MIT Libraries | DSpace@MIT

MIT Open Access Articles

A control theoretic framework for modular analysis and design of biomolecular networks

The MIT Faculty has made this article openly available. ***Please share*** how this access benefits you. Your story matters.

Citation: Del Vecchio, Domitilla. "A Control Theoretic Framework for Modular Analysis and Design of Biomolecular Networks." Annual Reviews in Control 37.2 (2013): 333–345.

As Published: http://dx.doi.org/10.1016/j.arcontrol.2013.09.011

Publisher: Elsevier

Persistent URL: http://hdl.handle.net/1721.1/108595

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Terms of use: Creative Commons Attribution-NonCommercial-NoDerivs License

Massachusetts Institute of Technology

A Control Theoretic Framework for Modular Analysis and Design of Biomolecular Networks

Domitilla Del Vecchio *

* *Department of Mechanical Engineering and Laboratory for Information and Decision Systems (LIDS) Massachusetts Institute of Technology, Cambridge (MA)*

Abstract: Control theory has been instrumental for the analysis and design of a number of engineering systems, including aerospace and transportation systems, robotics and intelligent machines, manufacturing chains, electrical, power, and information networks. In the past several years, the ability of *de novo* creating biomolecular networks and of measuring key physical quantities has come to a point in which quantitative analysis and design of biological systems is possible. While a modular approach to analyze and design complex systems has proven critical in most control theory applications, it is still subject of debate whether a modular approach is viable in biomolecular networks. In fact, biomolecular networks display context-dependent behavior, that is, the input/output dynamical properties of a module change once this is part of a network. One cause of context dependence, similar to what found in many engineering systems, is retroactivity, that is, the effect of loads applied on a module by downstream systems. In this paper, we focus on retroactivity and review techniques, based on nonlinear control and dynamical systems theory, that we have developed to quantify the extent of modularity of biomolecular systems and to establish modular analysis and design techniques.

1. INTRODUCTION

Modularity is the property that allows to predict the behavior of a complicated system from that of its components, guaranteeing that the salient properties of individual components do not change after interconnection. The assumption that is often made when designing or analyzing a network modularly is that the input/output behavior of a module does not change upon connection to other modules. Researchers in the systems biology community have hypothesized that modularity may be a property of biological systems, proposing functional modules as a critical level of biological organization (Hartwell et al. [1999], Alon [2007a]). This view has profound implications on evolution (Kirschner and Gerhart [2005]) and further suggests that biology, just like engineering, can be understood in a hierarchical fashion (Asthagiri and Lauffenburger [2000], Lauffenburger [2000]). Lauffenburger [2000] in fact elaborates that biological systems can be analyzed in a nested manner, similar to what is performed in engineering design, in which individual components are first characterized and tested in isolation prior to incorporation into incrementally larger and more sophisticated systems. However, modularity is still subject of intense debate and remains one of the most vexing questions in systems biology (Alexander et al. [2009], Cardinale and Arkin [2012], Kaltenbach and Stelling [2012]).

Modularity is also critical for the field of synthetic biology (Purnick and Weiss [2009]), in which researchers are pursuing a bottom-up approach to designing biomolecular networks (Andrianantoandro et al. [2006], Baker et al. [2006]). The past decade has seen tremendous advances to the point that creation of simple biomolecular networks, or "circuits", that control the behavior of living organisms is now possible (Becskei and Serrano [2000], Elowitz and Leibler [2000], Gardner et al. [2000], Atkinson et al. [2003], Stricker et al. [2008], Danino et al. [2009], Bleris et al. [2011], Moon et al. [2012]). A near future is envisioned in which re-engineered cells will perform a number of useful functions from turning feedstock into energy (Peralta-Yahya et al. [2012], Zhang et al. [2012]), to killing cancer cells in ill patients (Xie et al. [2011]), to detecting pathogens in the environment (Kobayashi et al. [2004]), to regulate cell differentiation in diseases like diabetes (Miller et al. [2012]). To meet this vision, one key challenge must be tackled, namely designing biomolecular networks that can realize substantially more complex functionalities than those currently available.

A main bottleneck in advancing in this direction is that modules display context-dependent behavior (Cardinale and Arkin [2012]). Hence, accurate prediction of a module's behavior once part of a network is still an open question. As a result, synthetic circuits need to be re-tuned/re-designed through a lengthy and *ad hoc* process, which takes years, every time they are inserted into a different context (Slusarczyk et al. [2012]). Context-dependence is due to a number of different factors. These include direct inter-modular interactions that result in load effects, a phenomenon known as *retroactivity* (Saez-Rodriguez et al. [2004, 2005], Del Vecchio et al. [2008]); interactions of synthetic circuits with the cell "chassis", a phenomenon broadly known as "metabolic burden" (Bentley et al.

⋆ This work was supported in part by the National Science Foundation and the Air Force Office of Scientific Research.

[1990]), which encompasses a number of different effects, such as effects on cell growth (Scott et al. [2010]) and competition for shared resources, such as ATP, transcription/translation machinery, proteases, etc. (N. A Cookson and et al. [2011], Yeung et al. [2013], Mather et al. [2013]); and perturbations in the cell environment, such as changes in temperature, acidity, nutrients' level, etc.

In this paper, we address the fundamental question of modularity in biology focusing on the problem of retroactivity. Retroactivity extends the notion of impedance or loading to non-electrical systems and, in particular, to biological systems (Saez-Rodriguez et al. [2004, 2005], Del Vecchio et al. [2008]). Because of retroactivity, the dynamical behavior of a system changes upon connection to other systems and, therefore, retroactivity effects imply a loss of modularity. We review some of our results towards (a) understanding the extent of modularity of biomolecular systems by quantifying the effects of retroactivity, (b) establishing a predictive modeling framework for modular analysis in the presence of retroactivity, and (c) uncovering design principles ensuring that the behavior of modules is, to some extent, robust to retroactivity effects.

This work is complementary to but different from studies focusing on partitioning biological networks into modules using graph-theoretic approaches (Saez-Rodriguez et al. [2005], Sridharan et al. [2011], Anderson et al. [2011]). Instead, the work presented here describes a theoretical framework to accurately predict both the quantitative and the qualitative behavior of interconnected modules from their behavior in isolation and from key physical properties. In this sense, our approach is closer to that of disciplines in biochemical systems analysis, such as metabolic control analysis (MCA) (Fell [1992], Heinrich and Schuster [1999], Sauro and Kholodenko [2004]). However, while MCA is primarily focused on steady state and near-equilibrium behavior, our approach focuses on global nonlinear dynamics evolving possibly far from equilibrium situations.

To make the retroactivity problem amenable for a mathematical analysis, we propose a system concept that explicitly incorporates retroactivity (Del Vecchio and Sontag [2009]). We then leverage the natural separation of time scales in biomolecular networks to provide an operative quantification of retroactivity for general networks as a function of measurable biomolecular parameters and network topology (Gyorgy and Del Vecchio [2012], Gyorgy and Del Vecchio [2013]). From this operative quantification and the dynamic model of a module in isolation, we can accurately predict how the behavior of the module will change after connection to other systems. The conceptual similarity of this framework with the electrical circuit theory based on equivalent impedances (Thevenin [1883]) is apparent. The key difference is that electrical circuit theory is largely based on linear systems, while our framework relies on tools from nonlinear system theory, due to the inherent nonlinearity of biomolecular networks, which makes linear systems theory inapplicable.

We then consider the problem of designing larger systems by composing smaller modules. We first briefly describe how the framework developed in Gyorgy and Del Vecchio [2013] allows to determine how the parameters of an upstream system and a downstream system should be tuned such that the effects of retroactivity are minimized. This approach to minimizing retroactivity effects implies that modules should be co-designed, a procedure that can become impractical as the number of modules composing a system increases. A different approach is that to design and optimize modules in isolation and then connect modules to each other through a special device called *insulation device*, such that the modules retain their isolated behavior. The insulation device should apply negligible loading to the upstream module while being able to keep the desired input/output response in the face of significant retroactivity to its output (Jayanthi and Del Vecchio [2011]). This approach has been used in electrical engineering with the introduction of operational amplifiers and their derived circuits such as inverting and non-inverting amplifiers (Schilling and Belove [1968]). We illustrate a design principle for insulation devices, which exploits the distinctive feature of the interconnection structure in biomolecular networks and the separation of time scales among key processes. We illustrate the implementation of this mechanism through covalent modification (Heinrich et al. [2002], Goldbeter and Koshland [1981]), such as phosphorylation (Jiang et al. [2011]).

This paper is organized as follows. In Section 2, we introduce the retroactivity problem through a motivating example. In Section 3, we introduce the system concept to model retroactivity, illustrate analogies with other engineering systems, provide an operative quantification of retroactivity for a simple system, and then for general gene transcription networks. Section 4 formulates the problem of designing insulation devices as a disturbance attenuation problem and tackles it by using singular perturbation theory. In this section, we also illustrate one specific biomolecular implementation of an insulation device through phosphorylation and discuss implications for natural systems.

2. MOTIVATING EXAMPLE

As a motivating example, consider the activator-repressor clock of Atkinson et al. [2003] showed in Figure 1(a). This oscillator is composed of an activator A activating itself and a repressor R, which, in turn, represses the activator A. Both activation and repression occur through transcription regulation. Transcription regulation is a common regulatory mechanism used in synthetic gene circuits. In transcription regulation, a protein, called a transcription factor, binds specific DNA sites, called operator sites, located on a region called promoter, which is upstream of the gene of interest and represses or activates the gene's expression. When the gene is expressed, the protein that the gene encodes is produced, through the process of transcription/translation (a process called the central dogma of molecular biology (Alberts et al. [2007])). This protein can, in turn act as an activator or repressor for other genes, creating circuits of activation and repression interactions (Alon [2007b]). It is common practice in synthetic biology to fabricate the desired DNA sequences of promoters and genes on plasmids (see Bleris et al. [2011], for example). Plasmids are circular pieces of DNA that come with a nominal copy number, from low copy (1-5 copies of plasmid per cell) to high copy (hundreds copies of plasmid per cell).

