

# Molecular Filters for Noise Reduction

L. Laurenti\* and M. Kwiatkowska  
*University of Oxford*

A. Csikasz-Nagy  
*King's College London & Pázmány Péter Catholic University*

L. Cardelli  
*Microsoft Research Cambridge & University of Oxford*

Living systems are inherently stochastic and operate in a noisy environment, yet despite all these uncertainties they perform their functions in a surprisingly reliable way. The biochemical mechanisms used by natural systems to tolerate and control noise are still not fully understood, and this issue also limits our capacity to engineer reliable quantitative synthetic biological circuits. We study how representative models of biochemical systems propagate and attenuate noise accounting for intrinsic as well as extrinsic noise. We investigate three molecular noise filtering mechanisms, study their noise reduction capabilities and limitations, and show that non-linear dynamics, such as complex formation, are necessary for efficient noise reduction. We further suggest that the derived molecular filters are widespread in gene expression and regulation and, particularly, that microRNAs can serve as such noise filters. Our results provide new insight into how biochemical networks control noise and could be useful to build robust synthetic circuits.

Biochemical processes such as gene expression are inherently stochastic and must control noise, which presents itself as stochastic fluctuations. These fluctuations can be extrinsic, arising from interactions occurring with other processes in the environment, or intrinsic, resulting from the random timing of the reactions themselves [4, 13, 15, 19, 25, 33]. Molecular processes transform noisy input signals from the environment into output signals through a number of stages, with signals represented by chemical species and each stage implemented by a molecular reaction network. Since each stage can compound the noise, in order to obtain a reliable final output natural systems must integrate mechanisms that, directly or indirectly, reduce noise or otherwise confine it. Examples include signaling cascades, which have been shown to reduce extrinsic fluctuations [46], and the role microRNA plays in attenuating noise of protein expression [38].

Noise reduction has been studied extensively in electronics [39] and certain noise filtering principles have been successfully applied to molecular systems [9, 37, 49]. For instance, negative feedback and feed-forward loops have been shown to reduce noise [7, 24], and fundamental limits for noise suppression of feedback loops have been derived using techniques from control and information theory [22]. Analogues of mechanisms from signal processing such as *low-pass filters*, which transduce low-frequency signals while attenuating high-frequency signals, have also been observed in biochemical systems [34]. However, these are not always true low-pass filters: a low-pass filter must preserve the low frequencies, and for example not amplify them while attenuating the high frequencies. Moreover, the classical theory of filters in

electronics does not account for intrinsic noise. Therefore, difficulties arise when implementing such filters in terms of stochastic biochemical networks, because it is not clear how intrinsic noise may affect their noise reduction performance [4, 13, 19, 33].

Inspired by the concept of low-pass filters, we study noise reduction capabilities of *molecular filters* in a stochastic setting. We propose three fundamental filter modules and their implementation as stochastic chemical reaction networks. We account for intrinsic as well as extrinsic noise, and derive principles holding for filters when embedded in a general biochemical network, which may include multiple feed-forward and feedback loops. First, we consider *linear* filters (Figure 1A), implemented by means of at most uni-molecular reactions, and we show that, in case of positive correlation between the elements of the network, they are limited by Poisson levels, that is, the variance of the output signal is lower bounded by its mean. We then show how the presence of feedback loops may improve the performance of linear filters below Poisson levels. Specific models of linear filters have already been studied in the context of closed models of gene expression [27, 47]. Here, while still providing exact analysis based on the solution of the Chemical Master Equation, we generalize the analysis in the referenced papers taking also into account the more general and common case where the dynamics of the other components of the networks are left unspecified, and possibly involving non-linearity and feedback loops. The limitations in the noise suppression capabilities of linear filters motivate us to consider *non-linear* filters. We introduce a *non-linear* filter mechanism given by the co-expression of two species that then bind together, called the *annihilation module* (Figure 1B), which we demonstrate is able to reduce the noise to below Poisson levels. We then propose the *annihilation filter* (Figure 1C), which combines the prop-

---

\* Corresponding author.

erties of the linear filter and annihilation module. We show how the annihilation filter in particular can greatly reduce molecular noise. We observe that co-expression and non-linear degradation are key requirements for such noise reduction. This is important to be stressed as theoretical analysis is usually restricted to linear degradation [22]. Using analytical, numerical and stochastic simulation techniques we demonstrate how the different filters improve robustness of the systems in which they are embedded.

Finally, we discuss that that the molecular filters we derive are prevalent in gene expression. For instance, the linear filter, implemented by simply producing and degrading a species at a slow enough rate is a low-pass filter mechanism widely deployed in gene expression to increase robustness, both at the transcription and translation level [47]. Moreover, we find that the annihilation module and annihilation filter are sound models of microRNA regulated gene expression, in the case of correlated expression of microRNAs with the target gene. This supports the hypothesis that microRNA may play a role in increasing robustness and precision of gene expression. We stress how the focus of this paper is not to offer new models of particular molecular processes, but to identify fundamental and general mechanisms that, at the molecular level, can reduce noise, and to understand their properties and limitations. Thus, our results provide new insight into how biochemical networks control noise.

## MATERIALS AND METHODS

Detailed information about the modelling framework and mathematical derivations can be found in the Supplementary Material. CRNs and LNA simulations have been performed using the Microsoft Visual GEC tool [30]. Details on the code can be found in [SI Section G](#).

## RESULTS

We first investigate filters composed from linear reactions and show their limitations; then we discuss non-linear filters, showing how non-linearities can improve the performances in terms of noise reduction. The techniques used are detailed in Materials and Methods.

### Linear filters and their limits in noise suppression

To model biological regulatory networks we focus on the underlying molecular interactions represented as Chemical Reaction Networks (CRNs). A CRN is a set of biochemical species that interact according to the reaction laws. Input and output signals are modeled as biochemical species. In this paper, species  $A$  will always represent the input signal/species. We assume  $A$  is a

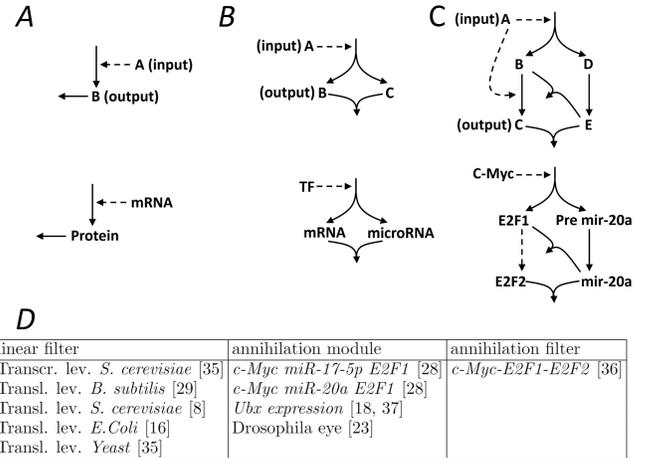
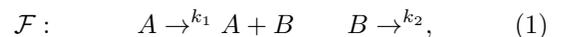


FIG. 1. Molecular filters (Top row) along with plausible biological examples where such components can be found (Middle row). We stress that for many of these examples it has yet to be verified experimentally that those systems behave as our filters. We simply present them here to facilitate understanding (see Discussion). Dashed arrows represent catalytic production/activation effects and continuous arrows molecular transitions. (A) First-order low-pass filter implemented by means of slow production and degradation reactions (linear filter, Eqn (1)). These networks are present both at the translation and transcription level to increase robustness of gene expression. (B) annihilation module (Eqn (7)) given by co-expression of two species that are then degraded together. These networks may be a model for a class of microRNA regulated gene expression, where mRNA and target protein are co-expressed. (C) annihilation filter (Eqn (10)) based on the annihilation module. Such a system may be a suitable model for the *c-Myc*—*E2F1*—*hsa-miR-20a* circuit with its extension to *E2F2* [36] under the assumption that *E2F1* activates *E2F2*, except for the feed-forward loop between  $A$  and  $C$  which has no influence on noise reduction. We stress that this assumption needs further experiments to be confirmed. (D) Biological systems where the filters may play a role in reducing noise.

noisy input with the noise identified by its Fano factor (ratio between variance and mean). Thus, a molecular filter is a CRN with input  $A$  and whose output has a reduced Fano factor compared to  $A$ , but still maintains certain features of its time evolution. In this paper, we focus on filters that maintain the same long term behaviour of  $A$ , while reducing its Fano factor.

The CRN  $\mathcal{F}$  (Eqn (1)), which we call a *linear filter*, is composed of a production and a degradation reaction with output species  $B$



where  $k_1, k_2 \in \mathbb{R}_{>0}$  are the *rate parameters*. We consider the general scenario where the linear filter  $\mathcal{F}$  (Eqn (1)) is embedded within an arbitrarily complex, possibly non-linear, reaction network with the only constraint that the output species  $B$  is changed only by the reactions in  $\mathcal{F}$ .

We do allow  $B$  to act as a catalyst in arbitrarily many reactions, and  $A$  to interact with the larger network with no constraint. This scenario is very general and includes the case where  $A$  is a function of  $B$  with feedback loops.

#### Classical frequency analysis

The transfer function of  $\mathcal{F}$  is obtained by applying the Fourier transform to the mass action rate equation corresponding to  $B$  [17, 43]. As  $B$  is changed only by reaction in  $\mathcal{F}$ , we obtain

$$\frac{d\Phi_B(t)}{dt} = k_1\Phi_A(t) - k_2\Phi_B(t),$$

where  $\Phi_A$  and  $\Phi_B$  are the deterministic signals modeling the time evolution of  $A$  and  $B$ . In the frequency domain, we get

$$i\omega\hat{\Phi}_B(\omega) = k_1\hat{\Phi}_A(\omega) - k_2\hat{\Phi}_B(\omega)$$

where  $\omega$  is the angular frequency, and  $\hat{\Phi}_B(\omega), \hat{\Phi}_A(\omega)$  are the Fourier transforms of signals  $\Phi_B$  and  $\Phi_A$ . For  $k_1 = k_2$ , we obtain

$$\frac{\hat{\Phi}_B(\omega)}{\hat{\Phi}_A(\omega)} = \frac{1}{1 + \frac{i\omega}{k_1}}, \quad (2)$$

Eqn (2) is the transfer function of a *first-order low-pass filter* (see [SI Section A](#)). This network attenuates frequencies higher than the cutoff frequency by introducing a delay and integrating the fast dynamics. From Eqn (2) the cutoff frequency is exactly  $\bar{\omega} = k_1$ . This means that, the higher the value of  $k_1$ , the less noise is filtered out, but the faster  $B$  tracks  $A$ .

#### Stochastic analysis

The classical frequency analysis of  $\mathcal{F}$ , based on the Fourier transform of the rate equations, does not take into account the intrinsic noise introduced by the reactions firing in  $\mathcal{F}$ : it only considers the extrinsic noise modeled as fast fluctuations of the input. However, the intrinsic noise cannot be simply neglected, as it may drive the behavior of biological systems [34]. This is the case in gene expression, where low molecular counts are often involved, and deterministic modeling is generally unsatisfactory [15, 26]. To resolve this, we need to consider the continuous time Markov chain (CTMC) induced by  $\mathcal{F}$ , whose transient evolution is described by the Chemical Master Equation (CME) [48]. The evolution of the moments of the CME can be described as a (possibly infinite) set of ODEs, so called *moment equations* [14, 42] ([SI Section B](#)). We quantify the noise with the *Fano factor* (ratio between expectation and variance) and, under the general scenario described in the previous section,

using moment equations the transient evolution of the expectation of  $B$  at time  $t$  can be computed exactly as

$$\frac{dE[B(t)]}{dt} = k_1 \cdot E[A(t)] - k_2 \cdot E[B(t)],$$

where  $E[A(t)]$  is the expectation of  $A$  at time  $t$ . We call  $\lim_{t \rightarrow \infty} E[A(t)] = E[A]_\infty$  the steady-state solution of  $E[A]$ , and we assume it exists and is finite. The steady-state solution of  $B$ ,  $E[B]_\infty$ , can then be derived by solving  $\frac{dE[B(t)]}{dt} = 0$ , which results in

$$E[B]_\infty = \frac{k_1}{k_2} E[A]_\infty \quad (3)$$

Eqn (3) guarantees that the expected value of  $B$  always tracks the expectation of  $A$ , no matter which biochemical system is producing  $A$  and what happens in the rest of the system. Importantly, for  $V[A]_\infty$ , variance of  $A$  at steady state, we can derive the following exact relation

$$V[B]_\infty = E[B]_\infty + \frac{k_1}{k_2} Cov[A, B]_\infty, \quad (4)$$

where  $Cov[A, B]_\infty = E[A \cdot B]_\infty - E[A]_\infty E[B]_\infty$  is the covariance of  $A$  and  $B$  at steady state, with  $E[A \cdot B]_\infty$  expectation of  $A \cdot B$  at steady state. Full derivation of Eqn (3) and (4) is shown in [SI Section C](#). The idea is that, even though  $B$  can participate in other reactions as a catalyst and  $A$  may be a non-linear function of  $B$ , in Eqn (3) all the non-linearities disappear, while in Eqn (4), these are included in the term  $Cov[A, B]_\infty$ . Eqn (4) shows that, for any input signal, the filtered signal  $B$  has variance that is equal to its mean plus the covariance between  $A$  and  $B$ . Assuming  $A$  and  $B$  are non-negatively correlated, then we have  $E[A \cdot B]_\infty \geq E[A]_\infty E[B]_\infty$ . As a result, in case of non-negative correlation between  $A$  and  $B$ , for any  $k_1, k_2 \in \mathbb{R}_{\geq 0}$ , the following *lower bound* holds

$$F_B \geq 1, \quad (5)$$

where  $F_B$  is the Fano factor of  $B$  at steady state. The above lower bound has already been observed and studied in the context of specific closed models of gene expression for which mRNA and protein are positively correlated [27, 31, 47]. However, in the more general scenario we consider, we can observe that Eqn (5) holds only in the case of non-negative correlation of the species, meaning that a simple birth-death process of a downstream component cannot reduce the noise of an input signal below Poisson levels.

