**Joint 48th IEEE Conference on Decision and Control and
28th Chinese Control Conference
Shanghai, P.R. China, December 16-18, 2009**

ThB11.2

On the Compromise between Retroactivity Attenuation and Noise Amplification in Gene Regulatory Networks

Shridhar Jayanthi and Domitilla Del Vecchio

Abstract— A bio-molecular system can be rendered insensitive to impedance-like effects, called retroactivity, at its downstream interconnections by implementing a large input amplification gain in a negative feedback loop. This type of design, however, relying on large amplifications, may have undesired effects on the internal noise of the system. We investigate this problem on a simple transcriptional component connected to downstream load by performing a stochastic analysis based on the Ω-expansion. While high gains increase the signal-to-noise ratio of the species in the upstream system and attenuate retroactivity, they also contribute to a shift toward high frequency of the internal noise of the system. We mathematically study this compromise by employing the Langevin equation and by analyzing the noise-to-state transfer function of the linearized system.

I. Introduction

A common approach to either analyzing or synthesizing a complex network is to decompose it into input/output modules. Then, one predicts the overall network function by the composition of the functions of its modules [1]. This approach implicitly assumes that the (dynamic) behavior of a component (or module), characterized in isolation, does not change upon interconnection. We have shown in our earlier work that, just as in many engineering systems, this assumption is not satisfied in transcriptional networks [5]. This effect, which we call *retroactivity*, arises at the interconnection between any two components. To counteract retroactivity and enable modular design, we have proposed to place *insulation devices* between any two modules. An insulation device does not apply retroactivity to its upstream system and is insensitive to retroactivity at its interconnection with downstream systems. As in electronic amplifiers, these devices attain this property by implementing a large input amplification coupled with a large negative feedback [5].

While large gains enable retroactivity attenuation, it is known that they can impact the noise properties of a biomolecular system [2], [3], [12]. Since biological processes are intrinsically stochastic [14], [16], [17], [20], it is essential to analyze and quantify the impact of these gains on noise. We thus employ stochastic analysis techniques, such as Ω-expansion [21] and the Langevin equation [10] to study these effects. Our results show that the signal-to-noise ratio increases with the gains, while it decreases when retroactivity increases. As a consequence, higher gains contribute to retroactivity attenuation *and* better signal-to-noise ratio.

However, data obtained from simulating our system through the Gillespie algorithm [9] shows that when gains are increased, the frequency content of realizations at higher frequencies also increases. To study this effect, also observed in [2], we analyze the power spectral density of the noise process [13], [18]. In line with results shown in [18], we employ the Langevin equation. Our studies point to the following compromise. We usually seek a small amplitude of the noise-to-state transfer function in the range of frequencies where we seek high retroactivity attenuation. However, we show that in the range of frequencies of interest attenuating retroactivity may also increase noise. This limitation, while being similar to the fundamental limitations implied by Bode's integral formula [15], cannot be derived by the latter because the relative degree of the system under study is not at least two. This limitation arises from the stochastic nature of the system and not from the structure of the corresponding deterministic system.

This paper is organized as follows. In Section II, the system studied in the paper is described and its deterministic model is introduced. In Section III, we introduce the stochastic model using the Master Equation, derive the Fokker-Planck equation obtained from the Linear Noise Approximation, and study the effect of retroactivity and system parameters on the signal-to-noise ratio. Section IV employs the Langevin equation to understand the frequency content of the noise in single realizations. Finally, in Section V we discuss the methods and the various results.

II. Deterministic Model

In this paper, we focus on a simple transcriptional component and study the internal noise of this system when adding a downstream component (load) and when increasing the rates of protein transcription and turnover. These rates correspond to the input amplification and negative feedback in the transcriptional component. Experimentally, these rates can be changed through mutations in the promoters of the gene and by adding a protease that targets the expressed protein for degradation. More sophisticated designs that increase the gains employing phosphorylation for input amplification and dephosphorylation for negative feedback can also be employed [5].

Consider the transcriptional component shown in Figure 1. This component takes a transcription factor U as an input and produces a transcription factor Z as output. The transcription rate of gene z, which expresses protein Z, is given by a time varying function $Gk(t)$ that depends on the transcription factor U. This dependency is not modeled, since it is not central to our discussion. The parameter G models the input amplification gain. The degradation rate of protein Z is also assumed to be tunable and thus identified by $G\delta$. The variable gain parameter G will be adjusted to
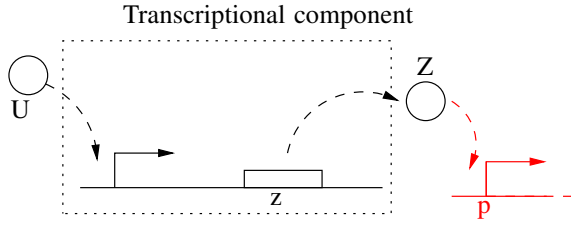
Transcriptional component

Fig. 1. The upstream transcriptional component takes the transcription factor U as input and produces the transcription factor Z as output. The downstream transcriptional component takes Z as input through its reversible binding to the promoter p.
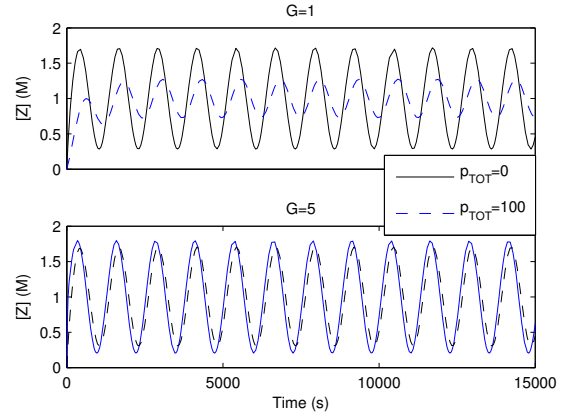

