Bounding the Effect of Retroactivity in the Presence of Parameter Uncertainty*

Thomas P. Prescott[1] and Andras Gyorgy[2]

***Abstract*— As the number of synthetic genetic modules grows, the issue of reliably predicting their behavior upon interconnection becomes more pressing. The trajectory of an upstream module changes once connected to a downstream module due to retroactivity. Here, we employ dissipativity analysis to provide an upper bound on the L_2 measure of this difference. To obtain this upper bound we formulate a Sum of Squares (SOS) optimization problem which we then solve using semi-definite programming. One particular strength of this approach is the ability to successfully handle parameter uncertainties while providing guaranteed upper bounds on the difference between the trajectories. We illustrate how to apply our method in the case of the most recurrent motif in gene networks.**

I. INTRODUCTION

Engineered systems often show a modular structure, thereby allowing the design and analysis of large-scale systems by focusing on their constituent smaller parts. The fact that natural biological networks also show modular structure is thus promising as it suggests that these complex networks can be comprehended and designed by following a similar strategy [1]. However, genetic modules display context-dependent behavior [2] as a result of dependence on the host organism and strain [3], growth-dependence [4], environmental dependence [5], and unwanted couplings due to the composition of modules [6], [7]. In this paper, we focus on this last source of context-dependence called retroactivity, a phenomenon in which a downstream module changes the behavior of an upstream module [6], [7]. Possible changes include an increase in response time [8], the alteration of both the frequency and amplitude of transcriptional oscillators [9], and the performance of signaling networks *in vitro* [10] and cascades *in vivo* [11].

In the case of transcription networks, the effects of retroactivity have been characterized in the general framework introduced in [7]. While [7] answers *how* the behavior of connected modules change, it lacks quantitative insight into *how much* a downstream module affects the performance of an upstream module. Here, we employ dissipativity analysis [12] to provide a bound on this effect through constructing storage functions. Unfortunately, it is usually an intractable task to find storage functions analytically. Instead, numerical methods are applied to a relaxed version [13] of the original problem. We use the sum-of-squares (SOS) relaxation to seek the structure of the storage function and semi-definite programming (SDP), implemented in SOSTOOLS [14], to find the coefficients of this storage function. Our results make it possible to provide a bound on the effect of downstream modules even in the case of uncertain parameters [15].

Our work presented here connects with the rapidly increasing effort to reliably predict the behavior of multi-module systems by accounting for the effects of interactions. For instance, the effects of sharing various resources by different modules have been investigated in [16], [17]. Additionally, [18] addresses how the expression of unneeded proteins decreases the growth rate of the cell, thereby affecting the allocation of cellular resources, which in turn affects the concentration of other proteins. Finally, [7] characterizes the effects of retroactivity in gene networks of arbitrary topology and provides guidelines to minimize these adverse effects, whereas [19] presents a load driver to facilitate the modular design of gene networks. The approach presented here differs from the above methods as we employ dissipativity analysis (previously applied to the model reduction problem [20]), allowing us to handle parameter uncertainty and provide a guaranteed upper bound on the effects of interaction.

This paper is organized as follows. In Section II, the standard model of gene transcription networks is presented, together with the description of how retroactivity from downstream modules changes the dynamics of an upstream module. This is followed by a dissipativity analysis of an auxiliary system, which is constructed to measure the difference between the isolated and connected systems. In particular, we describe storage functions, their SOS relaxation and how such storage functions can be found using SDP techniques. In Section IV we then illustrate how the tools can be applied in the case of the most recurrent motif in gene transcription networks, negative autoregulation. Finally, we discuss how our results can be further developed to handle high-dimensional problems more efficiently.