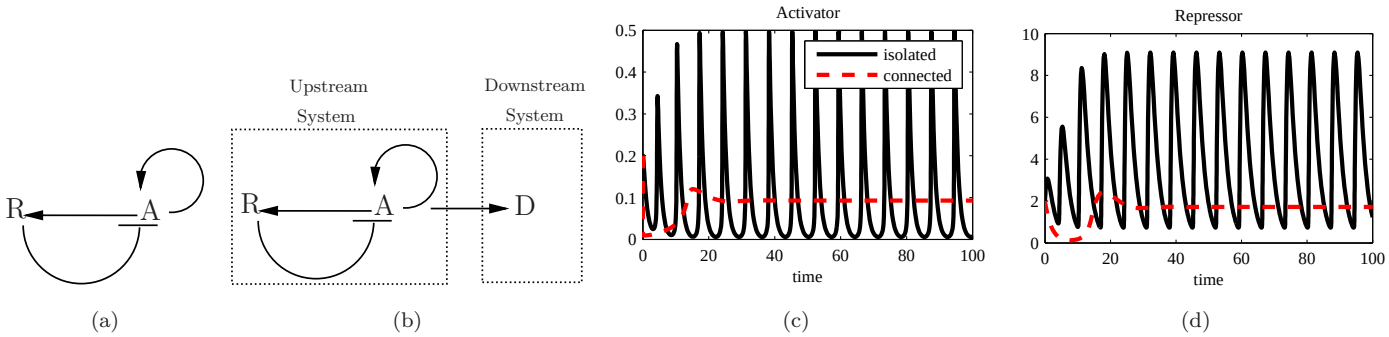


Fig. 1. (a) shows the activator-repressor clock when isolated and (b) when connected to a downstream system. (c)-(d) shows the trajectories of the activator and repressor concentrations when the clock is isolated (solid line) and when it is connected (dashed line). In all simulations, we have $\alpha_A = \alpha_R = 100$, $\alpha_{0,A} = .04$, $\alpha_{0,R} = .004$, $\delta_A = 1$, $\delta_R = 0.5$, $K_a = K_r = 1$, $n = 2$ and $m = 4$ (model (1)). In the connected clock configuration (b), we also have $k_{\text{on}} = k_{\text{off}} = 100$ and $p_T = 20$ (model where the dynamics of A modify to those modeled in equation (2)).

There are several different modeling frameworks that can be used to mathematically describe the dynamic evolution of biological networks, depending on the desired level of granularity in space and time. At the coarsest level, we have reaction rate equation models in the form of ordinary differential equations (ODE), in which individual states represent concentrations (number of molecules divided by the cell volume) of chemical species. These models provide a good approximation of the system dynamics when the number of molecules is sufficiently high and the space is "well stirred", that is, there are no significant spatial gradients in the number of molecules. When, instead, the number of molecules is low, stochastic models need to be employed. These models include the chemical master equation (Van Kampen [2007], Gillespie [1977]), which describes the evolution in time of the probability of the number of molecules of each species, and the chemical Langevin equation (Gillespie [2000]), which describes the dynamic evolution of the noise under the assumption of sufficiently large volumes and molecule numbers. Reaction-diffusion equations, which take the form of a parabolic partial differential equation, are instead employed when there is a significant spatial gradient in the concentration of chemical species (Shvartsman and Baker [2012]).

In this paper, we focus on ODE models under the assumption that the number of molecules in the cell volume is sufficiently high. Letting italics denote the concentration of species, the ODE model of the clock can be written as

$$\begin{aligned} \dot{A} &= \frac{\alpha_A (A/K_a)^n + \alpha_{0,A}}{1 + (A/K_a)^n + (R/K_r)^m} - \delta_A A, \\ \dot{R} &= \frac{\alpha_R (A/K_a)^n + \alpha_{0,R}}{1 + (A/K_a)^n} - \delta_R R, \end{aligned} \tag{1}$$

which describes the rate of change of the activator A and repressor R concentrations. In this model, δ_A and δ_R represent protein decay, due to dilution and/or degradation. The functions $\frac{\alpha_A (A/K_a)^n + \alpha_{0,A}}{1+(A/K_a)^n+(R/K_r)^m}$ and $\frac{\alpha_R (A/K_a)^n + \alpha_{0,R}}{1+(A/K_a)^n}$ are Hill functions (Alon [2007b]). The first one increases with A and decreases with R, since A is an activator and R is a repressor for A, while the second one increases with A as A is an activator for R.

It was shown in the work of Del Vecchio [2007] that the key mechanism by which this system displays sustained oscillations is a Hopf bifurcation with bifurcation parameter given by the relative time scale between the activator and the repressor dynamics. Specifically, as the activator dynamics become faster than the repressor dynamics, the system goes through a supercritical Hopf bifurcation and a periodic orbit appears (solid plots of Figure 1(c-d)).

Assume now that we would like to use the clock as a signal generator to, for example, provide the timing to or synchronize downstream systems. To do so, we need to take one of the proteins of the clock, say protein A, as an input for a downstream system (Figure 1(b)), in which A will activate the expression of another protein D, for example. In this case, one needs to add to the clock dynamics the description of the physical process by which information is transmitted from the upstream system to the downstream one. In any biomolecular system, information is transmitted through (reversible) binding reactions. In this case, n molecules of A will reversibly bind with the promoter p controlling the expression of protein D to form a transcriptionally active complex C (see Jayanthi and Del Vecchio [2012] for more details on the model). Letting p_T denote the total concentration of this promoter (given by the plasmid copy number divided by the cell volume) and k_{on} and k_{off} the association and dissociation rate constants, we have that the A dynamics modify to

$$\begin{aligned} \dot{A} &= \frac{\alpha_A (A/K_a)^n + \alpha_{0,A}}{1 + (A/K_a)^n + (R/K_r)^m} - \delta_A A \\ &\quad - n k_{\text{on}} A^n (p_T - C) + n k_{\text{off}} C, \\ \dot{C} &= k_{\text{on}} A^n (p_T - C) - k_{\text{off}} C, \end{aligned} \tag{2}$$

while the differential equation for R remains the same. Since A is an activator of D, we will also have that $\dot{D} = kC - \delta_D D$, in which k and δ_D are the production rate constant and the decay rate constant, respectively.

As a result of the interconnection, the clock stops functioning (dashed plots in Figure 1(c-d)). This effect has been called *retroactivity* to extend the notion of loading or impedance to non-electrical systems, and in particular to biomolecular systems (Del Vecchio et al. [2008]). It is due to the fact that the communicating species, A in this case, when occupied in the reactions of the downstream system cannot participate in the reactions of the upstream system and hence the clock behavior is affected.

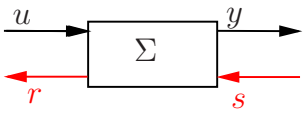


Fig. 2. System concept with retroactivity.

In the next sections, we illustrate a systems theory framework to explicitly model retroactivity so to make the problem of retroactivity amenable of theoretical study.

3. RETROACTIVITY

In order to model retroactivity, we propose to model systems as shown in Figure 2. Specifically, we explicitly add retroactivity as signals traveling from downstream to upstream, that is, in the direction opposed to that in which information is traveling (from upstream to downstream). Signal s is the retroactivity to the output and models the fact that whenever the output y of Σ becomes an input to a downstream system, this system affects the upstream system dynamics because of the physics of the interconnection mechanism. Similarly, r is called the retroactivity to the input and models the fact that whenever Σ receives signal u, it changes the dynamics of the sending system. Related system concepts include that of Paynter [1961] and that of Polderman and Willems [2007]. Differently from Paynter [1961], our framework does not require that signals r and u (s and y) are generalized effort and flow variables and hence that their product is the power flowing through the port. Differently from Willems [1999], we keep a directionality to signals as we consider upstream-to-downstream as the direction in which we think information is being transmitted. From a practical point of view, this is useful because a module is usually characterized by forcing input signals and measuring the consequent output signal. So, there is an intrinsic directionality already associated with the information transfer within a module.

The concept of retroactivity is not restricted to biomolecular systems and it applies in general to many engineering systems (Figure 3). For illustration, we provide two examples: an electrical and an hydraulic system.

Electrical Systems. Figure 3(b) depicts a voltage generator with internal resistor R_0 (upstream system) and the downstream system to which it is connected, a load resistor with value R_L. The output of the isolated system is $y = V_0$ while the output of the connected system is given by $y = V_0 - R_0 I$. The two outputs are the same when the current drawn by the resistor is $I = 0$. Hence, we can take $s = I$. Note that for the connected system, I can be rendered arbitrarily small by taking R_L very large, corresponding to a downstream system with high input impedance (a downstream system with low retroactivity to the input).

Hydraulic Systems. Figure 3(c) depicts the interconnection of two tanks. In the case in which the upstream tank is isolated, we have that the valve at the output pipe is closed and hence the outlet flow is zero, that is, $f = 0$. When the tank is connected to the downstream tank, we have that $f = \rho k\sqrt{(p - p_1)}$ for $p > p_1$, in which ρ is the fluid density and k is a parameter that depends on the output valve geometry. In this case, the retroactivity to the output can be taken as $s = f$. This additional flow will cause a change in the pressure p of the upstream tank. For the connected system this flow can be rendered small by decreasing k, which corresponds to decreasing the aperture of the valve.

Biomolecular Systems. Figure 3(d) depicts the biomolecular clock module with downstream system with sites p to which A binds. In the case in which the clock is not connected, A is not taken as an input to any downstream system, while when it is connected, A serves as an input to a downstream system. This interconnection takes place by having A bind to targets p through a reversible binding reaction. The corresponding reaction rate is given by $-n\ k_{\text{on}}\ A^n\ p + n\ k_{\text{off}}\ C$. Hence, we can take $s = -n\ k_{\text{on}}\ A^n\ p + n\ k_{\text{off}}\ C$. This "usage" of A can change the behavior of the clock and hence the value of the A concentration as discussed in Section 2.

Retroactivity physically has the form of a "flow" between systems: it is a current in the case of the electrical example, a flow of fluid in the case of the hydraulic system, and the rate of a chemical reaction in the biomolecular system. Also, note that retroactivity is fundamentally different from feedback. In fact, s can be removed only if the connection from the upstream system to the downstream system is broken, that is, y is not transmitted to the downstream system. If s were feedback, it could be removed while keeping the transmission of y from upstream to downstream.

3.1 Quantification of retroactivity in a simple biomolecular system

Because of retroactivity, the connected behavior of a module differs from the behavior of the same module in isolation ($s = 0$). We seek to predict how the behavior of a module will change upon interconnection as a function of measurable biochemical parameters. As a simple example, consider the dynamics of a protein X subject to production and decay:

$$\dot{X} = K(t) - \delta X, \tag{3}$$

in which the production rate $K(t)$ varies with the activity of the promoter controlling the expression of X. When this protein is used as an activator or repressor for another protein, such as protein A in the clock is used to activate protein D, we need to modify the dynamics of X to

$$\begin{aligned} \dot{X} &= K(t) - \delta X - k_{\text{on}} X(p_T - C) + k_{\text{off}} C, \\ \dot{C} &= k_{\text{on}} X(p_T - C) - k_{\text{off}} C, \end{aligned} \tag{4}$$

in which we have assumed for simplicity that X binds to DNA promoter sites p as a monomer (in one copy). In this system, we have that $s = -k_{\text{on}} X(p_T - C) + k_{\text{off}} C$ is the retroactivity to the output. How this additional reaction rate affects the X dynamics when compared to the isolated system (3) is the next question we address.

