{us} = \alpha$, $J_{ws} = \beta$, $J_{xu} = 2\gamma_u \langle u \rangle + 2\gamma_{uw} \langle w \rangle$, $J_{xw} = 2\gamma_w \langle w \rangle + 2\gamma_{uw} \langle u \rangle$.

The first two terms in Eq. (32) describe the gain due to the transmission of the input signal via the pathways containing U and W , respectively, while the third term describes the coherence of their interaction at the output X . It can be verified that when U and W are identical and equal to V , meaning that $\alpha = \beta$, $\alpha_v = 2\alpha$, $J_{us} = J_{ws} = J_{vs}/2$, $\langle v \rangle = \langle w \rangle + \langle u \rangle$, $\gamma_{uw} = \gamma_u = \gamma_w = \gamma_v = \gamma$, $J_{xu} = J_{xw} = J_{xv,n=2} = 2\gamma \langle v \rangle$, and $\mu_u = \mu_w = \mu_v$, the gain of the diamond motif equals that of the homodimer motif.

The noise of the diamond motif of Eq. (31) is

$$\begin{aligned} N(\omega) &= \frac{\overbrace{J_{xu}^2 \langle |\eta_u|^2 \rangle}^{N_{u \rightarrow x}(\omega)}}{(\omega^2 + \mu_u^2)(\omega^2 + \mu_x^2)} + \frac{\overbrace{J_{xw}^2 \langle |\eta_w|^2 \rangle}^{N_{w \rightarrow x}(\omega)}}{(\omega^2 + \mu_w^2)(\omega^2 + \mu_x^2)} + \frac{\overbrace{\langle |\eta_x|^2 \rangle}^{N_x(\omega)}}{(\omega^2 + \mu_x^2)} \\ &= g_{u \rightarrow x}^2(\omega) N_u(\omega) + g_{w \rightarrow x}^2(\omega) N_w(\omega) + N_x(\omega), \end{aligned} \quad (33)$$

where we have exploited that $\langle \eta_u(\omega) \eta_w(-\omega) \rangle = 0$ when U and W are different. If the properties of U and W are identical and equal to V , then $J_{xu} = J_{xw} = J_{xv,n=2}$, $\mu_u = \mu_w = \mu_v$, $\langle |\eta_u|^2 \rangle = \langle |\eta_w|^2 \rangle = \langle |\eta_v|^2 \rangle / 2$ and the noise of the diamond motif is indeed equal to that of the homodimer motif.

The gain-to-noise ratio for the diamond motif is

$$\Theta(\omega) = \frac{(J_{xu} J_{us})^2 H(\omega, \mu_w) + (J_{xw} J_{ws})^2 H(\omega, \mu_u) + 2J_{xu} J_{us} J_{xw} J_{ws} (\omega^2 + \mu_u \mu_w)}{H(\omega, \mu_w) J_{xu}^2 \langle |\eta_u|^2 \rangle + H(\omega, \mu_u) J_{xw}^2 \langle |\eta_w|^2 \rangle + H(\omega, \mu_u) H(\omega, \mu_w) \langle |\eta_x|^2 \rangle}, \quad (34)$$

where $H(x, y) = x^2 + y^2$.

We now compare Θ of the diamond motif to that of the homodimer motif. To compare on equal footing, we assume in

what follows below that the production rate of x is equal in the two motifs— $\gamma_v \langle v \rangle^2 = 2\gamma_{uw} \langle u \rangle \langle w \rangle + \gamma_u \langle u \rangle^2 + \gamma_w \langle w \rangle^2$ —and that the production rate of v in the homodimer motif equals

the sum of that of u and w in the diamond motif; $\alpha_v\langle s \rangle = \alpha\langle s \rangle + \beta\langle s \rangle$.

If the degradation rates of the intermediate components in the two motifs are equal— $\mu_v = \mu_u = \mu_w$ —and the coupling between the intermediate components and the output is the same— $\gamma_{uv} = \gamma_u = \gamma_w$ —then Θ of the diamond motif is equal to that of the homodimer motif, as it should, since there is no distinction between the components. If the components U and W are different, leading, for example, to different couplings ($\gamma_u \neq \gamma_w$, still assuming equal degradation rates), the homodimer motif has a larger Θ than the diamond motif at small $\omega < \mu_u, \mu_w, \mu_v$. In this case, the gain is equal for both processes because the concentration of X is taken to be the same in both networks. Hence, the difference in the Θ originates from the noise. The noise term $N_X(\omega)$ is equal in both networks, because the production and degradation rates of X are taken to be the same in the two networks. However, the extrinsic noise propagated from the intermediate components is larger in the diamond motif, $N_{u \rightarrow x} + N_{w \rightarrow x} > N_{v \rightarrow x}$. It can be shown that under the constraints that (a) the production of v equals the total production of u and w and (b) the production of x is the same in both motifs, the extrinsic noise is minimized when the coupling of u and w to x are identical.

We stress, however, that the gain-to-noise ratio of this diamond motif is higher than that of a simple two-step cascade, with one intermediate component V that does not activate X in a cooperative manner. Indeed, for a diamond motif the coherent interaction between the two pathways plays a crucial role. While for a motif with noncooperative activation of the output by the two pathways, this coherence between the two pathways exactly compensates for the decrease of the gain of each independent pathway, for a network with cooperative interaction, the coherence term increases Θ over that of a simple two-step cascade. In other words, splitting the input signal into two and then recombining them with AND logic at the output does increase the gain-to-noise ratio.

Finally, we study a motif in which only the heterodimer UW , and not the homodimers U_2 and W_2 , can activate the output X . We thus consider the case that $\gamma_w = \gamma_u = 0$, and consider what happens if the degradation rates of U and W are allowed to be different. With unequal degradation rates, a bandpass filter for information is possible, if the coupling of one pathway to x is stronger than the coupling of the other pathway. We take $\mu_w > \mu_u$, such that the pathway $S \rightarrow W \rightarrow X$ is capable of transmitting information on faster time scales than the pathway $S \rightarrow U \rightarrow X$. If $\beta \gg \alpha$, the input signal is relayed more strongly via the pathway containing W , and the other pathway acts as a noise source. Consequently, for small frequencies, $\omega < \mu_u$, signal transmission will be corrupted by noise originating at u , but for $\mu_u < \omega < \mu_w$ this noise is averaged out. We thus obtain a band-pass filter for information transmission [Fig. 5(c), thick red solid line]. By actively changing the degradation rates μ_u and μ_w the cell can tune the frequency range of the band.

IV. DISCUSSION

Our analysis relies on two key assumptions. First, it assumes that the network can be linearized. By comparing the results of our analysis with those of stochastic simulations of

the full nonlinear network, we found that this assumption is quite accurate. This is because the copy numbers are fairly large [$\mathcal{O}(100)$] and hence the fluctuations relative to the mean are small, such that the transfer function is indeed linear on the scale of the fluctuations. Second, our analysis ignores the effect of cross-correlations in the noise terms, it assumes that $\langle \eta_i \eta_j \rangle = 0$ for $i \neq j$. For networks consisting of birth-death reactions, such as those occurring in gene networks, the cross-correlations are, in fact, zero [17]. However, in general, for association-dissociation reactions and protein modification reactions these cross-correlations can be important. More in particular, these cross-correlations can significantly affect information transmission in push-pull networks [17], especially when the enzymatic reactions are in the zero-order regime; yet, when these reactions are in the linear regime (i.e., linear in the substrate concentration), the effect of cross-correlations are significantly reduced [17], and the results presented here thus directly apply to these networks as well. In the Appendix we discuss in more detail the importance of cross-correlations in the noise for the transmission of information in time-varying signals via ff motifs.

Our analysis reveals that ff motifs and diamond motifs are very rich information processing devices. More specifically, our study shows that both coherent and iff motifs can act as either low- or high-pass filters for information. This behavior can be understood by noting that while at high frequencies the direct pathway always dominates the output signal, at low frequencies the contribution of each pathway to the output varies between networks, depending on the coupling constants between the components in the network; moreover, at low frequencies, the output strongly depends on the nature of the interaction between the two pathways. If the interaction is coherent, as in cff motifs, then the frequency dependence of the gain, noise, and gain-to-noise ratio tends to have low-pass characteristics. If, however, not only at high frequencies, but also at low frequencies the direct pathway dominates the response, then the gain-to-noise ratio can have high-pass characteristics; in this case, the signal is predominantly transmitted via the direct pathway, while the indirect pathway acts as a noise source, masking this signal at low frequencies. For an iff motif, the gain is low at low frequencies because of the destructive interference between the two pathways; consequently, the frequency dependence of the gain-to-noise ratio of iff motifs tends to have high-pass characteristics. However, an incoherent motif can also act as a low-pass filter for information. This scenario arises when the direct pathway acts weakly on the output at low frequencies; then at low frequencies the output is dominated by the indirect pathway, which exhibits low-pass signal filtering, while at high frequencies it is dominated by the direct pathway. More generally, our analysis demonstrates that by changing the coupling constants between the components the frequency-dependence of the gain, noise, and gain-to-noise ratio can be sculpted in almost any desirable manner. Importantly, this shows that the topology by itself does not determine the qualitative behavior of the system.

For equal total production cost of all molecules in the network a cff motif has a lower information transmission capacity than a simple one-step motif. However, if we allow for a higher production cost in the ff motif (e.g., we require equal production of x , but allow for the additional production

of the intermediate component V), a cff motif in which the two pathways are combined following AND logic can have a larger Θ than the one-step cascade. Combining the pathways following OR logic has always a smaller Θ than the one-step cascade, even for larger total production. This demonstrates for these simple cascades a possible advantage of coincidence detection.

Our results also underscore the important observation that the power spectrum of the output signal is not a good measure for information transmission [5]. The results on the iff motif provide a concrete illustration of this idea: While the frequency dependence of the gain exhibits band-pass characteristics, the frequency dependence of the gain-to-noise ratio shows high-pass characteristics; indeed, at high frequencies not only the gain, and hence the output, decreases, but also the noise. The cff motif with a high-pass gain-to-noise ratio provides another striking example: While the gain and hence the output decreases with frequency, the gain-to-noise ratio increases; this is because the high gain at low frequencies is masked by the high noise. Our results thus show that in order to draw any conclusion on how reliably a network can transmit time-varying signals, one needs to measure not only the power spectrum of the output $P_{xx}(\omega)$, but also the power spectrum of the input $P_{ss}(\omega)$ and their cross-power spectrum $P_{sx}(\omega)$: From these quantities one can obtain the gain $g^2(\omega) \equiv |P_{sx}(\omega)|^2 / P_{ss}^2(\omega)$ and the frequency dependence noise $N(\omega)$, and hence the gain-to-noise ratio [see Eqs. (7) and (8)].

Finally, our analysis of the diamond motifs reveals that cooperatively activating the output can markedly enhance the gain, as well as the gain-to-noise ratio. The latter is due to coincidence detection: While variations in the input signal lead to correlated variations in the intermediate components that tend to boost the output, noise generates uncorrelated fluctuations in the intermediate components. We emphasize that this mechanism is very generic. Indeed, cooperative activation of the output can be implemented in many ways: via the formation of homo- or heteromultimers that then bind and activate the output component or via the individual binding of the intermediate components to the output component. While cooperative activation of the output via one and the same type of intermediate component, as in the case of homodimerization, increases the overall gain-to-noise ratio, cooperative output activation via components that are different, as in the case of heterodimerization of the intermediate components, makes it possible to mold the frequency dependence of the gain-to-noise ratio, even allowing for band-pass filters for information.

ACKNOWLEDGMENTS

We would like to thank Philippe Nghe for a critical reading of this paper. This work is part of the research program of the Foundation for Fundamental Research on Matter (FOM), which is financially supported by the Netherlands Organization for Fundamental Research (NWO).