II. SYSTEM MODEL AND PROBLEM FORMULATION

We first briefly introduce the standard model of gene transcription networks, together with a reduced order model that closely approximates the dynamics of protein concentrations. Then we characterize how downstream modules perturb the dynamics of upstream modules due to retroactivity, which allows us to formulate the primary goal of this paper: *how*

*This work was supported by AFOSR grant FA9550-12-1-0129 and NIGMS grant P50 GM098792, and by UK EPSRC projects EP/F500394/1, EP/I031944/1, and EP/M002454/1.

[1]TPP is with Department of Engineering Science, University of Oxford, Parks Road, Oxford, OX1 3PJ, UK thomas.prescott@eng.ox.ac.uk

[2]AG is with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139, USA gyorgy@mit.edu

to quantify the effects of retroactive perturbations on the trajectory of the upstream module.

A. Dynamic Model of Gene Transcription Networks

A gene transcription network is made up of a collection of N transcription factors (TFs) x_i, each of which may regulate the expression of one or more TFs (including its own expression). This network is usually represented as a directed graph where the nodes and edges denote TFs and regulatory interactions among them, respectively (Fig. 1A). In particular, there is a directed edge from x_j to x_i if x_j binds to the promoter of x_i to form a promoter complex, thereby regulating the expression of x_i.

Let $x = (x_1, x_2, \ldots, x_N)^T$ denote the concentration vector of TFs. Considering protein production/decay, together with the binding/unbinding reactions by which TFs regulate each other's expression, it is shown in [7] that the dynamics of x can be approximated by

$$\dot{x} = [I + R(x)]^{-1} h(x). \tag{1}$$

Here the vector $h(x)$ of length N encompasses the production and decay of the TFs, and the $N \times N$ matrix $R(x)$, called the *internal retroactivity*, measures the strength of regulatory interactions within the network. For a more detailed interpretation of each of these quantities, see [7].

B. Loads Change the Dynamics of Modules

Next, we consider the interconnection of two transcription networks. In particular, we focus on the case when edges between these two transcription networks point from one network (upstream module) to the other (downstream module), see Fig. 1B. Furthermore, we assume that each node in the downstream module has regulators from either the upstream or from the downstream module, but not from both (as in Fig. 1B).

For more details on the following derivation, see [7]. Referring to (1), introduce $f(x) = [I + R(x)]^{-1} h(x)$ so that $\dot{x} = f(x)$ describes the dynamics of the upstream module when it is not connected to the downstream module (Fig. 1A). Upon interconnection with the downstream module (Fig. 1B), the dynamics of the upstream module change to

$$\dot{x} = [I + (I + R(x))^{-1} \bar{S}(x)]^{-1} f(x). \tag{2}$$

Here the $N \times N$ matrix $\bar{S}(x)$, called the *scaling retroactivity* of the downstream module, measures the strength of the regulatory interactions between the modules. In other words, $\bar{S}(x)$ characterizes the load the downstream module applies to the upstream module. In particular, if $\bar{S}(x) = 0$ then the dynamics of the upstream module remains unaffected upon interconnection as the vector fields in (1) and (2) are equal. In general, the "greater" $\bar{S}(x)$ compared to $I + R(x)$, the greater the effect of the downstream module on the dynamics of the upstream module.

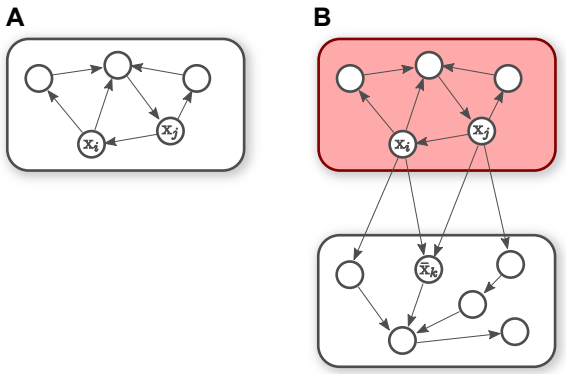

Fig. 1. Downstream modules change the dynamic behavior of upstream modules. (A) The dynamics of the isolated upstream modules are given by (1). (B) The upstream module is connected to the downstream modules, yielding a change in its dynamics according to (2).