To answer this question, we can exploit the natural timescale separation between protein production and decay (δ), of the order of minutes to hours, and binding/unbinding rates ($k_{\text{on}}/k_{\text{off}}$), of the order of seconds/subseconds (Alon [2007b]). By defining the small parameter $\epsilon = \delta/k_{\text{off}}$ and the slow variable $z = X + C$, system (4) can be taken to the standard singular perturbation form (Khalil [2002])

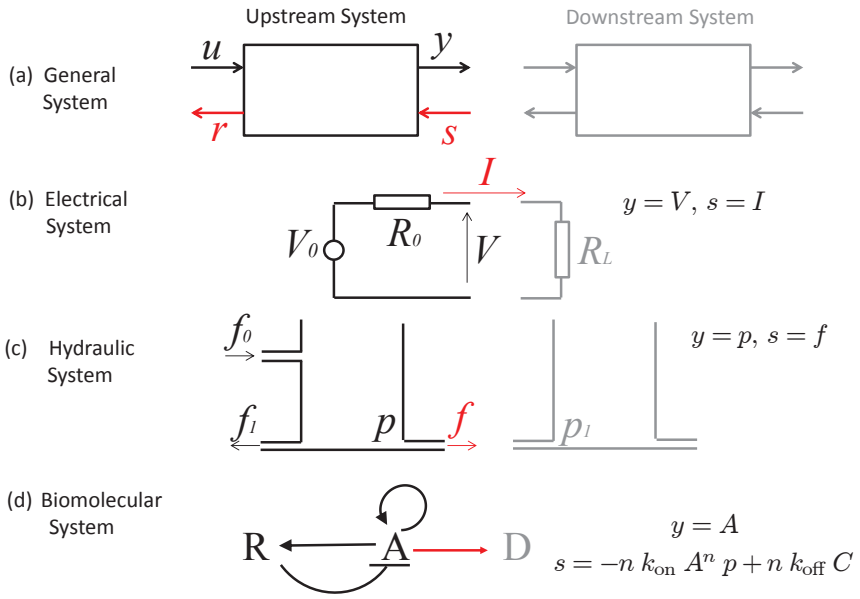


Fig. 3. Interconnection of two systems. (a) The upstream system is sending the signal y and the downstream system is receiving it. Information travels from upstream to downstream. The r and s arrows represent retroactivity to the input and retroactivity to the output, respectively. (b) V and I in the electrical system represent voltage and current, respectively; (c) f and p in the hydraulic system represent flow and pressure, respectively; (d) the upstream system is the activator/repressor clock described in Section 2. Retroactivity takes different forms depending on the specific physical domain, but it always has the physical meaning of a *flow*.

$$\begin{aligned} \dot{z} &= K(t) - \delta(z - C), \\ \epsilon\dot{C} &= \frac{\delta}{K_d}(z - C)(p_T - C) - \delta C, \end{aligned} \tag{5}$$

in which $K_d = k_{\text{off}}/k_{\text{on}}$ is the dissociation constant of the binding between X and p. One can show that the slow manifold of this system is locally exponentially stable (Del Vecchio et al. [2008]), so that system (5) can be approximated up to order ϵ by the reduced system where $\epsilon = 0$ given by $\dot{z} = K(t) - \delta X$ and $C(X) = \frac{p_T X}{X + K_d}$. Considering that

$$\dot{z} = \dot{X} + \dot{C} = \dot{X}\left(1 + \frac{dC}{dX}\right),$$

and that $\dot{z} = K(t) - \delta X$, we finally obtain

$$\dot{X} = (K(t) - \delta X)\left(\frac{1}{1 + \mathcal{R}(X)}\right),\ \mathcal{R}(X) = \frac{p_T/K_d}{(X/K_d + 1)^2}, \tag{6}$$

in which $\mathcal{R}(X) = \frac{dC}{dX}$. Comparing this system with the isolated system (3), we note that the net effect of retroactivity is that of decreasing the rate of change of the species X by a factor $1/(1 + \mathcal{R}(X))$. The expression of $\mathcal{R}(X)$ provides an operative quantification of retroactivity as a function of relevant biochemical parameters. Specifically, retroactivity increases when p_T increases (the load increases) and/or K_d decreases (the affinity of the binding increases), which is physically intuitive. Furthermore, since $\mathcal{R}(X) > 0$, we have that equation (6) implies that the dynamics of X slow down upon connection to the downstream system. Further, as K_d becomes smaller, this "slow down" becomes close to a finite-time delay (Figure 4).

The effects of retroactivity on the behavior of biomolecular systems have been experimentally demonstrated in the MAPK cascade in eukaryotic cells (Kim et al. [2011]), in gene circuits in bacterial cells (Jayanthi et al. [2013]), and in signaling systems *in vitro* (Ventura et al. [2010a], Jiang et al. [2011]).

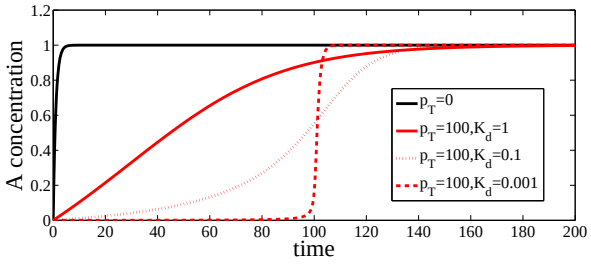


Fig. 4. Effect of retroactivity on the dynamics of X when subject to constant production rate $K(t) = 1$.

3.2 Revisiting the effects of retroactivity on the clock

Since retroactivity slows down the dynamics of the output species, it is natural that the clock in Section 2 stops oscillating if the load is sufficiently high. In fact, the addition of the load causes the activator dynamics to slow down compared to the repressor dynamics and hence the system moves to the "left" of the supercritical Hopf bifurcation so that the equilibrium point becomes stable. Using the results of the previous section, the clock dynamics with load on the activator become (see Jayanthi and Del Vecchio [2012] for the full derivation)

$$\begin{aligned} \dot{A} &= \left(\frac{\alpha_A (A/K_a)^n + \alpha_{0,A}}{1 + (A/K_a)^n + (R/K_r)^m} - \delta_A A\right)\left(\frac{1}{1 + \mathcal{R}'(A)}\right), \\ \dot{R} &= \frac{\alpha_R (A/K_a)^n + \alpha_{0,R}}{1 + (A/K_a)^n} - \delta_R R, \end{aligned} \tag{7}$$

in which we have $\mathcal{R}'(A) = np_T(A^{(n-1)}/K_a^n)/(1 + (A/K_a)^n)^2$, so that the activator dynamics become slower compared to the unloaded case (1). In particular, the effect of retroactivity is that of making the rate of change of A slower compared to the rate of change of R. As a consequence, when the activator load p_T is increased, the system goes through a Hopf bifurcation so that the stable

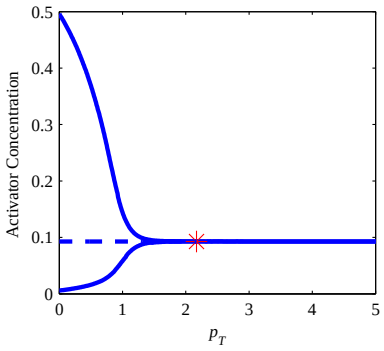


Fig. 5. Bifurcation diagram for the system in equation (7), in which activator load is p_T. A Hopf bifurcation occurs in correspondence to the asterisk. Parameters are the same as in Figure 1.

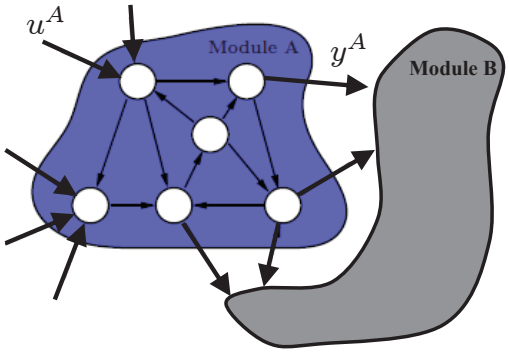


Fig. 6. Network A connects to network B.

limit cycle disappears and the system trajectories converge to a stable equilibrium point (Figure 5).

At this point, one may ask how the clock behavior would be affected if the load were applied to the repressor. Since retroactivity slows down the dynamics of the repressor, one should expect that a non-oscillating clock can be turned into an oscillating clock for sufficiently high load as the system moves through the Hopf bifurcation. This is in fact the case as formally shown in the work of Jayanthi and Del Vecchio [2012].

These results illustrate that downstream targets allow to tune the dynamics of a module without changing the parts of the module itself, hence providing an additional degree of freedom in design. Therefore, retroactivity does not always have a negative connotation and natural systems may actually use it in advantageous ways. It is very well known, for example, that transcription factors in natural systems can have large numbers of DNA binding sites, several of which do not even have regulatory functions (see Burger et al. [2010]). Our finding highlights an additional role that these DNA binding sites may have, that is, to tune effective kinetic rates in natural gene regulatory networks.

3.3 A Thevenin-like framework to quantify retroactivity in gene networks

The operative quantification of retroactivity outlined in the previous sections can be extended to the interconnection of any two general gene transcription networks (Figure (6)). In particular, given the isolated modules dynamics $\dot{x}^A = f_0^A(x^A, u^A)$ and $\dot{x}^B = f_0^B(x^B, u^B)$, in which x^A, x^B, u^A, u^B are vectors with $\dim(y^A) = \dim(u^B)$, one can demonstrate (see Gyorgy and Del Vecchio [2013] for the technical details) that the dynamics of connected module A will have the form

$$\dot{x}^A = (I + (I + R^A)^{-1} S^B)^{-1} (f_0^A(x^A, u^A) + (I + R^A)^{-1} M^B f_0^B(x^B, y^A)), \tag{8}$$

in which R^A, S^B, and M^B are state-dependent matrices. Specifically, R^A depends only on parameters (promoter amounts and dissociation constants) of module A, while S^B and M^B depend only on parameters of module B. Matrix R^A is called *internal retroactivity* and quantifies the effect of intramodular connections on the dynamics of module A in isolation. In fact, the vector field $f_0^A(x^A, u^A)$ contains in its expression also matrix R^A according to $f_0^A(x^A, u^A) = (I + R^A)^{-1}[(g^A(x^A, u^A) - Q^A \dot{u}^A]$, in which $g^A(x^A, u^A)$ is a vector field that models the dynamics of the nodes in module A neglecting retroactivity and Q^A is a matrix that determines how the rate of change of the input affects the state (see Gyorgy and Del Vecchio [2013] for details). Matrix M^B is called the *mixing retroactivity* and quantifies the "coupling" between the isolated dynamics of module A and the isolated dynamics of module B. Specifically, when $M^B = 0$, the isolated dynamics f_0^B of module B do not appear in the dynamics of module A, so that the two dynamics are not "mixed". From a physical point of view, this mixing occurs when nodes in module B have parents both from module B itself and module A, so that transcription factors from A and B interfere with each other while binding to promoter sites in module B. Matrix M^B models the phenomenon by which transcription factors in B can force transcription factors from A to bind/unbind promoter sites in module B, thus effectively changing the free concentration of transcription factors in A. When $M^B = 0$, the dynamics of module A are simply given by

$$\dot{x}^A = (I + (I + R^A)^{-1} S^B)^{-1} f_0^A(x^A, u^A), \tag{9}$$

so that the dynamics of module A are a "matrix-scaled" version of the dynamics of A in isolation. This is why matrix S^B is called the *scaling retroactivity* and quantifies the loading effect that module B has on module A due to transcription factors in A binding to promoter sites in module B. When also $S^B = 0$, the dynamics of module A are the same as in isolation.