APPENDIX: SUPPLEMENTARY INFORMATION

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

$$\frac{ds}{dt} = k_s - m_s s + \eta_s(t). \quad (\text{A1})$$

1. Simple cascades

Simple cascades, cascades without ff interaction, are described using roman symbols for the kinetic rates, while motifs with ff interactions, are described using Greek symbols. The network for the one-step simple cascade is described by

$$\frac{dx}{dt} = k_x s - m_x x + \eta_x(t). \quad (\text{A2})$$

Gain, noise, and gain-to-noise are

$$g^2(\omega) = \frac{k_x^2}{\omega^2 + m_x^2}, \quad (\text{A3})$$

$$N(\omega) = \frac{\langle |\eta_x|^2 \rangle}{\omega^2 + m_x^2}, \quad (\text{A4})$$

$$\Theta(\omega) = \frac{g^2(\omega)}{N(\omega)} = \frac{k_x^2}{\langle |\eta_x|^2 \rangle} = \frac{k_x}{2\langle s \rangle}. \quad (\text{A5})$$

The network for the two-step simple cascade is described by

$$\frac{dv}{dt} = k_v s - m_v v + \eta_v(t), \quad (\text{A6})$$

$$\frac{dx}{dt} = k_x v - m_x x + \eta_x(t). \quad (\text{A7})$$

Gain, noise, and gain-to-noise are

$$g^2(\omega) = \frac{(k_v k_x)^2}{(\omega^2 + m_x^2)(\omega^2 + m_v^2)}, \quad (\text{A8})$$

$$N(\omega) = \frac{k_x^2 \langle |\eta_v|^2 \rangle}{(\omega^2 + m_v^2)(\omega^2 + m_x^2)} + \frac{\langle |\eta_x|^2 \rangle}{(\omega^2 + m_x^2)}, \quad (\text{A9})$$

$$\Theta(\omega) = \frac{g^2(\omega)}{N(\omega)} = \frac{(k_v k_x)^2}{k_x^2 \langle |\eta_v|^2 \rangle + (\omega^2 + m_v^2) \langle |\eta_x|^2 \rangle}. \quad (\text{A10})$$

a. Production constraints

In the following sections we compare the one-step (os) and two-step (ts) cascades with different ff motifs. We always assume that degradation rates for proteins are equal, unless specified otherwise. We use three different constraints, such that the comparison is performed on an equal footing. These constraints are as follows (with ts = two-step, ff = feed-forward, os = one-step).

(1) Equal production of x , free production of v :

$$p_X^{\text{ts}} = p_X^{\text{ff}}, \quad p_X^{\text{os}} = p_X^{\text{ff}}. \quad (\text{A11})$$

We note that this also implies $\langle |\eta_x^{\text{ts}}|^2 \rangle = \langle |\eta_x^{\text{os}}|^2 \rangle = \langle |\eta_x^{\text{ff}}|^2 \rangle$.

(2) Equal total production of x and v :

$$p_X^{\text{ts}} + p_V^{\text{ts}} = p_X^{\text{ff}} + p_V^{\text{ff}}, \quad p_X^{\text{os}} = p_X^{\text{ff}} + p_V^{\text{ff}}. \quad (\text{A12})$$

(3) Equal production of x and v separately (this constraint has no meaning for the one-step cascade):

$$p_X^{\text{ts}} = p_X^{\text{ff}}, \quad p_V^{\text{ts}} = p_V^{\text{ff}}. \quad (\text{A13})$$

We note that this also implies $\langle |\eta_x^{\text{ts}}|^2 \rangle = \langle |\eta_x^{\text{ff}}|^2 \rangle$, $\langle |\eta_v^{\text{ts}}|^2 \rangle = \langle |\eta_v^{\text{ff}}|^2 \rangle$.

Unless specified otherwise, the degradation rates are equal for components in either cascade, for example, $m_v = \mu_v$.

2. The effect of cross-correlations in the noise

The Langevin equations for the cff motif with OR logic are

$$\frac{dv}{dt} = \alpha s - \mu_v v + \eta_v(t), \quad (\text{A14})$$

$$\frac{dx}{dt} = \beta s + \gamma v - \mu_x x + \eta_x(t), \quad (\text{A15})$$

and the most general expression for the noise $N(\omega)$, as defined through Eq. (8) (main text) is

$$N(\omega) = \frac{\langle |\eta_x|^2 \rangle}{\omega^2 + \mu_x^2} + \frac{\gamma^2 \langle |\eta_v|^2 \rangle}{(\omega^2 + \mu_x^2)(\omega^2 + \mu_v^2)} + \frac{\mu_v \gamma \langle \eta_v \eta_x \rangle}{(\omega^2 + \mu_x^2)(\omega^2 + \mu_v^2)}. \quad (\text{A16})$$

Comparing Eq. (A16) with Eq. (12) (main text) the difference lies in the appearance of the term consisting of the cross-correlations in the noise. This term, however, can be positive or negative, depending on the precise molecular interactions [3,5,17]. Indeed, for a molecular reaction where V transforms into X ,



the Langevin equations change into

$$\frac{dv}{dt} = \alpha s - (\gamma + \mu_v) v + \eta_v(t), \quad (\text{A18})$$

$$\frac{dx}{dt} = \beta s + \gamma v - \mu_x x + \eta_x(t). \quad (\text{A19})$$

For these molecular reactions the cross-correlation $\langle \eta_v \eta_x \rangle$ is negative, reducing the total noise. The physical intuition for this effect is that for this type of molecular interaction the production event of X is identical to the degradation event of V . Therefore, the noise, defined through to the summation of all individual production and degradation events, overestimates the actual noise. Indeed, the expression for the noise is

$$N(\omega) = \frac{2\langle S \rangle (\alpha \gamma + \beta (\gamma + \mu_v))}{(\mu_x^2 + \omega^2)(\gamma + \mu_v)}, \quad (\text{A20})$$

which if decay of V is assumed to be absent ($\mu_v = 0$) reduces to

$$N(\omega) = \frac{2\langle S \rangle (\alpha + \beta)}{(\mu_x^2 + \omega^2)}. \quad (\text{A21})$$

Comparing Eq. (A21) with Eq. (A42) we observe that the noise is reduced, especially at low frequencies, since the effective transmission of noise from V to X has decreased.

However, not only the noise term is affected, but also the gain has changed, leading to

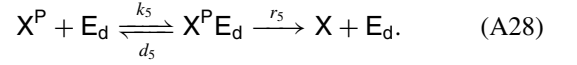
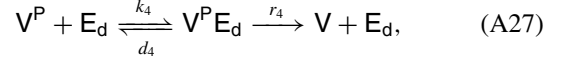
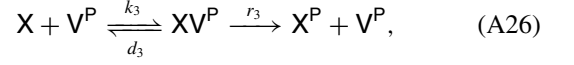
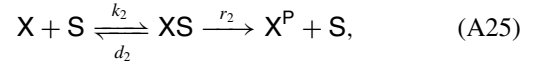
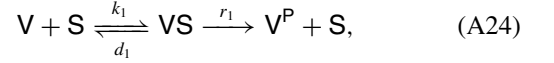
$$g^2(\omega) = \frac{(\alpha \gamma + \beta (\gamma + \mu_v))^2 + \beta^2 \omega^2}{(\mu_x^2 + \omega^2)(\omega^2 + (\gamma + \mu_v)^2)}. \quad (\text{A22})$$

If decay of V ($\mu_v = 0$) is assumed to be absent, this reduces to

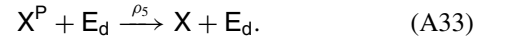
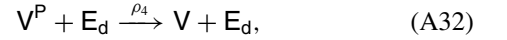
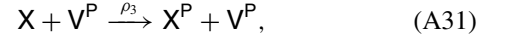
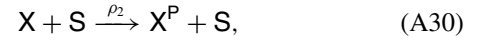
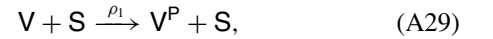
$$g^2(\omega) = \frac{(\alpha \gamma + \beta \gamma)^2 + \beta^2 \omega^2}{(\mu_x^2 + \omega^2)(\omega^2 + \gamma^2)}, \quad (\text{A23})$$

which is equivalent to the result in the main text for $\gamma = \mu_v$.

We now focus on an enzymatic push-pull network:



In the limit that $d_i \gg r_i$, the back-reactions can be integrated out and we obtain



This is a nonlinear coarse-grained push-pull network. Although it resembles a birth-death process, it is not, since in the push-pull network the cross-correlations between V, V^P and X, X^P are still present (because there is a constraint on the total substrate concentration). In order to study the influence of cross-correlations on gain, noise, and gain-to-noise, we compare this push-pull network with an equivalent birth-death process.



The expressions for the frequency-dependence of the gain, noise, and gain-to-noise ratio are unwieldy and we do not present them here, but in Fig. 6 we show the gain, noise, and Θ for different values of V_T and X_T , the total copy numbers of X and V , respectively, where the average of the components in each of the two networks has been made equal by changing the rate κ_i of the birth-death reactions.

It is seen that for large $V_T > \langle S \rangle$ [Figs. 6(c) and 6(d)] the difference in $g^2(\omega)$, $N(\omega)$, and Θ between the push-pull and birth-death network is larger than for small $V_T < \langle S \rangle$ [Figs. 6(a) and 6(b)] which is in agreement with earlier observations [17]. This effect is stronger if also X_T is large compared to its substrates $\langle S \rangle, \langle V^P \rangle$ [compare Figs. 6(c)

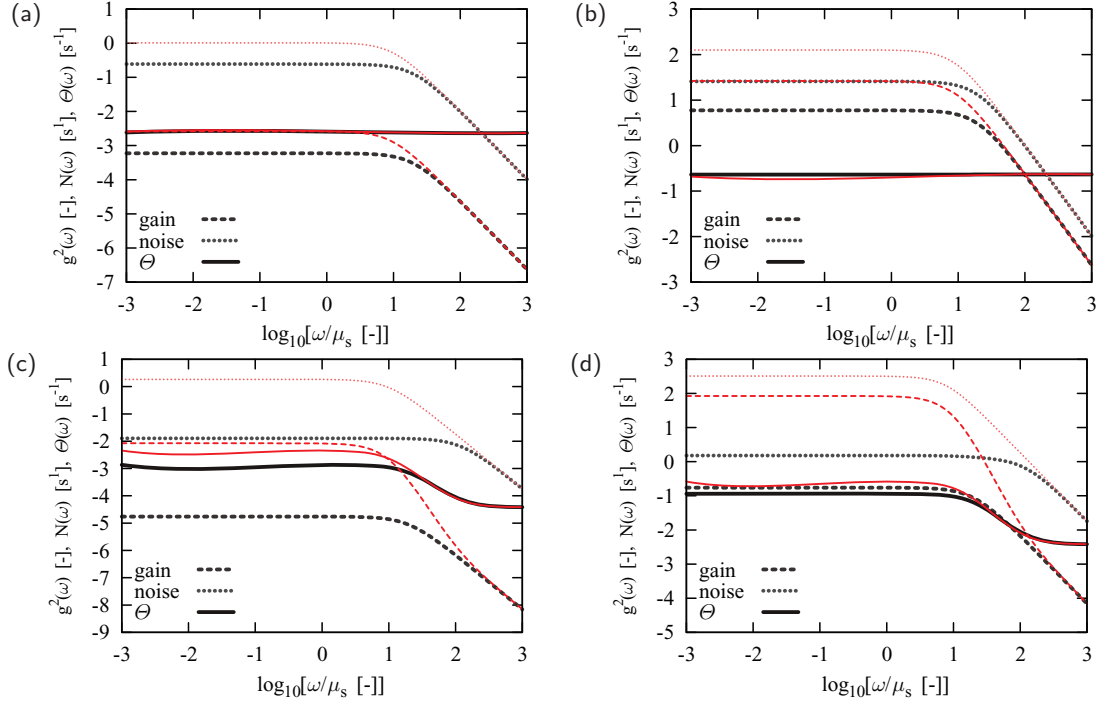


FIG. 6. (Color online) For different values of V_T [(a),(b) $V_T = 10$, (c),(d) $V_T = 2000$] and X_T [(a),(c) $X_T = 10$, (b),(d) $X_T = 1000$] the gain (dashed), noise (dotted), and gain-to-noise ratio (solid) are shown. The results from the push-pull network are shown in thick black, the birth-death system is shown in thin red. Results are shown on the basis of equal average (V^P), (X^P). For $V_T < \langle S_1 \rangle$ [(a),(b)] Θ of the push-pull network and birth-death system are very similar, while for $V_T > \langle S_1 \rangle$ [(c),(d)] Θ of the push-pull network is smaller than that of the birth-death network, due to the large effect of the cross-correlations in reducing the noise. Parameters: $k_s = 100$, $\rho_1 = 0.1$, $\rho_2 = 0.1$, $\rho_3 = 0.1$, $\rho_4 E_D = 10$, $\rho_5 E_D = 10$. μ_s sets the time scale.

and 6(d)]. The cross-correlations tend to reduce $N(\omega)$ and $g^2(\omega)$ in the push-pull network. The decrease due to the cross-correlations is more prominent in the gain, and as a result, Θ of the push-pull network is smaller than for the birth-death network.