C. Problem Formulation

The goal of this paper is to quantify the effect of the downstream module on the dynamics of the upstream module. In particular, we seek an upper bound on the difference between the trajectories of (1) and (2), allowing also for the possibility of parameter uncertainties in the downstream module. The following section clarifies how we intend to measure the retroactive effect, and introduces the computational techniques which can be used to bound this measure.

III. DISSIPATIVITY ANALYSIS

Suppose we are given the autonomous dynamical system

$$\dot{x} = f(x), \quad y = y(x), \quad x(0) = x_0, \tag{3}$$

and a comparable perturbation of the system, given by

$$\dot{\tilde{x}} = \tilde{f}(\tilde{x}), \quad \tilde{y} = \tilde{y}(\tilde{x}), \quad \tilde{x}(0) = \tilde{x}_0, \tag{4}$$

where the dimensions n and $\tilde{n}$ of the states x and $\tilde{x}$ may be different. Such a perturbation could arise from model order reduction [20], parameter perturbations [21], or in this case from retroactivity, where (1) and (2) correspond to (3) and (4), respectively, and $n = \tilde{n} = N$. By "comparable", we mean that y and $\tilde{y}$ measure the same output(s), and are therefore of the same dimension.

Suppose that both systems are stable, and that $\lim_{t\to\infty} y(t) = \lim_{t\to\infty} \tilde{y}(t)$. Without loss of generality, we can assume that $\lim_{t\to\infty} x(t) = \lim_{t\to\infty} \tilde{x}(t) = 0$. A measure of the perturbation introduced by the new dynamics is the L_2 norm of their difference given by

$$\|y - \tilde{y}\|_2^2 = \int_0^\infty |y(t) - \tilde{y}(t)|^2 \, \mathrm{d}t,$$

where we take the standard $\mathbb{R}^n$ Euclidean norm $|\cdot|$ in the integrand. It is this measure that we aim to find a bound for through using SOS techniques. Hence, in the context of the retroactive effect quantification discussed above, we will attempt to estimate the L_2 distance between the trajectories of the unloaded and loaded modules.

A. Dissipativity and Bounding Perturbations

Consider the combined system (3)–(4) with the output $z(t) = y(x(t)) - \tilde{y}(\tilde{x}(t))$ measuring the error at each time t between the two outputs $y(t)$ and $\tilde{y}(t)$.

Suppose there exists a positive semi-definite (storage) function $V(x, \tilde{x}) \geqslant 0$ of the combined original and perturbed state, where $V(0, 0) = 0$. Suppose further that $\dot{V} + z^T z \leqslant 0$, so that V certifies the dissipativity [12] of the combined system with respect to the supply rate $z^T z$. Integrating this inequality gives an upper bound on the error of

$$\int_0^T z^T z \, \mathrm{d}t \leqslant V(x_0, \tilde{x}_0) - V(x(T), \tilde{x}(T)).$$

As $T \to \infty$, we deduce the upper bound $z^T z = \|y - \tilde{y}\|^2 \leqslant V(x_0, \tilde{x}_0)$ for the L_2 norm of the perturbation in output.

B. SOS Relaxation of the Storage Function

If the dynamics in (3)–(4) are polynomial, or polynomial–rational, then the desired inequalities $V \geqslant 0$ and $\dot{V} + |z|^2 \leqslant 0$ can be relaxed. Any polynomial Ψ is SOS (denote $\Psi \in \Sigma$) if there exist r polynomials ψ_i for $i = 1, \ldots, r$ such that $\Psi = \sum_{i=1}^{r} \psi_i^2$. Clearly $\Psi \in \Sigma$ implies $\Psi \geqslant 0$, and it can be shown that the reverse implication does not hold [13]. Although $\Psi \in \Sigma$ is a stronger condition than $\Psi \geqslant 0$, it is a semi-definite constraint and therefore computationally tractable through semi-definite programming. Thus the two conditions to check can be relaxed to $V \in \Sigma$ and $-(\dot{V} + |z|^2) \in \Sigma$. These semi-definite constraints can be implemented using SOS programming, which has been implemented in MATLAB as the SOSTOOLS [14] toolbox.