$A$  and  $B$  being non-negatively correlated is natural since  $A$  catalyzes the production of  $B$ . In fact, in [SI Section C](#) we show that, for a large class of systems,  $A$  and  $B$  are effectively positively correlated. However, in the following example, we show that a negative feedback loop between  $B$  and  $A$  may change the sign of their correlation, potentially leading to noise reduction to below

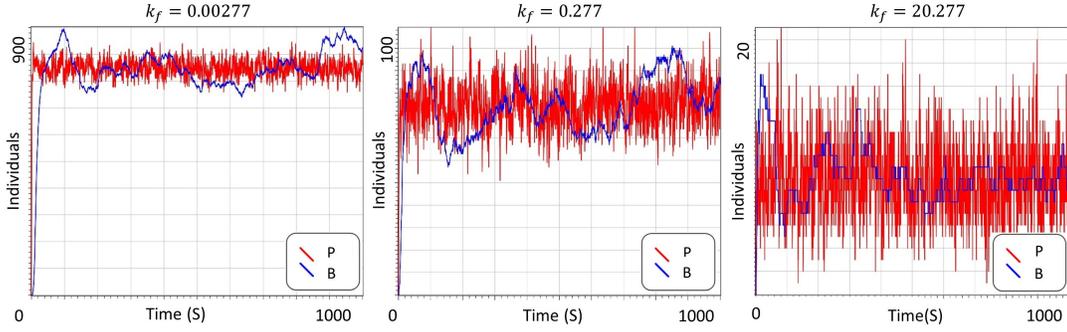
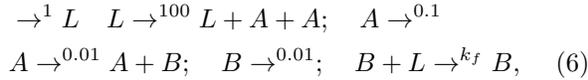


FIG. 2. For the CRN in Eqn (6) we plot a stochastic simulation of  $B$  for three different values of  $k_f$  and assuming initial condition of all the species are 0. In all the plots  $P$  is a signal with Poisson noise and such that  $E[P]_\infty = E[A]_\infty$ . When the feedback is weak ( $k_f = 0.00277$ ) the Fano factor of  $B$  is greater than one. For  $k_f = 0.277$  our analysis predicts a Fano factor for  $B$  of approximately 1. When the feedback is strong ( $k_f = 20.277$ ) the Fano factor of  $B$  is smaller than 1, but there is also a strong repression of  $E[B]_\infty$ .

Poisson levels. Thus, our analysis gives a further explanation of why negative feedback regulation in gene expression may be a widely selected mechanism to reduce noise and increase robustness [7] (see Discussion).

**Example 1** We consider the following CRN where there is a feedback between  $A$  and  $B$  and  $L$  is an auxiliary species.



where  $k_f > 0$  is a rate constant. That is, a feedback between  $A$  and  $B$  is present. The strength of the feedback can be controlled by changing the rate  $k_f$ .  $B$  can be thought as a protein that inhibits its expression. The above CRN meets the condition of validity of Eqns (3) and (4). Thus, Eqn (3) guarantees that for any possible initial condition and value of  $k_f$

$$E[B]_\infty = E[A]_\infty,$$

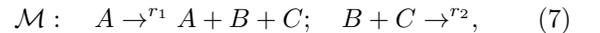
while Eqn (4) guarantees that  $F_B = E[A]_\infty + \text{Cov}[A, B]_\infty$ . Thus, to compute the Fano factor of  $B$  at steady state we need to estimate  $\text{Cov}[A, B]_\infty$ . The system is non-linear. As a consequence,  $\text{Cov}[A, B]_\infty$  cannot be computed exactly, but can be estimated using the LNA. We obtain that, for  $k_f > 0.277368$ ,  $\text{Cov}[A, B]_\infty < 0$ . Thus,  $F_B < 1$ . This shows how strong feedback can reduce the Fano factor of  $B$  to below Poisson levels. However, strong feedback means strong repression of the mean of  $B$  and  $A$  by Eqn (3). To confirm the mathematical analysis, in Figure 2, for different values of  $k_f$ , we plot a single stochastic simulation of  $B$  compared with a signal with the same mean of  $B$  but affected by Poisson noise.

One might think that greater noise reduction compared to the linear filter (Eqn (1)) can be obtained by considering *higher-order low-pass filters* (i.e. low-pass filters whose transfer function has order greater than 1). However, since such filters (in case of real and non-positive

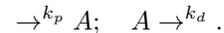
roots) can be implemented as cascades of linear filters (*SI Section C*), where all components are therefore limited by Poisson noise, their noise reduction performance is similarly limited. An example of such a mechanism can be found in multi-step models of gene expression [31], where protein expression is a sequence of linear reactions, thus limited by Eqn (4). Another example can be observed in signaling cascades, such as the mitogen-activated protein kinase (MAPK) cascade [40], where non-linear filtering mechanisms are necessary in order to reduce stochastic fluctuations of a downstream process below Poisson levels [46].

### Correlated production and degradation can reduce noise below Poisson levels

The noise filtering capability of linear filters can be improved by using *higher-order reactions*. The simplest second-order reaction is complex formation. We show how complex formation of two molecules that are positively correlated in their expression can indeed work as an efficient noise filter. The following network, which we call *annihilation module* (Figure 1B), is based on binding and degradation of two parallel synthesized molecules. The annihilation module can be described by the following two reactions



where  $A$  is the input and  $B$  (or equivalently  $C$ ) is the output. In this module,  $B$  and  $C$  are co-expressed and then they inhibit each other. To study how the annihilation module behaves with respect to intrinsic and extrinsic noise, we consider a general birth-death process  $A$  affected by Poisson noise. That is,  $A$  is generated and removed by the following reactions



Thus, we have  $E[A(t)] = V[A(t)]$ ,  $t \in \mathbb{R}_{\geq 0}$ . Since Eqn (7) is non-linear, a general and exact analysis, as in the

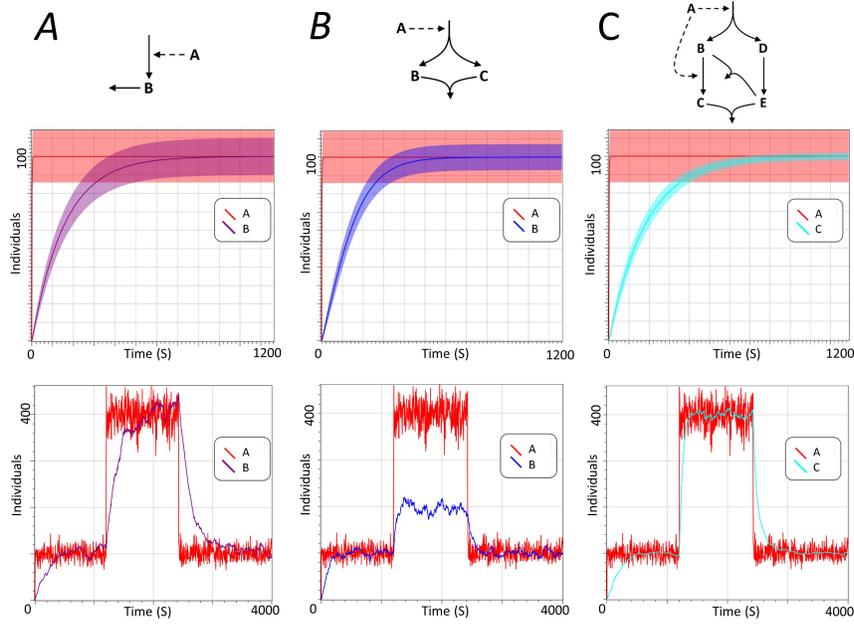


FIG. 3. Comparison of the actions of  $\mathcal{F}$  (linear filter, Eqn (1)),  $\mathcal{M}$  (annihilation module, Eqn (7)), and  $\mathcal{A}$  (annihilation filter, Eqn (10)) on noisy input  $A$  generated by reactions  $\rightarrow^{k_p} A + A$ ;  $A \rightarrow^{k_d}$ . (Top row) Diagrams of the respective species interactions. (Middle row) Plots of expectation and standard deviation of the respective species until time 1200, as estimated by the Linear Noise Approximation (LNA). (Bottom row) Plots of the effect of each filter obtained from a single stochastic simulation until time 4000. (A) The linear filter (Eqn (1)), for  $k_1 = 0.0064$ ,  $k_2 = 0.0064$ , introduces a delay to buffer stochastic fluctuations and reduce the noise, but only to Poisson levels. (B) For  $r_1 = 0.005$ ,  $r_2 = 0.00005$ , the annihilation module (Eqn (7)) improves noise reduction performance but cannot proportionally follow changes in the input. In fact, Eqn (8) predicts that the expectation of the output of the annihilation module changes with the square root of input. Thus, changes on the average value of the input are attenuated in the output. (C) For  $r_1 = 100$ ,  $r_2 = 1000$ ,  $r_3 = 0.000055$ , the annihilation filter (Eqn (10)) not only improves the noise reduction capabilities compared to the other modules, but also proportionally follows changes in the input.

linear case, cannot be performed, as the moment equations cannot be solved. Consequently, we make use of the Linear Noise Approximation (LNA) [11, 48] and derive analytical solution for the expectation and Fano factor of  $B$  at steady state for such an input process  $A$ . We get (see *SI Section D*)

$$E[B]_\infty = \sqrt{\frac{r_1 E[A]_\infty}{r_2}} \quad (8)$$

$$F_B = \frac{2r_1^{3/2} \sqrt{r_2 k_p k_d} + 4r_1 r_2 k_p - r_1 k_d^2 - k_d^3}{8r_1 r_2 k_p - 2k_d^3}, \quad (9)$$

where  $F_B$  stands for the Fano factor of  $B$  at steady state. Assume  $r_1 = r\gamma$  and  $r_2 = r$ , with  $r, \gamma \in \mathbb{R}_{>0}$ , then for  $r \rightarrow 0$  we have  $F_B = \frac{1}{2}$ , thus halving the variance with respect to Poisson noise. Moreover, for  $r \rightarrow \infty$ , we have  $F_B = \frac{2+n}{4}$  where  $n = \frac{E[B]_\infty}{E[A]_\infty}$ . This leads to a surprising result: for  $n = 1$ , that is,  $E[B]_\infty = E[A]_\infty$  (perfect tracking of the mean), the Fano factor is always smaller than 1 for arbitrarily large values of  $r_1$  and  $r_2$ . Thus,  $\mathcal{M}$  can reduce the noise even without introducing a delay in its buffering action. This can be justified because this architecture, where  $B$  and  $C$  are co-expressed, enables attenuation of the low-frequency components of the in-

put signal. Therefore, we obtain noise reduction even if the high-frequency components are not necessarily attenuated (*SI Section D*).

We note that, in the annihilation module, the steady-state value of the output signal is proportional not to the steady-state value of the input signal, but to its square root (Eqn (8)). This may be beneficial in molecular networks, where it may help maintain regulatory stability under changes in initial conditions. However, this mechanism would not be appropriate in cases where the long-term evolution of the upstream component should be followed, because changes to the input would not be followed proportionally (Figure 3B).

The annihilation module is closely related to the incoherent Feed Forward Loop motif (iFFL) [24], where two species are co-expressed and one inhibits the other. However, in *SI Section D*, we show that an iFFL with mass action kinetics cannot reduce the noise below Poisson levels. Hence, having  $B$  and  $C$  degraded together is essential for efficient noise reduction.

The annihilation module and the annihilation filter (see next section) are also related to the *antithetical integral feedback motif* [9]. The main similarity lies in the fact that all these modules have an annihilation reaction, thus suggesting a key role for such a reaction in dealing

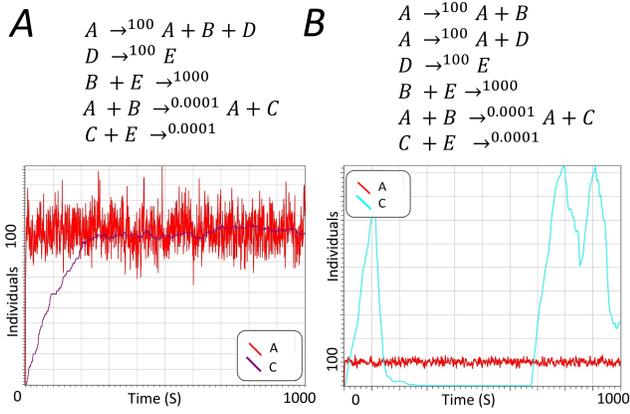


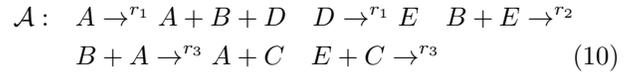
FIG. 4. We consider two networks with the same rate equations that differ only in how  $B$  and  $D$  are expressed. For both figures we plot a stochastic simulation of an input signal  $A$  affected by Poisson noise with the output of the filter. (A) considers the annihilation filter (Eqn (10)). (B) considers a network that is identical to the one in Figure A, except for the fact that  $B$  and  $D$  are not co-expressed.

with noisy dynamics. However, we stress here how the mechanisms differ: a key requirement of both our filters for efficient noise reduction is the co-expression of the molecules that will participate in the annihilation reaction. This requirement cannot be implemented in the antithetical integral feedback schema. Moreover, the antithetical integral feedback motif is known to increase the noise of the controlled network [10]. In fact, we argue that one of the reasons why in the antithetical integral feedback the noise increases is that the species that undergo an annihilation reaction are not co-expressed. In order to illustrate this point, in Figure 4, we compare two networks: one is the annihilation filter, the other is identical to the annihilation filter except for the fact that  $B$  and  $D$  are not co-expressed, but only positively correlated in their expression. The two networks, deterministically, behave identically (they have same rate equations). However, interestingly, the stochastic behavior is completely different, thus demonstrating the importance of co-expression.