Fig. 2. Effect of retroactivity from the load on system (2). Here, $\omega = 0.005\text{rad/s}$, $\delta = 0.01\text{s}^{-1}$, $k_{off} = 50\text{s}^{-1}$ and $k_d = 20\text{M}$. The input signal is $k(t) = \delta(1 + 0.8\sin{(\omega t)})\text{Ms}^{-1}$. The top plot shows how adding the load impacts $[Z]$. The bottom plot shows how increasing G can reduce the impact of the retroactivity from the load. For this plot we chose $G = 1 + R_l$ (see text).

improve the insulation properties. The transcription factor Z is in turn an input to the downstream load through the reversible binding of Z to promoter sites p. Neglecting the Z messenger RNA dynamics, which are typically much faster than transcription and decay, the system can be modeled by the chemical equations

$$\varnothing \underset{G\delta}{\overset{Gk(t)}{\rightleftharpoons}} \text{Z}, \ \text{Z} + \text{p} \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} \text{C}. \tag{1}$$

We assume that $k(t)$ and δ are of the same order and denote $k_d = k_{off}/k_{on}$. We also assume that the production and decay processes are slower than binding and unbinding reactions, that is, $k_{off} \gg G\delta$, $k_{on} \gg G\delta$ [1]. Let the total concentration of promoter be p_T. The deterministic ordinary differential equation model of system (1) is given by

$$[\dot{Z}] = Gk(t) - G\delta[Z] + k_{off}[C] - k_{on}(p_T - [C])[Z],$$
$$[\dot{C}] = -k_{off}[C] + k_{on}(p_T - [C])Z, \tag{2}$$

in which $[Z]$ and $[C]$ denote the concentrations of Z and C, respectively. Figure 2 shows that adding load to the system decreases the amplitude of the signal $[Z]$. This effect is due retroactivity to the output of the transcriptional component. The figure shows also how the retroactivity effect can be compensated by increasing the gains G. In our examples, the amount of load chosen is much higher than the output concentration to model the fact that the upstream system signal is designed to carry information and is independent of the load value. This, in turn, can become arbitrarily large as the number of different clients requiring the upstream signal increases.

To identify by what amounts G should be increased to compensate the retroactivity effect, we perform a linearized analysis of system (2) about $k(t) = \bar{k}$, and the corresponding equilibrium $[\bar{Z}] = \bar{k}/\delta$ and $[\bar{C}] = [\bar{Z}]p_T/([\bar{Z}] + k_d)$. The dynamics of small perturbations about the equilibrium $(\bar{k}, [\bar{Z}], [\bar{C}])$ are, with abuse of notation, given by

$$[\dot{Z}] = Gk(t) - (G\delta + k_{on}(p_T - [\bar{C}]))[Z]$$
$$+ (k_{off} + k_{on}[\bar{Z}])[C],$$
$$[\dot{C}] = k_{on}(p_T - [\bar{C}])[Z] - (k_{off} + k_{on}[\bar{Z}])[C]. \tag{3}$$

Since $k_{on} \gg \delta$ and $k_{off} = k_{on}k_d$, write

$$k_{on} = \delta/\epsilon \text{ and } k_{off} = \delta k_d/\epsilon, \tag{4}$$

in which $\epsilon \ll 1$. Let $y = [Z] + [C]$. Substituting (4) into (3) and writing the system in terms of y and $[C]$, one obtains the system in the standard singular perturbation form

$$\dot{y} = Gk(t) - ((G\delta(y - [C]))$$
$$\epsilon[\dot{C}] = \delta(p_T - [\bar{C}])y - \delta(p_T - [\bar{C}] + k_d + [\bar{Z}])[C]$$

Setting $\epsilon = 0$, one obtains the expression of the slow manifold as $[C] = k_d p_T[Z]/([\bar{Z}] + k_d)^2 =: \gamma([Z])$. It can be shown that this manifold is exponentially stable. Defining the constant

$$R_l = \frac{k_d p_T}{(\bar{k}/\delta + k_d)^2}, \tag{5}$$

the expression of the slow manifold can be rewritten as $\gamma([Z]) = R_l[Z]$. Letting $[Z] = y - \gamma([Z])$, we obtain that $[\dot{Z}] = \dot{y} - R_l[\dot{Z}]$, in which $\dot{y} = Gk(t) - G\delta[Z]$. The approximated dynamics of $[Z]$ on the slow manifold then becomes

$$[\dot{Z}] = \frac{G}{1 + R_l}(k(t) - \delta[Z]). \tag{6}$$

Thus, for small perturbations about the equilibrium, we should choose $G \approx 1 + R_l$ to compensate for retroactivity from the load.

III. STOCHASTIC MODEL

A. Master Equation

Let $\mathbf{P}(Z, C, p; t | Z_0, C_0, p_0; t_0 = 0)$ denote the conditional probability that at time t, the number of molecules of the species Z, C and p are Z, C and p respectively, given that the number of molecules were Z_0, C_0 and p_0 at the initial time $t_0 = 0$. Throughout this paper we omit the dependency on initial conditions and write $\mathbf{P}(Z, C, p; t)$ for convenience. Let Ω denote the volume of the system. Then, the relation between concentrations and number of molecules is given by $Z = \Omega[Z]$ and $C = \Omega[C]$. Define the step operator as $\mathbb{E}_X^a f(X) = f(X + a)$. We know that $p + C = \Omega p_T$ with

probability one. The resulting two-state Master Equation [21] for the system described in (1) is given by

$$\dot{\mathbf{P}}(Z,C;t) = \big(Gk(t)\Omega(\mathbb{E}_Z^{-1} - 1) + G\delta(\mathbb{E}_Z^{+1} - 1)Z$$
$$+k_{on}(\Omega^{-1}(\mathbb{E}_Z^{+1}\mathbb{E}_C^{-1} - 1)Z(\Omega p_T - C)$$
$$+k_{off}(\mathbb{E}_Z^{-1}\mathbb{E}_C^{+1} - 1)C\big)\mathbf{P}(Z,C;t). \qquad (7)$$

From the above Master Equation, one can see that when the input signal $k(t)$ is periodic, the resulting random process is cyclostationary [8].

B. Linear Noise Approximation

Since the coefficients of $\mathbf{P}(Z,C;t)$ in equation (7) are not linear functions of the states, we cannot obtain a closed set of exact equations for the moments [11]. To proceed with the analysis, we thus employ the Ω-expansion [21].