3. Discussion on Eq. (15) (main text) for the coherent feed-forward motif

In this section we provide some intuition on Eq. (15) (main text) for a cff motif. Θ of a cff motif either increases monotonically with frequency or decreases monotonically with frequency, as mentioned in the main text. If Eq. (15) (main text) is satisfied, it increases monotonically, and the motif acts as a high-pass filter for information. We can intuitively understand the terms in Eq. (15) (main text) as follows. A decrease in the ratio J_{vs}/J_{xs} means that the input signal s is relayed more to x directly than to x via v . However, while the direct pathway $S \rightarrow X$ contributes to information transmission at all frequencies—its Θ is flat—the indirect pathway $S \rightarrow V \rightarrow X$ only contributes at low frequencies— Θ of a two-step cascade falls off as ω^{-2} for high frequencies. The effect of the indirect pathway, both the gain and the noise, at high frequencies becomes negligible. Indeed, in the limit that J_{vs}/J_{xs} reaches zero, the signal is transmitted completely via the direct pathway only; yet, while the Θ of the direct pathway, a one-step cascade, is flat, Θ of the cff increases with frequency, because the indirect pathway

still adds noise to the signal, especially in the low-frequency regime.

The presence of the second term on the right-hand side of Eq. (15) (main text) can be understood by noting that it arises from the interplay between the noise [Eq. (12) (main text)] and the coherence term in the gain [Eq. (9) (main text)]. We can understand the dependence on $\mu_v = -J_{vv}$ by noting that at high frequencies $\omega \gg \mu_v$ the coherence and the noise coming from v hardly contribute to the gain and the total noise $N(\omega)$, respectively, while at $\omega = 0$ the coherence decreases with increasing degradation rate as μ_v^{-1} [Eq. (9) (main text)] while the noise coming from v decreases as μ_v^{-2} [Eq. (12) (main text)]. Similarly, at $\omega = 0$, the coherence increases with J_{xv} while the noise coming from v increases with J_{xv}^2 . Thus, at low frequencies decreasing the degradation rate μ_v and/or increasing the coupling between V to X , J_{xv} , increases the noise more than it does the gain, thus reducing Θ at low frequencies, while at high frequencies the influence on Θ is negligible.

4. OR-coherent feed-forward

The network is described by

$$\frac{dv}{dt} = \alpha s - \mu_v v + \eta_v(t), \quad (\text{A39})$$

$$\frac{dx}{dt} = \beta s + \gamma v - \mu_x x + \eta_x(t), \quad (\text{A40})$$

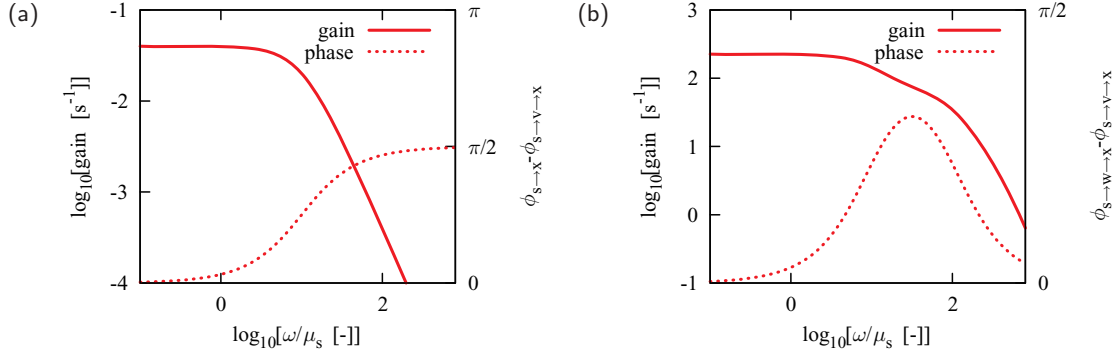


FIG. 7. (Color online) In these panels we show the gain (solid line, left axis) and phase (dotted line, right axis). In panel (a) we show the gain and the phase for an ocff motif. The decrease of the gain corresponds with the increase in the phase difference between the indirect $S \rightarrow V \rightarrow X$ and direct pathway $S \rightarrow X$. In panel (b) we show a similar plot, but now for the diamond motif with pathways of equal length, but with different degradation rates of the intermediate components. Again we observe that the initial decrease in the gain coincides with the increase in phase difference. However, although the phase difference decreases for larger ω again, the gain continues to decrease, due to the time averaging of the fluctuations over the lifetime of the intermediate components.

and the gain, noise, and gain-to-noise [see Eqs. (9), (12), and (14) (main text)],

$$g^2(\omega) = \frac{(\alpha\gamma + J_{vv}\beta)^2 + J_{xs}^2\omega^2}{(\omega^2 + \mu_v^2)(\omega^2 + \mu_x^2)}, \quad (\text{A41})$$

$$N(\omega) = \frac{\gamma^2\langle|\eta_v|^2\rangle + (\omega^2 + \mu_v^2)\langle|\eta_x|^2\rangle}{(\omega^2 + \mu_v^2)(\omega^2 + \mu_x^2)}, \quad (\text{A42})$$

$$\Theta(\omega) = \frac{(\alpha\gamma - \mu_v\beta)^2 + \beta^2\omega^2}{\gamma^2\langle|\eta_v|^2\rangle + (\omega^2 + \mu_v^2)\langle|\eta_x|^2\rangle}. \quad (\text{A43})$$

In Fig. 7 (left) the dependence of the gain and the phase difference between the direct $S \rightarrow X$ and indirect $S \rightarrow V \rightarrow X$ pathway as a function of frequency is shown. For $\omega < \mu_v$ the two pathways have no phase difference, but for an increase in frequency the phase difference increases and as a result the magnitude of the coherent term in the gain decreases. The magnitude of the coherent term also decreases due to the time-averaging over the finite lifetime of the intermediate

component V . Even for a constant phase-difference, at large ω the gain decreases.

In Tables I and II we list the expression for Θ for different production constraints. Θ of the one-step cascade given equal production of x is always larger than that of the ocff. Compared to the two-step cascade, for $\omega \rightarrow \infty$ the ocff motif always has a larger Θ , since the noise source $N_{v \rightarrow x}$ is averaged out, and only an effective one-step cascade with constant Θ remains. Thus, we only compare the ocff with the two-step cascade for $\omega \rightarrow 0$.

The various constraints lead to different results. The first scenario, where we independently constrain p_v and p_x and take $\mu_v = m_v$ in both cascades leads to an expression for Θ in the two-step cascade that is smaller than that for the ocff for all ω . The gain in both the ocff and the two-step cascade are precisely equal (not shown), and therefore the larger Θ is due to the increase in the noise ($N(\omega)$) in the two-step cascade. A constraint that allows for more freedom in the network is such that only the production of x is constrained, but there is free production of v . We observe that for $\omega = 0$, Θ for the ocff is

TABLE I. The results for the gain, noise, and gain-to-noise ratio for the one-step cascade and the two-step cascade in parameters of the ocff, given the constraint conditions. For the two-step only the $\omega \rightarrow 0$ limit is given, since for large ω the ocff always has a larger Θ . We assume always $\mu_v = m_v$. E.p.x, equal production of x ; E.t.p., equal total production; E.s.p., equal separate production.

Constraint	One-step	Θ^{os}
E.p.x	$k_x = \beta + \frac{\alpha\gamma}{\mu_v}$	$\Theta = \frac{(\mu_v\beta + \alpha\gamma)^2}{\mu_v^2\langle \eta_x ^2\rangle} > [\Theta]^{\text{ocff}}$
E.t.p.		Always larger
	Two-step	$\Theta^{\text{ts}}(\omega = 0)$
E.p.x	$k_x k_v = \beta + \frac{\alpha\gamma}{\mu_v}$	$\Theta = \frac{(\mu_v\beta + \alpha\gamma)^2}{\mu_v^2\langle \eta_x ^2\rangle + \frac{2(s)}{k_v}(\mu_v\beta + \alpha\gamma)^2}$
E.t.p.	$(k_x + \mu_v)k_v = (\alpha + \beta)\mu_v + \gamma\alpha$	$\Theta = \frac{k_x(\beta\mu_v + \alpha(\gamma + \mu_v))}{(k_x + \mu_v)^2}$
E.s.p.	$k_v = \alpha,$	$g^2 = \frac{(\mu_v\beta + \alpha\gamma)^2}{\mu_v^2\mu_x^2}$
	$k_x = \gamma + \frac{\beta\mu_v}{\alpha}$	$N = \frac{(\mu_v\beta + \alpha\gamma)^2\langle \eta_v ^2\rangle + \mu_v^2\langle \eta_x ^2\rangle}{\mu_v^2\mu_x^2}$
		$\Theta = \frac{(\mu_v\beta + \alpha\gamma)^2}{\mu_v^2\langle \eta_x ^2\rangle + \frac{2(s)}{\alpha^2}\langle \eta_v ^2\rangle + \mu_v^2\langle \eta_x ^2\rangle}$

TABLE II. Summary of the results for the ocff.

Constraint	One-step vs ocff	Two-step vs ocff
Equal production of x	$\Theta^{\text{os}} > \Theta^{\text{ocff}}$	Parameter dependent
Equal total production	$\Theta^{\text{os}} > \Theta^{\text{ocff}}$	Parameter dependent
Equal separate production		$\Theta^{\text{ts}} < \Theta^{\text{ocff}}$

larger than the two-step if

$$\gamma^2 \langle |\eta_v|^2 \rangle^{\text{ocff}} < \frac{2\langle s \rangle}{k_v} (\mu_v \beta + \alpha \gamma)^2, \quad (\text{A44})$$

$$2\alpha \langle s \rangle \gamma^2 < \frac{2\langle s \rangle}{k_v} (\mu_v \beta + \alpha \gamma)^2, \quad (\text{A45})$$

$$\alpha \gamma^2 < \frac{(\mu_v \beta + \alpha \gamma)^2}{k_v}. \quad (\text{A46})$$

This inequality is not satisfied if, for example, $k_v \rightarrow \infty$, since then the two-step cascade effectively becomes a one-step cascade and as a result, has a larger Θ than the ocff.