### Annihilation filter suppresses molecular noise

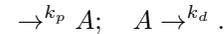
We propose a general architecture, called the *annihilation filter*, which is based on the annihilation module but also guarantees  $E[C]_\infty = sE[A]_\infty$  for a given constant  $s$ , independently of  $E[A]_\infty$  (in what follows, without any loss of generality, we assume  $s = 1$ ). We show how the annihilation filter can asymptotically reduce molecular noise to zero, that is, for appropriate limiting values for the rates, the Fano factor of the output converges to zero (Eqn (15)).

Our annihilation filter,  $\mathcal{A}$ , is illustrated in Figure 3C.  $\mathcal{A}$  is composed of the following reactions



$A$  is the input species and  $C$  is the output filtered species. The first three reactions are similar to the annihilation module, but with an additional delay introduced by the reaction  $D \xrightarrow{r_1} E$ . Eqn (3) guarantees that  $D$  is a copy of  $A$ , and the number of times that  $D$  molecules have been produced or destroyed is stored respectively in  $B$  and  $E$ . As the role of these reactions is to act as a sensor, high values of  $r_1$  are more informative than small ones. If  $r_2$  is large enough, the count of  $C$  is modified not any time a  $B$  or  $E$  molecule is produced, but just by their difference. The fourth and fifth reactions increase or decrease  $C$ . The rate  $r_3$  controls the delay introduced by the filter, and thus also the noise reduction.

Since the system is non-linear, to study the noise reduction capabilities of the annihilation filter we make use of the LNA. We assume  $A$  is a general input process, with extrinsic noise modeled by a Poisson process. That is,  $A$  is generated and removed by the following reactions



Thus, we have  $E[A(t)] = V[A(t)]$ ,  $t \in \mathbb{R}_{\geq 0}$ . Using the LNA equations we can derive the following conditions (*SI Section E*)

$$E[A]_\infty = E[D]_\infty \quad (11)$$

$$E[C]_\infty = E[D]_\infty \frac{E[B]_\infty}{E[E]_\infty} \quad (12)$$

$$0 = r_1 E[D]_\infty - r_2 E[B]_\infty E[E]_\infty - r_3 E[E]_\infty E[C]_\infty \quad (13)$$

$$E[B(t)] - E[E(t)] = E[B(0)] - E[E(0)] + E[D(t)] - E[C(t)]. \quad (14)$$

Assuming the same initial concentration of  $B, E$ , at steady state,  $C$  will always track  $A$  independently of the value of  $r_1, r_2, r_3$ . That is,  $E[A]_\infty = E[C]_\infty$ .

We can now study the Fano factor of  $C$  at steady state,  $F_C$ . To do that, we assume  $r_2 = \frac{\gamma}{r}$ ,  $r_3 = r$ , where  $\gamma, r$  are constants. In order for the annihilation filter to work as an efficient noise filter, as we discussed, we need large  $r_2$  and small  $r_3$ . Thus, we study  $F_C$  for  $r \rightarrow 0$ . Under this limit, we obtain the following elegant form for the Fano factor of  $C$

$$\lim_{r \rightarrow 0} F_C = \frac{k_p}{k_p + r_1}.$$

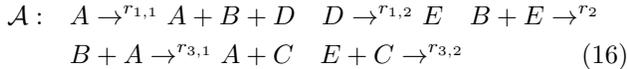
Hence, we have that

$$\lim_{r_1 \rightarrow \infty} (\lim_{r \rightarrow 0} F_C) = 0. \quad (15)$$

Thus, increasing  $r_1$ , the noise can be made arbitrarily small, showing how this architecture has ideal noise reduction capabilities: independently of the intrinsic noise

introduced the total noise can be made arbitrarily close to 0.

Note that in the CRN (10) it is assumed that some reactions have same rates. This assumption allows us to obtain simpler analytic results. In order to show that the above analysis remains valid also in the more general scenario where the reactions have different rates, we consider the following CRN, modifying (10)



In Figure 5 we compute  $F_C$  and  $E[C]_\infty$  for different values of  $r_{1,1}, r_{1,2}$  and  $r_{3,1}, r_{3,2}$ . It is easy to observe that Eqn (15) is confirmed: if  $r_{1,1}, r_{1,2}$  are big enough compared to  $r_{3,1}, r_{3,2}$  the Fano Factor will decrease converging to a value of 0. Nevertheless, if  $r_{1,1} \neq r_{1,2}$  or  $r_{3,1} \neq r_{3,2}$ , the noise will still be reduced, but this will affect  $E[C]_\infty$  which, may be different from  $E[A]_\infty$ .

The first two reactions of the annihilation filter (Eqn (10)) can be thought as a model for co-expression of molecules  $B$  and  $E$  at different rates. Such a model of co-expression can be generalized, having  $B$  and  $E$  co-expressed, and then interacting after a pathway of linear reactions. For instance, this is the case of mRNA and microRNA that, when co-expressed, undergo a series of maturation steps before of interacting [5, 31]. In Figure 6 we show that these auxiliary reactions do not influence the noise reduction capabilities of the annihilation filter.

### Numerical analysis

While the mathematical analysis performed on linear filters is exact and general, for annihilation module and annihilation filter, our claims are based on the LNA and for birth-death input processes. This is because those filters are non-linear, hence exact analysis based on the moment equations is not possible. Thus, we need to support our results about the noise suppression capabilities of such networks with stochastic simulations of such filters for different classes of input. In order to do that, in Figure 3, we consider a step-like perturbation, and in [SI Section F](#) we consider oscillatory inputs. In both cases, the annihilation filter outperforms the other filters: for the same delay introduced it suppresses more noise at the high frequencies while still maintaining similar long-term behaviour. Instead, the annihilation module does not follow the long term behaviour of the input proportionally. In fact, Eqn (8) predicts that the expectation of the output of the annihilation module changes with the square root of input. Thus, changes of the input are attenuated in the output.

To further confirm the mathematical analysis, using stochastic simulations we compare the power spectral density (PSD) (see Method Section) of the input species affected by Poisson noise with the PSD of the output

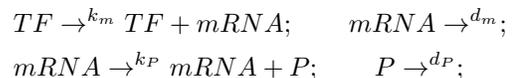
species of the annihilation filter (Eqn(10)), and of the linear filter (Eqn(1)), both for the same input (see Figure 7A). We can see that both filters are indeed low-pass filters in the sense that they attenuate the high frequencies. However, although they behave similarly at the high-frequencies, the linear filter is less robust to intrinsic noise, and such intrinsic noise amplifies the low frequencies, resulting in noise reduction being lower bounded by Poisson dynamics. That is, the reactions of the linear filter, introduce slow and medium time variations of the output leading to an amplification of the low frequency components of its spectrum. The annihilation filter enables a much better reduction of the intrinsic noise, leading to a smaller amplification of the low frequencies. The connection between the noise of a process and its PSD is explained in detail in the Supplementary Material ([SI Section A](#)).

In Figure 7B, we consider again input species  $A$  affected by Poisson noise, and, on this input, we compare the action of the annihilation filter (Eqn (10)), annihilation module (Eqn (7)), and linear filter (Eqn (1)). In order to reduce the number of free variables in the system, we constrain the output of the filters to have the same mean as  $A$ . Then, we plot the Fano factor as a function of the remaining free rate parameters. As expected, our key observations are that the annihilation module is the only mechanism that guarantees noise reduction for any value of the parameter rate, confirming the theoretical result of Eqn (9) that it is able to reduce the noise even without introducing a delay in its buffering action. Also, for an arbitrarily long delay, the annihilation filter converges to a Fano factor of 0, showing the ability of this network to remove all the noise (variance tends to 0 when delay tends to infinity). On the other hand, the linear filter converges to a Fano factor of 1, corresponding to Poisson levels, thus confirming Eqn (5).

## DISCUSSION

### *Gene and protein expression can work as linear filters*

Gene expression is often modeled as a two or three-stage process, where mRNA is transcribed from a transcription factor TF, and the protein P is translated from the mRNA [41]



Under this modeling assumption, the linear filtering mechanism (Eqn (1)) may be present both at the transcription (top two reactions) and translation level (bottom two reactions) to buffer noise and increase robustness by slowing down transcription or translation. At the transcription level, inefficient transcription that follows fast promoter activation is a mechanism that has been widely observed to buffer fluctuations in the mRNA time

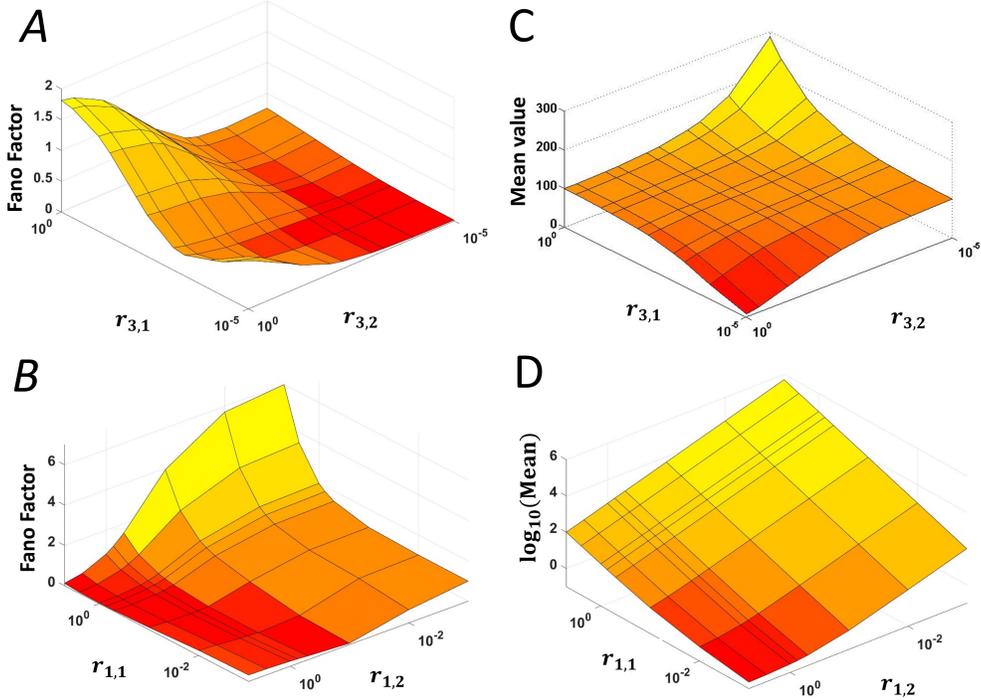


FIG. 5. For the CRN in Eqn (16), the figure shows how the Fano Factor of  $C$ ,  $F_C$ , and the mean value of  $C$  at steady state,  $E[C]_\infty$  depends on  $r_{1,1}, r_{1,2}$  and  $r_{3,1}, r_{3,2}$ . For all figures we consider an input species  $A$  such that  $E[A]_\infty = 100$  and  $A$  is affected by Poisson noise. It is possible to observe that, as predicted by the theoretical analysis, the Fano Factor tends to decrease either when both  $r_{1,1}, r_{1,2}$  increases or when both  $r_{3,1}, r_{3,2}$  decreases. Interestingly, note that  $F_C$  tends to be smaller when there is a strong suppression of the mean. Thus, when a low number of molecules is involved, the noise has more influence on the behaviour of the system. (A,C) We plot  $F_C$  and  $E[C]_\infty$  for  $r_{1,1} = r_{1,2} = 10$  and  $r_2 = 100$ . The non-linearity of the reactions involved is such that  $F_C$  and  $E[C]_\infty$  are robust with respect to parameters variation. (B,D) We plot  $F_C$  and  $E[C]_\infty$  for  $r_{3,1} = r_{3,2} = 0.001$  and  $r_2 = 100$ . In this case,  $E[C]_\infty$  changes linearly with the rates.  $F_C$  tends to increase when there is a strong amplification of  $E[C]_\infty$ .

evolution [35]. At the translation level, the linear filtering mechanism reduces the noise in protein expression, as supported by experimental evidence on *B. subtilis* and *S. cerevisiae* [8, 29]. Eqn (4) guarantees that, independently of the presence of arbitrarily many feed-forward or feed-back loops between  $P$  and transcription factor or mRNA, we have

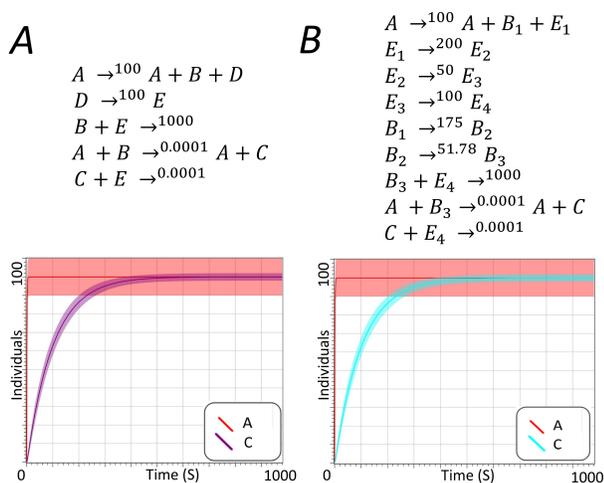
$$V[P]_\infty = E[P]_\infty + \frac{k_P}{d_P} Cov[mRNA, P]_\infty. \quad (17)$$

Thus, if mRNA and protein are positively correlated, then the stochastic fluctuations in protein expression cannot be reduced below Poisson levels, and this limit is approached for slow translation. Hence, Eqn (17) may explain why in yeast and *E. coli* the translation rates tend to be slower than the transcription rates [16, 35], and also suggests that mechanisms to induce a negative correlation between protein and mRNA may have been selected to enhance robustness. This is the case for negative feedback, which may thus enable noise reduction below Poisson levels, as again confirmed by experimental evidence [7]. We note that a more realistic model

of gene expression requires representation of transcription and translation as multi-step processes. However, since such processes can still be modeled as a sequence of first-order reactions [31], our analysis and the linear filter mechanism still applies.

