Postprint

This is the accepted version of a paper published in *Journal of the Royal Society Interface*. This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Stochastic focusing coupled with negative feedback enables robust regulation in biochemical reaction networks

Andreas Milias-Argeitis[1], Stefan Engblom[2], Pavol Bauer[2], and Mustafa Khammash[*1]

[1]Department of Biosystems Science and Engineering, ETH Zurich, Mattenstrasse 26, 4058 Basel, Switzerland

[2]Division of Scientific Computing, Department of Information Technology, Uppsala University, P.O. Box 337, SE-75105 Uppsala, Sweden

Abstract

Nature presents multiple intriguing examples of processes which proceed at high precision and regularity. This remarkable stability is frequently counter to modelers' experience with the inherent stochasticity of chemical reactions in the regime of low copy numbers. Moreover, the effects of noise and nonlinearities can lead to "counter-intuitive" behavior, as demonstrated for a basic enzymatic reaction scheme that can display *stochastic focusing* (SF). Under the assumption of rapid signal fluctuations, SF has been shown to convert a graded response into a threshold mechanism, thus attenuating the detrimental effects of signal noise. However, when the rapid fluctuation assumption is violated, this gain in sensitivity is generally obtained at the cost of very large product variance, and this unpredictable behavior may be one possible explanation of why, more than a decade after its introduction, SF has still not been observed in real biochemical systems.

In this work we explore the noise properties of a simple enzymatic reaction mechanism with a small and fluctuating number of active enzymes that behaves as a high-gain, noisy amplifier due to SF caused by slow enzyme fluctuations. We then show that the inclusion of a plausible negative feedback mechanism turns the system from a noisy signal detector to a strong homeostatic mechanism by exchanging high gain with strong attenuation in output noise and robustness to parameter variations. Moreover, we observe that the discrepancy between deterministic and stochastic descriptions of stochastically focused systems in the evolution of the means almost completely disappears, despite very low molecule counts and the additional nonlinearity due to feedback.

The reaction mechanism considered here can provide a possible resolution to the apparent conflict between intrinsic noise and high precision in critical intracellular processes.

1 Introduction

Random fluctuations due to low-copy number phenomena inside the microscopic cellular volumes have been an object of intense study in recent years. It is now widely recognized that deterministic modeling of chemical kinetics is in many cases inadequate for capturing even the mean behavior of stochastic chemical reaction networks, and several studies have explored the discrepancy between deterministic and stochastic system descriptions [1, 23, 33, 26].

---

[*]mustafa.khammash@bsse.ethz.ch

Despite the all-pervasive stochasticity, cellular processes and responses proceed with surprising precision and regularity, thanks to efficient noise suppression mechanisms also present within cells. The structure and function of these mechanisms has been a topic of great interest [7, 2, 32, 31, 15, 22, 3, 26], and in many cases still remains unknown.

Moreover, recent theoretical works on enzymatic reaction schemes with a single or a few enzyme molecules [18, 11, 13, 29] have repeatedly shown that low-copy enzymatic reactions demonstrate a stochastic behavior that can lead to markedly different responses in comparison to the predictions of deterministic enzyme kinetic models.

In this work we investigate the properties of a possible noise suppression mechanism for an enzymatic reaction with a small and fluctuating number of active enzymes. Under certain conditions, presented in [23], this system displays an increased sensitivity to enzyme fluctuations, a phenomenon that has been termed *stochastic focusing*.

Stochastic focusing has been presented as a possible mechanism for *sensitivity amplification*: compared to a deterministic model of a biochemical network, the mean output of the stochastic version of the system can display increased sensitivity to changes in the input, when the input species has sufficiently low abundance. Consequently, it has been postulated that stochastic focusing can act as a signal detection mechanism, that converts a graded input into a "digital" output.

The basic premise of [23] has been that fluctuations in the "input" species are sufficiently rapid, so that any rates that depend on the signaling species show minimal time-correlations. We show that if this condition fails, i.e. when the fluctuations in the input signal are slow compared to the average lifetime of a substrate molecule, stochastic focusing can result in a dramatic increase in substrate fluctuations, a fact also acknowledged in the original publication. Increased sensitivity to input changes does not only come at the cost of extremely high output noise levels; as we will demonstrate here, systems operating in this regime are also extremely sensitive to variations in reaction rates, which in fact precludes robust signal detection by stochastic focusing.

For the first time since its introduction we could study the steady-state behavior of this system analytically, by formulating and solving the equations for the conditional means and (co)variances [14]. Motivated by our observations on the open-loop, stochastically focused system, we investigated the system behavior in the presence of a plausible feedback mechanism. We treated the enzyme as a noisy "controller" molecule whose purpose is to regulate the outflux of a reaction product by – directly or indirectly – "sensing" the fluctuations in its substrate. For the sake of simplicity and clarity, we focused on very simple and highly abstracted mechanisms, but we should remark that several possible biochemical implementations of our feedback mechanism can be considered.

Our premise was that the great open-loop sensitivity of a stochastically focused system with relatively slow input fluctuations creates a system with very high open-loop gain, which in turn can be exploited to generate a very robust closed-loop system once the output is connected to the input. Our simulation results confirmed this intuition, revealing a dramatic decrease in noise levels and a significant increase in robustness in the steady-state mean behavior of the closed-loop system. Such a system no longer functions as a signal detector, but rather behaves as a strong homeostatic mechanism. Moreover, we observed that the steady-state behavior of the means in a stochastically focused system with feedback can be captured quite accurately by the corresponding deterministic system of reaction rate equations, despite the fact that the stochastic system still operates at very low copy numbers.

Noise attenuation through feedback and the fundamental limits of any feedback system implemented with noisy "sensors" and "controllers" have been studied theoretically in the recent years, and some fairly general performance bounds have been derived in [19]. We should note

that, despite its generality, the modeling framework assumed in [19] does not apply in our case, since our system contains a controlled degradation reaction, whereas [19] considers only control of production. More specifically, [19] examines the case where a given species regulates its own production through an arbitrary stochastic signaling network. In this setting, it is shown that, no matter the form or complexity of the intermediate signaling, the loss of information induced by stochasticity places severe fundamental limits on the levels of noise suppression that such feedback loops can achieve. On the other hand, it is still unclear what type of noise suppression limitations are present for systems such as the one studied here, and a complete analytical treatment of the problem of regulated substrate degradation seems very difficult at the moment.

A first attempt to analyze the noise properties of regulated degradation was presented in [7], which examined such a scheme using the Linear Noise Approximation (LNA) [9]. As the authors of that work pointed out, however, the LNA is incapable of correctly capturing the system behavior (i.e. means and variances) beyond the small-noise regime, due to the nonlinear system behavior. We verified this inadequacy, not only for LNA, but for other approximation schemes as well, such as the Langevin equations [21] and various moment closure approaches [30]. Perhaps this is the reason why, contrary to regulated production, the theoretical noise properties of regulated substrate degradation have received relatively little attention.