For $\omega \rightarrow \infty$, Θ of the ocff with high-pass (hp) filter characteristics can have a larger plateau value than an ocff with low-pass (lp) characteristics, under the constraint that the production of v and x separately are equal. This can be shown by solving for the inequality

$$\lim_{\omega \rightarrow \infty} \Theta^{\text{hp}} > \lim_{\omega \rightarrow \infty} \Theta^{\text{lp}} \rightarrow \left(\frac{\beta^{\text{hp}}}{\mu_v^{\text{hp}}} \right)^2 > \left(\frac{\beta^{\text{lp}}}{\mu_v^{\text{lp}}} \right)^2, \quad (\text{A47})$$

$$\text{low-pass (lp)} \rightarrow \beta^{\text{lp}} < \left(\beta^{\text{lp}} + \frac{\gamma^{\text{lp}} \alpha^{\text{lp}}}{\mu_v^{\text{lp}}} \right) \frac{\alpha^{\text{lp}}}{\beta^{\text{lp}}} + \frac{2\mu_v^{\text{lp}}}{\gamma^{\text{lp}}}, \quad (\text{A48})$$

$$\text{high-pass (hp)} \rightarrow \beta^{\text{hp}} < \left(\beta^{\text{hp}} + \frac{\gamma^{\text{hp}} \alpha^{\text{hp}}}{\mu_v^{\text{hp}}} \right) \frac{\alpha^{\text{hp}}}{\beta^{\text{hp}}} + \frac{2\mu_v^{\text{hp}}}{\gamma^{\text{hp}}}. \quad (\text{A49})$$

The first line describes the inequality. The second line describes the requirement for β^{lp} and the third line the requirement for β^{hp} , which are obtained following substitution of $\langle |\eta_v|^2 \rangle$ and $\langle |\eta_x|^2 \rangle$ in Eq. (15) (main text). From the constraints of equal separate production we obtain

$$\langle |\eta_v^{\text{lp}}|^2 \rangle = \langle |\eta_v^{\text{hp}}|^2 \rangle \rightarrow \alpha^{\text{lp}} \equiv \alpha^{\text{hp}}, \quad (\text{A50})$$

$$\langle |\eta_x^{\text{lp}}|^2 \rangle = \langle |\eta_x^{\text{hp}}|^2 \rangle \rightarrow \beta^{\text{lp}} + \frac{\gamma^{\text{lp}} \alpha^{\text{lp}}}{\mu_v^{\text{lp}}} \equiv \beta^{\text{hp}} + \frac{\gamma^{\text{hp}} \alpha^{\text{hp}}}{\mu_v^{\text{hp}}}. \quad (\text{A51})$$

Solving the system of inequalities Eqs. (A47)–(A49) with Eqs. (A50) and (A51) using MATHEMATICA, it can be shown that these can always be fulfilled. However, the full expressions are unwieldy to present here. Even if $\beta^{\text{lp}} \rightarrow \infty$, reflecting that the low-pass filter effectively is a one-step cascade, parameters can be found for which the high-pass filter can still have a larger Θ for large ω . However, the difference between the low-pass and high-pass filter is negligible.

5. AND-coherent feed-forward

The network is described by

$$\frac{dv}{dt} = \alpha s - \mu_v v + \eta_v(t), \quad (\text{A52})$$

$$\frac{dx}{dt} = \beta v s - \mu_x x + \eta_x(t), \quad (\text{A53})$$

and the gain, noise, and gain-to-noise [Eqs. (9), (12), and (14) (main text)] by

$$g^2(\omega) = \frac{\beta^2 \langle v \rangle^2 (\omega^2 + 4\mu_v^2)}{(\omega^2 + \mu_v^2)(\omega^2 + \mu_x^2)}, \quad (\text{A54})$$

$$N(\omega) = \frac{\beta^2 \langle s \rangle^2 \langle |\eta_v|^2 \rangle + (\omega^2 + \mu_v^2) \langle |\eta_x|^2 \rangle}{(\omega^2 + \mu_v^2)(\omega^2 + \mu_x^2)}, \quad (\text{A55})$$

$$\Theta(\omega) = \frac{g^2(\omega)}{N(\omega)} = \frac{\beta^2 \langle v \rangle^2 (\omega^2 + 4\mu_v^2)}{\beta^2 \langle s \rangle^2 \langle |\eta_v|^2 \rangle + (\omega^2 + \mu_v^2) \langle |\eta_x|^2 \rangle}. \quad (\text{A56})$$

For equal production of x and for $\omega \rightarrow \infty$, Θ of a one-step cascade is equal to Θ of the acff. For $\omega = 0$, we obtain the relation

$$\Theta^{\text{os}}(\omega = 0) < \Theta^{\text{acff}}(\omega = 0) \quad (\text{A57})$$

$$\frac{\beta^2 \langle v \rangle^2}{\langle |\eta_x|^2 \rangle} < 4 \frac{\beta^2 \langle v \rangle^2}{\langle |\eta_x|^2 \rangle + \left(\frac{\beta \langle s \rangle}{\mu_v} \right)^2 \langle |\eta_v|^2 \rangle}, \quad (\text{A58})$$

where the last line holds if $\beta \langle v \rangle / \alpha < 3 \langle |\eta_x|^2 \rangle / \langle |\eta_v|^2 \rangle$. This inequality can be sufficed for example in the limit that $\beta \rightarrow 0$. Then the acff has a larger Θ than the one-step cascade for $\omega = 0$. This limit corresponds to a situation where the noise $N_{v \rightarrow x}$ [Eq. (12) (main text)] is negligible, and the noise contributions $N(\omega)$ in both motifs are similar. The gain in the acff is larger than in the one-step cascade. This seems contradictory, but is a result of the production constraint. In the acff the production of x , p_x , is $\beta \langle s \rangle \langle v \rangle$, such that if $\langle v \rangle$ is very small, the production rate β becomes very large to ensure equal production of x . This is in contrast to the ocff, where the production of x is $\beta \langle s \rangle + \gamma \langle v \rangle$ and if $\langle v \rangle$ is nearly zero, production still is possible through the direct pathway $S \rightarrow X$. Due to this coincidence coupling between the two pathways, the gain for the acff is larger than for the one-step cascade.

As expected, the acff has a larger Θ than the simple two-step cascade if we equalize both p_v and p_x separately. More interesting is the behavior at low frequencies for the two other constraint types. If we equalize the total production the two-step cascade has a larger Θ than the acff if

$$\frac{\mu_v (\beta \langle v \rangle^{\text{acff}} + \alpha - k_v)}{2\langle s \rangle (\mu_v + k_x)} > \frac{\beta^2 \langle v \rangle^2 (\omega^2 + 4\mu_v^2)}{\beta^2 \langle s \rangle^2 \langle |\eta_v^{\text{acff}}|^2 \rangle + (\omega^2 + \mu_v^2) \langle |\eta_x^{\text{acff}}|^2 \rangle}, \quad (\text{A59})$$

$$\frac{\beta \langle s \rangle + \mu_v \beta \langle v \rangle^{\text{acff}} + \alpha - k_v}{k_x + \mu_v} \frac{\alpha - k_v}{k \langle v \rangle^{\text{acff}}} > 4, \quad (\text{A60})$$

where this is possible if $\beta \langle s \rangle \gg k_x$ or $\alpha \gg k_v$. From the constraint condition of total production we have the equality $(k_x + \mu_v) k_v = \mu_v \alpha + \beta \alpha \langle s \rangle$. Taking these relations together, we observe that for $\beta \langle s \rangle \gg k_x$ we require $k_v \gg 1$, while for

TABLE III. The results for the gain, noise, and gain-to-noise ratio for the one-step cascade and the two-step cascade in parameters of the acff, given the constraint conditions. For the two-step only the $\omega \rightarrow 0$ limit is given, since for large ω the ocff always has a larger Θ . E.p.x, equal production of x ; E.t.p., equal total production; E.s.p., equal separate production.

Constraint	One-step	Θ^{os}
E.p.x	$k_x = \beta \langle v \rangle$	$\Theta = \frac{\beta^2 \langle v \rangle^2}{\langle \eta_x ^2}$
E.t.p.	$k_x = \alpha + \beta \langle v \rangle$	$\Theta = \frac{(\alpha + \beta \langle v \rangle)^2}{\langle \eta_x^{\text{ts}} ^2} > \Theta^{\text{acff}}$
	Two-step ($\Theta^{\text{ts}} (\omega = 0)$)	
E.p.x	$k_x k_v = \beta \langle v^{\text{acff}} \rangle \mu_v$	$\Theta = \frac{(s) \alpha \beta^2}{\langle \eta_x ^2 \mu_v^2 + \frac{(s) \alpha \beta^2}{k_v^2} \langle \eta_v ^2}$
E.t.p.	$(k_x + \mu_v) k_v = \mu_v (\alpha + \beta \langle v^{\text{acff}} \rangle)$ $k_v = \alpha,$	$\Theta = \frac{k_x \alpha (\beta \langle s \rangle + \mu_v)}{2 \langle s \rangle (k_x + \mu_v)^2}$ $g^2 = \frac{(s) \alpha \beta^2}{\mu_v^2 \mu_x^2}$
E.s.p.	$k_x = \frac{\mu_v}{k_v} \beta \langle v \rangle$	$N = \frac{\langle \eta_x ^2 \mu_v^2 + (s) \beta^2 \langle \eta_v ^2}{\mu_v^2 \mu_x^2}$ $\Theta = \frac{(s) \alpha \beta^2}{\langle \eta_x ^2 \mu_v^2 + (s) \beta^2 \langle \eta_v ^2}$

$\alpha \gg k_v$ we require $k_x \gg 1$. In both cases the two-step cascade transforms into a one-step cascade because one of the two steps directly tracks the changes upstream.

The other option, where we constrain the production of x , but not v ,

$$k_x = \beta \langle s \rangle \frac{\langle v^{\text{acff}} \rangle}{\langle v^{\text{two}} \rangle}, \quad (\text{A61})$$

leads to the following expression for the gain-to-noise ratio at low frequencies for the two-step cascade:

$$\Theta(\omega = 0) = \frac{(\mu_v \beta \langle v^{\text{acff}} \rangle)^2}{\beta \langle s \rangle \frac{\langle v^{\text{acff}} \rangle}{\langle v^{\text{ts}} \rangle} \langle \eta_v^{\text{ts}} |^2 + \mu_v^2 \langle \eta_x |^2}. \quad (\text{A62})$$

The ratio of Θ 's is

$$\frac{\Theta^{\text{acff}}(\omega = 0)}{\Theta^{\text{ts}}(\omega = 0)} = 4 \frac{(\beta \langle s \rangle \langle v^{\text{acff}} \rangle)^2 \langle \eta_v^{\text{ts}} |^2 + \mu_v^2 \langle v^{\text{ts}} \rangle^2 \langle \eta_x |^2}{(\beta \langle s \rangle \langle v^{\text{ts}} \rangle)^2 \langle \eta_v^{\text{acff}} |^2 + \mu_v^2 \langle v^{\text{ts}} \rangle^2 \langle \eta_x |^2} \quad (\text{A63})$$

$$= 4 \frac{\frac{\alpha^2 \langle s \rangle^2 k_v}{\mu_v^2} + \frac{\mu_v^2 \langle v^{\text{ts}} \rangle^2 \langle \eta_x |^2}{2 \langle s \rangle (\beta \langle s \rangle)^2}}{\frac{\alpha \langle s \rangle^2 k_v^2}{\mu_v^2} + \frac{\mu_v^2 \langle v^{\text{ts}} \rangle^2 \langle \eta_x |^2}{2 \langle s \rangle (\beta \langle s \rangle)^2}} \quad (\text{A64})$$

$$= 4 \frac{\alpha + F}{k_v + F}, \quad \text{where} \quad F = \frac{k_v \mu_v^2 \langle \eta_x |^2}{2 \alpha \beta^2 \langle s \rangle}. \quad (\text{A65})$$

If $k_v \gg \alpha$ the two-step cascade has a larger Θ . This reflects again a situation where the two-step cascade has one very fast step and acts effectively as a single one-step cascade (see Tables III and IV).

6. Comparison of the coherent feed-forward AND and OR motifs

We study if Θ of the acff is larger or smaller than Θ of the ocff. We equalize production of both v and x such that we have

$$p_v^{\text{ocff}} = p_v^{\text{acff}} \rightarrow \alpha^{\text{ocff}} = \alpha^{\text{acff}} = \alpha, \quad (\text{A66})$$

TABLE IV. Summary of the results for the acff.