#### *MicroRNAs can serve as annihilation filters*

Slow translation/transcription is a very simple mechanism of noise reduction. Since gene expression involves low molecular counts and highly stochastic signals, different (and more complex) network architectures may have been selected to deal with scenarios where greater noise reduction is needed [18]. A simple example can be found in microRNA-regulated post-transcriptional regulation. MicroRNAs (miRNAs) are short RNAs that are widely conserved in biological networks [6]. In animals, it is common that miRNAs and their target mRNAs are co-expressed or positively correlated in their expression [12, 32, 36]. For example, c-myc induces the expression of the microRNAs miR-17-5p and miR-20a together with



their target E2F1 [28]. Furthermore, the system where miRNAs repress gene expression by binding with the target mRNAs, and either inhibiting translation of mRNA or promoting mRNA decay [1, 21], leads to a pattern that can be modelled with the annihilation module (Figure 1B). Although it is well accepted that microRNAs confer robustness on gene expression [12, 20, 38], it is still not clear what aspects of their inhibitory mechanisms are used to gain efficient noise reduction [3], and previous analysis has focused on microRNAs that are not co-expressed with the target proteins [38]. One hypothesis is that co-expression of microRNAs with their targets has a role in increasing robustness of gene expression [18]. This is also supported by experimental evidence [45]. Our mathematical analysis confirms such a hypothesis, and shows that correlated expression of microRNAs and mRNA, followed by translational repression of the mRNA when bound to the microRNA, may lead to noise reduction below Poisson levels. This result suggests that microRNA regulation may have been selected to post-regulate highly noisy genes. One specific example of such a pattern can be found in the *Drosophila eye*, where miR-7 and its target protein are co-expressed, and experimental studies have suggested the role of miR-7 in buffering fluctuations [23, 32]. A particular network involving miR-7 and verified experimentally in [23] is shown in Figure 8. This network is responsible for sensory organ precursor (SOP) fate. An annihilation module between Ato, miR-7, and E(spl) genes may have a role in increasing robustness. We stress that it has yet to be verified experimentally that miR-7 works as an annihilation filter, especially due to the lack of experiments concerning the joint degradation of micro-RNA and mRNA in different organisms [2].

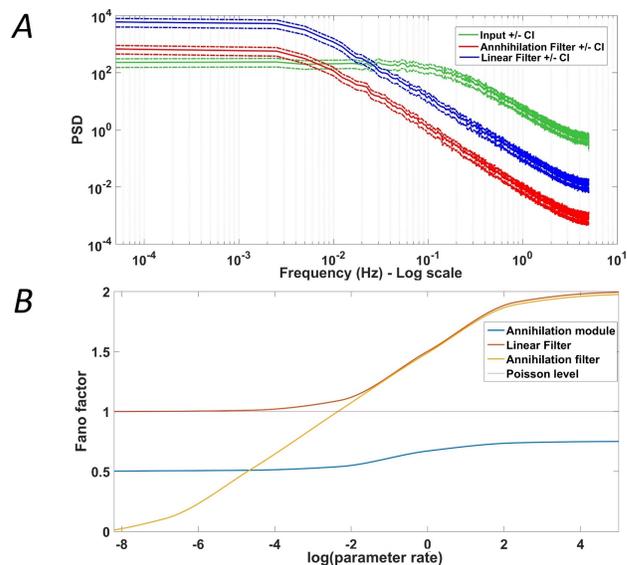


FIG. 8. The network controlling sensory organ precursor (SOP) fate [23]. miR-7 participates in an annihilation module highlighted in black. The annihilation module is also interconnected with a double-negative feedback loop between Ato and E(spl), with miR-7 as an effector of Ato, and E(spl) directly inhibiting Ato.

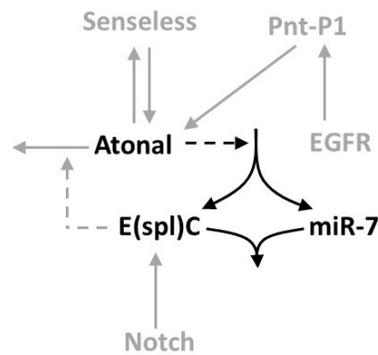


FIG. 8. The network controlling sensory organ precursor (SOP) fate [23]. miR-7 participates in an annihilation module highlighted in black. The annihilation module is also interconnected with a double-negative feedback loop between Ato and E(spl), with miR-7 as an effector of Ato, and E(spl) directly inhibiting Ato.

We have also found that the annihilation filter (Eqn (10)) may be a suitable model for the *c-Myc/E2F1/hsa-miR-20a* circuit, with its extension to E2F2 in the case when E2F1 activates E2F2 [28, 36] (Figure 1C). This may suggest that miR-20a, by repressing both E2F1 and E2F2, confers greater robustness on E2F2 expression. We note, however, that there is no experimental evidence yet for whether E2F1 activates or represses E2F2. Note that the annihilation filter link between A (*c-Myc*) and C (E2F2) is not present here: its role is to regulate the mean of the output, but it has no influence on noise reduction.

In conclusion, we analyzed three simple molecular noise filters and derived their properties and limitations. These filters can be related to biological mechanisms. We show, for example, that gene expression with slow translation/transcription can reduce noise, but only down to Poisson levels (Eqn (17)), even when considering linear multiple-step models [31]. In contrast, the annihilation module (Eqn (7)), which relies on co-expression and joint degradation, can drive the noise below Poisson levels. Such mechanisms can be related to microRNA regulated systems, where a key requirement for effective noise reduction is the co-expression of mRNA with its microRNAs. Our analysis suggests how a trade-off between performance and resources arises: simpler circuits can reduce less noise, but also require fewer resources. In this sense, it is interesting to emphasize that complex noise reduction mechanisms, such as the annihilation module, tend to be found in highly regulated systems. In fact, co-

expression of mRNA and microRNA, followed by translation inhibition, is a pattern that is common in animals but much less prevalent in plants [1].

While biological systems deal with noise in a variety of ways, in this paper we focused on scenarios where noise should be controlled. It remains an interesting endeavor to similarly discover and analyze the basic principles that allow biological systems to exploit noise functionally and use it to their benefit [13]. We believe that a systematic analysis of noise reduction in molecular systems, together with evidence of widespread noise reduction capabilities in biological systems, such as in gene expression, are fundamental to obtain new insights into the structure and evolutionary origin of noise reduction mechanisms.

## AUTHOR CONTRIBUTIONS

L.C, A.C and L.L. designed the research; L.L. performed the research; All authors contributed to analysis, interpretation of the outcomes and preparation of the manuscript

## ACKNOWLEDGMENTS

The authors thank Alessandro Abate and Neil Dalchau for comments on a draft of the manuscript. L.C. is partially funded by a Royal Society Professorship.

- 
- [1] Ameres, S. L. and Zamore, P. D., *Nature reviews Molecular cell biology* **14**, 475 (2013).
  - [2] Baccarini, A., Chauhan, H., Gardner, T. J., Jayaprakash, A. D., Sachidanandam, R., and Brown, B. D., *Current biology* **21**, 369 (2011).
  - [3] Bakker, R. and Carthew, R. W., *Developmental Cell* **40**, 515 (2017).
  - [4] Balázs, G., van Oudenaarden, A., and Collins, J. J., *Cell* **144**, 910 (2011).
  - [5] Bartel, D. P., *cell* **116**, 281 (2004).
  - [6] Bartel, D. P., *Cell* **136**, 215 (2009).
  - [7] Becskei, A. and Serrano, L., *Nature* **405**, 590 (2000).
  - [8] Blake, W. J., Kærn, M., Cantor, C. R., and Collins, J. J., *Nature* **422**, 633 (2003).
  - [9] Briat, C., Gupta, A., and Khammash, M., *Cell systems* **2**, 15 (2016).
  - [10] Briat, C., Gupta, A., and Khammash, M., arXiv preprint arXiv:1711.08291 (2017).
  - [11] Cardelli, L., Kwiatkowska, M., and Laurenti, L., *Biosystems* **149**, 26 (2016).
  - [12] Ebert, M. S. and Sharp, P. A., *Cell* **149**, 515 (2012).
  - [13] Eldar, A. and Elowitz, M. B., *Nature* **467**, 167 (2010).
  - [14] Engblom, S., *Applied Mathematics and Computation* **180**, 498 (2006).
  - [15] Ethier, S. N. and Kurtz, T. G., *Markov processes: characterization and convergence*, Vol. 282 (John Wiley & Sons, 2009).
  - [16] Fraser, H. B., Hirsh, A. E., Giaever, G., Kumm, J., and Eisen, M. B., *PLoS Biol* **2**, e137 (2004).
  - [17] Horn, F. and Jackson, R., *Archive for rational mechanics and analysis* **47**, 81 (1972).
  - [18] Hornstein, E. and Shomron, N., *Nature genetics* **38**, S20 (2006).
  - [19] Kærn, M., Elston, T. C., Blake, W. J., and Collins, J. J., *Nature Reviews Genetics* **6**, 451 (2005).
  - [20] Kasper, D. M., Moro, A., Ristori, E., Narayanan, A., Hill-Teran, G., Fleming, E., Moreno-Mateos, M., Vejnar, C. E., Zhang, J., Lee, D., *et al.*, *Developmental Cell* **40**, 552 (2017).
  - [21] Krol, J., Loedige, I., and Filipowicz, W., *Nature Reviews Genetics* **11**, 597 (2010).
  - [22] Lestas, I., Vinnicombe, G., and Paulsson, J., *Nature* **467**, 174 (2010).
  - [23] Li, X., Cassidy, J. J., Reinke, C. A., Fischboeck, S., and Carthew, R. W., *Cell* **137**, 273 (2009).
  - [24] Mangan, S. and Alon, U., *Proceedings of the National Academy of Sciences* **100**, 11980 (2003).
  - [25] McAdams, H. H. and Arkin, A., *Proceedings of the National Academy of Sciences* **94**, 814 (1997).
  - [26] McAdams, H. H. and Arkin, A., *Trends in genetics* **15**, 65 (1999).
  - [27] Munsky, B., Neuert, G., and Van Oudenaarden, A., *Science* **336**, 183 (2012).

- [28] O'donnell, K. A., Wentzel, E. A., Zeller, K. I., Dang, C. V., and Mendell, J. T., *Nature* **435**, 839 (2005).
- [29] Ozbudak, E. M., Thattai, M., Kurtser, I., Grossman, A. D., and Van Oudenaarden, A., *Nature genetics* **31**, 69 (2002).
- [30] Pedersen, M. and Phillips, A., *Journal of the Royal Society Interface* , rsif (2009).
- [31] Pedraza, J. M. and Paulsson, J., *Science* **319**, 339 (2008).
- [32] Posadas, D. M. and Carthew, R. W., *Current opinion in genetics & development* **27**, 1 (2014).
- [33] Raj, A. and van Oudenaarden, A., *Cell* **135**, 216 (2008).
- [34] Rao, C. V., Wolf, D. M., and Arkin, A. P., *Nature* **420**, 231 (2002).
- [35] Raser, J. M. and O'shea, E. K., *Science* **309**, 2010 (2005).
- [36] Re, A., Corá, D., Taverna, D., and Caselle, M., *Molecular BioSystems* **5**, 854 (2009).
- [37] Samoilov, M., Arkin, A., and Ross, J., *The Journal of Physical Chemistry A* **106**, 10205 (2002).
- [38] Schmiedel, J. M., Klemm, S. L., Zheng, Y., Sahay, A., Blüthgen, N., Marks, D. S., and van Oudenaarden, A., *Science* **348**, 128 (2015).
- [39] Sedra, A. S. and Smith, K. C., *Microelectronic circuits*, Vol. 1 (New York: Oxford University Press, 1998).
- [40] Seger, R. and Krebs, E. G., *The FASEB journal* **9**, 726 (1995).
- [41] Shahrezaei, V. and Swain, P. S., *Proceedings of the National Academy of Sciences* **105**, 17256 (2008).
- [42] Singh, A. and Hespanha, J. P., *IEEE Transactions on Automatic Control* **56**, 414 (2011).
- [43] Stein, E. M. and Weiss, G., *Introduction to Fourier analysis on Euclidean spaces (PMS-32)*, Vol. 32 (Princeton university press, 2016).
- [44] Stoica, P. and Moses, R. L., *Spectral analysis of signals*, Vol. 452 (Pearson Prentice Hall Upper Saddle River, NJ, 2005).
- [45] Strovas, T. J., Rosenberg, A. B., Kuypers, B. E., Muscat, R. A., and Seelig, G., *ACS synthetic biology* **3**, 324 (2014).
- [46] Thattai, M. and van Oudenaarden, A., *Biophysical journal* **82**, 2943 (2002).
- [47] Thattai, M. and Van Oudenaarden, A., *Proceedings of the National Academy of Sciences* **98**, 8614 (2001).
- [48] Van Kampen, N. G., *Stochastic processes in physics and chemistry*, Vol. 1 (Elsevier, 1992).
- [49] Zechner, C., Seelig, G., Rullan, M., and Khammash, M., *Proceedings of the National Academy of Sciences* **113**, 4729 (2016).