With the rapid advancement of single-molecule enzymatic assays [27, 20], we expect that the study of noise properties of various low-copy enzymatic reactions, including the proposed feedback mechanism described here, will soon be amenable to experimental verification. It also remains to be seen whether the proposed feedback mechanism is actually employed by cells to achieve noise attenuation in enzymatic reactions. At any rate, the noise attenuation scheme presented here could be tried and tested in synthetic circuits through enzyme engineering [24, 4].

2 Results

2.1 A slowly fluctuating enzymatic reaction scheme that exhibits stochastic focusing

In this section we formulate and analyze a simple biochemical reaction network capable of exhibiting the dynamic phenomenon of stochastic focusing. It is shown that in the stochastic focusing regime, the system acts as a noisy amplifier with an inherent strong sensitivity to perturbations. It follows that without modifications, the network cannot be used under conditions requiring precision and regularity.

2.1.1 Modeling

We consider the simple branched reaction scheme studied in [23] and shown schematically in Fig. 1(a). In this scheme, substrate molecules C enter the system at a constant influx, and can either be converted into a product P or degraded under the action of a low-copy enzyme (or, equivalently, converted into a product that leaves the system). While the number of enzymes in the system is assumed constant, enzyme molecules can spontaneously fluctuate between an active (E^*) and an inactive (E) form. The generality of this model and its sensitivity to variations in the active enzyme levels is further discussed in the Supplement (sections 1 and 2).

Recent single-enzyme turnover experiments have shown that single enzyme molecules typically fluctuate between conformations with different catalytic activities, a phenomenon called *dynamic disorder* [18, 11, 13, 29]. In the simple model considered here, the enzyme randomly switches between two activity states. The stationary distribution of E^* in this case is known to

3

be binomial [35]; that is, $E^* \sim \mathrm{B}(N, p)$, where $N = E + E^*$ is the total number of enzymes in the system and $p = \alpha_E/(\alpha_E + \mu_E)$, such that the mean $\langle E^* \rangle = Np$.

The basic (empirically derived) conditions for stochastic focusing [23] are that the magnitude of active enzyme fluctuations is significant compared to the mean number of active enzymes, while the total number of enzymes is low. Moreover, it is assumed that the level of E^* fluctuates rapidly compared to the average lifetime of C and P molecules. Without this assumption, the noise in E^* can be greatly amplified by C and transmitted to P.

The first assumption (large enzyme fluctuations and low abundance) is maintained in our setup. However, we shall dispense with the second assumption. We further postulate that C (possibly a product of upstream enzymatic reactions) enters the system at a high input flux (large α_C) and that there exists a strong coupling between C and E^*, in the sense that a few active enzymes can strongly affect the degradation of C.
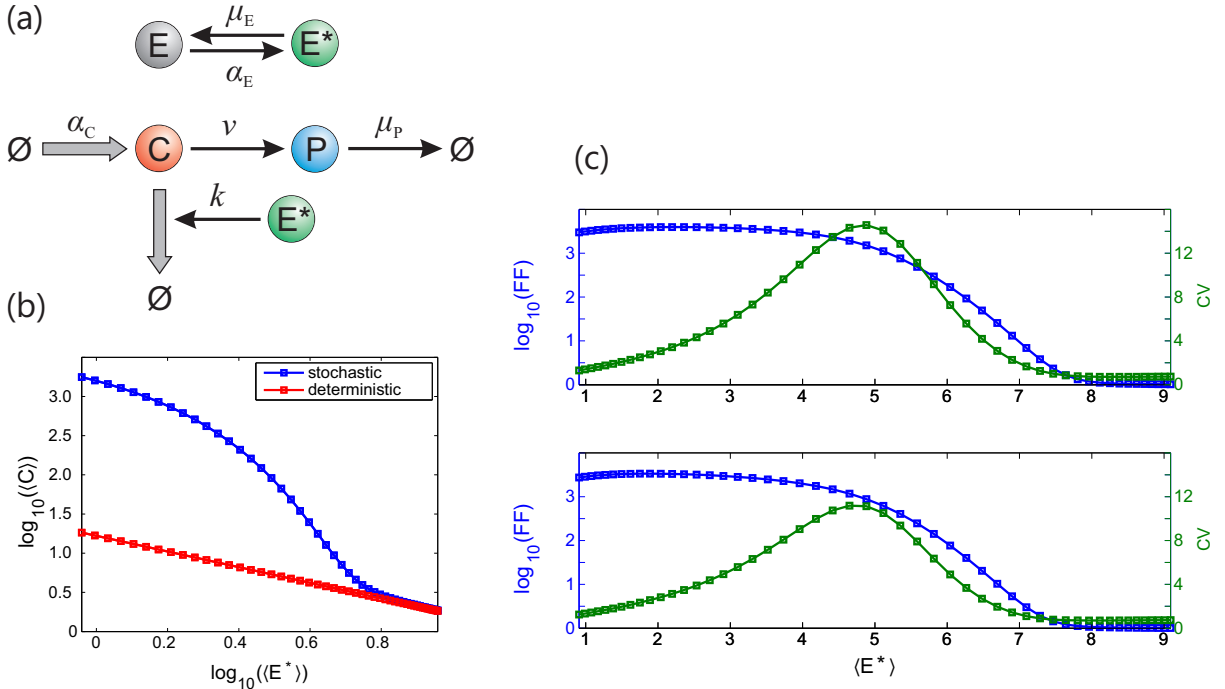

Figure 1: **(a)** The reaction scheme studied in this work. Thick arrows represent high-flux reactions. **(b)** Blue line: Steady-state mean of C (denoted by $\langle C \rangle$ throughout) as a function of the average number of active enzyme molecules ($N\alpha_E/(\alpha_E + \mu_E)$) for $N = 10$, $\mu_E = 0.1$, $\alpha_C = 5010$, $k = 300$, $\nu = 1$, $\mu_P = 1$. Since $\nu = \mu_P$, the average of P displays the same behavior as the substrate. Red line: Steady-state of the ODE model for the same parameter values. The large difference (notice the logarithmic scale) between the blue and red lines is a consequence of stochastic focusing [23]. **(c)** Upper: Substrate (C) noise statistics as a function of the average number of active enzymes. Blue: Steady-state Fano factor (variance/mean; notice the logarithmic scale). Green: Steady-state coefficient of variation (standard deviation/mean). Lower: Product (P) noise statistics as a function of the average number of active enzymes. Color coding same as in upper panel. Both plots were obtained for $N = 10$, $\mu_E = 0.1$, $\alpha_C = 5010$, $k = 300$, $\nu = 1$, $\mu_P = 1$. All calculations were performed analytically, using the conditional moment equations and the known stationary distribution of E^*.

More concretely, the previously stated assumptions imply that the reaction rates must satisfy the following conditions:

4

1. $\alpha_C \gg 1$ (high influx of C)

2. $k \gg \alpha_E + \mu_E$ (enzyme fluctuations are slow compared to the average lifetime of a substrate molecule)

3. $k \gg 1$ (strong coupling between enzyme and substrate)

4. $Np = N\dfrac{\alpha_E}{\alpha_E + \mu_E}$ is small (e.g. below 10).