Define the change of variables

$$Z = \Omega\phi_Z + \Omega^{1/2}\zeta \text{ and } C = \Omega\phi_C + \Omega^{1/2}\xi, \qquad (8)$$

in which ζ and ξ are random variables and ϕ_Z, ϕ_C are deterministic quantities. Let also $\Pi(\zeta,\xi;t) := \mathbf{P}(\Omega\phi_Z + \Omega^{1/2}\zeta, \Omega\phi_C + \Omega^{1/2}\xi;t)$. Then, the left hand side of equation (7) becomes

$$\dot{P}(Z,C;t) = \partial_t\Pi(\zeta,\xi;t) - \Omega^{1/2}\dot{\phi}_Z\partial_\zeta\Pi(\zeta,\xi;t)$$
$$- \Omega^{1/2}\dot{\phi}_C\partial_\xi\Pi(\zeta,\xi;t), \qquad (9)$$

in which $\partial_x := \partial/\partial x$. By performing Taylor expansion of the step operator in the new variables ζ and ξ, we obtain the identities

$$\mathbb{E}_Z^a = \sum_{k=0}^{\infty} \frac{(a\Omega^{-1/2})^k}{k!}\partial_\zeta^k, \quad \mathbb{E}_C^a = \sum_{k=0}^{\infty} \frac{(a\Omega^{-1/2})^k}{k!}\partial_\xi^k. \quad (10)$$

Substituting (7) and (10) in equation (9), solving for $\partial_t\Pi(\zeta,\xi;t)$ and collecting the terms in powers of Ω up to order Ω^0, we obtain a differential equation. As shown in [21], it is possible to obtain from this equation the macroscopic laws by setting the coefficient of $\Omega^{1/2}$ to zero, and the Fokker-Planck equation by taking the volume to be large enough. The resulting macroscopic laws for the deterministic variables ϕ_Z and ϕ_C are given by

$$\dot{\phi}_Z = Gk(t) - \delta\phi_Z - k_{on}\phi_Z(p_T - \phi_C) + k_{off}\phi_C$$
$$\dot{\phi}_C = k_{on}\phi_Z(p_T - \phi_C) - k_{off}\phi_C. \qquad (11)$$

and the resulting Fokker-Planck equation is

$$\partial_t\Pi(\zeta,\xi;t) = \Big[\partial_\zeta\big((G\delta + k_{on}(p_T - \phi_C))\zeta + (-k_{on}\phi_Z$$
$$- k_{off})\xi\big) + \partial_\xi\big(-k_{on}(p_T - \phi_C)\zeta + (k_{on}\phi_Z + k_{off})\xi\big)$$
$$+ \frac{1}{2}\partial_\zeta^2\big(Gk(t) + G\delta\phi_Z + k_{on}\phi_Z(p_T - \phi_C) + k_{off}\phi_C\big)$$
$$+ \frac{1}{2}\partial_\xi^2\big(k_{on}\phi_Z(p_T - \phi_C) + k_{off}\phi_C\big)$$
$$+ \partial_\zeta\partial_\xi\big(-k_{on}\phi_Z(p_T - \phi_C) - k_{off}\phi_C\big)\Big]\Pi(\zeta,\xi;t). \quad (12)$$

The above procedure, often referred to as Linear Noise Approximation, takes the jump Markov process defined by equation (7) and approximates it by the continuous Markov process that solves equation (12) [7]. This approximation is valid for (i) large volumes and (ii) when changes in the number of molecules caused by single events (jumps) of the original process are small compared to the total number of molecules in the system [7]. Assumption (i) is satisfied by the fact that the number of proteins and promoter sites are of the order of tens to hundrends. The fact all the reactions have stoichiometry 1, along with assumption (i) makes assumption (ii) relevant in our model.

Given a general Fokker-Planck equation of the form

$$\partial P_t(\mathbf{x};t) = -\sum_i \partial_i A_i(\mathbf{x},t)P(\mathbf{x},t) +$$
$$\frac{1}{2}\sum_i\sum_j \partial_i\partial_j B_{ij}(\mathbf{x},t)P(\mathbf{x},t), \qquad (13)$$

it is possible to derive differential equations for the expectancy of any polynomial function $f(\mathbf{x})$. Multiplying both sides of equation (13) by $f(\mathbf{x})$ and integrating both sides over the state space one obtains the differential equation [7]

$$\langle\dot{f(\mathbf{x})}\rangle = \sum_i\langle A_i\partial_i f(\mathbf{x})\rangle + \frac{1}{2}\sum_i\sum_j\langle B_{ij}\partial_i\partial_j f(\mathbf{x})\rangle (14)$$

Repeating this process with the Fokker-Planck equation (12), we obtain the differential equations for the first order moments as

$$\dot{\langle\zeta\rangle} = -\delta\langle\zeta\rangle - k_{on}(p_T - \phi_C)\langle\zeta\rangle + k_{on}\phi_Z\langle\xi\rangle + k_{off}\langle\xi\rangle,$$
$$\dot{\langle\xi\rangle} = k_{on}(p_T - \phi_C)\langle\zeta\rangle - k_{on}\phi_Z\langle\xi\rangle - k_{off}\langle\xi\rangle. \qquad (15)$$

Setting the initial conditions of the macroscopic equations (11) to correspond to the initial values of the number of species, that is, setting $\phi_Z(0) = \Omega^{-1}Z_0$ and $\phi_C(0) = \Omega^{-1}C_0$, then $\langle\zeta(0)\rangle = 0$ and $\langle\xi(0)\rangle = 0$. From this, $\langle\zeta(t)\rangle = 0$ and $\langle\xi(t)\rangle = 0$ for all time. Therefore, $\zeta(t)$ and $\xi(t)$ are zero-mean random processes.