## SUPPLEMENTARY INFORMATION

L. Laurenti, A. Csikasz-Nagy, M. Kwiatkowska, and L. Cardelli

In Section A we introduce low-pass filters. In Section B we formally define Chemical Reaction Networks (CRNs) and the mathematical results we use in the main paper. Linear filters are discussed in Section C, where we give the mathematical details of their analysis. In Sections D and E we include detailed analysis of the annihilation module and annihilation filter. In Section F we study our molecular filters against time-varying inputs, while in Section G we report details on the software and on the code used to generate the figures in the text.

### Appendix A: SI Low Pass Filters

A *low-pass filter* is a filter that preserves signals with a frequency lower than a certain cutoff frequency and attenuates signals with frequencies higher than such a frequency. Low-pass filters can be analyzed in the frequency domain by considering their transfer function, which describes the relationship between the input and the output in the Fourier domain [43]. The frequency response can be described by means of the Bode diagram, which is usually a combination of the Bode magnitude diagram, expressing the magnitude of the frequency response, and the Bode phase diagram, expressing the phase shift. Given an output signal  $o(t)$  and an input signal  $i(t)$ , *first-order low-pass filters* are characterized by the following transfer function

$$\frac{\hat{o}(\omega)}{\hat{i}(\omega)} = \frac{1}{1 + k \cdot i\omega}, \quad (\text{A1})$$

where  $\hat{i}(\omega), \hat{o}(\omega)$  are the Fourier transform of  $i(t)$  and  $o(t)$ , and  $k \in \mathbb{R}_{>0}$ . The cutoff frequency is defined as the frequency for which the signal has an attenuation of 3 decibel (see Figure S1), corresponding to halving the amplitude. For a filter with the above transfer function, the cutoff frequency is  $\bar{\omega} = \frac{1}{k}$ . *Higher-order* low-pass filters are characterized by a higher-order transfer function, and they perform a sharper attenuation of the high frequency of a signal [43].

### Noise filtering and low-pass filters

Consider a stochastic process  $X(t), t \in \mathbb{R}_{\geq 0}$ , with values in  $\mathbb{R}_{\geq 0}$ . The autocorrelation function of  $X$  at time  $t$ , for  $\tau \in \mathbb{R}_{\geq 0}$ , is defined as

$$R_X(t, \tau) = E[X(t)X(t + \tau)].$$

Note that  $R_X(t, 0)$  is the second moment of  $X$  at time  $t$ . Under the assumption that  $X$  is stationary,  $R_X(t, \tau)$  does

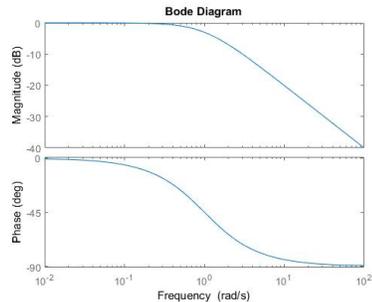


FIG. S1. Plots of the Bode diagram of the transfer function of a first-order low-pass filter with cutoff frequency  $\bar{\omega} = k_1$ . If an input signal has a frequency lower than the cutoff frequency, then the output of the filter has the same amplitude. If instead the input signal has a higher frequency, then the output signal is attenuated and a delay is introduced.

not depend on  $t$ , so we write directly  $R_X(\tau)$ . Provided that  $R_X(\tau)$  is absolutely integrable, the Power Spectral Density (PSD) of  $X$  is defined for each angular frequency  $\omega$  in  $\mathbb{R}$  as

$$S_X(\omega) = \int_{-\infty}^{+\infty} R_X(\tau) \exp^{-i\omega\tau} d\tau.$$

That is, the PSD of a signal is the Fourier transform of its autocorrelation function. From the definition of PSD the following dual relation can be derived

$$R_x(\tau) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S_X(\omega) \exp^{i\omega\tau} d\omega.$$

Setting  $\tau = 0$ , we have

$$R_x(0) = E[X^2(t)] = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S_X(\omega) d\omega.$$

If  $X$  is zero mean, we have  $R_x(0) = V[X(t)] = E[X^2(t)] = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S_X(\omega) d\omega$ . That is, the variance of  $X$  at time  $t$  is given by the summation of all the frequency components of its spectrum. As a result, given two signals  $X_1, X_2$  with the same spectrum up to frequency  $\omega^*$ , but with all frequencies  $\bar{\omega} > \omega^*$  such that  $S_{X_1}(\bar{\omega}) < S_{X_2}(\bar{\omega})$ , then, necessarily,  $V[X_1(t)] < V[X_2(t)]$ . Therefore, attenuating some frequency components of a signal without introducing any amplification at the other frequencies automatically reduces the variance.

### Appendix B: SI Chemical Reaction Network (CRN)

A *chemical reaction network (CRN)*  $\mathcal{C} = (\Lambda, \mathcal{R})$  is a pair of finite sets, where  $\Lambda$  is a set of *species*,  $|\Lambda|$  denotes its size, and  $\mathcal{R}$  is a set of reactions. Species in  $\Lambda$  interact according to the reactions in  $\mathcal{R}$ . A *reaction*  $\tau \in \mathcal{R}$  is a triple  $\tau = (r_\tau, p_\tau, k_\tau)$ , where  $r_\tau \in \mathbb{N}^{|\Lambda|}$  is the *reactant complex*,  $p_\tau \in \mathbb{N}^{|\Lambda|}$  is the *product complex* and

$k_\tau \in \mathbb{R}_{>0}$  is the coefficient rate.  $r_\tau$  and  $p_\tau$  represent the stoichiometry of reactants and products. Given a reaction  $\tau_1 = ([1, 1, 0], [0, 0, 2], k_1)$ , where  $\Lambda = \{A, B, C\}$ , we often refer to it as  $\tau_1 : A + B \xrightarrow{k_1} 2C$ . The *state change* associated to  $\tau$  is defined by  $v_\tau = p_\tau - r_\tau$ . For example, for  $\tau_1$  as above, we have  $v_{\tau_1} = [-1, -1, 2]$ .

### 1. SI Stochastic Model of CRNs

Given a CRN  $\mathcal{C} = (\Lambda, \mathcal{R})$ , the stochastic model of  $\mathcal{C}$  is given by a continuous-time discrete-space Markov process (CTMC)  $(X(t), t \geq 0)$ , whose transient evolution can be described by the Chemical Master Equation (CME) [48]. The state of the system at time  $t$  is given by the number of molecules of each species at that time, and it stays constant until a reaction  $\tau$  happens, when the state jumps by  $v_\tau$ . Given  $A \in \Lambda$ , with an abuse of notation we identify the number of molecules of  $A$  at time  $t$  as  $A(t)$ . Computing the transient evolution of the CME requires solving a system of ODEs whose size is equal to the number of possible configurations. This number grows exponentially with the number of species if the species counts are bounded, and is infinite when the species counts are finite but unbounded. As a consequence, obtaining a solution of the CME is generally infeasible, even numerically.

Given a CRN  $\mathcal{C} = (\Lambda, \mathcal{R})$  and a polynomial function  $\mathcal{T} : \mathbb{N}^{|\Lambda|} \rightarrow \mathbb{R}$  over the species of  $\Lambda$ , it is possible to show that the evolution of the expectation of  $\mathcal{T}$  over time can be described by the following system of ODEs, which can be derived from the CME [42]

$$\frac{d(E[\mathcal{T}(X(t))])}{dt} = \sum_{\tau \in \mathcal{R}} E[\alpha_\tau(X(t))(\mathcal{T}(X(t) + v_\tau) - \mathcal{T}(X(t)))]. \quad (\text{B1})$$

where  $\alpha_\tau(X(t))$  is the propensity rate of reaction  $\tau$  in state  $(X(t))$ . We assume mass action kinetics. That is, for  $\tau : \rightarrow^k A$ , we have  $\alpha_\tau(X(t)) = k$ , for  $A \rightarrow^k$ , we have  $\alpha_\tau(X(t)) = kA(t)$ , and for  $A + B \rightarrow^k$ , we have  $\alpha_\tau(X(t)) = kA(t)B(t)$ , assuming  $A \neq B$ , while for  $A + A \rightarrow^k$ , we have  $\alpha_\tau(X(t)) = k\frac{A(t)^2}{2}$ . Eqn (B1) can be used to describe the moments of  $X$ ; for instance, for  $\mathcal{T}(X(t)) = A^2(t)$ , Eqn (B1) describes the time evolution of the second moment of  $A$ . If all the reactions in  $\mathcal{R}$  are at most uni-molecular, then the equations for the  $k$ th moment depend on moments of order smaller than or equal to  $k$ . This leads to a system of ODEs in number polynomial with the respect to the number of species, and independent of the molecular counts, ensuring much greater scalability with respect to solving the CME. If, instead, there is at least one multi-molecular reaction, then the equations for the  $k$ th moment will depend on the higher-order moments. This leads to an infinite system of ODEs. However, approximations of the solutions can still be obtained by using moment closure techniques

[42], additionally guaranteeing much greater scalability with respect to numerical solutions of the CME.

### 2. SI Deterministic Model for CRNs

Given a CRN  $\mathcal{C} = (\Lambda, \mathcal{R})$  its time evolution is often described as a deterministic function  $\Phi : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^{|\Lambda|}$ , which is given by the solution of the following ODEs, called the *rate equations*

$$\frac{d\Phi(t)}{dt} = \sum_{\tau=(p,r,k)} k \prod_{A \in \Lambda} \Phi_A(t)^{r_A} \quad (\text{B2})$$

where  $\Phi_A$  is the component of  $\Phi$  relative to species  $A$ , and  $r_A$  is the component of  $r$  relative to  $A$ .  $\Phi$  approximates the evolution of  $\mathcal{C}$  as a deterministic system, neglecting stochastic fluctuations. It can be shown that this is accurate in the limit of large populations, where  $X$ , the CTMC induced by  $\mathcal{C}$ , tends to behave deterministically, and the CME solution converges to  $\Phi$  [15].

### 3. SI linear Noise Approximation (LNA)

The *linear Noise Approximation* (LNA) is a continuous state space approximation of the CTMC induced by a CRN in terms of a Gaussian process. As a Gaussian process depends just on its first two moments, the LNA can also be seen as a moment closure technique for Eqn (B1). Given a CRN  $\mathcal{C} = (\Lambda, \mathcal{R})$ , the LNA at time  $t$  approximates the distribution of  $X(t)$  with the distribution of the random vector  $Y(t)$  such that

$$X(t) \approx Y(t) = N\Phi(t) + N^{\frac{1}{2}}G(t) \quad (\text{B3})$$

where  $N$  is the volume of the system,  $G(t) = (G_1(t), G_2(t), \dots, G_{|\Lambda|}(t))$  is a random vector, independent of  $N$ , representing the stochastic fluctuations at time  $t$  and  $\Phi(t)$  is the solution of the rate equations. The probability distribution of  $G(t)$  is then given by the solution of a linear Fokker-Planck equation. As a consequence, for every time instant  $t$ ,  $G(t)$  has a multivariate normal distribution whose expected value  $E[G(t)]$  and covariance matrix  $C[G(t)]$  are the solution of the following differential equations:

$$\frac{dE[G(t)]}{dt} = J_F(\Phi(t))E[G(t)] \quad (\text{B4})$$

$$\frac{dC[G(t)]}{dt} = J_F(\Phi(t))C[G(t)] + C[G(t)]J_F^T(\Phi(t)) + W(\Phi(t)) \quad (\text{B5})$$

where  $J_F(\Phi(t))$  is the Jacobian of  $F(\Phi(t))$ ,  $J_F^T(\Phi(t))$  its transpose,  $W(\Phi(t)) = \sum_{\tau \in \mathcal{R}} v_\tau v_\tau^T \alpha_\tau(\Phi(t))$  and  $F_j(\Phi(t))$  the  $j$ th component of  $F(\Phi(t))$ . Note that, since  $G$  is a Gaussian process,  $Y$  is still Gaussian with expectation and variance given by

$$E[X(t)] \approx E[Y(t)] = N\Phi(t)$$

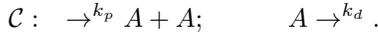
$$C[X(t)] \approx C[Y(t)] = NC[G(t)]$$

Thanks to the Central Limit Theorem, it is possible to show that, by increasing  $N$ ,  $X(t)$  converges in distribution to  $Y(t)$  [15].