These retroactivity matrices can be calculated once the measurable parameters (promoter amounts and dissociation constants) and the interconnection graph are known. Specifically, they can be calculated as:

$$R^A(x^A, u^A) = \sum_{i \in \text{ nodes in A}} V_i^T R_i(x^A, u^A) V_i$$

$$S^B(x^A, x^B) = \sum_{i \in \text{ input nodes in B}} W_i^T R_i(x^A, u^A) W_i$$

$$M^B(x^A, x^B) = \sum_{i \in \text{ input nodes in B}} W_i^T R_i(x^A, u^A) D_i,$$

in which $R_i(x^A, u^A)$ is the retroactivity of the node and depends on the promoter concentration at the node and

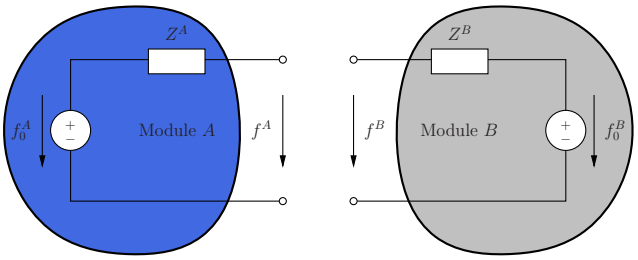


Fig. 7. Equivalent representation of an upstream electrical network (Module A) and a downstream electrical network (Module B).

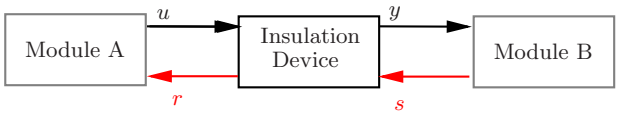


Fig. 8. An insulation device is placed between an upstream module A and a downstream module B in order to protect these systems from retroactivity. An insulation device should have $r \approx 0$ and the dynamic response of y to u should be practically independent of s.

on the dissociation constant of the same promoter. For example, if the node has only one parent, R^i is a scalar quantity and it has the expression given in equation (6). If the node has two parents, R^i is a two by two matrix whose off-diagonal entries are zero if the parents bind independently to the promoter, while they are non-zero when the parents bind non-independently (cooperatively or competitively). The specific expressions can be found in the work of Gyorgy and Del Vecchio [2012]. The matrices V_i, W_i, and D_i have entries equal to 0 or 1 and encode the interconnection graph among nodes. Specifically, V_i encodes the topology of the interconnection graph internal to module A, while W_i and D_i encode the topology of the interconnection graph between nodes in module A and nodes in module B.

The conceptual analogy between equation (8) and Thevenin's theorem for electrical networks is instructive. Specifically, Thevenin's result shows that a linear electrical network between two terminals can be reduced, independently of its complexity, to an equivalent ideal energy source and an equivalent impedance (Agarwal and Lang [2005]). In particular, an upstream module A has an equivalent voltage source f_0^A and an equivalent output impedance Z^A, while its downstream module B can be represented by an equivalent voltage source f_0^B and an equivalent input impedance Z^A (Figure 7). Referring to Figure 7, when module A and module B are not connected to each other the voltage at the output terminals of module A is given by $f^A = f_0^A$. When these modules are connected, the output voltage of module A becomes

$$f^A = \frac{1}{(Z^A/Z^B)+1} f_0^A + \frac{Z^A/Z^B}{(Z^A/Z^B)+1} f_0^B,$$

so that when $Z^A/Z^B \approx 0$, that is, the output impedance of module A is much smaller than the input impedance of module B, then we have that $f^A \approx f_0^A$. Comparing this equation with equation (8), it is apparent that the ratio Z^A/Z^B plays a conceptually similar role to the "ratios" $(I+R^A)^{-1}S^B$ and $(I+R^A)^{-1}M^B$: as S^B and M^B become "small" compared to $(I+R^A)$, equivalently, Z^A becomes "small" compared to Z^B, the output of module A is the same as that in isolation (not connected to module B). Here, we have placed quotes on "small" since we are dealing with matrices for the case of gene networks and with complex numbers for the case of electrical circuits. The notion of small in the case of the gene network can be made precise by calculating the matrix 2-norm (see Gyorgy and Del Vecchio [2013] for more details). Therefore, from a conceptual point of view the internal retroactivity R^A plays a similar role to the output admittance $1/Z^A$ of module A in the electrical circuit, while S^B plays a similar role to the input admittance $1/Z^B$ of module B in the electrical circuit. This analogy is purely conceptual since the structure of the two systems (gene network and electric network) is fundamentally different especially since one (the electrical network) is linear while the other (gene network) is nonlinear.

We can further quantify the difference between trajectories of connected and isolated systems as function of measurable biochemical parameters. Specifically, let $x(t)$ and $\hat{x}(t)$ denote the trajectory of the state of Module A when in isolation and when connected to module B, respectively. When $M^B = 0$, it can be shown under mild assumptions that the difference between the isolated and connected module trajectories satisfies

$$\|x(t) - \hat{x}(t)\|_2 = O(\mu),\ \mu = \frac{\sigma_M(S^B)/\sigma_m(I+R^A)}{1-\sigma_M(S^B)/\sigma_m(I+R^A)},$$

in which $\sigma_M(S^B)/\sigma_m(I+R^A) < 1$, and $\sigma_M(A)$ and $\sigma_m(A)$ denote the largest and smallest singular values of A, respectively. Parameter μ provides a metric of robustness to interconnections. It can be employed to either partition a natural network into modules whose behavior can be isolated (for modules with small μ), to some extent, from that of the surrounding modules. At the same time, it can also be employed in the bottom-up design of synthetic circuits by matching the output retroactivity of module A to the scaling retroactivity of module B so that μ is small, leading to a module A that preserves its isolated behavior after connection to module B.

In summary, this framework can be employed to predict the behavior of a transcription network once it is connected into a larger system, that is, to understand how the input/output properties of a module depend on its context.

4. INSULATION

From a design point of view, it is often desirable that the behavior of the upstream system does not change when it is connected to a downstream one. However, we have seen in the previous sections that retroactivity causes a potentially dramatic change in the dynamics of the upstream system. If one cannot design the downstream system to have low scaling and mixing retroactivity compared to the internal retroactivity of the upstream system, a different approach is required to connect the two systems while preserving their isolated input/output response. A viable option is to design a device that, once placed between module A and module B, insulates them from retroactivity effects (Figure 8). We call this device an *insulation device* and we formally specify its properties by requiring that (a) $r \approx 0$ and (b) the effect of s on y is completely attenuated. The first requirement can be satisfied by picking the

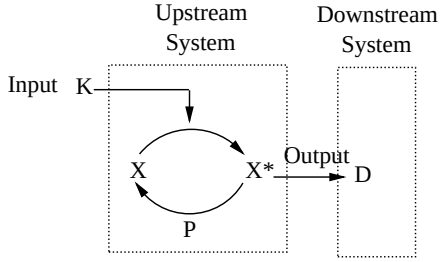


Fig. 9. A phosphorylation cycle is a protein modification mechanism in which an inactive protein X is converted by a kinase K to an active form X*, which is converted back to X by a phosphatase P.

parts of the input nodes of the device so that they have low retroactivity to the input (low scaling and mixing retroactivity). The second requirement can be formulated as a disturbance attenuation problem and it is the focus of the next section.

4.1 Retroactivity attenuation through high-gain feedback

One established technique for disturbance attenuation is high-gain negative feedback (Young et al. [1977]). We illustrate how this idea applies to our problem by considering again the dynamics of X in the isolated

$$\dot{X} = K(t) - \delta X$$

and connected

$$\dot{X} = (K(t) - \delta X)(1/(1 + \mathcal{R}(X))$$

upstream system configuration. In this case, the idea of high-gain feedback is to apply a negative feedback gain G and, in order not to attenuate the signal $X(t)$, to apply a similarly large gain $G' = \alpha G$ for some $\alpha > 0$ to the input $K(t)$. In this case, the dynamics of the isolated and connected systems become

$$\begin{aligned} \dot{X} &= G'K(t) - \delta X - GX \text{ (isolated)} \\ \dot{\bar{X}} &= (G'K(t) - \delta \bar{X} - G\bar{X}) \left(\frac{1}{1 + \mathcal{R}(\bar{X})} \right) \text{ (connected)} \end{aligned} \tag{10}$$

so that when G increases, we have that $|X(t) - \bar{X}(t)| \to O(1/G)$ (Del Vecchio et al. [2008]), which implies that the effect of retroactivity on $\bar{X}(t)$ is attenuated as G increases.

The next question is how to implement this mechanism through a biomolecular system. To this end, consider a phosphorylation cycle (Figure 9), in which protein X is converted to the active form X* through an enzymatic reaction with a kinase K (the input of the system) and converted back through another enzymatic reaction with a phosphatase P (Klipp et al. [2005]). Only when active, protein X can act as a transcription factor and activate or repress the expression of D. The basic idea is that amplification of the input K should occur through the forward cycle reaction while the negative feedback should occur through the reverse cycle reaction. In order to understand how this can be explained mathematically, we consider a simple model of the cycle, in which enzymatic reactions are modeled through one-step reactions:

$$\begin{aligned} \mathrm{K} + \mathrm{X} &\xrightarrow{k_1} \mathrm{K} + \mathrm{X}^* \\ \mathrm{P} + \mathrm{X}^* &\xrightarrow{k_2} \mathrm{P} + \mathrm{X}, \end{aligned}$$

in which we let X_T denote the total amount of cycle protein. Along with these equations, we need to model also the binding reaction of X* with sites in the downstream system so that, after applying singular perturbation as performed before, we obtain

$$\dot{X}^* = (k_1 X_T K(t)(1 - X^*/X_T) - k_2 P X^*) \left(\frac{1}{1 + \mathcal{R}(X^*)} \right). \tag{11}$$

Comparing this equation with equation (10), we see that the gain on the input is given by $G' = k_1 X_T$, while the negative feedback gain is given by $G = k_2 P$. As a consequence, we can conclude that as X_T and P are increased, the behavior of X^* should be minimally affected by the retroactivity to the output of the phosphorylation cycle. This result implies that phosphorylation cycles can function as insulation devices, suggesting another reason why these cycles are ubiquitous in natural signal transduction: they can enforce unidirectional signal propagation, which is certainly desirable in any (human-made or natural) signal transmission system. Related works have proposed that adding an explicit negative feedback to the cycle should further improve its robustness (Sauro and Ingalls [2007]). Here, we focus on the cycle without explicit negative feedback due to simpler experimental implementation.