Similarly, the dynamics of the second order moments are given by

$$\dot{\langle\zeta^2\rangle} = -2G\delta\langle\zeta^2\rangle - 2k_{on}(p_T - \phi_C)\langle\zeta^2\rangle + 2k_{on}\phi_Z\langle\zeta\xi\rangle$$
$$+2k_{off}\langle\zeta\xi\rangle + k_{on}\phi_Z(p_T - \phi_C) + k_{off}\phi_C$$
$$+Gk(t) + G\delta\phi_Z$$
$$\dot{\langle\zeta\xi\rangle} = -G\delta\langle\zeta\xi\rangle + k_{on}\phi_Z\langle\xi^2\rangle + k_{off}\langle\xi^2\rangle$$
$$+k_{on}(p_T - \phi_C)\langle\zeta^2\rangle - k_{on}\phi_Z\langle\zeta\xi\rangle - k_{off}\zeta\xi$$
$$-k_{on}\phi_Z(p_T - \phi_C) - k_{off}\phi_C$$
$$\dot{\langle\xi^2\rangle} = 2k_{on}(p_T - \phi_C)\langle\zeta\xi\rangle - 2k_{on}\phi_Z\langle\xi^2\rangle - 2k_{off}\langle\xi^2\rangle$$
$$+k_{on}\phi_Z(p_T - \phi_C) + k_{off}\phi_C. \qquad (16)$$

To validate the Fokker-Planck approximation (12), we compare the time-dependent mean and standard deviation of the concentrations predicted by numerical integrations of equations (11) and (16) with the mean and standard-deviation from sample realizations given by a Stochastic Simulation Algorithm (SSA) implementation [9]. Let the means be denoted by $\mu_{[Z]}$ and $\mu_{[C]}$ and the standard deviations denoted

by $\sigma_{[Z]}$ and $\sigma_{[C]}$. To obtain these quantities from the Fokker-Planck approximation, recall that ζ and ξ are zero-mean random variables and that ϕ_Z and ϕ_C are deterministic. Then, from the substitution of variables (8) the mean and the standard deviation of the concentration of Z and C are given by $\mu_{[Z]} = \phi_Z$, $\mu_{[C]} = \phi_C$, $\sigma_{[Z]} = \sqrt{\Omega^{-1} \langle \zeta^2 \rangle}$ and $\sigma_{[C]} = \sqrt{\Omega^{-1} \langle \xi^2 \rangle}$ respectively. To obtain these quantities from N realizations Z_i and C_i of the SSA we used the sample mean and the biased sample variance estimator.

Figure 3 shows that the means and standard deviations of Z and C predicted by the Fokker-Planck equation are very close to the values obtained from the samples of the SSA.
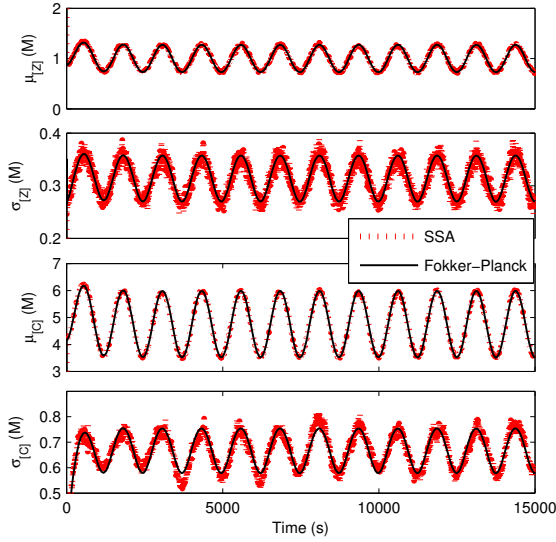

Fig. 3. The sample means and variances from SSA and from Fokker-Planck equation are shown to be very close to each other. For these plots, $\delta = 0.01\mathrm{s}^{-1}$, $k_d = 20\mathrm{M}$, $k_{off} = 50\mathrm{s}^{-1}$ with input signal $k(t) = \delta(1 + 0.8 \sin \omega t)\mathrm{Ms}^{-1}$ and volume $\Omega = 10\mathrm{M}^{-1}$. To simulate the time varying input in the SSA, we imposed a deterministic time-varying concentration of a Z protein messenger with concentration $k(t) = \delta(1 + 0.8 \sin(\omega t))\mathrm{Ms}^{-1}$. Means from SSA were calculated using 500 realizations. Note that due to the sinusoidal input, the random process is cyclostationary.

C. Signal to Noise Ratio Analysis

A traditional metric used to assess noise in many electrical engineering applications is the signal-to-noise ratio. This quantity is usually defined by taking the ratio between the power of the signal and the power of the noise and gives a measure of how much the noise corrupts the signal.

For this study, we consider periodic input signals and characterize the signal-to-noise ratio as a function of the input frequency. Since concentrations are always positive, we consider inputs of the form $k(t) = \bar{k} + \tilde{k}(t)$, in which $\bar{k}$ is a constant bias and $\tilde{k}(t) = A_0 sin(\omega t)$ is a periodic signal with amplitude $A_0 < \bar{k}$ and frequency ω. We assume that all the information transmitted is contained in the signal $\tilde{k}(t)$. Therefore, to obtain a signal-to-noise figure of merit, the power of a signal is taken to be the square of its amplitude. The power of the noise is quantified by the steady-state variance calculated when the input is constant and equal to

the bias, that is, $k(t) = \bar{k}$. Denoting A the amplitude of a signal and $\bar{\sigma}^2$ the steady-state variance, the figure of merit for the noise is given by

$$SNR := \frac{A^2}{\bar{\sigma}^2}. \qquad (17)$$

To calculate the values of $\bar{\sigma}^2$, set $k(t) = \bar{k}$ in equations (11) and (16). The corresponding equilibrium values of the deterministic variables ϕ_Z and ϕ_C are obtained from equations (11) as

$$\bar{\phi}_Z = \frac{\bar{k}}{\delta} \text{ and } \bar{\phi}_C = \frac{p_T \bar{\phi}_Z}{1 + \delta k_d / \bar{k}} = \frac{p_T / k_d}{1 + \bar{\phi}_Z / k_d}. \qquad (18)$$