Constraint	One-step vs acff	Two-step vs acff
Equal production of x	Parameter dependent	Parameter dependent
Equal total production	$\Theta^{\text{os}} > \Theta^{\text{acff}}$	Parameter dependent
Equal separate production		$\Theta^{\text{ts}} < \Theta^{\text{acff}}$

$$p_x^{\text{ocff}} = p_x^{\text{acff}} \rightarrow \beta^{\text{ocff}} + \frac{\alpha \gamma}{\mu_v} = \beta^{\text{acff}} \langle v \rangle, \quad (\text{A67})$$

and using these constraints and Eqs. (A43) and (A56), we obtain the ratio

$$\frac{\Theta^{\text{acff}}(\omega)}{\Theta^{\text{ocff}}(\omega)} = \frac{(\beta^{\text{ocff}} + \frac{\alpha \gamma}{\mu_v})^2 (\omega^2 + 4\mu_v^2)}{(\beta^{\text{ocff}})^2 (\omega^2 + \mu_v^2) + (\alpha \gamma)^2 + 2\alpha \beta \gamma \mu_v} \quad (\text{A68})$$

$$\times \frac{\gamma^2 \langle \eta_v |^2 + (\omega^2 + \mu_v^2) \langle \eta_x |^2}{(\frac{\beta^{\text{ocff}} \mu_v}{\alpha} + \gamma)^2 \langle \eta_v |^2 + (\omega^2 + \mu_v^2) \langle \eta_x |^2}. \quad (\text{A69})$$

In the limit $\omega \rightarrow \infty$ we obtain

$$\frac{\Theta^{\text{acff}}(\omega \rightarrow \infty)}{\Theta^{\text{ocff}}(\omega \rightarrow \infty)} = \frac{(\beta^{\text{ocff}} + \frac{\alpha \gamma}{\mu_v})^2}{(\beta^{\text{ocff}})^2} > 1, \quad (\text{A70})$$

such that for large ω the acff has a larger Θ . In the other limit $\omega = 0$, we obtain

$$\frac{\Theta^{\text{acff}}(\omega = 0)}{\Theta^{\text{ocff}}(\omega = 0)} = \frac{(\beta^{\text{ocff}} + \frac{\alpha \gamma}{\mu_v})^2 4\mu_v^2}{(\beta^{\text{ocff}})^2 \mu_v^2 + (\alpha \gamma)^2 + 2\alpha \beta \gamma \mu_v} \quad (\text{A71})$$

$$\times \frac{\gamma^2 \langle \eta_v |^2 + \mu_v^2 \langle \eta_x |^2}{(\frac{\beta^{\text{ocff}} \mu_v}{\alpha} + \gamma)^2 \langle \eta_v |^2 + \mu_v^2 \langle \eta_x |^2}$$

$$= 4 \frac{\gamma^2 \langle \eta_v |^2 + \mu_v^2 \langle \eta_x |^2}{(\frac{\beta^{\text{ocff}} \mu_v}{\alpha} + \gamma)^2 \langle \eta_v |^2 + \mu_v^2 \langle \eta_x |^2}, \quad (\text{A72})$$

such that if

$$4 (\gamma^2 \langle \eta_v |^2 + \mu_v^2 \langle \eta_x |^2) < \left(\frac{\beta^{\text{ocff}} \mu_v}{\alpha} + \gamma \right)^2 \langle \eta_v |^2 + \mu_v^2 \langle \eta_x |^2, \quad (\text{A73})$$

the ocff has a larger Θ . This condition [Eq. (A73)] can be satisfied if in the ocff motif the direct pathway $\mathbf{S} \rightarrow \mathbf{X}$ couples (large β) much stronger to \mathbf{X} than the indirect pathway $\mathbf{S} \rightarrow \mathbf{V} \rightarrow \mathbf{X}$ (small α) (Table V). Interestingly, the parameter values for which this condition is satisfied, result in an Θ for the acff motif with a high-pass filter. This we show by comparing the two extrema for the acff ($\omega \rightarrow \infty$ and $\omega = 0$),

$$\Theta^{\text{acff}}(\omega = 0) = \frac{4\mu_v^2 (\beta^{\text{ocff}} + \frac{\alpha \gamma}{\mu_v})^2}{(\frac{\beta^{\text{ocff}} \mu_v}{\alpha} + \gamma)^2 \langle \eta_v |^2 + \mu_v^2 \langle \eta_x |^2}, \quad (\text{A74})$$

$$\lim_{\omega \rightarrow \infty} \Theta^{\text{acff}}(\omega) = \frac{(\beta^{\text{ocff}} + \frac{\alpha \gamma}{\mu_v})^2}{\langle \eta_x |^2}, \quad (\text{A75})$$

TABLE V. Summary of the results for the comparison between the acff and the ocff.

Constraint	$\omega \rightarrow 0$	$\omega \rightarrow \infty$
Equal separate production	Parameter dependent	$\Theta^{\text{acff}} > \Theta^{\text{ocff}}$

where it can be shown that, given the condition in Eq. (A73) is satisfied, the acff has a larger Θ for large frequencies than for small frequencies.

7. Incoherent feed-forward motif

The network is described by

$$\frac{dv}{dt} = \alpha s - \mu_v v + \eta_v(t), \quad (\text{A76})$$

$$\frac{dx}{dt} = \frac{vKv}{K+s} - \mu_x x + \eta_x(t). \quad (\text{A77})$$

The gain, noise, and gain-to-noise [Eqs. (9), (12), and (14) (main text)] in general terms are

$$g^2(\omega) = \frac{(J_{vs}J_{xv} + J_{vv}J_{xs})^2 + J_{xs}^2\omega^2}{(\omega^2 + \mu_v^2)(\omega^2 + \mu_x^2)}, \quad (\text{A78})$$

$$N(\omega) = \frac{J_{xv}^2\langle|\eta_v|^2\rangle + (\omega^2 + \mu_v^2)\langle|\eta_x|^2\rangle}{(\omega^2 + \mu_v^2)(\omega^2 + \mu_x^2)}, \quad (\text{A79})$$

$$\Theta(\omega) = \frac{(J_{vs}J_{xv} + J_{vv}J_{xs})^2 + J_{xs}^2\omega^2}{J_{xv}^2\langle|\eta_v|^2\rangle + (\omega^2 + \mu_v^2)\langle|\eta_x|^2\rangle}. \quad (\text{A80})$$

First, we study the influence of the topology. For the iff, four different topologies exist with respect to the regulation. It is known that different topologies have influence on the noise characteristics. Here we show that Θ indeed depends on the specific topology, comparing both a different position for the negative regulation, and the type of integration of the two pathways at the X component.

Instead of negatively regulating X by V or S, we can also negatively regulate V by S, such that we have

$$\frac{dv}{dt} = \frac{\alpha K_V}{K_V + x} - \mu_v v + \eta_v(t), \quad (\text{A81})$$

$$\frac{dx}{dt} = \beta s + \gamma v - \mu_x x + \eta_x(t), \quad (\text{A82})$$

where we have assumed the signals combine at X following OR strategy. In the previous section, we equalized production and studied the difference in the coupling parameters J_{ij} . Here, instead of equalizing production, we equalize the Jacobian coefficients (as a result the gain is equal). For these coupling parameters, we compare the production of the components (and thus the noise terms). We refer to the topology with superscript V for negative regulation of V and superscript X for negative regulation of X:

$$|J_{vs}^V| = J_{vs}^X \rightarrow \alpha^V \frac{K_V}{(K_V + \langle s \rangle)^2} = \alpha^X, \quad (\text{A83})$$

$$J_{xs}^V = J_{xs}^X \rightarrow \beta^V = v^X \frac{K \langle v \rangle}{(K + \langle s \rangle)^2}, \quad (\text{A84})$$

$$J_{xv}^V = |J_{xv}^X| \rightarrow \gamma^V = v^X \frac{K}{K + \langle s \rangle}. \quad (\text{A85})$$

Following these equalities, we write for the noise terms

$$\eta_V^V = 2\alpha^V \frac{K_V}{K_V + S} = 2(K_V + \langle s \rangle)\alpha^X > \eta_V^X, \quad (\text{A86})$$

$$\eta_X^V = 2(\beta \langle s \rangle + \gamma \langle v \rangle) = 2v^X \frac{K \langle v \rangle}{K + \langle s \rangle} \left(1 + \frac{\langle s \rangle}{K + \langle s \rangle}\right) > \eta_X^X. \quad (\text{A87})$$

For equal coupling constants the noise terms (or equivalently the production terms) in the iff motif are larger for negative regulation on V than on X.

Second, we discuss the effect of linearization. It is interesting to note that the Hill-functional prescribing the effect of the negative interaction has a saturated regime. Indeed, rewriting the Hill function [Eq. (23b)] as $1/(s/K + 1)$, we observe that if $s \gg K$ or $s \ll K$ the negative interaction is saturated. One could expect that for fluctuations that drive the system from the unsaturated regime to the saturated regime the differences between the linearized and nonlinearized system are large. To study this in detail, we focus on Fig. 4 and Eq. (23) (main text) and we discuss all three regimes, $K \gg \langle S \rangle$, $K \approx \langle S \rangle$, and $K \ll \langle S \rangle$.

If $K \gg \langle S \rangle$, we can approximate the Hill function by a linear dependence of X on V, so no effect should be observed. Similarly, for $K \approx \langle S \rangle$, no saturation effect is expected. The Hill function saturates if due to fluctuations $s \pm \delta s \gg K$ or $s \pm \delta s \ll K$. However, the size of the fluctuations in s are on the same order of $\sqrt{\langle S \rangle}$, since S is driven by a birth-death process. The fluctuations therefore do not drive the system to the saturated regime, as long as $\langle S \rangle > \mathcal{O}(10)$ and the linearization should give reasonable results. The last scenario is the one in which $K \ll \langle S \rangle$, indicating that the negative regulation is maximal. In this regime, the response can only track the variations in s with very small amplitude, since due to the negative regulation, the fluctuations are greatly suppressed. This effect can also be observed in the linearized system, since for this small K value, the gain (the amplitude of the transmitted fluctuations) is considerably smaller than for systems with larger K . The amplitude of the fluctuations is not large enough to drive the system out of the completely suppressed state, and as a result the linearization is a good approximation.

If the pathways combine following AND logic we have

$$\frac{dx}{dt} = \gamma v s - \mu_x x + \eta_x(t). \quad (\text{A88})$$

Now, we equalize production terms

$$p_V^V = \alpha^V \frac{K_V}{K_V + \langle s \rangle} \quad p_V^X = \alpha^X \langle s \rangle, \quad (\text{A89})$$

$$p_X^V = \gamma \langle v \rangle \langle s \rangle \quad p_X^X = v \frac{K \langle v \rangle}{K + \langle s \rangle}. \quad (\text{A90})$$

and obtain $\alpha^X = \alpha^V K_V [(K_V + \langle s \rangle) \langle s \rangle]^{-1}$ and $\gamma = v K [(K + \langle s \rangle) \langle s \rangle]^{-1}$. We then compare the Jacobian

terms

$$|J_{vs}^V| = \alpha^V \frac{K_V}{(K_V + \langle s \rangle)^2}; \quad J_{vs}^X = \alpha^X = \alpha^V \frac{K_V}{(K_V + \langle s \rangle) \langle s \rangle}, \quad (\text{A91})$$

$$J_{xs}^V = \gamma \langle v \rangle = v \frac{K \langle v \rangle}{(\langle s \rangle + K) \langle s \rangle}; \quad J_{xs}^X = v \frac{K \langle v \rangle}{(\langle s \rangle + K)^2}, \quad (\text{A92})$$

$$J_{xv}^V = \gamma \langle s \rangle = v \frac{K}{(\langle s \rangle + K)}; \quad J_{xv}^X = v \frac{K}{(\langle s \rangle + K)}, \quad (\text{A93})$$

and we observe that $J_{vs}^X > |J_{vs}^V|$, while $J_{xs}^V > J_{xs}^X$. Therefore, we conclude that at large frequencies the AND integration with negative regulation on V has a larger Θ than for negative regulation on X. For small frequencies Θ depends on the specific parameters.