These conditions are motivated via a short theoretical and numerical analysis in the Supplement (section 1). When they hold, we expect the amount of C to fluctuate wildly as E^* varies over time and these fluctuations to propagate to P. In the rest, we will refer to this motif as the (open-loop) *slowly fluctuating enzyme* (SFE) system.

The computational analysis of this and similar systems has thus far been hindered by the presence of the bimolecular reaction, which leads to statistical moment equations that are not closed [30], while the presence of stochastic focusing presents further difficulties for any moment closure method. In this work, we circumvent these difficulties by formulating and solving the conditional moment equations [14] for the means and (co)variances of C and P conditioned on the enzyme state (whose steady-state distribution is known). This enables for the first time the analytical study of the steady-state behavior of this system (more details can be found in the Supplement, section 3). Chiefly, the equations for the first two conditional moments of the SFE system are in fact closed, i.e. they do not depend on moments of order higher than 2, and thus do not require a moment closure approximation despite the fact that the unconditional moment equations themselves are open. We next use these analytic equations to shed new light on the properties of the network under consideration.

2.1.2 The SFE system functions as a noisy amplifier

According to the method of conditional moments (MCM) [14], the chemical species of a given system are divided into two classes. Species of the first class, collectively denoted by Y, are treated fully stochastically, while species of the second class, denoted by Z, are described through their conditional moments given Y. More analytically, the MCM considers a chemical master equation (CME) [36] for the marginal distribution of Y, $p(Y,t) = \sum_{Z \geq 0} p(Y,Z,t)$, and a system of conditional means ($\mu_Z(Y,t) := \mathbb{E}_{p(Z|Y,t)}[Z|Y,t]$) and higher-ordered centered moments (e.g. conditional variances $C_Z(Y,t) := \mathbb{E}_{p(Z|Y,t)}[(Z - \mu_Z(Y,t))^2|Y,t]$) for the Z species. In the case of the SFE network, by taking $Y = E^*$ and $Z = (C,P)$, we see that Y is independent of Z and its evolution is described by a CME whose stationary solution is known. Moreover, the system of conditional means and (co)variances of Z given Y turns out to be closed.

We thus begin by examining the behavior of the system as the enzyme activation rate is varied while keeping other parameters unchanged. Assuming a fixed μ_E, α_E is directly related to $\langle E^* \rangle$, the average number of active enzymes. Thus, any changes in $\langle E^* \rangle$ are assumed to be driven by α_E, which means that the two can be used interchangeably. We present the performance of the open-loop system with respect to $\langle E^* \rangle$ wherever possible, as we find this more intuitive.

The results in Fig. 1(b) show that the stationary means of C and P (denoted by angle brackets throughout the paper) depend very sensitively on $\langle E^* \rangle$, as one would expect from a stochastically focused system. Moreover, owing to the relatively slow switching frequency of enzyme states, the stationary distributions of substrate and product are greatly over-dispersed, as shown in Fig. 1(c).

Apart from the enzyme activation rate, the catalytic degradation rate (k) is also expected to affect noise in the system, as it controls both the timescale and magnitude of substrate fluctuations: as k increases, the rate of substrate consumption grows as well. On the other hand, the impact of a change in the number of active enzymes is also magnified. We can study the interplay of α_E and k by varying both simultaneously, as shown in Fig. 2(a). Although α_E has a much more pronounced effect on substrate and product means and variances, the interplay of k and α_E is what determines the overall noise strength in the system, as the third row of plots shows.
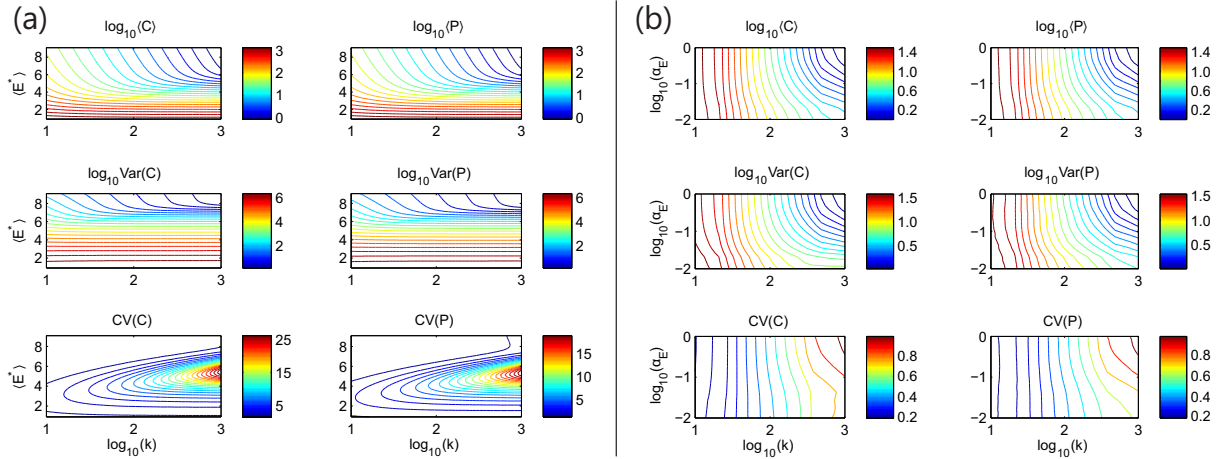

Figure 2: **(a)** Steady-state means, variances and CVs of substrate and product as a function of the average active enzymes ($\alpha_E N/(\alpha_E + \mu_E)$) and k for $N = 10$, $\mu_E = 0.1$, $\alpha_C = 5010$, $\nu = 1$, $\mu_P = 1$ (notice the logarithmic scales). **(b)** Closed-loop SFE system with substrate feedback: steady-state means, variances and CVs of substrate and product as a function of α_E (which determines the average number of active enzymes) and k for $K = 3$, $C_0 = 5$, $N = 10$, $\mu_E = 0.1$, $\alpha_C = 5010$, $\nu = 1$, $\mu_P = 1$. Logarithmic scales are preserved to make comparisons with Fig. 2 easier, although the range of variation is much smaller in this case.

From the above analysis, we deduce that the open-loop motif amplifies both small changes in the average number of active enzymes (Fig. 1(b)), as well as temporal fluctuations in the active enzyme levels (Fig. 1(c)): for intermediate values of α_E, the CV and FF of C and P are much greater than zero. This implies that the instantaneous flux of substrate through the two alternative pathways experiences very large fluctuations, which would propagate to any reactions downstream of C.