Substituting (18) in equations (16) and setting the time derivatives to zero, the equilibrium values for the second-order moments become $\langle \bar{\zeta}^2 \rangle = \bar{k}/\delta$, $\langle \bar{\zeta}\bar{\xi} \rangle = 0$ and $\langle \bar{\xi}^2 \rangle = \bar{\phi}_Z R_l$, in which R_l is the same constant defined in expression (5). From the change of variables (8), and since ζ and ξ are zero-mean, we have that $\bar{\sigma}_{[Z]}^2 = \Omega^{-1}\langle \bar{\zeta}^2 \rangle$ and $\bar{\sigma}_{[C]}^2 = \Omega^{-1}\langle \bar{\xi}^2 \rangle$ leading to expressions

$$\bar{\sigma}_{[Z]}^2 = \bar{\phi}_Z / \Omega \text{ and } \bar{\sigma}_{[C]}^2 = \bar{\phi}_Z R_l / \Omega. \qquad (19)$$

For small amplitudes of the signal $\tilde{k}(t)$, $\bar{\sigma}^2$ approximates the time-average value of the time-dependent variance $\sigma(t)$ when the system is subject to the input $k(t) = \bar{k} + \tilde{k}(t)$.

Due to the fact that ζ and ξ are zero-mean random variables, $\langle Z \rangle = \Omega \phi_Z$ and $\langle C \rangle = \Omega \phi_C$. Therefore, the amplitude of the mean concentration signals $[Z]$ and $[C]$ are equal to the amplitude of ϕ_Z and ϕ_C, respectively. To calculate this amplitude, we proceed by linearizing equation (11) about the equilibrium corresponding to the fixed input $k(t) = \bar{k}$. With abuse of notation, the linear system becomes

$$\dot{\phi}_Z = G\tilde{k}(t) - (G\delta - k_{on}(p_T - \bar{\phi}_C))\phi_Z$$
$$+(k_{on}\bar{\phi}_Z + k_{off})\phi_C,$$
$$\dot{\phi}_C = k_{on}(p_T - \bar{\phi}_C)\phi_Z - (k_{on}\bar{\phi}_Z + k_{off})\phi_C. \qquad (20)$$

In order to obtain the amplitude of the signals ϕ_Z and ϕ_C, we compute the transfer functions from $\tilde{k}$ to ϕ_Z and ϕ_C. Let $\Phi_Z(s)$, $\Phi_C(s)$ and $\tilde{K}(s)$ denote the Laplace transforms of $\phi_Z(t)$, $\phi_C(t)$ and $\tilde{k}(t)$, respectively. From (20), applying substitution (4) and setting $\epsilon = 0$, we obtain the transfer functions

$$G_1(s) = \frac{\Phi_Z(s)}{\tilde{K}(s)} = \frac{G}{s(1 + R_l) + G\delta}$$
$$G_2(s) = \frac{\Phi_C(s)}{\tilde{K}(s)} = \frac{GR_l}{s(1 + R_l) + G\delta}.$$

Therefore, for a input signal $\tilde{k}(t)$ with frequency ω and amplitude A_0, the amplitude of the concentrations are

$$A_{[Z]} = |G_1(j\omega)|A_0 = \sqrt{\frac{G^2}{G^2\delta^2 + \omega^2(1 + R_l)^2}} A_0,$$

$$A_{[C]} = |G_2(j\omega)|A_0 = \sqrt{\frac{G^2 R_l^2}{G^2\delta^2 + \omega^2(1 + R_l)^2}} A_0. \qquad (21)$$

Substituting expressions (19) and (21) in the definition (17), the signal-to-noise ratios obtained for an input $\tilde{k}(t)$ with amplitude A_0 and frequency ω are

$$SNR_Z(\omega) = \frac{\Omega}{k\delta} \frac{G^2}{G^2 + \frac{\omega^2}{\delta^2}(1+R_l)^2} A_0^2, \qquad (22)$$

$$SNR_C(\omega) = \frac{\Omega}{k\delta} \frac{G^2 R_l}{G^2 + \frac{\omega^2}{\delta^2}(1+R_l)^2} A_0^2. \qquad (23)$$

Recalling from (5) that R_l is monotonically increasing with p_T, expression (22) shows that for a signal with non-zero frequency, addition of load p_T leads to a lower value of SNR_Z. Notice also that higher input frequencies increase the sensitivity of the SNR_Z to the load. Increasing the gain G improves SNR_Z and in the limit when $G \to \infty$, $SNR_Z \to \Omega A_0^2/(\bar{k}\delta)$ giving a theoretical upper bound on the SNR_Z. Equation (23) shows that the effect of increasing G on SNR_C is similar to the effect on SNR_Z: as the gain increases, the signal-to-noise ratio increases, and in the limit when $G \to \infty$, $SNR_C \to \Omega R_l A_0^2/\bar{k}\delta$. The effect of increasing the load, however is not trivial. If we are able to increase the gain to $G \approx 1 + R_l$ to compensate for the retroactivity, then SNR_C decreases with a higher load. If, instead, the value of G cannot be large enough to reach $1 + R_l$, then increasing p_T will reduce SNR_C. This is a consequence of the fact that when G cannot compensate for the retroactivity the amplitude $A_{[Z]}$ decreases and consequently so does $A_{[C]}$.

IV. FREQUENCY ANALYSIS OF DISTURBANCES AND THE LANGEVIN APPROACH

In Section III, we have shown that increasing the gain G is beneficial for both attenuating retroactivity and decreasing the noise-to-signal ratio. However, as shown in Figure 4, increasing the gain G impacts the frequency content of the noise in a single realization. For low values of G, the error signal between a realization and the mean is of lower frequency when compared to a higher gain. This effect can only be captured by autocorrelation based metrics and, therefore, is not evident from the signal-to-noise ratio.