Compared with a one-step cascade for equal production of x (so v is not constrained), we have for Θ of the incoherent motif

$$p_x^I = p_x^{\text{os}} \rightarrow \frac{vK \langle v \rangle}{K + \langle s \rangle} = k_x^{\text{os}} \langle s \rangle \rightarrow \frac{vK}{K + \langle s \rangle} = \frac{k_x^{\text{os}} \mu_v}{\alpha}, \quad (\text{A94})$$

leading to

$$\Theta(\omega) = \frac{\left(\frac{K k_x^{\text{os}} \mu_v}{K + \langle s \rangle}\right)^2 + \left(\frac{k_x^{\text{os}} \langle s \rangle}{K + \langle s \rangle}\right)^2 \omega^2}{\left(\frac{k_x^{\text{os}} \mu_v}{\alpha}\right)^2 \langle |\eta_v|^2 \rangle + (\omega^2 + \mu_v^2) \langle |\eta_x|^2 \rangle} < \frac{k_x^{\text{os}}}{\langle |\eta_x|^2 \rangle}, \quad (\text{A95})$$

such that the incoherent motif always has a smaller Θ than the one-step motif for equal production of x .

Compared to the two-step motif and equal total production we distinguish two scenarios. First, we assume equal production of v and obtain

$$p_x^{\text{iff}} = p_x^{\text{two}} \rightarrow \frac{vK}{K + \langle s \rangle} = \frac{k_x k_v}{\alpha}, \quad (\text{A96})$$

and for Θ of the iff cascade we obtain

$$\Theta(\omega) = \frac{(k_x k_v)^2 \left[\left(\frac{K}{K + \langle s \rangle}\right)^2 + \left(\frac{\langle s \rangle}{(K + \langle s \rangle) \mu_v}\right)^2 \omega^2 \right]}{(k_x)^2 \langle |\eta_v|^2 \rangle + (\omega^2 + \mu_v^2) \langle |\eta_x|^2 \rangle}. \quad (\text{A97})$$

Compared to Θ of a two-step cascade, for small ω Θ of the iff is smaller due to the negative interference in the iff, while for large ω this is larger than Θ of a two-step cascade due to the direct pathway.

TABLE VI. Summary of the results for the comparison between the iff and the two-step cascade.

Constraint	$\omega \rightarrow 0$	$\omega \rightarrow \infty$
Equal total production	$\Theta^{\text{iff}} < \Theta^{\text{ts}}$	$\Theta^{\text{iff}} > \Theta^{\text{ts}}$
Equal production of x	Parameter dependent	$\Theta^{\text{iff}} > \Theta^{\text{ts}}$

In the other scenario we only require $p_x^{\text{iff}} = p_x^{\text{two}}$ to obtain

$$\Theta(\omega) = \frac{(k_x k_v)^2 \left[\left(\frac{K}{K + \langle s \rangle}\right)^2 + \left(\frac{\langle s \rangle}{(K + \langle s \rangle) \mu_v}\right)^2 \omega^2 \right]}{2 \left(\frac{k_x k_v}{\alpha}\right)^2 \alpha \langle s \rangle + (\omega^2 + \mu_v^2) \langle |\eta_x|^2 \rangle}. \quad (\text{A98})$$

Due to the direct pathway the gain-to-noise ratio for the iff is larger for $\omega \rightarrow \infty$. For $\omega = 0$ we obtain

$$\Theta^{\text{iff}}(\omega = 0) > \Theta^{\text{ts}}(\omega = 0), \quad (\text{A99})$$

$$\frac{(k_x k_v)^2 \left(\frac{K}{K + \langle s \rangle}\right)^2}{2 \left(\frac{k_x k_v}{\alpha}\right)^2 \alpha \langle s \rangle + \mu_v^2 \langle |\eta_x|^2 \rangle} > \frac{(k_v k_x)^2}{2 (k_x)^2 k_v \langle s \rangle + m_v^2 \langle |\eta_x|^2 \rangle}, \quad (\text{A100})$$

$$\begin{aligned} & \left(\frac{K}{K + \langle s \rangle}\right)^2 \frac{1}{2 \left(\frac{1}{\alpha}\right) \langle s \rangle + \frac{\mu_v^2}{(k_x k_v)^2} \langle |\eta_x|^2 \rangle} \\ & > \frac{1}{2 \frac{1}{k_v} \langle s \rangle + \frac{\mu_v^2}{(k_x k_v)^2} \langle |\eta_x|^2 \rangle}, \end{aligned} \quad (\text{A101})$$

where the above inequality is valid in the following two scenarios. (1) If $k_v < \alpha$, the iff motif has a larger Θ , since the signal is transmitted to v with larger gain than in the two-step cascade. (2) If $K \gg \langle s \rangle$, the negative feedback is greatly suppressed and we observe a general two-step cascade instead of an iff motif. If the first condition is not satisfied, the two-step simple cascade has a larger Θ for $\omega = 0$. The second condition is required if the difference in the gain due to the first condition is not large enough (see Table VI).

8. Multimerization

The network for multimerization is described by

$$\frac{dv}{dt} = \alpha s - \mu_v v + \eta_v(t), \quad (\text{A102})$$

$$\frac{dx}{dt} = \gamma v^n - \mu_x x + \eta_x(t), \quad (\text{A103})$$

where n is the number of proteins that jointly activate X. The gain, noise, and gain-to-noise are

$$g^2(\omega) = \frac{n^2 k_s^{2(n-1)} \alpha^{2n} \gamma_v^2}{\mu_s^{2(n-1)} \mu_v^{2(n-1)} (\omega^2 + \mu_v^2) (\omega^2 + \mu_x^2)}, \quad (\text{A104})$$

$$N(\omega) = \frac{n^2 k_s^{2(n-1)} \alpha^{2(n-1)} \gamma_v^2 \langle |\eta_v|^2 \rangle + \mu_s^{2(n-1)} \mu_v^{2(n-1)} (\omega^2 + \mu_v^2) \langle |\eta_x|^2 \rangle}{\mu_s^{2(n-1)} \mu_v^{2(n-1)} (\omega^2 + \mu_v^2) (\omega^2 + \mu_x^2)}, \quad (\text{A105})$$

$$\Theta(\omega) = \frac{n^2 k_s^{2(n-1)} \alpha^{2n} \gamma_v^2}{n^2 k_s^{2(n-1)} \alpha^{2(n-1)} \gamma_v^2 \langle |\eta_v|^2 \rangle + \mu_s^{2(n-1)} \mu_v^{2(n-1)} (\omega^2 + \mu_v^2) \langle |\eta_x|^2 \rangle}. \quad (\text{A106})$$

For equal production of x (v is by construction in all cascades equal), we have

$$\gamma_{v,n} = \gamma_{v,n=1} \frac{\mu_s^{n-1} \mu_v^{n-1}}{k_s^{n-1} \alpha^{n-1}}, \quad (\text{A107})$$

which after we substitute this in Eq. (A106), results in Eq. (30) (main text).

The diamond motif is described by

$$\begin{aligned} \frac{du}{dt} &= \alpha s(t) - \mu_u u(t) + \eta_u(t), & \frac{dw}{dt} &= \beta s(t) - \mu_w w(t) + \eta_w(t), \\ \frac{dx}{dt} &= \gamma_u u^2(t) + \gamma_w w^2(t) + 2\gamma_{uw} u(t)w(t) - \mu_x x(t) + \eta_x(t). \end{aligned} \quad (\text{A108})$$

For $\mu_w = \mu_u = \mu_v$, $\alpha = C_I \beta$, and $\gamma_u = \gamma_w = C_{II} \gamma_w$, where we have introduced the coefficients C_I and C_{II} to study the form of the gain, noise, and Θ for differences between the intermediates **U** and **W** in a general context, the expressions for the gain, noise, and gain-to-noise are

$$g^2(\omega) = \frac{4k_s^2 \beta^4 (1 + 2C_I + C_I^2 C_{II})^2 \gamma_u^2}{\mu_s^2 \mu_w^2 H(\mu_w, \omega) H(\mu_x, \omega)}, \quad (\text{A109})$$

$$N(\omega) = \frac{4k_s^2 \beta^2 \gamma_u^2 (A^2 \langle |\eta_u|^2 \rangle + (1 + C_I)^2 \langle |\eta_w|^2 \rangle) + \mu_s^2 \mu_w^2 H(\mu_w, \omega) \langle |\eta_x|^2 \rangle}{\mu_s^2 \mu_w^2 H(\mu_w, \omega) H(\mu_x, \omega)}, \quad (\text{A110})$$

$$\Theta(\omega) = \frac{4k_s^2 \beta^4 (1 + 2C_I + C_I^2 C_{II})^2 \gamma_u^2}{4k_s^2 \beta^2 \gamma_u^2 (A^2 \langle |\eta_u|^2 \rangle + (1 + C_I)^2 \langle |\eta_w|^2 \rangle) + \mu_s^2 \mu_w^2 H(\mu_w, \omega) \langle |\eta_x|^2 \rangle}, \quad (\text{A111})$$

where $A = (1 + C_I C_{II})$ and $H(x, y) = x^2 + y^2$.

We take $C_{II} = 1$ and compare this cascade to a cascade with only homodimerization of the component **U** ($\beta = \gamma_w = \gamma_{uw} = 0$). For readability, we refer to this cascade as if it has an intermediate component **V** and, if required, subscripts denote **U**, **V**, and **W**. Equal production at the intermediate level $p_V = p_U + p_W$ and at the level of **X** ($p_X^{\text{dm}} = p_X^{\text{homo}}$) gives

$$(1 + C_I) \beta = \alpha_v, \quad (\text{A112})$$

$$\gamma_{uw} = \gamma_v \frac{\langle v \rangle^2}{U^2 + 2\langle u \rangle \langle w \rangle + \langle w \rangle^2} = \gamma_v \frac{(\langle u \rangle + \langle w \rangle)^2}{U^2 + 2\langle u \rangle \langle w \rangle + \langle W \rangle^2} = \gamma_v \frac{(\alpha_u + \beta_w)^2}{\alpha_u^2 + 2\alpha_u \beta_w + \beta_w^2} = \gamma_v. \quad (\text{A113})$$

Inserting these equalities in Eq. (A109), we observe that the gain is equal to Eq. (A104) for $n = 2$,

$$g^2(\omega) = \frac{4k_s^2 \alpha_v^4 \gamma_v^2}{\mu_s^2 \mu_v^2 (\omega^2 + \mu_v^2) (\omega^2 + \mu_x^2)}, \quad (\text{A114})$$

$$= \frac{4k_s^2 (1 + C_I)^4 \beta_w^4 \left(\frac{\alpha_u^2 + 2\alpha_u \beta_w + \beta_w^2}{(\alpha_u + \beta_w)^2} \right)^2 \gamma_{uw}^2}{\mu_s^2 \mu_v^2 (\omega^2 + \mu_v^2) (\omega^2 + \mu_x^2)}, \quad (\text{A115})$$

$$= \frac{4k_s^2 (1 + C_I)^4 \beta_w^4 \gamma_{uw}^2}{\mu_s^2 \mu_v^2 (\omega^2 + \mu_v^2) (\omega^2 + \mu_x^2)}. \quad (\text{A116})$$

and the noise is

$$N(\omega) = \frac{4k_s^2 (1 + C_I)^2 \beta_w^2 \gamma_v^2 (\langle |\eta_u|^2 \rangle + \langle |\eta_w|^2 \rangle) + \mu_s^2 \mu_v^2 (\omega^2 + \mu_v^2) \langle |\eta_x|^2 \rangle}{\mu_s^2 \mu_v^2 (\omega^2 + \mu_v^2) (\omega^2 + \mu_x^2)}, \quad (\text{A117})$$

where Eq. (A117) is exactly Eq. (A110). Therefore, unequal production of u and w is not important, as long as the sum is constrained (and all γ 's are equal).