The hypothesis that phosphorylation cycles function as insulation devices when both the amounts of cycle protein and phosphatase are sufficiently large is appealing since it can be experimentally tested. However, model (11), while providing an intuition behind the mechanisms responsible for retroactivity attenuation, is overly simplified and hides many important details that may be relevant to the overall cycle robustness. Specifically, covalent modification cycles can be modeled considering the enzymatic reactions as two-step processes (Goldbeter and Koshland [1981]) and can even include all of the details of the units that makeup a protein, some (or all) of which can be modified by the enzymes (Ventura et al. [2010b]). In the latter case, the ODE model of the upstream system in Figure 9 can even have several dozens of state variables. Hence, a mathematical framework to study insulation from retroactivity is required to handle models of arbitrary dimension. This is illustrated in the next section.

4.2 Retroactivity attenuation through time scale separation

We have developed a technique to analyze and design retroactivity attenuation, which holds for arbitrarily complex biomolecular network models. The basic idea exploits separation of time scales and the specific structure of the interconnection between biomolecular modules. This can be illustrated through the following simplified treatment. For the general results, the reader is referred to the work of Jayanthi and Del Vecchio [2011].

For the insulation device of Figure 8, which we refer to as system Σ, we seek to determine conditions under which the dynamic response of y to u is minimally affected by retroactivity s. In order to do so, we write the model of the system in its isolated configuration ($s = 0$) and in its connected configuration ($s \neq 0$) and quantify the difference between the trajectories of the isolated and connected systems. Specifically, letting u be a vector variable and assuming for simplicity that $y \in \mathbb{R}^n$ is the

state of Σ, we can write the dynamics of the isolated system as

$$\begin{aligned}\dot{u} &= f_0(t,u) + r(u,y)\\ \dot{y} &= G_1 f_1(u,y),\end{aligned} \tag{12}$$

in which $G_1 > 0$ is a positive constant, and we can write the dynamics of the connected system as

$$\begin{aligned}\dot{\bar{u}} &= f_0(t,\bar{u}) + r(\bar{u},\bar{y})\\ \dot{\bar{y}} &= G_1 f_1(\bar{u},\bar{y}) + G_2 M s(\bar{y},v)\\ \dot{v} &= -G_2 N s(\bar{y},v),\end{aligned} \tag{13}$$

in which $G_2 > 0$ is also a positive constant and M and N are matrices called stoichiometry matrices (Klipp et al. [2005]). Here, v is a vector variable that models the dynamics of the downstream system to which Σ is connected. The distinctive structure of the interconnection comes in the fact that matrices M and N are such that there is a non-singular $n \times n$ matrix B and a matrix T such that $BM - TN = 0$. This is the case because the entries of s physically represent the rate of reversible binding between two species and it always affects with opposite signs the species involved in the binding, one of which belongs to the upstream system and the other belongs to the downstream system. The constant G_2 models the fact that binding reactions are among the fastest reactions in biomolecular networks, so that $G_2 \gg 1$.

Now, assume that we can take $G_1 \gg 1$ probably not as large as G_2 but still sufficiently larger than 1. This can be achieved, for example, by letting y be driven by protein modification reactions, such as phosphorylation or allosteric modification, which are usually much faster than protein production and decay processes (Klipp et al. [2005]). In this case, we can re-write the dynamics of the connected system (13) by using the change of variables $z = B\bar{y} + Tv$, $\epsilon_1 = 1/G_1$, and $\epsilon_2 = 1/G_2$ as

$$\begin{aligned}\dot{\bar{u}} &= f_0(t,\bar{u}) + r(\bar{u},\bar{y})\\ \epsilon_1 \dot{z} &= B f_1(\bar{u},\bar{y})\\ \epsilon_2 \dot{v} &= -s(\bar{y},v),\end{aligned}$$

which is in standard singular perturbation form with two small parameters. Under the assumption that the slow manifold is locally exponentially stable (the technical conditions can be found in Jayanthi and Del Vecchio [2011]) the above system can be well approximated by one in which $\epsilon_1 = 0$ and $\epsilon_2 = 0$. This leads to $\bar{y} = \gamma(\bar{u})$ (the locally unique solution of $f_1(\bar{u},\bar{y}) = 0$), which is the same solution as found in the isolated system (12) when $\epsilon_1 = 0$. As a consequence, we can conclude that $\|y(t) - \bar{y}(t)\| = O(\epsilon_1)$ for $0 < t_b \le t < T$ for some $t_b > 0$ and $T > 0$, independently of the value of G_2. This result implies that if the time scale of Σ is sufficiently faster than the time scale of the input and suitable stability conditions are satisfied, then Σ attenuates the effects of retroactivity s on the response of y to u.

4.3 Implementation through phosphorylation cycles

In view of the result of the previous section, we can consider a more realistic model of the phosphorylation cycle and exploit the natural time scale separation between phosphorylation and gene expression (controlling $K(t)$) to show retroactivity attenuation.

Consider a two-step reaction model for the phosphorylation reactions, given by

$$\mathrm{K} + \mathrm{X} \underset{\beta_2}{\overset{\beta_1}{\rightleftharpoons}} \mathrm{C}_1 \xrightarrow{k_1} \mathrm{X}^* + \mathrm{K}$$

$$\mathrm{P} + \mathrm{X}^* \underset{\alpha_2}{\overset{\alpha_1}{\rightleftharpoons}} \mathrm{C}_2 \xrightarrow{k_2} \mathrm{X} + \mathrm{P}$$

with conservation laws $P_T = P + C_2$, $A_T = X + X^* + C_1 + C_2 + C$, along with the binding of X^* with downstream sites p, that is,

$$\mathrm{X}^* + \mathrm{p} \underset{k_{\mathrm{off}}}{\overset{k_{\mathrm{on}}}{\rightleftharpoons}} \mathrm{C}.$$

The resulting ODE model is given by

$$\begin{aligned}\dot{K} &= k(t) - \delta K - \beta_1 X_T K \left(1 - \frac{X^*}{X_T} - \frac{C_1}{X_T} - \frac{C_2}{X_T} - \frac{C}{X_T}\right)\\ &\quad + (\beta_2 + k_1) C_1\\ \dot{C}_1 &= -(\beta_2 + k_1) C_1 + \beta_1 X_T K \left(1 - \frac{X^*}{X_T} - \frac{C_1}{X_T} - \frac{C_2}{X_T}\right.\\ &\quad \left. - \frac{C}{X_T}\right)\\ \dot{C}_2 &= -(k_2 + \alpha_2) C_2 + \alpha_1 P_T X^* \left(1 - \frac{C_2}{P_T}\right)\\ \dot{X}^* &= k_1 C_1 + \alpha_2 C_2 - \alpha_1 P_T X^* \left(1 - \frac{C_2}{P_T}\right)\\ &\quad + k_{\mathrm{off}} C - k_{\mathrm{on}} X^* (p_T - C)\\ \dot{C} &= -k_{\mathrm{off}} C + k_{\mathrm{on}} X^* (p_T - C).\end{aligned} \tag{14}$$

To take this system in the form (13), we can define the gain G_1 by considering the separation of time scales between gene expression and protein phosphorylation, so that

$$G_1 := \frac{\beta_1 X_T}{\delta} \gg 1,$$

$b_2 := \frac{\beta_2}{G_1}$, $a_2 := \frac{\alpha_2}{G_1}$, $a_1 := \frac{\alpha_1 P_T}{G_1}$, and $\kappa_i := \frac{k_i}{G_1}$ for $i = 1, 2$. Similarly, we can define the gain G_2 by considering the separation of time scales between gene expression and binding reactions, so that

$$G_2 := \frac{k_{\mathrm{off}}}{\delta} \gg 1$$

and $K_d := \frac{k_{\mathrm{off}}}{k_{\mathrm{on}}}$. By using the change of variables $z = K + C_1$ and ensuring that $X_T \gg p_T$ so that $C/X_T \ll 1$, we can re-write the system as

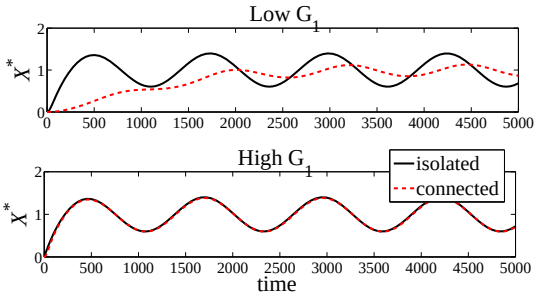


Fig. 10. Simulation of the phosphorylation cycle in (14) with low gain G_1 and high gain G_1 when $K(t)$ is a periodic signal. Specifically, we have $\delta = 0.01$, $X_T = 5000$, $P_T = 5000$, $\alpha_1 = \beta_1 = 2 \times 10^{-6} G_1$, and $\alpha_2 = \beta_2 = k_1 = k_2 = 0.01 G_1$, in which $G_1 = 10$ (upper panel), and $G_1 = 1000$ (lower panel). The downstream system parameters are $k_{\text{on}} = 100$, $k_{\text{off}} = 100$ and, thus, $G_2 = 10000$. Simulations for the connected system ($s \neq 0$) correspond to $p_T = 100$ while simulations for the isolated system ($s = 0$) correspond to $p_T = 0$.

$$
\begin{aligned}
\dot{z} &= k(t) - \delta(z - C_1) \\
\dot{C}_1 &= G_1 \left[-(b_2 + \kappa_1) C_1 + \delta(z - C_1) \left(1 - \frac{X^*}{X_T} - \frac{C_1}{X_T} \right. \right. \\
&\qquad \left. \left. - \frac{C_2}{X_T} \right) \right] \\
\dot{C}_2 &= G_1 \left[-(\kappa_2 + a_2) C_2 + a_1 X^* \left(1 - \frac{C_2}{P_T} \right) \right] \\
\dot{X}^* &= G_1 \left[\kappa_1 C_1 + a_2 C_2 - a_1 X^* \left(1 - \frac{C_2}{P_T} \right) \right] \\
&\qquad + G_2 \left[\delta C - \frac{\delta}{K_d} X^* (p_T - C) \right] \\
\dot{C} &= -G_2 \left[\delta C - \frac{\delta}{K_d} X^* (p_T - C) \right],
\end{aligned}
$$

which is in the form of system (13) with

$$
\begin{gathered}
u = z,\ y = (C_1, C_2, X^*)',\ v = C, \\
s = k_{\text{off}} C - k_{\text{on}} X^* (p_T - C), \\
M = (0, 0, 1)',\ N = 1,\ r = 0.
\end{gathered}
$$

Therefore, the main result of the previous section applies with $B = I$ and $T = (0, 0, 1)'$, so that as G_1 increases, the response of X^* to K becomes insensitive to retroactivity s, as also reflected by the simulation results of Figure 10.