2.1.3 The SFE system is very sensitive to parameter perturbations

The increased sensitivity of the SFE network to fluctuations in the active enzyme would suggest sensitivity with respect to variations in reaction rates. To verify this, we generated 10000 uniformly distributed joint random perturbations of all system parameters that reach up to 50% of their nominal values. That is, every parameter was perturbed according to the following scheme:

$$p_{pert} = p_{nom} + (n - 0.5)p_{nom}, \ n \sim \mathcal{U}([0, 1]). \tag{1}$$

For each perturbed parameter set, the steady-state conditional moment equations were solved to obtain the means, variances and noise measures for both the substrate and product. The results are summarized in Fig. 3 (dashed lines), where the large parametric sensitivity of the

system can be clearly seen. Global sensitivity analysis of the mean, variance and CV histograms [25] reveals that the total number of enzymes (N) has the largest effect on all these quantities, with the enzyme activation/deactivation rates (α_E, μ_E) coming at second and third place. Although one could argue that α_E and μ_E are biochemical rates that are uniquely determined by molecular features of the enzyme, the total number of enzyme molecules would certainly be variable across a cellular population.
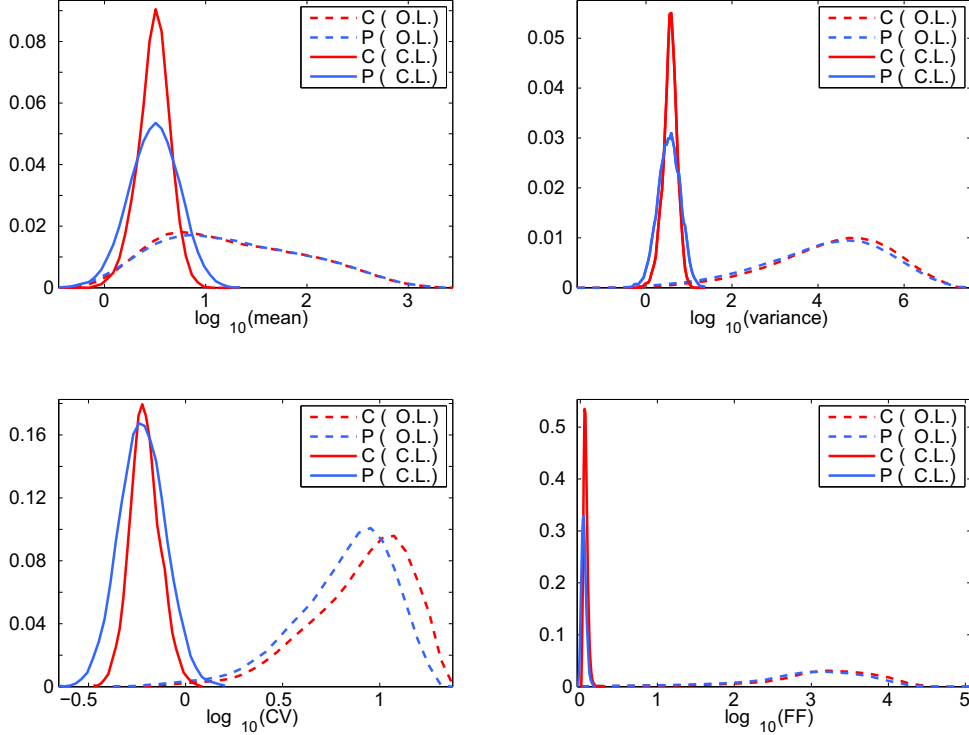

Figure 3: **Dashed lines**: Histograms of steady-state means, variances, CVs and Fano Factors of substrate and product, obtained from 10000 randomly sampled parameters, following (1). Nominal values of perturbed parameters: $N = 10$, $\alpha_E = 0.08$, $\mu_E = 0.1$, $\alpha_C = 5010$, $k = 300$, $\nu = 1$, $\mu_P = 1$. The black line on the top left plot denotes the (common) mean of substrate and product for the nominal parameters. On the top right plot, black lines mark the nominal variance for substrate (solid) and product (dashed). **Continuous lines**: Closed-loop SFE system ($K = 3$, $C_0 = 5$) with substrate feedback: Histograms of steady-state means, variances, CVs and FFs of substrate and product, obtained from the same 10000 randomly sampled parameters used for the dashed line histograms. The great reduction in sensitivity of the closed-loop SFE system in comparison to the open-loop can be easily observed.

Taken together, the results of this section and the previous one suggest that the operation of the SFE reaction scheme in Fig. 1(a) as a signal detection mechanism (the original point made in [23]) is severely compromised when the system operates in the regime defined by our set of assumptions: besides amplifying enzyme fluctuations, the system responds very sensitively to parametric perturbations. These features render the enzyme a highly non-robust controller of the substrate and product outfluxes, which can fluctuate dramatically in time. In addition, reaction rates have to be very finely tuned to achieve a certain output behavior, for example a given mean and variance, or a given average substrate outflux.

2.2 Closing the loop: The SFE network with negative feedback

It is a well-known fact in control theory that negative feedback results in a reduction of the closed-loop system gain [6, 17]. However, this reduction is exchanged for increased stability and robustness to input fluctuations, and a more predictable system behavior that is less dependent on parameter variations. Systems with large open-loop gain tend to also display extreme sensitivity to input and parametric perturbations, and can thus benefit the most from the application of negative feedback.

We shall examine the operation of the SFE network under feedback by assuming that C (or P) affects the rate of activation of the enzyme, for example by controlling its activation rate. We will call this new motif the *closed-loop* SFE system, to differentiate it from the open-loop system presented above.

According to the closed-loop reaction scheme (Fig. 4(a)), the activation rate of E becomes $\alpha_E(1+f(x))$ (x being C or P), where f models activation by x, thus creating a negative feedback loop between the system input and output. Our only requirement for f is to be nondecreasing (e.g. a Hill function). To facilitate our simulation-based analysis, we will assume that f arises from the local, piecewise linear approximation of a Hill function, as shown in Fig. 4(b). In this case, the form of f is controlled by two parameters: K (the "gain") and x_0 (the point beyond which feedback is activated). Finally, α_E can be thought of as the "basal" activation rate in the absence of the regulating molecule.