Next we compare to a two-step cascade. We first assume the diamond motif to be completely symmetric for the intermediate components **U**, **W**, $\alpha = \beta$, $\mu_u = \mu_w$, and $\gamma_u = \gamma_w = \gamma_{uw}$.

We constrain the individual production of each step in the cascade:

$$p_v^{\text{two}} = p_u^{\text{dm}} + p_w^{\text{dm}} \rightarrow k_v = 2\alpha, \quad (\text{A118})$$

$$p_x^{\text{two}} = p_x^{\text{dm}} \rightarrow k_x k_v = 4\gamma \frac{\alpha^2 \langle s \rangle}{\mu_v}. \quad (\text{A119})$$

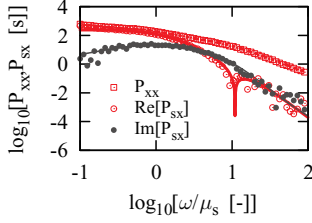


FIG. 8. (Color online) The results (symbols) of the Gillespie simulations for the coff. Since this motif is linear, we do not expect any deviations between the numerical simulations and the analytical results. We only show the result for one set of parameters: $k_s = 100$, $\alpha = 11$, $\beta = 1$, $\gamma = 10.1$, $\mu_v = 10$, $\mu_x = 10$. μ_s sets the time scale.

In the limit $\omega \rightarrow \infty$ the diamond motif has a larger Θ independent of any kinetic rates, since

$$\frac{\Theta^{\text{dm}}}{\Theta^{\text{two}}} = \frac{4\mu_v^2}{\mu_w^2} = 4. \quad (\text{A120})$$

In the opposite limit, $\omega \rightarrow 0$, we have

$$\frac{\Theta^{\text{dm}}}{\Theta^{\text{two}}} = \frac{(8\langle s \rangle \alpha \gamma_{uw} + 4\mu_w^2)}{(8\langle s \rangle \alpha \gamma_{uw} + \langle s \rangle \mu_w^2)} > 1. \quad (\text{A121})$$

Thus, the symmetric diamond motif has a larger Θ for $\omega \rightarrow 0$ and $\omega \rightarrow \infty$. For equal $\mu_u = m_v$, Θ of the diamond motif is larger for all frequencies.

The above case is for $\alpha = \beta$. Here we use $\alpha = C_I \beta$, but $\mu_v = \mu_w = m_v$. The ratio of Θ 's for $\omega \rightarrow \infty$ is simply 4 [Eq. (A120)]. For the limit $\omega \rightarrow 0$ we have

$$\frac{\Theta^{\text{dm}}}{\Theta^{\text{two}}} = \frac{4((1 + C_I) \langle s \rangle \alpha \gamma_{uw} + \mu_w^2)}{4(1 + C_I) \langle s \rangle \alpha \gamma_{uw} + \mu_w^2} > 1. \quad (\text{A122})$$

We do observe that the more asymmetric the diamond motif becomes, the ratio of Θ 's decreases. The performance of the diamond motif then becomes more similar to the two-step cascade.

More interesting is to have different degradation rates $\mu_v = C_{III} \mu_w$, where C_{III} is an arbitrary constant. The equal production constraints then result in the following two expressions:

$$\begin{aligned} p_v^{\text{two}} &= p_v^{\text{dm}} + p_w^{\text{dm}} \rightarrow k_v = 2\alpha, \\ p_x^{\text{two}} &= p_x^{\text{dm}} \rightarrow k_x k_v = m_v \gamma \frac{\alpha^2 \langle s \rangle}{C_{III} \mu_w^2}. \end{aligned} \quad (\text{A123})$$

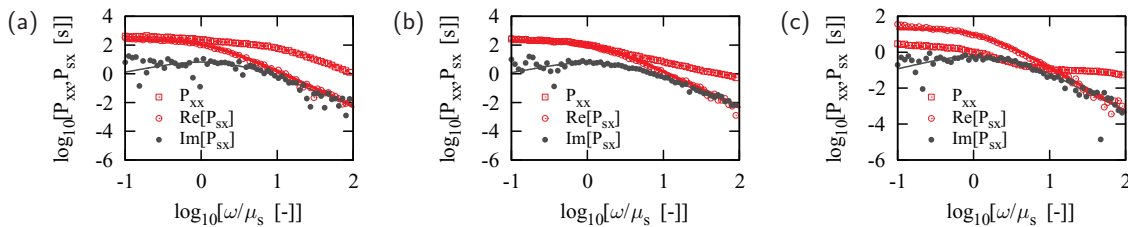


FIG. 9. (Color online) The results (symbols) of the Gillespie simulations for the coff. Although this motif is nonlinear, the results of the simulations are in good agreement with those of the linear analysis (lines). In P_{xx} and $\text{Re}[P_{sx}]$ the symbols of the simulations results are on top of the lines in the analytical results. Parameters: (a) as in Fig. 3(c), red solid line; (b) Fig. 3(c), red dashed line; (c) Fig. 3(c), red dotted line.

For high frequencies the ratio of the Θ of the two-step and diamond motif is

$$\frac{\Theta^{\text{dm}}}{\Theta^{\text{two}}} = 16 \frac{C_{III}^2 \mu_w^2}{(1 + C_{III})^2 m_v^2}, \quad (\text{A124})$$

which does not lead to a unique conclusion whether which of the two is larger. For $\omega = 0$, we have

$$\frac{\Theta^{\text{dm}}}{\Theta^{\text{two}}} = \frac{2(1 + C_{III})^2 \langle s \rangle \alpha \gamma_{uw} + 2C_{III}^2 \mu_w^2}{4(1 + C_{III}^2) \langle s \rangle \alpha \gamma_{uw} + C_{III}^2 \mu_w^2}. \quad (\text{A125})$$

Again, the ratio of Θ depends on the kinetic rates.

9. Influence of the phase

In this section we provide more detail on the effect of the phase difference on the gain [Eq. (9), main text]. For this reason we rewrite Eq. (9) in the main text, while keeping explicitly the phase dependence in the Fourier Transforms of $\hat{s}(t)$ and $\hat{v}(t)$. To make the analysis intuitive, we take an example system: the cff with OR logic [Eq. (17), main text, and Eqs. (A39) and (A40)]. In the general form of the Fourier transform, the expressions for \hat{s} and \hat{v} are

$$\hat{s}(\omega) = \frac{\Gamma}{i\omega + \mu_s} = A_s(\omega) e^{i\phi_s(\omega)}, \quad (\text{A126})$$

$$\hat{v}_d(\omega) = \frac{\alpha \hat{s}(\omega)}{i\omega + \mu_v} = A_v(\omega) e^{i\phi_v(\omega)}, \quad (\text{A127})$$

where the subscript d indicates we only take the deterministic part, since the noise term η_v does not appear in the gain $g^2(\omega)$, but only in the noise $N(\omega)$ (note that Γ is the stochastic driving force which generates an ensemble of trajectories). The coefficients are defined as

$$A_s^2(\omega) = \frac{\langle |\Gamma|^2 \rangle}{\mu_s^2 + \omega^2}, \quad (\text{A128})$$

$$A_{v_d}^2(\omega) = \frac{\alpha^2 \langle |\Gamma|^2 \rangle}{(\mu_s^2 + \omega^2)(\omega^2 + \mu_v^2)}, \quad (\text{A129})$$

$$\phi_s(\omega) = -\arctan \left[\frac{\omega}{\mu_s} \right], \quad (\text{A130})$$

$$\phi_v(\omega) = \arctan \left[\frac{(\mu_s + \mu_v)\omega}{\omega^2 - \mu_v \mu_s} \right]. \quad (\text{A131})$$

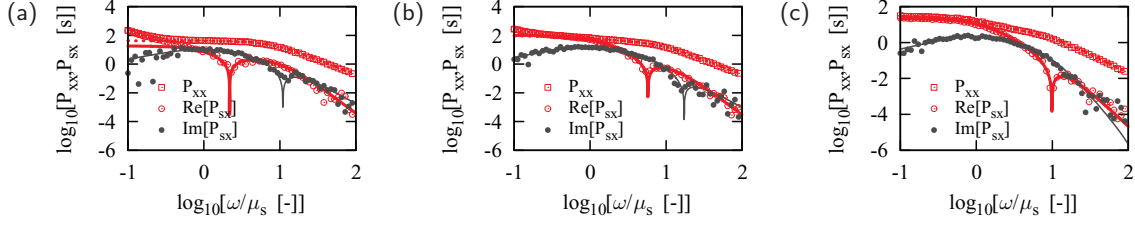


FIG. 10. (Color online) The results (symbols) of the Gillespie simulations for the iff motif. Although this motif is nonlinear, the results of the simulations are in good agreement with those of the linear analysis (lines). In P_{xx} and $\text{Re}[P_{sx}]$ the symbols of the simulations results are on top of the lines in the analytical results. Parameters: (a) as in Fig. 4(c), red solid line; (b) Fig. 4(c), red dashed line; (c) Fig. 4(c), red dotted line.

The Fourier transform of $\tilde{x}(t)$ is (again ignoring η_x)

$$\hat{x}_d(\omega) = \frac{\beta \hat{s}(\omega) + \gamma \hat{v}(\omega)}{i\omega + \mu_x}. \quad (\text{A132})$$

Using Eqs. (A126) and (A127) we write the gain as

$$g^2(\omega) = \frac{\langle \hat{x}_d(\omega) \hat{x}_d(-\omega) \rangle}{P_{SS}(\omega)} = \frac{(\alpha\beta)^2 + \gamma^2(\mu_v^2 + \omega^2) + 2\alpha\beta\gamma\sqrt{\mu_v^2 + \omega^2} \cos[\phi_s(\omega) - \phi_v(\omega)]}{(\mu_v^2 + \omega^2)(\mu_x^2 + \omega^2)} \quad (\text{A133})$$

$$= \frac{(\alpha\beta)^2}{(\mu_v^2 + \omega^2)(\mu_x^2 + \omega^2)} + \frac{\gamma^2}{\mu_x^2 + \omega^2} + \underbrace{\frac{2\alpha\beta\gamma\sqrt{\mu_v^2 + \omega^2} \cos[\phi_s(\omega) - \phi_v(\omega)]}{(\mu_v^2 + \omega^2)(\mu_x^2 + \omega^2)}}_{\text{Coherence effect: } A_{\text{coh}}}. \quad (\text{A134})$$

Now, the last term directly involves the phase difference [compare with Eq. (9) in the main text]. Using the trigonometric identities for the arctan,

$$-\arctan[u] = \arctan[-u], \quad (\text{A135})$$

$$\arctan[u] + \arctan[v] = \arctan\left[\frac{u+v}{1-uv}\right], \quad (\text{A136})$$

$$\cos[\arctan[x]] = \frac{1}{\sqrt{1+x^2}}. \quad (\text{A137})$$

and the expressions in Eqs. (A130) and (A131), we can rewrite the third term in Eq. (A134) (A_{coh}) as a multiplication of two effects, phase-difference (A_{pd}) and time-averaging (A_{ta}):

$$A_{\text{coh}} = A_{\text{ta}} A_{\text{pd}} = \overbrace{\frac{2\alpha\beta\gamma\sqrt{\mu_v^2 + \omega^2}}{(\mu_v^2 + \omega^2)(\mu_x^2 + \omega^2)}}^{\text{time-averaging}} \overbrace{\cos[\Delta\phi(\omega)]}^{\text{phase difference}}. \quad (\text{A138})$$

Interestingly, we have obtained a mathematical expression for the effect of the phase difference on the coherent term. First the phase difference itself is

$$\begin{aligned} \Delta\phi = \phi_s - \phi_v &= \arctan\left[\frac{-\omega}{\mu_s}\right] + \arctan\left[\frac{-(\mu_s + \mu_v)\omega}{\omega^2 - \mu_v\mu_s}\right] \\ &= \arctan\left[\frac{\omega}{\mu_v}\right]. \end{aligned} \quad (\text{A139})$$