To study how this problem impacts retroactivity attenuation, we employ the Langevin equation derived from the Master Equation (7), as performed in [18]. As shown in [10], a Master Equation of the form $\frac{dP(\mathbf{X};t)}{dt} = \sum_{j=1}^{M} (\prod_{i=1}^{N} \mathbb{E}_{x_i}^{v_{ij}} - 1) a_j(\mathbf{X}) P(\mathbf{X};t)$ can be approximated by a Langevin system of equations of the form $\frac{dX_i}{dt} = \sum_{j=1}^{M} v_{ij} a_j(\mathbf{X}(t)) + \sum_{j=1}^{M} v_{ij} a_j^{1/2}(\mathbf{X}(t)) \Gamma_j(t)$, in which $\Gamma_j(t)$ are independent Gaussian white noise processes. Applying the above approximation to the Master Equation (7), one obtains the system of Langevin equations

$$\dot{Z} = Gk(t) - G\delta Z - k_{on}(p_T - C)Z + k_{off}C$$
$$+ \sqrt{Gk(t)}\,\Gamma_1(t) - \sqrt{G\delta Z}\,\Gamma_2(t)$$
$$- \sqrt{k_{on}(p_T - C)Z}\,\Gamma_3(t) + \sqrt{k_{off}C}\,\Gamma_4(t),$$
$$\dot{C} = k_{on}(p_T - C)Z - k_{off}C + \sqrt{k_{on}(p_T - C)Z}\,\Gamma_3(t)$$
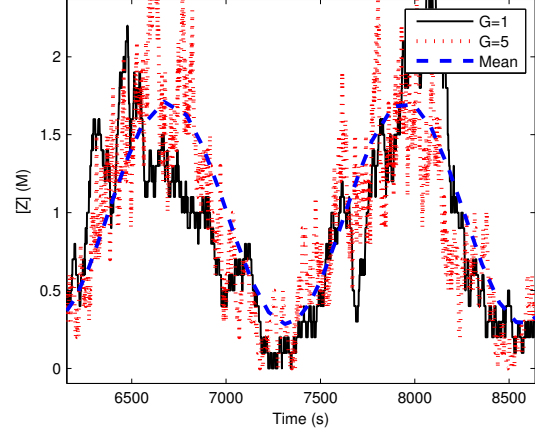$$- \sqrt{k_{off}C}\,\Gamma_4(t). \qquad (24)$$

Fig. 4. Increasing the value of G produces a disturbance signal of higher frequency. Two realizations are shown with different values for G without load. The parameters used in the simulations are $\delta = 0.01\text{s}^{-1}$, $k_d = 20\text{M}$, $k_{off} = 50\text{s}^{-1}$, $\omega = 0.005\text{rad/s}$ and $\Omega = 10\text{M}^{-1}$. The input signal used is $k(t) = \delta(1 + 0.8\sin\omega t)\text{Ms}^{-1}$. The mean of the signal is given as reference.

The above system can be viewed as a non-linear system with five inputs, $k(t)$ and $\Gamma_i(t)$ for $i = 1, 2, 3, 4$. Let $k(t) = \bar{k}$, $\Gamma_1(t) = \Gamma_2(t) = \Gamma_3(t) = \Gamma_4(t) = 0$ be constant inputs and let $\bar{Z}$ and $\bar{C}$ be the corresponding equilibrium points. As we are considering small amplitude signals $\tilde{k}(t)$ we work with a linearization of the system (24) around the fixed input $\bar{k}$ with corresponding equilibrium points $\bar{Z}$ and $\bar{C}$. We can simplify the system further by noting that in equilibrium $\delta\bar{Z} = G\bar{k}$ and $k_{on}(p_T - \bar{C})\bar{Z} = k_{off}\bar{C}$. Also, since Γ_j are independent identical Gaussian white noises, we can write $\Gamma_1(t) - \Gamma_2(t) = \sqrt{2}N_1(t)$ and $\Gamma_3(t) - \Gamma_4(t) = \sqrt{2}N_2(t)$, in which $N_1(t)$ and $N_2(t)$ are independent Gaussian white noises identical to $\Gamma_j(t)$. With these simplifications, the system becomes

$$\dot{Z} = G\tilde{k}(t) - G\delta Z - k_{on}(p_T - \bar{C})Z + k_{on}\bar{Z}C + k_{off}C$$
$$+ \sqrt{2G\bar{k}}N_1(t) - \sqrt{2k_{off}\bar{C}}N_2(t),$$
$$\dot{C} = k_{on}(p_T - \bar{C})Z - k_{on}\bar{Z}C - k_{off}C$$
$$+ \sqrt{2k_{off}\bar{C}}N_2(t). \qquad (25)$$

This is a system with three inputs: the deterministic input $\tilde{k}(t)$ and two independent white noise sources $N_1(t)$ and $N_2(t)$. One can interpret N_1 as the source of the fluctuations caused by the production and degradation reactions while N_2 is the source of fluctuations caused by binding and unbinding reactions. Since the system is linear, we can analyze the different contributions of each noise source separately and independent from the signal $\tilde{k}(t)$.

The transfer function from N_1 to Z, after employing substitutions(4) and setting $\epsilon = 0$ is

$$T_1(s) = \frac{Z(s)}{N_1(s)} = \frac{\sqrt{2G\bar{k}}}{s(1+R_l) + G\delta}. \qquad (26)$$

The DC gain of this transfer function is equal to $T_1(0) = \sqrt{2\bar{k}}/\sqrt{G}\delta$. Thus, as G increases, the DC gain decreases. But for large enough frequencies ω, $j\omega(1 + R_l) + G\delta \approx j\omega(1+R_l)$, and the amplitude $|T_1(j\omega)| \approx \sqrt{2\bar{k}G}/\omega(1 + R_l)$ becomes a monotone function of G. This effect is illustrated in the upper plot of Figure 5. The frequency at which the amplitude of $|T_1(j\omega)|$ computed with $G = 1$ intersects the amplitude $|T_2(j\omega)|$ computed with $G > 1$ is given by the expression $\omega_e = \delta\sqrt{G}/(1 + R_l)$. Thus, when increasing the gain from 1 to $G > 1$, we reduce the noise at frequencies lower than ω_e but we increase it at frequencies larger than ω_e.