We should note that the proposed form of feedback regulation is fairly abstract and general enough to have many alternative biochemical implementations. It is possible, for example, for the enzyme activity to be allosterically enhanced by the cooperative binding of C or P (termed substrate and product activation respectively in the language of enzyme kinetics), giving rise to a Hill-like relation between effector abundance and enzyme activity [28]. In this work we will work with the abstract activation rate function defined above.
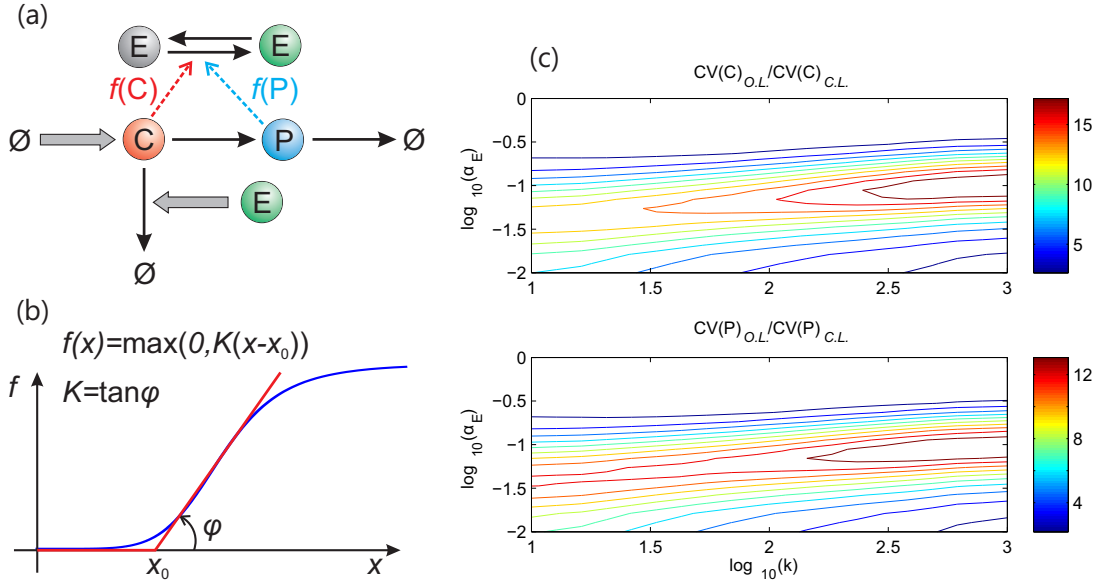

Figure 4: **(a)** The SFE network with feedback (two possible configurations). **(b)** Red line: the class of feedback functions considered in this work, which can approximate a Hill function around its lower end. **(c)** Ratio of open-loop vs. closed-loop CVs as a function of α_E (which determines the average number of active enzymes) and k for $K = 3$, $C_0 = 5$, $N = 10$, $\mu_E = 0.1$, $\alpha_C = 5010$, $\nu = 1$, $\mu_P = 1$.

In the following we analyze the closed-loop SFE network behavior by studying how the SFE network properties described in the previous sections are transformed under feedback.

2.2.1 Feedback results in a dramatic noise reduction and increased robustness to parameter variation

Here we study the SFE network under the influence of negative feedback. We should point out that we characterize the open- and closed-loop systems with respect to the same features (noise and robustness), not to directly compare them, but because these features play an important role in the function of both mechanisms. Whenever we use the open-loop system as a baseline for assessing closed-loop system properties, scale-independent measures are used since this allows for the evaluation of relative distribution spreads. This principle is only disregarded in Fig. 5 below.

For our first test, we use the same settings and parameters as those of Fig. 2(a), only this time we add a feedback term from the substrate to the enzyme activation rate. Increasing the gain K or shifting the activation point x_0 to the left results in a decrease of both means and variances of substrate and product. For the ranges of K and x_0 values considered, the means change by at most a factor of 2.5, while the variances by about 5 times. At the same time, the CVs vary by about 50% and the corresponding Fano Factors by a factor of 5. Moreover, as the analysis in the Supplement (section 5) shows, the CVs of both species become relatively flat as K and x_0 increase, while the Fano Factor gets very close to 1 as K increases for small values of x_0, indicating that the resulting substrate and product stationary distributions are approximately Poissonian in this regime. For the analysis that follows, we fix $K = 3$ and $x_0 = 5$ in the feedback function $f(\cdot)$.

With the above choice of feedback parameters, we first study the sensitivity of the closed-loop SFE system to variations of the two key parameters, α_E and k. As Fig. 2(b) demonstrates, means and variances (and, consequently, CVs) of substrate and product become largely independent of α_E, except for very large values of k (similar results are obtained for product feedback). Moreover, noise of substrate and product is dramatically reduced in comparison to the open-loop SFE system, while the variation of means and variances is now quite small, despite the large ranges of α_E and k values considered. It is also worth noting that the Fano Factors of both substrate and product are very close to one for a large range of parameters.

Interestingly, if we quantify noise reduction by the ratio of open-loop vs. closed-loop CV, we observe that noise reduction is maximal where the open-loop SFE system noise is greatest, as demonstrated by comparing Fig. 4(c) with Fig. 2(a).

Another striking effect of feedback regulation of enzyme activity is that the closed-loop SFE system becomes much less sensitive to parameter variations in comparison to the open-loop case. Applying the same parametric perturbations described in §2.1.3, we obtain the histograms of Fig. 3 (continuous lines). As it becomes apparent, the histograms corresponding to the closed-loop SFE system are several orders of magnitude narrower compared to the open-loop. Moreover, despite the relatively large parametric perturbations, variability in substrate and product statistics of the closed-loop SFE system is largely contained within an order of magnitude.

As it was pointed out in §2.1.3, variability in biochemical reaction rates can be considered "artificial", however changes in the number of enzymes, N, are to be expected in a cellular population. It is therefore interesting to study how variations in N alone are propagated to the substrate and product statistics. Assuming that both the open- and closed-loop systems operate with the same average number of active enzymes for the "nominal" value of $N = 8$, Fig. 5(a,b) shows how the substrate mean and variance vary as N is perturbed around this value, both

in the open- and closed-loop systems (with substrate feedback). To achieve the same average number of active enzymes for $N = 8$, the closed-loop SFE system was simulated first, and the mean number of active enzymes was recorded. This number was then used to back-calculate an appropriate α_E value (keeping μ_E fixed) for the open-loop SFE system. Panel (c) also shows how the distribution of active enzymes differs in the two systems for $N = 8$. The cyan line corresponds to a binomial distribution, $B(8, p)$, where $p = \alpha_E/(\alpha_E + \mu_E)$ is determined by the α_E and μ_E values of the open loop. The red distribution is obtained from simulation of the closed loop and is markedly different from a binomial. The difference is especially significant at the lower end, as small values of E^* lead to fast accumulation of C. Similar results are obtained for product feedback.
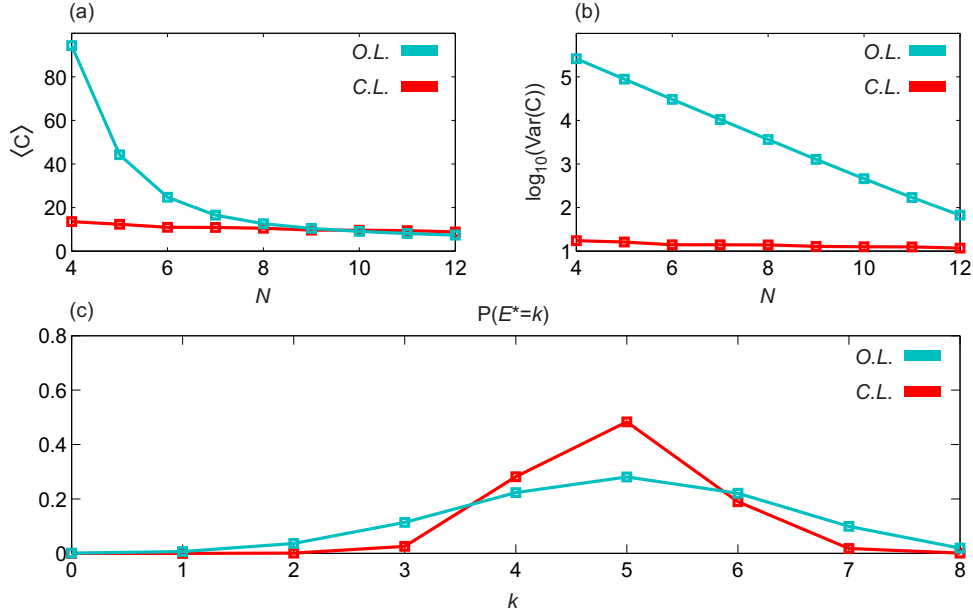