This technique is applicable to large models and was utilized for designing experiments on a covalent modification cycle *in vitro* (Jiang et al. [2011]) and for designing a buffer between an *in vitro* biomolecular oscillator and a load (Franco et al. [2009, 2011]).

From a design point of view, the gain G_1 can be made larger by increasing the total concentration of substrate X_T and of phosphatase P_T in comparable amounts. Note, however, that in doing so the loading applied to the input kinase K increases, so that the retroactivity to the input r of the insulation device also increases. As a consequence, by increasing G_1 this way, the retroactivity attenuation property improves to the expense of applying increased retroactivity to the input K. As a result, the output is not the desired one as it mirrors an input $K(t)$ that has been distorted due to retroactivity r. In this case, it becomes necessary to perform an optimal design in which

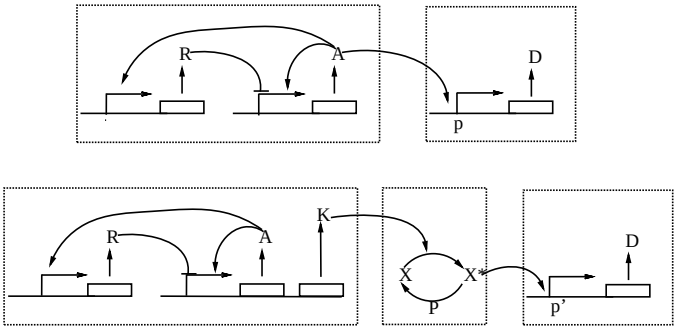


Fig. 11. Illustration of how one can interconnect the clock to its downstream system through the insulation device of Figure 9. The top diagram illustrates a simplified genetic layout of the activator-repressor clock of Figure 1(a). The boxes represent the genes expressing R, A, and D, while the arrows upstream of the genes represent the promoters that control these genes. The bottom diagram illustrates how the genetic layout of the clock should be modified such that it can connect to the phosphorylation cycle that takes as input the kinase K. In this case, the downstream system still expresses protein D, but its expression is controlled by a different promoter that is activated by X* as opposed to being activated by A.

the amounts of substrate and phosphatase are optimized to minimize the overall error due to both the contribution of r and of s. This optimal design can be performed by explicitly computing bounds on the errors due to r and s and my minimizing these, which, in turn, can be obtained by employing robustness results from contraction theory. For details on these, the reader is referred to the work of Rivera-Ortiz and Del Vecchio [2013].

A further downside in overexpressing proteins is that the cost of the circuit in terms of cellular resources increases (Stoebel et al. [2008]). Of particular concern in a phosphorylation cycle is the consumption of ATP, the cellular energy currency. An extensive analysis of how increased gains (increased substrate and phosphatase concentrations) reflect in an increased energy cost of the cycle can be found in the work of Barton and Sontag [2013]. Finally, it is well known from the Bode's integral formula for linear systems that increased gains can lead to undesired sensitivity increase in certain frequency ranges, ultimately potentially leading to noise amplification. This limitation is found also in the presented insulation design, for which a simplified analysis using the chemical Langevin equation can be found in the work of Jayanthi and Del Vecchio [2009]. All these constraints should be taken into account when implementing the insulation device design.

4.4 Connecting the clock to its downstream system through an insulation device

In this section, we illustrate how the insulation device implemented through phosphorylation of Figure 9 can be employed to connect the clock to the desired downstream system such that (a) the clock behavior is not significantly perturbed and (b) the signal from the clock is transmitted to the downstream system robustly with respect to the load that this downstream system applies. To this end, consider the diagrams depicted in Figure 11. The top diagram illustrates a simplified genetic layout of the clock. The activator A is expressed from a gene under the control

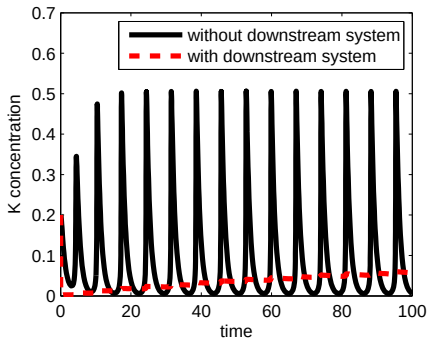


Fig. 12. Simulation results for the concentration of protein K in Figure 11 in the case in which this were used directly as an input to the downstream system, thus binding sites p' (dashed red plot).

of a promoter activated by A and repressed by R, while the repressor is expressed from a gene under the control of a promoter activated by A. Protein A, in turn, activates the expression of protein D in the downstream system. In this case, the promoter p controlling the expression of D contains a sequence that A recognizes, so that A can bind to it.

When the insulation device of Figure 9 is employed to interconnect the clock to the downstream system, two modifications need to be made to enable the connections. Since A is not a kinase, we need to insert in sequence to the gene expressing A the gene expressing the kinase K (bottom diagram of Figure 11). Since both A and K are under the control of the same promoter, they will be produced at the same rates and hence the concentration of K should mirror that of A. Note that a solution in which we insert in sequence to the gene expressing A a transcription factor K that directly binds to downstream promoter sites p' to produce D (without the insulation device in between) would not solve the problem. In fact, while the clock behavior would be preserved in this case, the behavior of the concentration of K would not mirror that of A since protein K would be loaded by the downstream promoter sites p' (Figure 12). As a consequence, we would still fail to transmit the clock signal to protein D. The second modification that needs to be made is to change the promoter p to a new promoter p' that has a sequence that the protein X* recognizes (bottom diagram of Figure 11).

In the case of the bottom diagram of Figure 11, the dynamics of the clock remain that of model (1). To these, we need to add the dynamics of $K(t)$, which, when the phosphorylation cycle is not present, will be given by

$$\dot{K} = \frac{\alpha_A (A/K_a)^n + \alpha_{0,A}}{1 + (A/K_a)^n + (R/K_r)^m} - \delta_K K. \qquad (15)$$

When the phosphorylation cycle is present, this ODE changes to

$$\begin{aligned} \dot{K} &= \frac{\alpha_A (A/K_a)^n + \alpha_{0,A}}{1 + (A/K_a)^n + (R/K_r)^m} - \delta_K K \\ &-\beta_1 X_T K \left(1 - \frac{X^*}{X_T} - \frac{C_1}{X_T} - \frac{C_2}{X_T} - \frac{C}{X_T}\right), \end{aligned} \qquad (16)$$

in which the term $r = -\beta_1 X_T K \left(1 - \frac{X^*}{X_T} - \frac{C_1}{X_T} - \frac{C_2}{X_T} - \frac{C}{X_T}\right)$ represents the retroactivity to the input of the insulation device realized by the phosphorylation cycle. The ODE model of the insulation device with the downstream system remains the same as before and given by the second to fifth equations of system (14), in which we replace p_T by p'_T, so that the retroactivity to the output of the insulation device is given by $s = k_{\text{off}} C - k_{\text{on}} X^* (p'_T - C)$. Figure 13 shows the trajectories of $A(t)$, $K(t)$, and $X^*(t)$ for the system of Figure 11. As desired, the signal $X^*(t)$, which drives the downstream system, closely tracks that of $A(t)$ despite the retroactivity due to load applied by the downstream sites p'. Note that because of a non-zero retroactivity to the input r of the insulation device, the trajectory of $K(t)$ is slightly different from the same trajectory in the absence of the insulation device (Figure 13(b)). The retroactivity to the output s only slightly affects the output of the insulation device (Figure 13(c)). The plot of Figure 13(c), showing the signal that drives the downstream system, can be directly compared to the signal that would drive the downstream system in the case in which the insulation device would not be used (Figure 12). In the latter case, the downstream system would not be properly driven, while with the insulation device it is.

4.5 Natural mechanisms for retroactivity attenuation

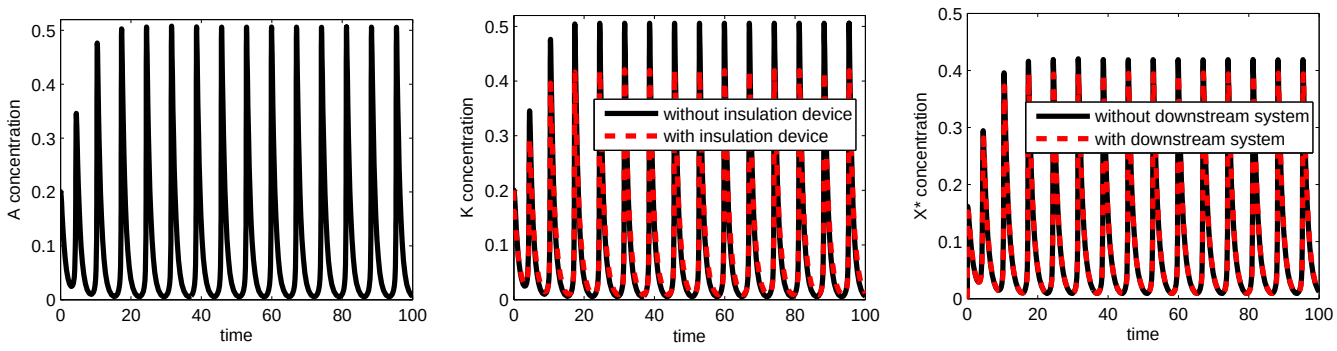
In natural signal transduction systems, phosphorylation cycles often appear in cascades, such as in the MAPK cascades (Seger and Krebs [1995], Rubinfeld and Seger [2005]), where signals from outside the cell are transmitted down to the gene expression machinery. Cascades often intersect with each other by sharing common substrates, so that perturbations at the bottom or at intermediate stages of a cascade are frequently applied by the intersecting cascade(s) (Roux and Blenis [2004], Müller [2004]). Since, we have seen that a load applied to the output protein of a cycle can affect the dynamic state of the cycle itself, a natural question is how these downstream perturbations propagate backward toward the upstream stages of the cascade. Specifically, the role of the length of a cascade has been subject of many studies in the systems biology community, indicating that it has specific functions in signal amplification, signal duration, and signaling time (Heinrich et al. [2002], Chaves et al. [2004]). Here, we investigate whether the length of the cascade has any role in how perturbations applied downstream of a cascade propagate backward.