The transfer function from the second white noise source N_2 to Z, after employing substitutions (4) and multiplying numerator and denominator by ϵ, is given by

$$T_2(s) = \left[\sqrt{\epsilon}\sqrt{2\delta\bar{C}}s\right] / \left[\epsilon s^2 + (\epsilon G\delta + \delta(p_T - \bar{C}) + \delta\bar{Z} \right.$$
$$\left. +\delta k_d)s + G\delta(\delta\bar{Z} + \delta k_d)\right]. \quad (27)$$

This transfer function has one zero at $s = 0$ and two poles s_- and s_+ such that $s_- \to -\infty$ and $s_+ \to -G\delta/(1 + R_l)$ as $\epsilon \to 0$. Thus, the contribution of $N_2(t)$ to Z is relevant at high frequencies due to the high-pass nature of the transfer function. Furthermore, increasing the gain G increases the cutoff frequency given by the pole s_+. It is also important to note that $N_2(s)$ is scaled by $\sqrt{\epsilon}$, making the noise at low-frequency caused by $N_2(t)$ negligible when compared to that caused by $N_1(t)$. The Bode plot of the transfer function $T_2(s)$ is shown in the lower plot of Figure 5.

Fig. 6. Increasing only the bias leads to a degraded signal. This is not reflected by the coefficient of variation which decreases when we increase the bias. Mean signals are shown for reference. The parameters used in the simulations are $\delta = 0.01\text{s}^{-1}$, $k_d = 20\text{M}$, $k_{off} = 50\text{s}^{-1}$, $\omega = 0.005\text{rad/s}$ and $\Omega = 10\text{M}^{-1}$. The input signals for the plots are $k(t) = \delta(\bar{k} + 0.8\sin(\omega t))\text{Ms}^{-1}$ in which $\bar{k}$ denote the bias.
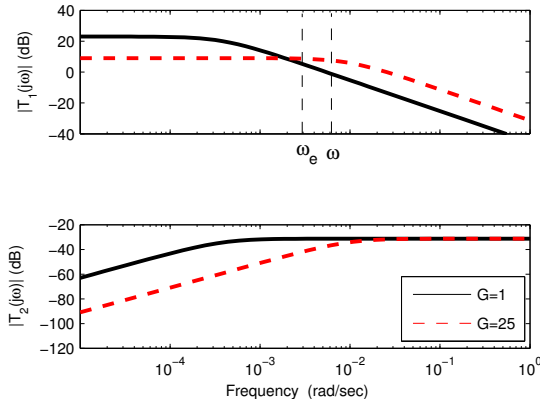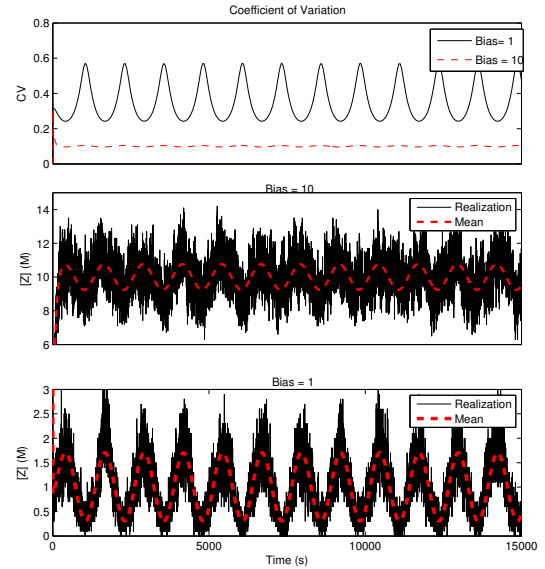

Fig. 5. Magnitude of the transfer functions $T_1(s)$ and $T_2(s)$. The parameters used in this plot are $\delta = 0.01\text{s}^{-1}$, $k_d = 1\text{M}$, $k_{off} = 50\text{s}^{-1}$, $\omega = 0.005\text{rad/s}$, $p_T = 100\text{M}$. When G increases from 1 to $1 + R_l = 25$, contribution from N_1 decreases but it now spreads to a higher range of the spectrum. Note that there was an increase on the noise at the frequency of interest ω. Increasing G reduces the contribution from N_2 in the low frequency range, leaving the high frequency range unaffected. Note also that the amplitude of T_2 is significantly smaller than that of T_1.

V. DISCUSSION

A. Figures of Merit for Noise

The literature presents a number of methods to quantify noise. Popular figures of merit include the Fano factor [6],

[19], [20] and the coefficient of variation [3], [12], [14].

Our interest in this work, in analogy to what is performed in electronic circuits, is to study the behavior of a system under a periodic input, resulting in a cyclostationary random process after the transient. Therefore, the mean $\mu(t)$ and variance $\sigma(t)$ are periodic time-varying quantities. As a result, the coefficient of variation is also a periodic quantity $CV(t) = \sigma(t)/\mu(t)$. In a previous work [4], the maximal coefficient of variation $CV^{max} := \max_t CV(t)$ was used as figure of merit. However, this metric does not capture the degradation of the signal we are interested in transmitting. In fact, Figure 6 shows that we can artificially decrease the coefficient of variation while reducing the signal quality when increasing the bias.

The coefficient of variation is, therefore, not suitable for our needs. In this paper, we adopt as a metric, the signal-to-noise ratio as defined in equation (17). Signal-to-noise ratios are defined by taking the power of the signal and dividing it by the power of the noise. Here, the power of the signal is given by the square of the amplitude A^2 and the power of the noise is given by the steady-state variance $\bar{\sigma}^2$.

B. Compromise between Retroactivity Attenuation and Noise Amplification

In our previous work, it was suggested that the noise increases when the gains of the system increase [4]. However, as discussed in Section V-A this effect seen in the figure of merit CV^{max} does not reflect the signal quality. The analysis in this paper using the SNR figure of merit leads to a different conclusion. In particular, equations (22) and (23) show that increasing the gain helps improving the

signal-to-noise ratio. Equation (22) points that the addition of load increases noise in the upstream component, while equation (23) shows that the addition of load impacts the downstream SNR in a frequency-dependent way. For inputs at high frequencies, where the system suffers from retroactivity, increasing the load will decrease the quality of the signal. However, if the gain is sufficiently large to overcome retroactivity, then increasing the load decreases the noise of the downstream component. Equations (22) and (23) also show that the signal-to-noise ratio cannot be made arbitrarily large by increasing the gain, presenting a limitation. It is also important to notice that in real biological systems, G cannot be arbitrarily large as it is limited by the rates of the fastest available reactions (binding and unbinding of small molecules to transcription factors) [1].