Figure 5: **(a, b)** Mean and variance of substrate in the open- and closed-loop SFE system with substrate feedback (notice the logarithmic variance scale). The closed loop was simulated with $\alpha_E = 0.04$, $\mu_E = 0.1$, $\alpha_C = 5010$, $\nu = 1$, $k = 100$, $\mu_P = 1$ and feedback parameters $K = 10$, $C_0 = 15$. In the open-loop SFE system all parameters were kept the same, except for α_E, which was set to 0.1572 to achieve the same mean of E^* that the closed-loop SFE system achieves for $N = 8$.

2.2.2 Open-loop stochasticity vs. closed-loop determinism

A further remarkable by-product of feedback in the closed-loop SFE system is the fact that the mean of the stochastic model ends up following very closely the predictions of the ODE equations for the deterministic system. This behavior becomes more pronounced as the number of available enzymes (N) grows, while the average number of active enzymes ($N\alpha_E/(\alpha_E + \mu_E)$) remains small. Under this condition, one can think of enzyme activation in the original system as a zeroth-order reaction with rate $N\alpha_E$, and the active enzyme abundance to be described by a birth-death process with birth rate $N\alpha_E$, death rate μ_E and Poisson stationary distribution with parameter $\lambda = N\alpha_E/\mu_E$.

The accuracy of the ODE approximation to the mean substrate levels in the case of substrate feedback can be demonstrated using the same type of parametric perturbations with those employed in §2.1.3. All nominal parameters were kept the same for this test, except for μ_E,

$N\alpha_E$ and k, which were set equal to 1, 5 and 100 respectively ($N = 100$, $\alpha_E = 0.05$). We compare the mean of the stochastic model with the ODE prediction in the case of substrate feedback with $K = 0.4$, $x_0 = 5$, and define the relative error

$$R.E. := \frac{|\langle C \rangle - C_{ss}|}{\langle C \rangle} \cdot 100\%, \tag{2}$$

where C_{ss} is the equilibrium solution of the ODE model. In this setting, a set of 5000 random perturbations leads to an average relative error of 1.5% with standard deviation 0.96%, which clearly shows that the ODE solution captures the mean substrate abundance with very good accuracy indeed (note that the same holds for the mean of P, since it depends linearly on the mean of C). Very similar results are obtained in the case of product feedback.

The above observations are even more striking, if we take into account 1) the fact that the closed-loop SFE system is still highly nonlinear and 2) the intrinsic property of stochastically focused systems to display completely different mean dynamics when compared to the ODE solutions. An explanation of this behavior can be given by examining the moment differential equations. In the limiting case considered in this section, denoting $\alpha_E(C)$ the production rate of active enzyme, we obtain

$$\frac{d\langle C \rangle}{dt} = \alpha_C - \nu\langle C \rangle - k\langle E^* C \rangle = \alpha_C - \nu\langle C \rangle - k\mathrm{Cov}(E^*, C) - k\langle C \rangle\langle E^* \rangle \tag{3}$$

$$\frac{d\mathrm{Cov}(E^*, C)}{dt} = -(\mu_E + \nu)\mathrm{Cov}(E^*, C) - \mathrm{Cov}(kE^* C, E^*) + \mathrm{Cov}(\alpha_E(C), C) \tag{4}$$

The last two terms on the right-hand side of (4) denote the covariance of the substrate enzymatic degradation rate with active enzyme and the covariance of the enzyme activation rate with the substrate. Both covariances are expected to be positive at steady state, which implies that the terms act against each other in determining the steady-state covariance of substrate and active enzyme. In turn, a small value of this covariance (compared to the product $\langle C \rangle\langle E^* \rangle$) implies that the mean of C can be approximately captured by a mean-field equation, where $\langle E^* C \rangle$ has been replaced by $\langle C \rangle\langle E^* \rangle$. This is indeed the case in our simulations, where $\mathrm{Cov}(E^*, C)$ turns out to be $\sim$20–30 times smaller than $\langle C \rangle\langle E^* \rangle$. On the other hand, the open-loop value of $\mathrm{Cov}(E^*, C)$ is about 30 times larger than the closed-loop one. This comes as no surprise, as one expects substrate and active enzyme to display a strong negative correlation, which is the cause of the discrepancy between stochastic and deterministic descriptions of stochastically focused systems.

Similar observations can be made when N is small (e.g. around 10), however the relative errors become at least one order of magnitude larger. We believe that this can be attributed to the fact that the enzyme activation propensity depends both on the abundance of inactive enzyme and the substrate/product abundance, which increases the inaccuracy of the ODEs.

2.2.3 The feedback mechanism is intrinsically robust

As we have already demonstrated, the closed-loop SFE system is remarkably robust to parametric perturbations of the open-loop model. However, in all of our numerical experiments we have kept the parameters of the feedback function f fixed to a few different values. Here we examine the opposite situation, in which only the controller parameters are free to vary while the rest are held constant. We therefore consider the problem of regulating the mean of C around a fixed value with feedback from C. The problem can be posed as follows:

$$\min_{K, C_0} \quad h(K, C_0) := \langle (C - C_{target})^2 \rangle = (\langle C \rangle_{ss} - C_{target})^2 + \mathrm{Var}(C)_{ss},$$

where both the mean and variance of C depend on the feedback function parameters. Fig. 6 shows the contour lines of h, obtained via stochastic simulation over a wide range of K and C_0 values for $C_{target} = 10$. It can be observed that h is more sensitive to C_0 than K: beyond a certain K value, the function quickly levels off.