### Appendix C: SI linear Filters

We begin this section with an illustrative example to show the noise reduction capabilities of linear filters.

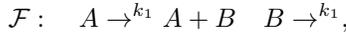
**Example 2** Consider the following CRN



This may model a highly noisy signal  $A$  from the environment, with production rate  $k_p$  and degradation rate  $k_d$ . Expectation and variance of  $A$  at steady state can be derived by solving Eqn (B1) at steady state, and we get

$$E[A]_\infty = 2\frac{k_p}{k_d}; \quad V[A]_\infty = 3\frac{k_p}{k_d}.$$

Note that the fact that expectation and variance stops changing does not imply that a stochastic realization of  $A$  will eventually stop changing. We ask if it is possible to somehow generate a noiseless version of  $A$ , that is, a copy of the  $A$  signal preserving the mean but with reduced variance. In order to answer this question, consider the following CRN, which we call  $\mathcal{F}$



where  $k_1$  is the rate constant. That is,  $A$  catalyzes the production of  $B$ , which is then degraded with the same rate. Then, composing  $\mathcal{F}$  and  $\mathcal{C}$ , we have

$$E[B]_\infty = 2\frac{k_p}{k_d}; \quad V[B]_\infty = \frac{k_p(5k_1 + 2k_d)}{k_d(k_1 + k_d)}.$$

The steady state expectation of  $B$  does not depend on the choice of  $k_1$ : it is always equal to  $E[A]_\infty$ . However,  $V[B]_\infty$  does depend on  $k_1$ . As a consequence, we ask if there are values of  $k_1$  for which  $B$  has less variance than  $A$ . This would yield a filter that is able to reduce the variance of a signal while still maintaining the same average value. It is easy to show that

$$\text{If } 2k_1 < k_d \text{ then } V[B]_\infty < V[A]_\infty.$$

An interesting question is what is the minimum value of  $V[B]_\infty$  that can be obtained. By minimizing  $V[B]_\infty$  with respect to  $k_1$  we have

$$\text{Inf}_{k_1 \in \mathbb{R}_{\geq 0}} (V[B]_\infty) = 2\frac{k_p}{k_d},$$

which is obtained for  $k_1 \rightarrow 0$ . This means that the best that we can achieve with  $\mathcal{F}$  is to produce an output species  $B$  with  $V[B]_\infty = E[A]_\infty$ , hence behaving as a Poisson process.

We can also derive an analytic solution for the transient evolution of the expectation of  $A$  and  $B$ . For  $A$  we have that

$$E[A(t)] = 2\frac{k_p}{k_d}(1 - \exp^{-k_d t}),$$

while for  $B$

$$E[B(t)] = 2\frac{k_p}{k_d(k_1 - k_d)}(-k_d + k_d \exp^{-k_1 t} + k_1 - k_1 \exp^{-k_d t}).$$

As expected, for  $t \rightarrow \infty$  we have  $E[B(t)] = E[A(t)]$ . Moreover,  $k_1$  and  $k_d$  control how fast  $E_{\mathcal{D}}[B(t)]$  tracks  $E[A(t)]$ . It is also easy to see that for  $k_1 \rightarrow \infty$  we have  $E[A(t)] = E[B(t)]$ , for any  $t \geq 0$ . This shows a trade-off between how fast  $B$  keeps tracking  $A$  and the quantity of noise that is filtered out.  $\mathcal{F}$  acts as a buffer, which introduces a delay to filter the fast dynamics of  $A$ .  $\mathcal{F}$  is indeed a molecular implementation of a first-order low-pass filter, that is, a molecular circuit that enables attenuating the high frequency components of a signal by introducing a delay. However, despite its simplicity, the performance of  $\mathcal{F}$  is limited by Poisson noise, in the sense that  $V[B(\infty)] \geq E[A(\infty)]$ . More specifically, noise reduction is lower bounded by the expectation of  $A$ .

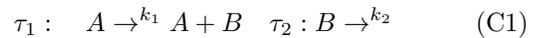
By using a more complex filter with non-linear dynamics (in the form of bi-molecular reactions), we can obtain much better performance in term of noise reduction at the cost of a more complex circuit. This annihilation filter,  $\mathcal{A}$ , is given by the following CRN with mass action kinetics



where  $A$  is the input species and  $C$  is the output filtered species. The filter is composed of two feed-forward loops [24]: a coherent one and an incoherent one. The two feed-forward loops are used to make  $C$  follow  $A$  with a delay. The annihilation reaction between  $B$  and  $E$ , which is supposed to be fast compared to the other reactions, and the fact that the production of  $B$  and  $E$  are correlated, enables noise reduction. In fact, consuming both  $B$  and  $E$ , this reaction acts as a selector which selectively enables only one feed-forward loop at the time according to whether  $C$  must be increased or decreased. The resulting network is therefore able to follow the long term evolution of a signal and, at the same time, to filter the noise. In Figure S2 we compare the performance of  $\mathcal{F}$  and  $\mathcal{A}$  for  $k_1 = 0.025$ ,  $r_1 = 1$ ,  $r_2 = 10$ ,  $r_3 = 0.0008$ .

#### 1. SI Stochastic Analysis of First-Order Low-Pass Filters

Consider the CRN  $\mathcal{F}$  given by the following reactions



where  $A$  is the input species and  $B$  is the output species. Suppose  $\mathcal{F}$  is embedded in an arbitrary CRN  $\mathcal{C} = (\Lambda, R)$ , with the restriction that  $B$  is changed only through  $\mathcal{F}$ . That is,  $\{A, B\} \subseteq \Lambda, v_{\tau, B} = 0$  for  $\tau \neq \tau_1$  and  $\tau \neq \tau_2$ , where  $v_{\tau, B}$  is the state change of reaction  $\tau$  relative to species  $B$ . We do allow the use of  $B$  as a catalyst in arbitrarily many reactions (i.e.  $C + B \rightarrow B + D$ ). Assuming the steady state solution of the CME exists, the steady state solution of the  $k$ th moment of  $B$  can be derived using Eqn (B1) as

$$\begin{aligned} 0 &= E\left[\sum_{\tau \in R} \alpha_{\tau}(X)((B + v_{\tau, B})^k - B^k)\right]_{\infty} = \\ &E\left[\sum_{\tau \in R \text{ s.t. } v_{\tau, B} \neq 0} \alpha_{\tau}(X)((B + v_{\tau, B})^k - B^k)\right]_{\infty} + \\ &E\left[\sum_{\tau \in R \text{ s.t. } v_{\tau, B} = 0} \alpha_{\tau}(X)((B)^k - B^k)\right]_{\infty} = \\ &E[k_1 A((B + 1)^k - B^k)]_{\infty} + E[k_2 B((B - 1)^k - B^k)]_{\infty}. \end{aligned}$$

For  $k = 1$ , we can derive the analytical form of the expectation of  $B$  at steady state as

$$E[B]_{\infty} = \frac{k_1}{k_2} E[A]_{\infty}.$$

That is,  $B$  follows  $A$ , but with an amplification factor dependent on the parameter rates. The second moment of  $B$  at steady state can be derived using  $k = 2$ . Moreover, as  $E[B]_{\infty} = \frac{k_1}{k_2} E[A]_{\infty}$ , we have

$$E[B^2]_{\infty} = \frac{k_1}{k_2} E[AB]_{\infty} + E[B]_{\infty},$$

which leads to the following form for the variance of  $B$  at steady state

$$V[B]_{\infty} = E[B^2]_{\infty} - E[B]_{\infty}^2 = E[B]_{\infty} + \frac{k_1}{k_2} Cov[A, B]_{\infty} \quad (\text{C2})$$

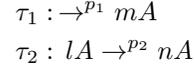
where  $Cov[A, B]_{\infty} = E[AB]_{\infty} - E[A]_{\infty}E[B]_{\infty}$  is the covariance between  $A$  and  $B$ . Assuming,  $A$  and  $B$  are non-negatively correlated, this leads to the following lower bound for the variance of  $B$ : for any  $A$ , the variance of  $B$  is such that

$$V[B]_{\infty} \geq E[B]_{\infty}.$$

The assumption that the correlation between  $A$  and  $B$  is positive seems natural, as  $A$  catalyzes the production of  $B$ . In fact, in the next subsection we show that this is true for a large class of inputs. However, we then show that, if a negative feedback between  $A$  and  $B$  is present in the larger network, then the correlation between  $A$  and  $B$  may become negative, and the noise of  $B$  can be reduced below Poisson levels.

## 2. Systems Without Feedback Loops

Consider a molecular signal  $A$  produced by the following class of reactions



where  $m, l, n \in \mathbb{N}$  and  $n < l$ . That is,  $A$  is co-expressed with a certain rate  $p_1$ , and can degrade non-linearly with rate  $p_2$ . For instance,  $\mathcal{C}_1 : \rightarrow^{p_1} A; A \rightarrow^{p_2}$  yields a Poisson process, while  $\mathcal{C}_2 : \rightarrow^{p_1} A + A; A \rightarrow^{p_2}$  yields a super-Poisson process (variance greater than mean), and  $\mathcal{C}_3 : \rightarrow^{p_1} A; A + A \rightarrow^{p_2}$  yields a sub-Poisson process (variance smaller than mean). The resulting CRN is in general non-linear. As a consequence, Eqn (B1) cannot be used directly, because it would lead to an infinite system of ODEs. Therefore, to analyze the system we consider the LNA. The rate equations of  $A$  (Eqn (B2)) can be written as

$$\frac{d\Phi_A(t)}{dt} = F_A(t) = mp_1 - (l - n)p_2\Phi_A(t)^l \quad (\text{C3})$$

The partial derivative of  $F_A(t)$  with respect to  $\Phi_A(t)$  is

$$J_A(t) = -l(l - n)p_2\Phi_A(t)^{l-1}. \quad (\text{C4})$$

Assume  $A$  is the input of the linear filter  $\mathcal{F}$  (Eqn (C1)), then using Eqn (B5) we can describe the covariance of the system,  $Cov(t)$ , as solution of the following ODEs

$$Cov(t) = J(t)Cov(t) + Cov(t)J^T + W(t) \quad (\text{C5})$$

where

$$Cov(t) = \begin{pmatrix} V[A(t)] & Cov[A(t), B(t)] \\ Cov[A(t), B(t)] & V[B(t)] \end{pmatrix},$$

$$J(t) = \begin{pmatrix} J_A(t) & 0 \\ k_1 & -k_2 \end{pmatrix}$$

$W(t) =$

$$\begin{pmatrix} m^2 p_1 + (l - n)^2 (p_2 \Phi_A(t)^l) & 0 \\ 0 & k_1 \Phi_A(t) + k_2 \Phi_B(t) \end{pmatrix}$$

Solving Eqn (C5) at steady state, we obtain the following equations for  $Cov[A, B]_{\infty}$

$$0 = Cov[A, B]_{\infty} J_A + V[A]_{\infty} k_1 - Cov[A, B]_{\infty} k_2 \quad (\text{C6})$$

where  $J_A = \lim_{t \rightarrow \infty} J_A(t)$ . Eqn (C6) can be rewritten as

$$Cov[A, B]_{\infty} = V[A]_{\infty} \frac{k_1}{k_2 - J_A}.$$

$J_A$  is always non positive. As a consequence, we have that, for any  $k_1, k_2 \in \mathbb{R}_{\geq 0}$ ,  $Cov[A, B]_{\infty} \geq 0$ . Moreover, as

$$V[B]_{\infty} = E[B]_{\infty} + \frac{k_1}{k_2} Cov[A, B]_{\infty},$$

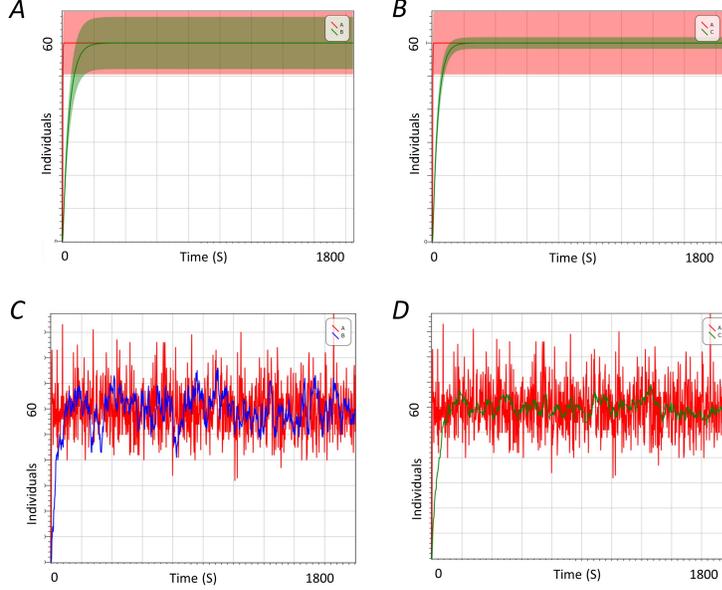


FIG. S2. Comparison of the actions of  $\mathcal{F}$  and  $\mathcal{A}$  on CRN  $\mathcal{C}$  for  $k_1 = 0.025$ ,  $r_1 = 1$ ,  $r_2 = 10$ ,  $r_3 = 0.0008$ . Plots compare expected value and variance of  $A$  with species  $B$  in  $\mathcal{F}$  (A) and with species  $C$  in  $\mathcal{A}$  (B). (C and D) Stochastic simulation plots of  $A$  (red) with  $B$  (blue) and  $C$  (green).

we have that, as expected

$$V[B]_\infty \geq E[B]_\infty \quad \text{for any } k_1, k_2 \in \mathbb{R}_{\geq 0}.$$

Also, assume  $k_1 = k$  and  $k_2 = km$ ,  $m \in \mathbb{R}_{\geq 0}$ , then we have

$$\lim_{k \rightarrow 0} V[B]_\infty = E[B]_\infty$$

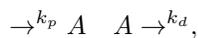
and

$$\lim_{k \rightarrow \infty} V[B]_\infty = E[B]_\infty + \frac{k_1}{k_2} V[A]_\infty = E[B]_\infty + \frac{1}{m} V[A]_\infty.$$

This shows that the noise is lower bounded by Poisson dynamics, which is approached for rates slow enough. Instead, if the rates of the filter are fast, then a noise term is added dependent on the variance of the input process.