For $\omega \ll \mu_v$ the phase difference $\Delta\phi \sim 0$, while for $\omega = \mu_v$, $\Delta\phi = \pi/4$ and for $\omega \rightarrow \infty$, $\Delta\phi \rightarrow \pi/2$. Note that the phase difference, as expected, only depends on $\mu_v = -J_{vv}$. With the phase difference we obtain an expression for the amplitude A_{pd} as

$$A_{\text{pd}} = \cos[\Delta\phi] = \frac{\mu_v}{\sqrt{\mu_v^2 + \omega^2}}. \quad (\text{A140})$$

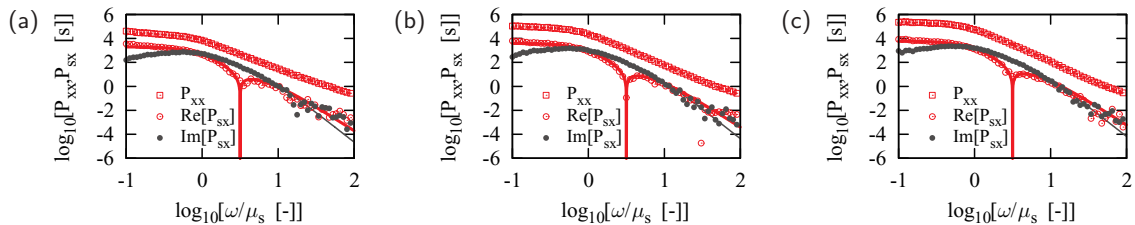


FIG. 11. (Color online) The results (symbols) of the Gillespie simulations for the dimerization process. Here the motif is nonlinear. We model the propensity function following a Hill function, $k(V) = k_{\text{max}} V^n / (V^n + K^n)$. Next, we take $K \gg V$, such that propensity is approximately $k(V) \approx k_{\text{max}} (V/K)^n$, as discussed in the main text. The results of the simulations are in good agreement with those of the linear analysis (lines). In P_{xx} and $\text{Re}[P_{sx}]$ the symbols of the simulations results are on top of the lines in the analytical results. Parameters: (a) two-step process ($n = 1$), (b) dimer process ($n = 2$), (c) trimer process ($n = 3$). Parameters are as in Fig. 5(b).

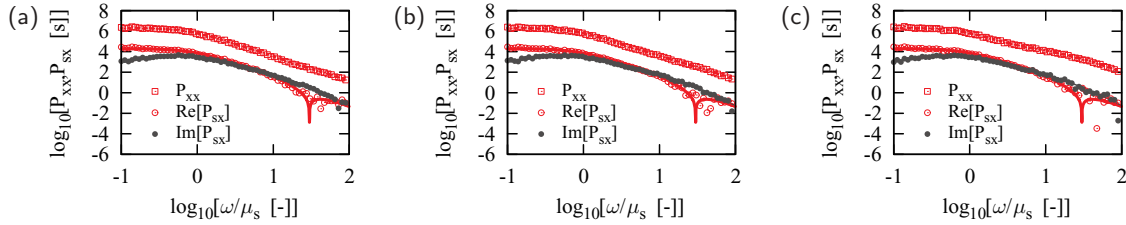


FIG. 12. (Color online) The results (symbols) of the Gillespie simulations for the heterodimerization motif, with $\gamma_u = \gamma_w = 0$. Although this motif is nonlinear, the results of the simulations are in good agreement with the results of the linear analysis (lines). In P_{xx} and $\text{Re}[P_{sx}]$ the symbols of the simulations results are on top of the lines in the analytical results. Parameters: (a) as in Fig. 5(c), red solid line; (b) Fig. 5(c), red dashed line; (c) Fig. 5(c), red dotted line.

For $\omega = 0$, $A_{pd} = 1$, since the two signals are in phase, but for an increase in the frequency of the signal $A_{pd} \rightarrow 0$, due to the increase in phase difference.

Next to the effect of the phase difference between the pathways, also the effect of time averaging is important for the amplitude of the coherence term. Indeed, A_{coh} is greatly reduced if the signal variations are much faster than the lifetime of the components in the cascade. The effective amplitude of the time-averaging A_{ta} , can mathematically be expressed as [see Eq. (A138)]

$$A_{ta} = \frac{2\alpha\beta\gamma\sqrt{\mu_v^2 + \omega^2}}{(\mu_v^2 + \omega^2)(\mu_x^2 + \omega^2)}. \quad (\text{A141})$$

We first note that the decrease in the amplitude due to the time averaging over the finite response time of V scales only with the inverse square root of the response time of V , $(\mu_v^2 + \omega^2)^{-1/2}$. Next, for signal oscillations that are much faster than the response time of V (μ_v^{-1}) or X (μ_x^{-1}), the amplitude $A_{ta} \rightarrow 0$, since then the individual components can not track the variations and start to time-average the variations, thereby losing the specific information in the high frequencies.

Combining A_{ta} and A_{pd} we obtain for A_{coh}

$$A_{coh} = A_{ta}A_{pd} = \frac{2\alpha\beta\gamma\mu_v}{(\mu_v^2 + \omega^2)(\mu_x^2 + \omega^2)}. \quad (\text{A142})$$

As required, Eq. (A142) is identical to the expression for the coherence in Eq. (9).

In Fig. 7 we show in some more detail the precise influence of the phase difference between two pathways that combine at a downstream component.

10. Numerical validation

The linearization used in the derivation can change the characteristics of the frequency response, since a linear(ized) system does not change the frequency of the transmitted signal. This may not be the case for a nonlinear system. In this section we show the comparison between our analytical results, following the linear noise approximation, for the power spectrum and the result from numerical simulations of the full nonlinear network. For the numerical simulations we use the Gillespie algorithm. The negative regulation as present in the iff motif, or the positive regulation as in the dimerization process are calculated through Hill-like interactions between the components. In the Gillespie simulation we calculated the propensities for every reaction using the coarse-grained Hill expressions for the propensities, such that



where r is

$$r = \frac{kK}{S + K}, \quad (\text{A144})$$

where here the actual copy number of S is used, and not $\langle s \rangle$, as in the linearized expressions [Eq. (A80)].

The power spectra are calculated using 2^{10} (1024) exponentially distributed frequencies from $\omega = 10^{-1}$ to $\omega = 10^2$ and averaged over 16 neighboring frequencies to obtain a single data point. In total we have 64 data points. Fourier transforms and power spectra are directly integrated during run time. We simulated a minimum of 27 blocks of 5000 s. The results are shown in Figs. 8–12.

[1] Wenzhe Ma, Ala Trusina, Hana El-Samad, Wendell A. Lim, and Chao Tang, *Cell* **138**, 760 (2009).
 [2] Leon O. Murphy, J. P. MacKeigan, and John Blenis, *Mol. Cell Biol.* **24**, 144 (2004).
 [3] Filipe Tostevin and Pieter Rein ten Wolde, *Phys. Rev. Lett.* **102**, 218101 (2009).
 [4] C. E. Shannon, *Bell Syst. Tech. J.* **27**, 379 (1948); **27**, 623 (1948).
 [5] Wiet Hendrik de Ronde, Filipe Tostevin, and Pieter Rein ten Wolde, *Phys. Rev. E* **82**, 031914 (2010).
 [6] Shai S. Shen-Orr, Ron Milo, Shmoolik Mangan, and Uri Alon, *Nat. Genet.* **31**, 64 (2002).

[7] Uri Alon, *An Introduction to Systems Biology: Design Principles of Biological Circuits*, 3rd ed. (CRC, Boca Raton, FL, 2007).
 [8] Shmoolik Mangan and Uri Alon, *Proc. Natl. Acad. Sci. USA* **100**, 11980 (2003).
 [9] Lea Goentoro, Oren Shoval, Marc W. Kirschner, and Uri Alon, *Mol. Cell* **36**, 894 (2009).
 [10] Mary J. Dunlop, Robert Sidney Cox, Joseph H. Levine, Richard M. Murray, and Michael B. Elowitz, *Nat. Genet.* **40**, 1493 (2008).
 [11] Vahid. Shahrezaei, Julien F. Ollivier, and Peter S. Swain, *Mol. Syst. Biol.* **4**, 196 (2008).

- [12] Bhaswar Ghosh, Rajesh Karmakar, and Indrani Bose, *Phys. Biol.* **2**, 36 (2005).
- [13] Aleksandra M. Walczak, Gašper Tkačik, and William Bialek, *Phys. Rev. E* **81**, 041905 (2010).
- [14] Filipe Tostevin and Pieter Rein ten Wolde, *Phys. Rev. E* **81**, 061917 (2010).
- [15] Daniel T. Gillespie, *J. Chem. Phys.* **113**, 297 (2000).
- [16] Johan Paulsson, *Nature (London)* **427**, 415 (2004).
- [17] Sorin Tănase-Nicola, Patrick B. Warren, and Pieter Rein ten Wolde, *Phys. Rev. Lett.* **97**, 068102 (2006).
- [18] Matthew Scott, Terence Hwa, and Brian P. Ingalls, *Proc. Natl. Acad. Sci. USA* **104**, 7402 (2007).
- [19] Etay Ziv, Ilya Nemenman, and Chris H. Wiggins, *PLoS one* **2**, e1077 (2007).
- [20] Gašper Tkačik, Curtis G Callan, and William Bialek, *Proc. Natl. Acad. Sci. USA* **105**, 12265 (2008).
- [21] Gašper Tkačik, Aleksandra M. Walczak, and William Bialek, *Phys. Rev. E* **80**, 031920 (2009).
- [22] Johan Paulsson, *Phys. Life Rev.* **2**, 157 (2005).
- [23] Patrick B. Warren, Sorin Tănase-Nicola, and Pieter Rein ten Wolde, *J. Chem. Phys.* **125**, 144904 (2006).
- [24] Nico G. van Kampen, *Stochastic Processes in Physics and Chemistry*, 3rd ed. (Elsevier, North-Holland, Amsterdam, 2007).
- [25] Shiraz Kalir, Shmoolik Mangan, and Uri Alon, *Mol. Syst. Biol.* **1**, 2005.0006 (2005).
- [26] Mark Kittisopikul and Gürol M. Süel, *Proc. Natl. Acad. Sci. USA* **107**, 13300 (2010).
- [27] Tolga Çağatay, Marc Turcotte, Michael B. Elowitz, Jordi García-Ojalvo, Gürol M. Süel, and Jordi Garcia-Ojalvo, *Cell* **139**, 512 (2009).
- [28] Tamar Friedlander and Naama Brenner, *Proc. Natl. Acad. Sci. USA* **106**, 22558 (2009).
- [29] J. M. G. Vilar and L. Saiz, *Nucleic Acids Res.* **39**, 6854 (2011).
- [30] Mariko Hatakeyama, Shuhei Kimura, Takashi Naka, Takuji Kawasaki, Noriko Yumoto, Mio Ichikawa, Jae-Hoon Kim, Kazuki Saito, Mihoro Saeki, Mikako Shirouzu, Shigeyuki Yokoyama, and Akihiko Konagaya, *Biochem. J.* **373**, 451 (2003).
- [31] Shalev Itzkovitz, Reuven Levitt, Nadav Kashtan, Ron Milo, Michael Itzkovitz, and Uri Alon, *Phys. Rev. E* **71**, 016127 (2005).
- [32] Nadav Kashtan and Uri Alon, *Proc. Natl. Acad. Sci. USA* **102**, 13773 (2005).
- [33] Arnold J. Levine, Wenwei Hu, Zhaohui Feng, and German Gil, *Ann. NY Acad. Sci.* **1115**, 32 (2007).
- [34] Axel Cournac and Jacques-Alexandre Sepulchre, *BMC Syst. Biol.* **3**, 29 (2009).