Based on our simulation runs, the optimal feedback parameters turned out to be $K = 30$ (the maximum K value considered for the plot) and $C_0 \simeq 16$ (given the inevitable uncertainty in h due to sampling variability, the true optimal value should be close to this). The optimal feedback function therefore resembles a "barrier": for $x < C_0$ it is zero, while it rises very steeply beyond C_0.
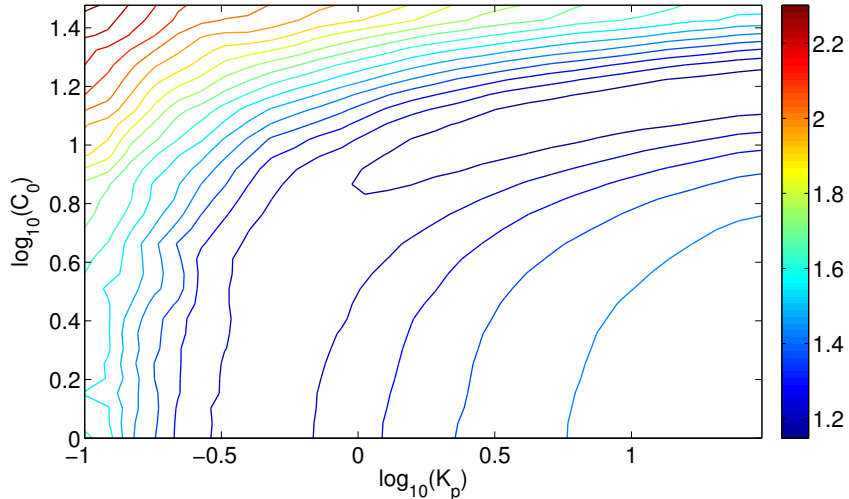

Figure 6: Contour lines of $\log_{10}(h(K, C_0))$ for $C_{target} = 10$, obtained by evaluating the function via stochastic simulation on a logarithmic grid in the parameter space.

Note that the mean and variance of C both depend on (K, C_0), but neither quantity is available in closed form as a function of the feedback parameters or obtainable from a closed set of moment equations. Thus, h had to be evaluated on a grid with the help of stochastic simulation. Alternatively, as we show in the Supplement (section 6), one can exploit the behavior presented in the previous Section, and optimize a similar objective function by directly evaluating the required moments of C using a simple moment closure approximation based on the *method of moments*. This scheme, introduced in [10], provides very accurate approximations of the mean and variance at a fraction of the computational effort, thus allowing optimization to be carried out very efficiently. The optimal parameters for the approximate system can be used as starting points in the optimization of h.

3   Discussion

In this work we have examined the behavior of a branched enzymatic reaction scheme. This system has already been shown to display stochastic focusing, a sensitivity amplification phenomenon that arises due to nonlinearities and stochasticity whenever only a few enzyme molecules are present in the system. We have additionally shown that when the enzyme activity evolves on a slow timescale compared to that of the substrate, very large fluctuations can be generated in the system. Moreover, the dynamics of the system is extremely sensitive to variations in its reaction rates. Both these observations imply that this simple model is not appropriate for

robust signal detection.

We asked how the system behavior would change in the presence of a feedback mechanism, so that the "controller" molecule (E^*) could sense the fluctuations in C (its substrate) or P (the product of the alternative reaction branch). We have shown that noise decreases dramatically in the presence of feedback, while the robustness of the average system behavior is boosted significantly. Consequently, the focused system with feedback ends up behaving almost as predictably as a mean-field ODE model, even when the number of active enzymes is very small.

There exist several biochemical systems which in certain aspects match the main ideas behind the SFE motif, i.e. display stochastic fluctuations in enzyme activity/abundance and substrate/product feedback activation. For example, it was recently discovered [16] that the guanine nucleotide exchange factor SOS and its substrate, the Ras GTPase, are involved in a feedback loop where SOS converts Ras-GDP (inactive) to Ras-GTP (active) and, in turn, active Ras allosterically stabilizes the high-activity state of SOS. Another prominent example is microRNA post-transcriptional regulation of gene expression [34], where a microRNA may mediate the degradation of a target mRNA, while the protein arising from this mRNA in turn activates the microRNA transcription.

Yet another instance is the heat shock response system in *E. coli* [8]: here the σ^{32} factor, which activates the heat-shock responsive genes, is quickly turned over under the action of the protease FtsH at normal growth temperatures. After a shift to high temperature, σ^{32} is rapidly stabilized and at the same time it also activates the synthesis of FtsH. The FtsH-mediated degradation of σ^{32} is under negative feedback. Finally, it is known that mRNA decapping (a process that triggers mRNA degradation) is controlled by the decapping enzyme Dcp2, which fluctuates between an open (inactive) and closed (active) form. Experimental evidence suggests that the closed conformation of the enzyme is promoted by the activator protein Dcp1 together with the mRNA substrate itself [12, 5]. We should stress, however, that it is still unclear if any of the aforementioned examples display all the dynamic features of the motif considered in this work. Speaking generally, due to the required levels of measurement accuracy, it is difficult to find examples that exactly match the conditions considered here with current experimental techniques. However, it is certainly conceivable that this will be achieved in the future.

An interesting feature of our system is that homeostasis is achieved with a very small number of controller molecules (in the order of 10), which are able to maintain the output at a very low level with fluctuations that are – to a good approximation – Poissonian. We have considered two alternative feedback schemes: in the first the substrate directly affects the activation rate of the enzyme (a case of substrate activation), while in the second the product of the alternative reaction branch is used as a "proxy" for the substrate abundance. Note that the flux through the C-P branch is many orders of magnitude smaller than the flux in the E-regulated branch. The C-P branch can thus be thought to act as a "sensor mechanism", used to control the high-flux branch of the system. To the best of our knowledge, this type of feedback has not yet been observed in naturally occurring reactions.

Finally, it is worth to note that one could achieve this type of regulation with an "unfocused" system, in which the coupling between enzyme and substrate (parameter k) would be much smaller. This would imply, however, that the number of enzyme molecules needed to achieve the same substrate levels would have to be much greater, and this could entail an added cost for a living cell.

In summary, to regulate a low-copy, high-flux substrate via an enzymatic mechanism such as the one considered here, there are three possibilities: a) use of a low number of controller enzymes and strong coupling between enzyme and substrate (which results in stochastic focusing and noise), b) use of a high-copy enzyme and weak coupling (with the associated production cost) or c) use of a low-copy controller with feedback: an alternative which, as we have demon-

strated, leads to a remarkably well-behaved closed-loop system.

It is thus conceivable that cellular feedback mechanisms have evolved to exploit the nature of stochastically focused systems to achieve regulation of low-copy substrates with the minimal number of controller molecules. We expect that the rapid development of experimental techniques in single-molecule enzymatics will soon enable the experimental verification of our findings and possibly the discovery of similar noise attenuation mechanisms inside cells. Finally, our results can be seen as a first step towards the rational manipulation of noise properties in low-copy enzymatic reactions.

Author's contributions

Main ideas were conceived jointly by A.M.A., S.E., and M.K following a preliminary numerical study by S.E. and P.B. A.M.A. developed the material and made the initial draft of the manuscript together with S.E. The final writing and revision was performed jointly by A.M.A., S.E., and M.K.

Funding

References

[1] BARKAI, N., AND LEIBLER, S. Circadian clocks limited by noise. *Nature 403* (2000), 267–268. doi:10.1038/35002258.