### 3. Higher-Order linear Filters

We have considered first-order low-pass filters and demonstrated that their molecular implementation is surprisingly simple, yet they have many appealing properties. However, the price to pay for their simplicity is that the performance of noise reduction may be unsatisfactory in some cases, and it may also happen that the output signal is more noisy than input. For instance, if  $A$  is produced as a birth-death process



then for the output species  $B$  of  $\mathcal{F}$  we have that  $E[B(t)] \geq E[A(t)]$ ,  $t \in \mathbb{R}_{\geq 0}$ , for any possible rate parameters. Thus, a simple first-order filter cannot reduce the stochastic fluctuations of  $A$  in this case.

A possible way to improve first-order filters could be to consider second-order filters, that is, filters with transfer function of the type

$$F(s) = \frac{c}{(s+a)(s+b)}.$$

In fact, these filters are known to guarantee better performance because of their sharper Bode diagram. A possible molecular implementation of such filters is given by a cascade of first-order low-pass filters



In fact, the transfer function of such a CRN is

$$\frac{k_1 \cdot k_2}{(s + k_1)(s + k_2)}$$

However, a frequency characterization of such systems does not take into account the intrinsic noise introduced by the filter itself. In fact, each of the filters in the cascade, being linear, is still limited by Poisson noise. Therefore, higher-order filters, if linear, do not improve the performance of first-order filters (see Figure S3) since they are limited by Poisson noise.

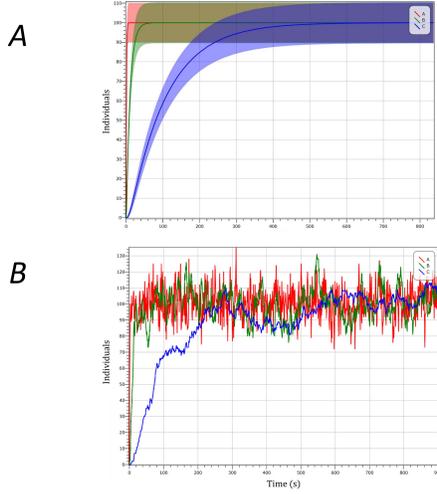
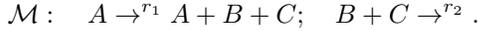


FIG. S3. Comparison of the evolution of species  $A, B, C$  for  $A$  a birth-death process with mean at steady state of 100 and  $C, B$  internal and output species of a second-order low-pass filter for  $k_1 = 0.1$  and  $k_2 = 0.01$ . (A) Plots of the time evolution of expectation and variance. Note that the variance of  $C$  is not reduced with respect to the variance of  $B$ . (B) Comparison of the species in terms of stochastic simulation.

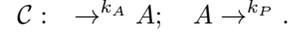
#### Appendix D: SI annihilation module

The *annihilation module*,  $\mathcal{M}$ , is composed of the following reactions:



In order to study the stochastic properties of such a system, we consider an input species  $A$  that implements a

Poisson process. That is,  $A$  is generated by the following reactions



Using Eqn (B1), it is easy to see that

$$E[A]_\infty = V[A]_\infty = \frac{k_A}{k_P}.$$

To study the variance of  $C$ , we cannot use Eqn (B1) directly, because it would lead to an infinite system of ODEs. Instead, we make use of the LNA (Eqn (B4) and Eqn (B5)), which allows us to derive a closed form system for mean and variance of  $B$  and  $C$ . For the expectation we have the following relation, which holds at steady state

$$r_1 E[A]_\infty = r_2 E[B]_\infty E[C]_\infty \quad (D1)$$

In order to obtain a finite number of solutions we introduce the following additional constraint, which holds assuming  $B$  and  $C$  start from the same initial condition

$$E[B]_\infty = E[C]_\infty.$$

Now, it is easy to see that

$$E[B]_\infty = E[C]_\infty = \sqrt{\frac{r_1}{r_2} E[A]_\infty}.$$

For the variance, solving Eqn (B5) at steady state, we obtain the following equations

$$\begin{aligned} -E[B]_\infty \text{Cov}[A, C]_\infty r_2 - E[C]_\infty \text{Cov}[A, B]_\infty r_2 + V[A]_\infty r_1 - \text{Cov}[A, B]_\infty k_P &= 0 \\ -E[B]_\infty \text{Cov}[A, C]_\infty r_2 - E[C]_\infty \text{Cov}[A, B]_\infty r_2 + V[A]_\infty r_1 - \text{Cov}[A, C]_\infty k_P &= 0 \\ E[A]_\infty r_1 + E[B]_\infty E[C]_\infty r_2 - E[B]_\infty V[B]_\infty r_2 - E[B]_\infty V[C]_\infty r_2 - \\ \text{Cov}[B, C]_\infty r_2 - E[C]_\infty \text{Cov}[B, C]_\infty r_2 + \text{Cov}[A, B]_\infty r_1 + \text{Cov}[A, C]_\infty r_1 &= 0 \\ Ar_1 + E[B]_\infty E[C]_\infty r_2 - 2E[B]_\infty \text{Cov}[B, C]_\infty r_2 - 2E[C]_\infty V[B]_\infty r_2 + 2\text{Cov}[A, B]_\infty r_1 &= 0 \\ E[A]_\infty r_1 + E[B]_\infty E[C]_\infty r_2 - 2E[B]_\infty V[C]_\infty r_2 - 2E[C]_\infty \text{Cov}[B, C]_\infty r_2 + 2\text{Cov}[A, C]_\infty r_1 &= 0 \end{aligned}$$

In order to derive a system of ODEs with a finite number of solutions we need to introduce an additional constraint, which will depend on the initial conditions. Since initially all the variances and covariances are 0 (we start from a known initial condition), from Eqn (B5) we can add the following constraint

$$2V[B]_\infty = \text{Cov}[B, C]_\infty + E[C]_\infty.$$

Solving the resulting system of equations, we obtain the following solution for the Fano factor of  $C$  and  $B$  at

steady state

$$F_C = F_B = \frac{2r_1^{3/2} \sqrt{r_2 k_A k_P} + 4r_1 r_2 k_A - r_1 k_P^2 - k_P^3}{8r_1 r_2 k_A - 2k_P^3} \quad (D2)$$

Assume  $r_1 = r, r_2 = rm, r, m \in \mathbb{R}_{\geq 0}$ , then for  $r \rightarrow 0$  we obtain

$$F_C = F_B = \frac{-k_P^3}{-2k_P^3} = \frac{1}{2},$$

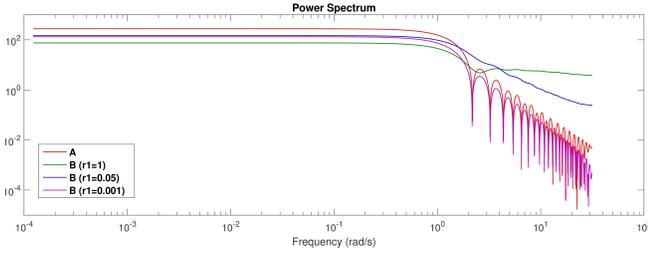


FIG. S4. Comparison of the PSD of the input signal  $A$  affected by Poisson noise for different implementations of the annihilation filter, varying the parameter rates, and for  $E[A]_\infty = E[B]_\infty$ . The power spectrum is estimated using the Blackman-Tukey Spectral Estimate algorithm with frequency deep resolution, as implemented in Matlab, over 300000 data points with sampling time of 0.1 seconds. In all the different cases, the annihilation module attenuates the low frequency components of the input signal, while it can amplify the high frequency if the rates are faster than those of the input signal.

while for  $r \rightarrow \infty$ , we get

$$F_C = F_B = \frac{2r_1^{3/2} \sqrt{r_2 k_A k_P} + 4r_1 r_2 k_A}{8r_1 r_2 k_A} = \frac{1}{4} \sqrt{\frac{k_1 k_P}{k_2 k_A}} + \frac{1}{2} = \frac{2+n}{4},$$

where  $n = \frac{E[C]_\infty}{E[A]_\infty} = \sqrt{\frac{k_1 k_A k_P}{k_2 k_P k_A}}$ .

### 1. Noise Reduction Performance of the annihilation module

In the previous subsection we show that the annihilation module can reduce the Fano factor of an input signal below Poisson levels (Fano Factor smaller than 1) and independently of the delay introduced. In order to explain why this is possible, we consider an input process

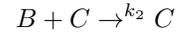
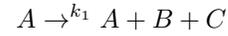
$A$  described by the following reactions



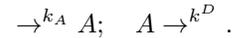
that is,  $A$  is affected by Poisson noise. In Figure S4, we compare the Power Spectral Density (PSD) of the input  $A$  with that of the output  $B$  of the annihilation filter, for different rate parameters. We select  $r_1, r_2$  such that  $E[B]_\infty = E[A]_\infty$ . For any choice of  $r_1$  we have that  $B$  attenuates the low frequency components of the input signal. This causes a reduction of the Fano factor independently of the value of  $r_1$ . However, if  $r_1$  is large, the high frequency components tend to be amplified, while for  $r_1$  small enough these are attenuated. This explains why noise attenuation is greater for small rates. Therefore, the annihilation module reduces the noise by attenuating the low frequency components of a signals, and, if its dynamics are slow enough, it reduces the noise further by integrating the fast dynamics.

### 2. SI Incoherent Feed-Forward Loop is Limited by Poisson Levels

The annihilation module has many similarities with the incoherent feed-forward module [24], which, in its linear form, can be described by the following reactions



where  $A$  is the input species and  $B$  is the output. That is,  $B$  and  $C$  are co-expressed, and then  $C$  represses  $B$ . We assume  $A$  is affected by Poisson noise, that is,  $A$  is produced by the following reactions



Then, using the LNA, we can calculate the expectation and Fano factor of  $B$  at steady state as

$$E[B]_\infty = \frac{k_3}{k_2}$$

$$F_B = \frac{k_1^2 k_2 k_A k_D n + k_1^2 k_D^3 n + k_1^2 k_D^3 + k_1 k_2^2 k_A^2 n^3 + 2k_1 k_2 k_A k_D^2 n^2}{(k_2 k_A n + k_D^2) (k_1^2 k_D + k_1 k_2 k_A n^2 + k_1 k_D^2 n + k_2 k_A k_D n^3)} + \frac{k_1 k_D^4 n + k_2^2 k_A^2 k_D n^4 + k_2 k_A k_D^3 n^3}{(k_2 k_A n + k_D^2) (k_1^2 k_D + k_1 k_2 k_A n^2 + k_1 k_D^2 n + k_2 k_A k_D n^3)}$$

where, in the Fano factor, we constrain  $E[B]_\infty = nE[A]_\infty$  for a constant  $n \in \mathbb{R}_{>0}$ . From now on, for

simplicity, and without loss of generality, we assume  $\frac{k_A}{k_D} = 100$ , and in Figure S5 we plot the Fano factor of

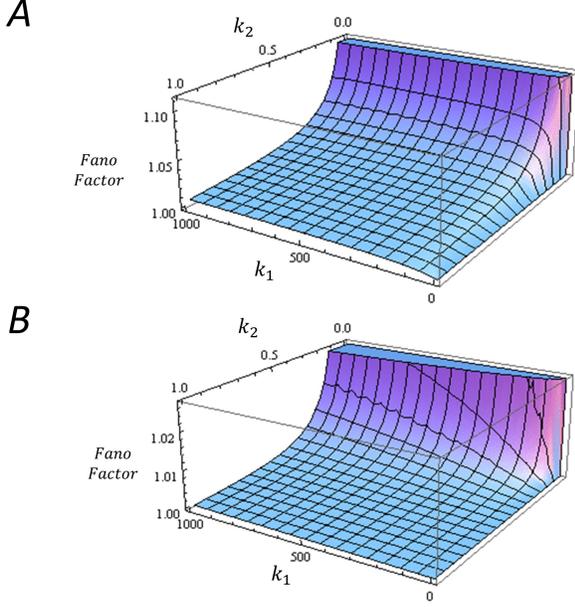


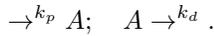
FIG. S5. Plot of  $F_B$  for  $n = 1$  (A) and  $n = 10$  (B)

$B$  at steady state as a function function of  $k_1$  and  $k_2$  for different values of  $n$ . The Fano factor is lower bounded by 1 (Poisson dynamics). This shows the importance of having  $B$  and  $C$  degraded together, in order to reduce fluctuations below Poisson statistics. Note also that in this case the expectation of the output does not depend on the input.