Note that if f and/or $\tilde{f}$ are rational, rather than polynomial, then $\dot{V} = \nabla V \cdot [f^T, \tilde{f}^T]^T$ will be a rational function. If we denote the common denominator of $\dot{V}$ as $d(x, \tilde{x})$, then satisfying the constraint

$$-d(x, \tilde{x}) \left[\dot{V} + |z|^2 \right] \in \Sigma \tag{5}$$

implies that $\dot{V} + |z|^2 \leqslant 0$ in the region $\mathcal{D} = \{[x^T, \tilde{x}^T]^T \mid d(x, \tilde{x}) > 0\}$ where the denominator is positive (assume $d = 1$ in the case of polynomial dynamics).

C. Local Results

Other considerations, mainly centered around the non-global stability of the dynamical systems in question, will mean that we cannot show dissipativity globally. However, we can relax the constraint (5) to enforce it only in a region of the state space defined as

$$\mathcal{G} = \{[x^T, \tilde{x}^T]^T \mid g_i(x, \tilde{x}) \leqslant 0 \text{ for } i = 1, 2, \ldots, k\}, \tag{6}$$

where $g_1, \ldots, g_k$ are polynomials. If there exist SOS multipliers $\sigma_1, \ldots, \sigma_k \in \Sigma$ such that V satisfies

$$-d(x, \tilde{x}) \left[\dot{V} + z^T z \right] + \sum_{i=1}^{k} \sigma_i g_i \in \Sigma,$$

then, for all $[x^T, \tilde{x}^T]^T \in \mathcal{G} \cap \mathcal{D}$, it follows that V satisfies

$$\dot{V} + |z|^2 \leqslant 0$$

so that (3)–(4) has the dissipativity property in the region $\mathcal{G}$.

D. Parameter Uncertainty

Suppose f or $\tilde{f}$ depend on an uncertain parameter μ. This uncertainty can be implemented by introducing a corresponding auxiliary variable v with dynamics $\dot{v} = 0$. Let μ_l and μ_u denote the lower and upper bounds on μ (i.e., $\mu \in [\mu_l, \mu_u]$), which can be encoded by including the polynomial inequality $g_\mu = (v - \mu_l)(v - \mu_u) \leqslant 0$ in the definition of $\mathcal{G}$ in (6). We still seek a polynomial storage function $V = V(x, \tilde{x})$ as a function of x and $\tilde{x}$, but assume that it does not depend on v. The solution $V \in \Sigma$ to the SOS program with constraint

$$-d(x, \tilde{x}) \left[\dot{V} + z^T z \right] + \sum_{i=1}^{k} \sigma_i g_i + \sigma_\mu g_\mu \in \Sigma,$$

for the SOS multiplier $\sigma_\mu \in \Sigma$, will return an upper bound on the error as a function of the initial conditions. Importantly, this bound is valid for all $\mu \in [\mu_l, \mu_u]$.

IV. APPLICATION EXAMPLE

We now demonstrate the use of SOS programming to answer the retroactivity quantification problem, by applying it to a loaded autoregulation module, which is described below. First, we seek an *a priori* upper bound for the L_2 norm of the difference between the trajectories $x(t)$ and $\tilde{x}(t)$ for $t \geqslant 0$, given certain fixed downstream parameters η_L and k_L (introduced next). We will then allow for bounded uncertainty in one, and then in both downstream parameters, and seek an upper bound on the worst-case effect for a fixed initial condition. In the case of uncertain parameters, we are thus constructing an upper bound on the worst-case retroactive loading effect produced by a *space* of possible downstream loads.