[2] BECSKEI, A., AND SERRANO, L. Engineering stability in gene networks by autoregulation. *Nature 405*, 6786 (2000), 590–593.

[3] BURGER, A., WALCZAK, A. M., AND WOLYNES, P. G. Influence of decoys on the noise and dynamics of gene expression. *Physical Review E 86*, 4 (2012), 041920.

[4] CROSS, P. J., ALLISON, T. M., DOBSON, R. C. J., JAMESON, G. B., AND PARKER, E. J. Engineering allosteric control to an unregulated enzyme by transfer of a regulatory domain. *Proceedings of the National Academy of Sciences 110*, 6 (2013), 2111–2116.

[5] DESHMUKH, M. V., JONES, B. N., QUANG-DANG, D.-U., FLINDERS, J., FLOOR, S. N., KIM, C., JEMIELITY, J., KALEK, M., DARZYNKIEWICZ, E., AND GROSS, J. D. mRNA decapping is promoted by an RNA-binding channel in Dcp2. *Molecular cell 29*, 3 (2008), 324–336.

[6] DORF, R. C., AND BISHOP, R. H. *Modern control systems*. Pearson, 2011.

[7] EL-SAMAD, H., AND KHAMMASH, M. Regulated degradation is a mechanism for suppressing stochastic fluctuations in gene regulatory networks. *Biophysical journal 90*, 10 (2006), 3749–3761.

[8] EL-SAMAD, H., KURATA, H., DOYLE, J., GROSS, C., AND KHAMMASH, M. Surviving heat shock: control strategies for robustness and performance. *Proc. Natl. Acad. Sci. USA 102*, 8 (2005), 2736–2741.

[9] Elf, J., and Ehrenberg, M. Fast evaluation of fluctuations in biochemical networks with the linear noise approximation. *Genome research 13*, 11 (2003), 2475–2484.

[10] Engblom, S. Computing the moments of high dimensional solutions of the master equation. *Applied Mathematics and Computation 180*, 2 (2006), 498 – 515.

[11] English, B. P., Min, W., van Oijen, A. M., Lee, K. T., Luo, G., Sun, H., Cherayil, B. J., Kou, S., and Xie, X. S. Ever-fluctuating single enzyme molecules: Michaelis-Menten equation revisited. *Nature Chemical Biology 2*, 2 (2005), 87–94.

[12] Floor, S. N., Jones, B. N., and Gross, J. D. Control of mRNA decapping by Dcp2: An open and shut case? *RNA biology 5*, 4 (2008), 189–192.

[13] Grima, R., Walter, N. G., and Schnell, S. Single-molecule enzymology à la michaelis–menten. *FEBS Journal 281*, 2 (2014), 518–530.

[14] Hasenauer, J., Wolf, V., Kazeroonian, A., and Theis, F. Method of conditional moments (MCM) for the chemical master equation. *Journal of mathematical biology* (2013), 1–49.

[15] Hornung, G., and Barkai, N. Noise propagation and signaling sensitivity in biological networks: a role for positive feedback. *PLoS computational biology 4*, 1 (2008), e8.

[16] Iversen, L., Tu, H.-L., Lin, W.-C., Christensen, S. M., Abel, S. M., Iwig, J., Wu, H.-J., Gureasko, J., Rhodes, C., Petit, R. S., et al. Ras activation by SOS: Allosteric regulation by altered fluctuation dynamics. *Science 345*, 6192 (2014), 50–54.

[17] Khammash, M. Robustness in engineering and biology. *BMC Biology* (submitted).

[18] Kou, S., Cherayil, B. J., Min, W., English, B. P., and Xie, X. S. Single-molecule Michaelis-Menten equations. *The Journal of Physical Chemistry B 109*, 41 (2005), 19068–19081.

[19] Lestas, I., Vinnicombe, G., and Paulsson, J. Fundamental limits on the suppression of molecular fluctuations. *Nature 467*, 7312 (2010), 174–178. doi:10.1038/nature09333.

[20] Mashanov, G. I., and Batters, C. *Single Molecule Enzymology*. Methods in Molecular Biology. Springer, 2011.

[21] Mazza, C., and Benaim, M. *Stochastic Dynamics for Systems Biology*. CRC Press, 2014.

[22] Osella, M., Bosia, C., Corá, D., and Caselle, M. The role of incoherent microrna-mediated feedforward loops in noise buffering. *PLoS computational biology 7*, 3 (2011), e1001101.

[23] Paulsson, J., Berg, O. G., and Ehrenberg, M. Stochastic focusing: Fluctuation-enhanced sensitivity of intracellular regulation. *Proc. Natl. Acad. Sci. USA 97*, 13 (2000), 7148–7153. doi:10.1073/pnas.110057697.

[24] Raman, S., Taylor, N., Genuth, N., Fields, S., and Church, G. M. Engineering allostery. *Trends in Genetics 30*, 12 (2014), 521–528.

[25] Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S. *Global sensitivity analysis: the primer.* John Wiley & Sons, 2008.

[26] Samoilov, M., Plyasunov, S., and Arkin, A. P. Stochastic amplification and signaling in enzymatic futile cycles through noise-induced bistability with oscillations. *Proc. Natl. Acad. Sci. USA 102*, 7 (2005), 2310–2315.

[27] Sauer, M., Hofkens, J., and Enderlein, J. *Handbook of Fluorescence Spectroscopy and Imaging: From Ensemble to Single Molecules.* John Wiley & Sons, 2010.

[28] Sauro, H. M. *Enzyme kinetics for systems biology.* Ambrosius Publishing, 2012.

[29] Schwabe, A., Maarleveld, T. R., and Bruggeman, F. J. Exploration of the spontaneous fluctuating activity of single enzyme molecules. *FEBS letters 587*, 17 (2013), 2744–2752.

[30] Singh, A., and Hespanha, J. P. Approximate moment dynamics for chemically reacting systems. *Automatic Control, IEEE Transactions on 56*, 2 (2011), 414–418.

[31] Swain, P. S. Efficient attenuation of stochasticity in gene expression through post-transcriptional control. *Journal of molecular biology 344*, 4 (2004), 965–976.

[32] Thattai, M., and van Oudenaarden, A. Attenuation of noise in ultrasensitive signaling cascades. *Biophysical Journal 82*, 6 (2002), 2943–2950.

[33] Togashi, Y., and Kaneko, K. Molecular discreteness in reaction-diffusion systems yields steady states not seen in the continuum limit. *Phys. Rev. E 70*, 2 (2004), 020901–1. doi:10.1103/PhysRevE.70.020901.

[34] Tsang, J., Zhu, J., and van Oudenaarden, A. MicroRNA-mediated feedback and feedforward loops are recurrent network motifs in mammals. *Molecular Cell 26*, 5 (2007), 753–767.

[35] Ullah, M., and Wolkenhauer, O. *Stochastic approaches for systems biology.* Springer, 2011.

[36] Van Kampen, N. *Stochastic Processes in Physics and Chemistry (Third Edition).* Elsevier, 2007.