## Appendix E: annihilation Filter

### 1. Stochastic Analysis of annihilation Filter

To study the noise reduction capabilities of the annihilation filter, we make use of the LNA. We assume  $A$  is a general input process, with extrinsic noise modeled by a Poisson process. That is,  $A$  is generated and removed by the following reactions



Thus, we have  $E[A(t)] = V[A(t)], t \in \mathbb{R}_{\geq 0}$ . Using the LNA equations for the mean we can describe the time

evolution of the species in  $\mathcal{F}_p$  with the following ODEs

$$\begin{aligned} \frac{dE[D(t)]}{dt} &= r_1 E[A(t)] - r_1 E[D(t)] \\ \frac{dE[B(t)]}{dt} &= r_1 E[A(t)] - r_2 E[B(t)]E[E(t)] - r_3 E[B(t)]E[A(t)] \\ \frac{dE[E(t)]}{dt} &= r_1 E[D(t)] - r_2 E[B(t)]E[E(t)] - r_3 E[E(t)]\Phi_C(t) \\ \frac{dE[C(t)]}{dt} &= r_3 E[B(t)]E[A(t)] - r_3 E[E(t)]E[C(t)] \end{aligned}$$

From these equations we can derive the following relations at steady state:

$$E[A]_\infty = E[D]_\infty \quad (\text{E1})$$

$$E[C]_\infty = E[D]_\infty \frac{E[B]_\infty}{E[E]_\infty} \quad (\text{E2})$$

$$0 = r_1 E[D]_\infty - r_2 E[B]_\infty E[E]_\infty - r_3 E[E]_\infty E[C]_\infty. \quad (\text{E3})$$

It is easy to see that we have an infinite number of possible solutions depending on the initial conditions. By integrating the differential equations for  $E[B]$  and  $E[E]$ , we get

$$\begin{aligned} E[B(t)] - E[E(t)] &= \\ (E[B(0)] - E[E(0)]) + \int_0^t r_1 E[A(s)] - r_1 E[D(s)] ds - \\ \int_0^t r_3 E[B(s)]E[A(s)] - r_3 E[E(s)]E[C(s)] ds, \end{aligned}$$

which, assuming  $E[D(0)] = E[C(0)]$ , leads to the following relation

$$\begin{aligned} E[B(t)] - E[E(t)] &= \\ E[B(0)] - E[E(0)] + E[D(t)] - E[C(t)] \quad (\text{E4}) \end{aligned}$$

By solving Eqn (E4), Eqn (E1), Eqn (E2), Eqn (E3) at steady state for  $E[B(0)] - E[E(0)] = 0$ , we obtain that  $E[D]_\infty = E[C]_\infty = EA]_\infty$ , that is, perfect tracking of the  $A$ .

Now, from the moment equations (Eqn (B1)), we get the following exact expression for  $E[C(t)]$ ,

$$\frac{dE[C(t)]}{dt} = r_3 E[B(t)]A(t) - r_3 E[E(t)]C(t),$$

which leads to

$$E[BA]_\infty = E[EC]_\infty.$$

The above relation, for  $E[D(0)] = E[C(0)] = E[B(0)] = E[E(0)] = 0$  with Eqn (E2), which holds under the LNA assumption, leads to

$$\text{Cov}[B, A]_\infty = \text{Cov}[E, C]_\infty. \quad (\text{E5})$$

Solving (E5) with Equations (B5), we can derive a closed form for the Fano Factor of  $C$  at steady state. However, the resulting expression is too complex to be useful. Nevertheless, we can assume  $r_2 = \frac{T}{r}$ ,  $r_3 = r$ , where  $T, r$  are constants. Then, for  $r \rightarrow 0$ , we obtain the following form for the Fano factor of  $C$

$$\lim_{r \rightarrow 0} F_C = \frac{k_p}{k_p + r_1}.$$

As  $r_1$  acts as a sensor, we have that increasing  $r_1$ , the noise can be made arbitrarily small.

#### a. Frequency Analysis of annihilation Filter

For  $C$ , output of the annihilation filter, the following linearized ODE can be derived

$$\frac{d\Phi_C(t)}{dt} = -r_3\Phi_E(t)C^{eq} - r_3E^{eq}\Phi_C(t) + r_3\Phi_B(t)A^{eq} + r_3B^{eq}\Phi_A(t)$$

where  $A^{eq}, C^{eq}, B^{eq}, E^{eq}$  are the values of the species at the equilibrium point. In the equilibrium point we have that  $A^{eq} = C^{eq}$ , that is,  $C$  tracks  $A$  and  $B^{eq} = E^{eq} = b$ . As a consequence, for dynamics near to the equilibrium, we can write

$$\frac{d\Phi_C(t)}{dt} = -r_3E^{eq}\Phi_C(t) + r_3B^{eq}\Phi_A(t)$$

Fourier transforming both terms, we get

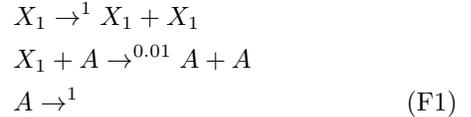
$$C(\omega) = \frac{1}{1 + \frac{i\omega}{r_3b}} A(\omega),$$

where  $C(\omega)$  is the Fourier transform of  $\Phi_C$ . Therefore,  $C$  acts as a first-order low-pass filter on the input  $A$  with cutoff frequency  $\bar{\omega} = r_3b$ . The spectrum of the transfer function between  $C$  and  $A$ , near to the equilibrium point,

is as shown in Figure 1 in the main text: a delay dependent on  $r_3$  is introduced in order to reduce the noise at the high frequencies.

## Appendix F: SI Time-Varying Input Analysis

In Figure S6 we compare the action of linear filter (CRN (1)), annihilation module (CRN (7)), and annihilation filter (CRN (10)) for an oscillatory input  $A$ , described by the following reactions



where  $X_1$  is an auxiliary species for initial condition of  $X_1$  and  $A$  of respectively 200 and 100 molecules. The above reactions implement a Lotka-Volterra oscillator. Hence, the Fano factor of  $A$  is due to two components: phase and amplitude of the oscillations, which may change at any oscillation period, and the fast fluctuations that corrupt the oscillatory behaviour. In this case, the goal of the filtering process should be to produce an output process with the same oscillatory behaviour of  $A$ , but with reduced fast fluctuations. Hence, the alone annihilation module cannot perform well for such a goal. In fact, we showed that the annihilation module reduces the low frequency components of the spectrum of  $A$ . This implies that the output of the annihilation module will have attenuated oscillations compared to  $A$ . The annihilation filter, instead, correctly reduces the fast fluctuations while still maintaining similar profile of the oscillations and not introducing any other noise, as it happens for the linear filter.

## Appendix G: SI Code

In Figure S7 we report the code used for stochastic simulations and LNA computation in Figure 3. In Figure S8 we report the code used for stochastic simulations and LNA computation in Figure S6. The figures are obtained using the Microsoft Visual GEC tool [30].

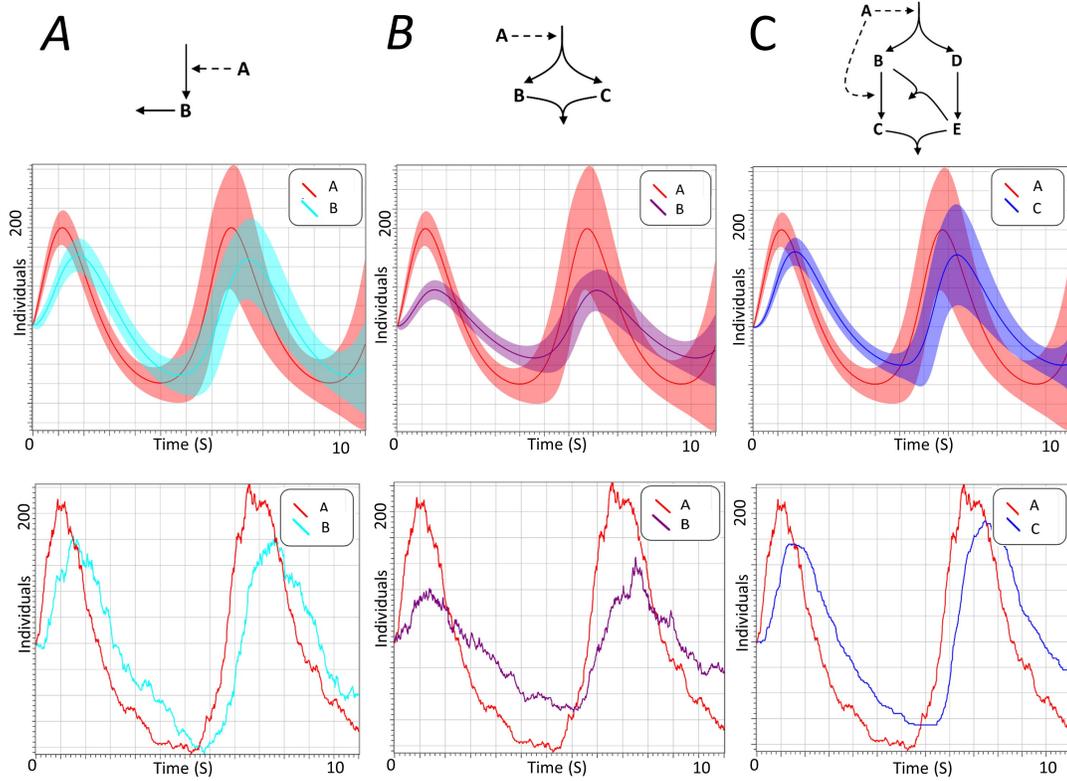


FIG. S6. Comparison of linear filter (A), annihilation module (B), and annihilation filter (C) for an oscillatory input as described in Eqns (F1). Details of the parameters used are shown in Figure S8. In the second row we plot the expectation and standard deviation of input and output of each filter according to the LNA. In the third row we show a stochastic simulation of input and output of each filter.

```

A
directive sample 4000.0 1000
directive simulation stochastic
directive plot A;B_Annh_Mod;C;B_Lin_Filt;C_Annh_Filt

//Produce X at time 1200
directive event X 1.0 @ 1200.0

//Input Generation
X->{0.0005} |
X ->{100} X + A + A + A |
->{33.33333} A + A + A |
A ->{1} |

//Linear Filter
A ->{0.0064} A + B_Lin_Filt |
B_Lin_Filt ->{0.0064} |

//Annihilation Module
A ->{0.005} A + B_Annh_Mod +C |
B_Annh_Mod + C->{0.0005} |

//Annihilation Filter
A ->{100} A + B1 + B2 |
B2 ->{100} B22 |
B1 + B22 ->{1000} |
B1 + A ->{0.000055} A + C_Annh_Filt |
C_Annh_Filt + B22 ->{0.000055}

B
directive sample 1200.0 1000
directive simulation lnastiff
directive plot A;B_Annh_Mod;C;B_Lin_Filt;C_Annh_Filt

//Produce X at time 1200
//directive event X 1.0 @ 1200.0

//Input Generation
X->{0.0005} |
X ->{100} X + A + A + A |
->{33.33333} A + A + A |
A ->{1} |

//Linear Filter
A ->{0.0064} A + B_Lin_Filt |
B_Lin_Filt ->{0.0064} |

//Annihilation Module
A ->{0.005} A + B_Annh_Mod +C |
B_Annh_Mod + C->{0.0005} |

//Annihilation Filter
A ->{100} A + B1 + B2 |
B2 ->{100} B22 |
B1 + B22 ->{1000} |
B1 + A ->{0.000055} A + C_Annh_Filt |
C_Annh_Filt + B22 ->{0.000055}

```

FIG. S7. Microsoft VisualGEC code for stochastic simulations (A) and LNA computation (B) of Figure 3.

<b>A</b>	<b>B</b>
<pre> directive sample 10 1000 directive simulation stochastic directive plot A;AnnihilationFilter;AnnihilationModule;LinearFilter  //Input init X1 200   init A 100    X1 -&gt;{1} X1 + X1   X1 + A -&gt;{0.01} A + A   A -&gt;{1}    //Annihilation filter A -&gt;{100} A + B + D   D -&gt;{100} E   E + B -&gt;{10000}   B + A -&gt;{0.008} A + AnnihilationFilter   AnnihilationFilter + E -&gt;{0.008}    //Linear filter A -&gt;{1.2} A + LinearFilter   LinearFilter -&gt;{1.2}    //Annihilation module A -&gt;{1} A + Bi + AnnihilationModule   Bi + AnnihilationModule -&gt;{0.01} </pre>	<pre> directive sample 10 1000 directive simulation lnastiff directive plot A;AnnihilationFilter;AnnihilationModule;LinearFilter  //Input init X1 200   init A 100    X1 -&gt;{1} X1 + X1   X1 + A -&gt;{0.01} A + A   A -&gt;{1}    //Annihilation filter A -&gt;{100} A + B + D   D -&gt;{100} E   E + B -&gt;{10000}   B + A -&gt;{0.008} A + AnnihilationFilter   AnnihilationFilter + E -&gt;{0.008}    //Linear filter A -&gt;{1.2} A + LinearFilter   LinearFilter -&gt;{1.2}    //Annihilation module A -&gt;{1} A + Bi + AnnihilationModule   Bi + AnnihilationModule -&gt;{0.01} </pre>

FIG. S8. Microsoft VisualGEC code for stochastic simulations (*A*) and LNA computation (*B*) of Figure S6.