To investigate this question, we refer to the cascade depicted in Figure 14. Here, the downstream substrate D has total available amount D_T, which we assume is subject to perturbations due to other intersecting cascades. Letting d_T be a small perturbation of D_T about a steady state value, the ODE system can be linearized about the steady state corresponding to D_T, and a static gain from d_T to the total phosphorylated ith cycle protein perturbation z_i can be obtained (see the work of Ossareh et al. [2011] for details). Let

$$\Phi_i := \frac{z_i}{z_{i+1}} \text{ and } \Phi_{tot} := \Pi_{i=1}^{n-1} \Phi_i$$

represent the gain from downstream stage i to upstream stage $i+1$ and the gain from stage n to stage 1, respectively. Then, it is possible to show that

$$|\Phi_i| < 1 \text{ and } \Phi_i < 0,$$



(a) (b) (c)

Fig. 13. Simulation results for the system of Figure 11. Panel (a) shows the concentration of the clock activator protein A. Panel (b) shows the concentration of the kinase K with and without ($r = 0$ in equation (16)) the insulation device. Panel (c) shows the behavior of the output of the insulation device without ($s = 0$) and with the downstream system. The clock parameters are the same as those in Figure 1 and $\delta_K = \delta_A$. The phosphorylation cycle parameters are as follows: $k_1 = 5000$, $k_2 = 5000$, $\alpha_1 = 1$, $\beta_1 = 1$, $\beta_2 = 1000$, $\alpha_2 = 1000$, $P_T = 1500$, and $X_T = 1500$. The load parameters are given by $k_{\text{on}} = k_{\text{off}} = 100$ and $p'_T = 100$.

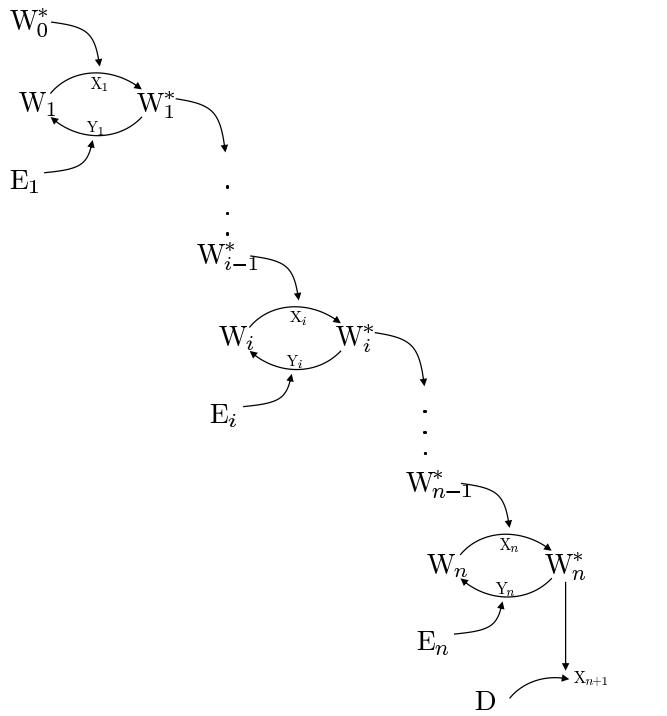


Fig. 14. Cascade of phosphorylation cycles. Proteins W_i^* are kinases that lead to phosphorylation of the substrates at the next downstream stage $i+1$. Proteins E_i are phosphatases for stage i. D represents a downstream substrate, where we assume a perturbation is applied.

that is, the perturbation from stage $i+1$ to stage i is attenuated in magnitude and there is a reversal of the sign as the perturbation travels from one stage to the next upstream. As a consequence, the total gain Φ_{tot} is also strictly less than one and it becomes smaller as the number of stages increases. That is, the perturbation in the total phosphorylated protein is attenuated as it propagates upstream in the cascade. Hence, longer cascades present enhanced retroactivity attenuation ability. For a more detailed study also including the free phosphorylated protein and a computational study with parameters from established databases, the reader is referred to the work of Ossareh et al. [2011].

This finding illustrates that natural systems may mitigate retroactivity effects by employing multiple stages of covalent modification, allowing each stage to have smaller gains. In fact, the same attenuation of retroactivity can be obtained by employing one stage with high gains or two stages, each with lower gain. The fact that natural systems may prefer using multiple stages with lower gains is not surprising. In fact, high gains imply significant expenditure both in terms of energy (ATP) and in terms of gene expression machinery, such as ribosomes, RNA polymerase, and amino acids (Barton and Sontag [2013]).

5. CONCLUSIONS AND DISCUSSION

In this review paper, we have illustrated how problems of loading, called retroactivity, are found in biomolecular systems just like they are found in many engineering systems. These effects alter the behavior of a module upon interconnection and hinder modular analysis and design approaches. Differently from electrical circuits, which can be analyzed to a large extent through linear systems theory, biomolecular network models are highly nonlinear and hence their study requires nonlinear systems theory. We have illustrated how, using singular perturbation theory, we can analyze and quantify retroactivity effects by obtaining equivalent system representations, just like Thevenin's theorem does for electrical circuits. Inspired by the approach used in electrical circuit design, we analyzed the question of how to design devices that attenuate the retroactivity to the output and have low retroactivity to the input (insulation devices). We described an approach to answer this question by exploiting the structure of the interconnection found in biomolecular networks and employing singular perturbation theory after a suitable change of variables. We illustrated on a simple example, how an insulation device could be used in synthetic biology circuits to connect systems to each other, while keeping the isolated systems behaviors.

Many system-level challenges remain to be addressed for the analysis and design of biological networks. First of all,

we have not touched on the problem of noise in biology, which is a central issue since all the reactions that make the circuits are probabilistic in nature. Unfortunately, while many analysis and design tools for nonlinear dynamical systems are available, only a limited pool of similar results are available for stochastic systems (Khammash and ElSamad [2005]). Hence, problems such as designing the probability distribution of a system's trajectories, as opposed to designing their mean value, remain mostly open, while being very important for engineering biological circuits. We have briefly mentioned the interactions that arise between biomolecular circuits and the cell machinery, such as due to loading of cellular resources or to interfering directly with important cellular pathways. These interactions are rarely modeled due to the fact that most of them are unknown. As a result, it is currently very difficult to design circuits that behave robustly once in the cellular chassis. Designing circuits such that their function is robust to external perturbations, parameter uncertainty, unwanted interactions, and competition for shared resources is a central problem, which is currently open. Robust design approaches have been extensively developed in the control community, but most of these tools are hardly applicable to biomolecular systems, which are highly nonlinear and often rely on nonlinearities for proper functioning. Finally, another hard question that remains open is the "realization" question, that is, how to implement a given function with the available biological parts. Addressing this question is especially challenging as it requires, even more than the other questions, a synergistic interplay of system-level analytical expertise and a deep knowledge of the available biological mechanisms.

REFERENCES

A. Agarwal and J. H. Lang. *Foundations of Analog and Digital Electronic Circuits*. Morgan Kaufmann Publishers, 2005.

B. Alberts, A. Johnson, J. Lewis, M. Raff, and K. Roberts. *Molecular Biology of the Cell*. Garland Science, fifth edition, 2007.

R. P. Alexander, P. M. Kim, T. Emonet, and M. B. Gerstein. Understanding modularity in molecular networks requires dynamics. *Sci. Signal.*, 2(81):pe44, 2009.

U. Alon. Network motifs: theory and experimental approaches. *Nature*, 8:450–461, June 2007a.

U. Alon. *An introduction to systems biology. Design principles of biological circuits*. Chapman-Hall, 2007b.

James Anderson, Yo-Cheng Chang, and Antonis Papachristodoulou. Model decomposition and reduction tools for large-scale networks in systems biology. *Automatica*, 47(6):1165–1174, 2011.

E. Andrianantoandro, S. Basu, D. K. Karig, and R. Weiss. Synthetic biology: New engineering rules for an emerging discipline. *Molecular Systems Biology*, pages 1–14, 2006.

A. Asthagiri and D. Lauffenburger. Bioengineering models of cell signaling. *Annual Review of Biomedical Engineering*, 2:31–53, 2000.

M. R. Atkinson, M. A. Savageau, J. T. Meyers, and A. J. Ninfa. Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in *Escherichia coli*. *Cell*, 113:597–607, 2003.

D. Baker, G. Church, J. Collins, D. Endy, J. Jacobson, J. Keasling, P. Modrich, C. Smolke, and R. Weiss. ENGINEERING LIFE: Building a FAB for biology. *Scientific American*, pages 44–51, June 2006.

J. Barton and E.D. Sontag. The energy costs of insulators in biochemical networks. *Biophysical Journal*, 104:1390–1380, 2013.

A. Becskei and L. Serrano. Engineering stability in gene networks by autoregulation. *Nature*, 405:590–593, 2000.

W. E. Bentley, N. Mirjalili, D. C. Andersen, R. H. Davis, and D. S. Kompala. Plasmid-encoded protein: the principal factor in the "metabolic burden" associated with recombinant bacteria. *Biotechnol Bioeng*, 35(7):668–81, 1990.

L. Bleris, Z. Xie, D. Glass, A. Adadey, E. Sontag, and Y. Benenson. Synthetic incoherent feedforward circuits show adaptation to the amount of their genetic template. *Molecular Systems Biology*, 7:519, 2011.

A. Burger, A.M. Walczak, and P.G. Wolynes. Abduction and asylum in the lives of transcription factors. *Proc. Natl. Acad. Sci.*, 107: 401621, 2010.

S. Cardinale and A. P. Arkin. Contextualizing context for synthetic biology identifying causes of failure of synthetic biological systems. *Biotechnology Journal*, 7:856866, 2012.

M. Chaves, E. D. Sontag, and R. J. Dinerstein. Optimal length and signal amplification in weakly activated signal transduction cascades. *J. Phys. Chem.*, 108:15311–15320, 2004.

T. Danino, O. Mondragón-Palomino, L. Tsimring, and J. Hasty. A synchronized quorum of genetic clocks. *Nature*, 463:326–330, 2009.

D. Del Vecchio. Design and analysis of an activator-repressor clock in *E. coli*. In *Proc. American Control Conference*, pages 1589–1594, 2007.

D. Del Vecchio and E. D. Sontag. Engineering principles in biomolecular systems: From retroactivity to modularity. *European Journal of Control, Special Issue*, 15(3):389, 2009.

D. Del Vecchio, A. J. Ninfa, and E. D. Sontag. Modular cell biology: Retroactivity and insulation. *Molecular Systems Biology*, 4:161, 2008.