The effect of the gains on the signal-to-noise ratio of the upstream component is tightly linked to the problem of retroactivity. From the first equation in (21), the amplitude $A_{[Z]}$ is attenuated by the addition of load and is recovered by increasing G. However, the steady-state variance $\bar{\sigma}_Z^2$ does not change with G or p_T. According to this, one should choose G to be large enough to compensate for retroactivity in the desired frequency range. However, the frequency of the noise in single realizations is shifted to higher frequencies when G increases, as seen in Figure 4. This effect has been studied in [18] in the context of auto-regulated gene circuits, where it is shown that increasing the negative feedback leads to increase in noise at high frequencies.

In our study, the analysis employing the linearized Langevin approximation (25) shows that when increasing the gain from 1 to G we reduce the noise in the frequency ranges below $\omega_e = \delta\sqrt{G}/(R_l + 1)$, but the noise at frequencies above ω_e increases. Starting from the original system suffering from retroactivity, it is possible to increase the gain up to $G = R_l + 1$ to recover the isolated system behavior. In this case we obtain $\omega_e = \delta/\sqrt{(R_l + 1)} < \delta$. This implies that if the input signal is on the same time-scale of the natural dilution rates, which is very likely, we increase the noise in the frequencies of interest. This effect is showed on the top plot of Figure 5. In theory, one could increase G beyond $(R_l + 1)^2$ so that $\omega_e > \delta$, but in practice the value of G is limited.

VI. CONCLUSIONS AND FUTURE WORK

In this paper, we employed two tools to study the relationship between retroactivity and noise. Moment dynamics were calculated using the Fokker-Planck approximation to obtain the signal-to-noise ratio. This study shows that increasing gains decreases the overall signal-to-noise ratio. However, the Langevin approximation shows that the frequency content of intrinsic noise in realizations is shifted to higher frequencies when gains are increased. This points to a trade-off between retroactivity attenuation and noise amplification. We are studying the application of these techniques to more complex insulation devices designed from phosphorylation systems. We are also extending the approach presented to study the effect of external (extrinsic) noise. Finally,

we are seeking methods to validate the theoretical results from the Langevin approach with data from SSA by using metrics based on sample autocorrelation functions for the cyclostationary random processes under study.

REFERENCES

[1] U. Alon. Network motifs: theory and experimental approaches. *Nature*, 8:450–461, June 2007.
[2] D. W. Austin, M. S. Allen, J. M. McCollum, R. D. Dar, J. R. Wilgus, G. S. Sayler, N. F. Samatova, C. D. Cox, and M. L. Simpson. Gene network shaping of inherent noise spectra. *Nature*, 439(7076):608–611, February 2006.
[3] A. Becskei and L. Serrano. Engineering stability in gene networks by autoregulation. *Nature*, 405:590–593, 2000.
[4] D. Del Vecchio and S. Jayanthi. Retroactivity attenuation in transcriptional networks: Design and analysis of an insulation device. *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*, pages 774–780, Dec. 2008.
[5] D. Del Vecchio, A. J. Ninfa, and E. D. Sontag. Modular cell biology: retroactivity and insulation. *Mol Syst Biol*, 4, February 2008.
[6] J. Elf and M. Ehrenberg. Fast evaluation of fluctuations in biochemical networks with the linear noise approximation. *Genome research*, 13(11):2475–2484, November 2003.
[7] C. W. Gardiner. *Handbook of Stochastic Methods: For Physics, Chemistry and the Natural Sciences (Springer Series in Synergetics)*. Springer, November 1996.
[8] W. Gardner and L. Franks. Characterization of cyclostationary random signal processes. *Information Theory, IEEE Transactions on*, 21(1):4–14, 1975.
[9] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81:2340–2361, 1977.
[10] D. T. Gillespie. The chemical Langevin equation. *J. Chem. Physics*, 113(1):297–306, 2000.
[11] J. Hespanha. Moment closure for biochemical networks. In *Communications, Control and Signal Processing, 2008. ISCCSP 2008. 3rd International Symposium on*, pages 142–147, 2008.
[12] S. Hooshangi and R. Weiss. The effect of negative feedback on noise propagation. *Chaos*, 16, 2006.
[13] I. Lestas, J. Paulsson, N. E. Ross, and G. Vinnicombe. Noise in gene regulatory networks. *Automatic Control, IEEE Transactions on*, 53(Special Issue):189–200, 2008.
[14] J. Paulsson. Summing up the noise in gene networks. *Nature*, pages 415–418, 2004.
[15] K. J. Åström and R. M. Murray. *Feedback Systems: An Introduction for Scientists and Engineers*. Princeton University Press, illustrated edition edition, April 2008.
[16] N. Rosenfeld, J. W. Young, U. Alon, P. S. Swain, and M. B. Elowitz. Gene regulation at the single-cell level. *Science*, pages 1962–1965, 2005.
[17] M. L. Simpson, C. D. Cox, M. S. Allen, J. M. Mccollum, R. D. Dar, D. K. Karig, and J. F. Cooke. Noise in biological circuits. *Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology*, 1(2):214–225, 2009.
[18] M. L. Simpson, C. D. Cox, and G. S. Sayler. Frequency domain analysis of noise in autoregulated gene circuits. *Proc Natl Acad Sci U S A*, 100(8):4551–4556, April 2003.
[19] Y. Tao, Y. Jia, and G. Dewey. Stochastic fluctuations in gene expression far from equilibrium: Ω expansion and linear noise approximation. *Journal of Chemical Physics*, 122:124108, 2005.
[20] M. Thattai and A. van Oudenaarden. Intrinsic noise in gene regulatory networks. *PNAS*, pages 8614–8619, 2001.
[21] N. G. Van Kampen. *Stochastic Processes in Physics and Chemistry, Third Edition (North-Holland Personal Library)*. North Holland, April 2007.