M. B. Elowitz and S. Leibler. A synthetic oscillatory network of transcriptional regulators. *Nature*, 403:339–342, 2000.

David A. Fell. Metabolic control analysis: a survey of its theoretical and experimental development. *Biochemical Journal*, 286:313–330, 1992.

E. Franco, D. Del Vecchio, and R. M. Murray. Design of insulating devices for in vitro synthetic circuits. In *Proc. IEEE Conf. Decision and Control*, 2009.

E. Franco, E. Friedrichs, J. Kim, R. Jungmann, R. Murray, E. Winfree, and F. C. Simmel. Timing molecular motion and production with a synthetic transcriptional clock. *Proc. Natl. Acad. Sci*, doi: 10.1073/pnas.1100060108, 2011.

T.S. Gardner, C.R. Cantor, and J.J. Collins. Construction of the genetic toggle switch in *Escherichia Coli*. *Nature*, 403:339–342, 2000.

D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81:2340–2361, 1977.

D. T. Gillespie. The chemical langevin equation. *The Journal of Chemical Physics*, 113:297–306, 2000.

A. Goldbeter and D. E. Koshland. An amplified sensitivity arising from covalent modification in biological systems. *Proc. National Academy of Sciences*, 78:6840–6844, 1981.

A. Gyorgy and D. Del Vecchio. Retroactivity to the input in complex gene transcription networks. In *Proc. IEEE Conf. Decision and Control*, pages 3595–3601, 2012.

A. Gyorgy and D. Del Vecchio. How slaves affect a master module in gene transcription networks. In *Proc. of IEEE Conference on Decision and Control*, 2013.

L.H. Hartwell, J.J. Hopfield, S. Leibler, and A.W. Murray. From molecular to modular cell biology. *Nature*, 402:47–52, 1999.

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

Reinhart Heinrich and Stefan Schuster. *The regulation of cellular systems*. Chapman & Hall, 1999.

S. Jayanthi and D. Del Vecchio. On the compromise between retroactivity attenuation and noise amplification in gene regulatory networks. In *Proc. Conference on Decision and Control*, pages 4565–4571, 2009.

S. Jayanthi and D. Del Vecchio. Retroactivity attenuation in biomolecular systems based on timescale separation. *IEEE Trans.*

Aut. Control, 56:748–761, 2011.

S. Jayanthi and D. Del Vecchio. Tuning genetic clocks employing DNA binding sites. *PLoS ONE*, 7:e41019, 2012.

S. Jayanthi, K. Nilgiriwala, and D. Del Vecchio. Retroactivity controls the temporal dynamics of gene transcription. *ACS Synthetic Biology*, DOI: 10.1021/sb300098w, 2013.

P. Jiang, A. C. Ventura, E. D. Sontag, A. J. Ninfa, and D. Del Vecchio. Load-induced modulation of signal transduction networks. *Science Signaling*, 4:ra67, 2011.

H. M. Kaltenbach and J. Stelling. Modular analysis of biological networks. *Adv. Exp. Med. Biol.*, 736:3–17, 2012.

H. Khalil. *Nonlinear Systems*. Prentice Hall, 2002.

M. Khammash and H. ElSamad. Stochastic modeling and analysis of genetic networks. In *Proc. Conference on Decision and Control*, pages 2320–2325, 2005.

Y. Kim, Z. Paroush, K. Nairz, E. Hafen, G. Jiménez, and S. Y. Shvartsman. Substrate-dependent control of mapk phosphorylation in vivo. *Mol. Sys. Biol.*, 7:467, 2011.

M. W. Kirschner and J. C. Gerhart. *The Plausibility of Life: Resolving Darwin's Dilemma*. Yale University Press, 2005.

E. Klipp, R. Herwig, A. Kowald, C. Wierling, and H. Lehrach. *Systems Biology in Practice*. Wiley-VCH, 2005.

H. Kobayashi, M. Krn, M. Araki, K. Chung, T. S. Gardner, C. R. Cantor, and J. J. Collins. Programmable cells: Interfacing natural and engineered gene networks. *Proc. National Academy of Sciences*, 101:8414–8419, 2004.

D. A. Lauffenburger. Cell signaling pathways as control modules: complexity for simplicity? *Proc. Natl. Acad. Sci.*, 97(10):5031–5033, May 2000.

W. H. Mather, J. Hasty, L. S. Tsimring, and R. J. Williams. Translational cross talk in gene networks. *Biophysical Journal*, 104:25642572, 2013.

M. Miller, M. Hafner, E. Sontag, N. Davidsohn, S. Subramanian, P. Purnick, D. Lauffenburger, and R. Weiss. Modular design of artificial tissue homeostasis: robust control through synthetic cellular heterogeneity. *PLoS computational biology*, 8:e1002579, 2012.

T. S. Moon, C. Lou, A. Tamsir, B. C. Stanton, and C.A. Voigt. Genetic programs constructed from layered logic gates in single cells. *Nature*, 491:249–253, 2012.

R. Müller. Crosstalk of oncogenic and prostanoid signaling pathways. *Journal of Cancer Research and Clinical Oncology*, 130(8):429–444, 2004.

W. H Mather N. A Cookson and, T. Danino, O. Mondragón-Palomino, R. J. Williams, L. S. Tsimring, and J. Hasty. Translational cross talk in gene networks. *Queueing up for enzymatic processing: correlated signaling through coupled degradation*, 7 (516), 2011.

H. R. Ossareh, A. C. Ventura, S. D. Merajver, and D. Del Vecchio. Long signaling cascades tend to attenuate retroactivity. *Biophysical Journal*, 100(7):1617–1626, 2011.

H. Paynter. *Analysis and design of engineering systems*. M.I.T. Press, Cambridge, Mass., 1961.

P. P. Peralta-Yahya, F. Zhang, S. B. del Cardayre, and J. D. Keasling. Microbial engineering for the production of advanced biofuels. *Nature*, 488:320–328, 2012.

J. W. Polderman and J. C. Willems. *Introduction to Mathematical Systems Theory. A Behavioral Approach. 2nd ed.* Springer-Verlag, New York, 2007.

P. Purnick and R. Weiss. The second wave of synthetic biology: from modules to systems. *Nature reviews. Molecular cell biology*, 10: 410–22, 2009.

P. M. Rivera-Ortiz and D. Del Vecchio. Optimal design of phosphorylation-based insulation devices. In *Proc. American Control Conference*, page to appear, 2013.

P.P. Roux and J. Blenis. ERK and p38 MAPK-Activated Protein Kinases: a family of protein kinases with diverse biological functions. *Microbiology and Molecular Biology Reviews*, 68(2):320–344, 2004.

H. Rubinfeld and R. Seger. The ERK cascade: a prototype of MAPK signaling. *Mol Biotechnol*, 31(2):151–174, 2005.

J. Saez-Rodriguez, A. Kremling, H. Conzelmann, K. Bettenbrock, and E. D. Gilles. Modular analysis of signal transduction networks. *IEEE Control Systems Magazine*, pages 35–52, 2004.

J. Saez-Rodriguez, A. Kremling, and E.D. Gilles. Dissecting the puzzle of life: modularization of signal transduction networks. *Computers and Chemical Engineering*, 29:619–629, 2005.

H. M. Sauro and B. Ingalls. MAPK cascades as feedback amplifiers. Technical report, http://arxiv.org/abs/0710.5195, Oct 2007.

H. M. Sauro and B. N. Kholodenko. Quantitative analysis of signaling networks. *Progress in Biophysics & Molecular Biology*, 86:5–43, 2004.

D.L. Schilling and C. Belove. *Electronic Circuits: Discrete and Integrated*. McGraw Hill, 1968.

M. Scott, C. W. Gunderson, E. M. Mateescu, Z. Zhang, and T. Hwa. Interdependence of cell growth and gene expression: Origins and consequences. *Science*, 330:1099–1202, 2010.

R. Seger and E. G. Krebs. The MAPK signaling cascade. *The FASEB Journal*, 9:726–735, 1995.

S. Y. Shvartsman and R. E. Baker. Mathematical models of morphogen gradients and their effects on gene expression. *Wires Developmental Biology*, page DOI: 10.1002/wdev.55, 2012.

A. L. Slusarczyk, A. Lin, and R. Weiss. Foundations for the design and implementation of synthetic genetic circuits. *Nature Reviews Genetics*, 13:406–420, 2012.

G. Vivek Sridharan, Soha Hassoun, and Kyongbum Lee. Identification of biochemical network modules based on shortest retroactive distances. *PLoS Computational Biology*, 7(11):e1002262, 2011.

D. M. Stoebel, A. M. Dean, and D. E. Dykhuizen. The cost of expression of escherichia coli lac operon proteins is in the process, not in the products. *Genetics*, 178:1653–1660, 2008.

J. Stricker, S. Cookson, M. R. Bennett, W. H. Mather, L. S. Tsimring, and J. Hasty. A fast, robust and tunable synthetic gene oscillator. *Nature*, 456:516–519, 2008.

L. C. Thevenin. Extension de la loi dohm aux circuits electromoteurs complexes [extension of ohms law to complex electromotive circuits]. *Annales Telegraphiques*, 10:222–224, 1883.

N. G. Van Kampen. *Stochastic Processes in PHysics and Chemistry*. Elsevier, third edition, 2007.

A. C. Ventura, P. Jiang, L. Van Wassenhove, D. Del Vecchio, S. D. Merajver, and A. J. Ninfa. The signaling properties of a covalent modification cycle are altered by a downstream target. *Proc. Natl. Acad. Sci. USA*, 107(22):10032–10037, 2010a.

A. C. Ventura, P. Jiang, L. Van Wassenhove, D. Del Vecchio, S. D. Merajver, and A. J. Ninfa. Signaling properties of a covalent modification cycle are altered by a downstream target. *Proc. Natl. Acad. Sci., USA*, 107:10032–10037, 2010b.

J.C. Willems. Behaviors, latent variables, and interconnections. *Systems, Control and Information*, 43:453–464, 1999.

Z. Xie, L. Wroblewska, L. Prochazka, R. Weiss, and K. Benenson. Multi-input rnai-based logic circuit for identification of specific cancer cells. *Science*, 333:1307–11, 2011.

E. Yeung, J. Kim, and R. M. Murray. Resource competition as a source of non-minimum phase behavior in transcription-translation systems. In *Proc. of IEEE Conf. Decision and Control*, 2013.

K. D. Young, P. V. Kokotovic, and V. I. Utkin. A singular perturbation analysis of high-gain feedback systems. *IEEE Trans. Aut. Control*, AC 22(6):931–938, 1977.

F. Zhang, J. M. Carothers, and J. D. Keasling. Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nat. Biotechnol.*, 30:354–359, 2012.