A. Negative Autoregulation

In the case of negative autoregulation, TF x represses its own production by binding to its own promoter c′, forming the TF-promoter complex c″ according to

$$\mathrm{c}' + \mathrm{x} \underset{k^-}{\overset{k^+}{\rightleftharpoons}} \mathrm{c}''.$$

The production and degradation of x are then modeled by the reactions

$$\mathrm{c}' \xrightarrow{\pi} \mathrm{c}' + \mathrm{x} \qquad \text{and} \qquad \mathrm{x}_i \xrightarrow{\delta_i} \varnothing. \tag{7}$$

With the dissociation constant $k = k^-/k^+$ and DNA copy number η, we obtain from (1) referring to [7] that the dynamics of x are given by

$$\dot{x} = \frac{1}{1 + R(x)} \left(\frac{\pi \eta}{1 + x/k} - \delta x \right) = f(x), \tag{8}$$

with

$$R(x) = \frac{\eta/k}{(1 + x/k)^2}. \tag{9}$$

Next, consider the case when this module drives another downstream module represented by the reactions

$$\mathrm{c}'_L + \mathrm{x} \underset{k_L^-}{\overset{k_L^+}{\rightleftharpoons}} \mathrm{c}''_L,$$

such that $\eta_L = c'_L + c''_L$ denotes the DNA copy number of the downstream module and $k_L = k_L^-/k_L^+$ is the dissociation constant of x binding to these sites. From (2) and referring to [7], the dynamics of the loaded variable $\tilde{x}$ in this case are given by the perturbed system

$$\dot{\tilde{x}} = \frac{1}{1 + R(\tilde{x}) + \bar{S}(\tilde{x})} \left(\frac{\pi\eta}{1 + \tilde{x}/k} - \delta\tilde{x} \right) = \tilde{f}(\tilde{x}). \quad (10)$$

where

$$\bar{S}(\tilde{x}) = \frac{\eta_L/k_L}{(1 + \tilde{x}/k_L)^2}. \quad (11)$$

Since $f(x)$ and $\tilde{f}(\tilde{x})$ are rational instead of polynomial functions of x and $\tilde{x}$, respectively, we need to make sure that, for the SOS decision variable $V(x, \tilde{x}) \in \Sigma$, the denominator $d(x, \tilde{x})$ of

$$\frac{\mathrm{d}V}{\mathrm{d}t} = \frac{\partial V}{\partial x} f(x) + \frac{\partial V}{\partial \tilde{x}} \tilde{f}(\tilde{x}) \quad (12)$$

is positive for $x, \tilde{x} \geqslant 0$. Indeed, from (8)–(11) we have that $d(x, \tilde{x}) > 0$ for $x, \tilde{x} \geqslant 0$, independently of the value of $k_L > 0$ and $\eta_L \geqslant 0$.

In what follows, we take the unloaded and loaded dynamics in (8) and (10), respectively, with initial condition $x(0) = \tilde{x}(0) = 2$ and with common upstream parameters taking values $\pi = 2$, $\delta = k = 1$ and $\eta = 10$. Both the unloaded and loaded module approache the same steady state $\lim_{t\to\infty} x = \lim_{t\to\infty} \tilde{x} = 4$. Since $\dot{x}, \dot{\tilde{x}} \geqslant 0$ are strictly increasing for $x, \tilde{x} \in [2, 4)$ from (8) and (10), both states take values only in the interval $[2, 4]$. To have the steady state at zero, we translate x and $\tilde{x}$ and (with some abuse of notation) redefine $x := x - 4$ and $\tilde{x} := \tilde{x} - 4$. To represent the fact that $x, \tilde{x} \in [-2, 0]$, we define the state space $\mathcal{G}$ in (6) by the polynomial conditions $g_1 = x(x + 2) \leqslant 0$ and $g_2 = \tilde{x}(\tilde{x} + 2) \leqslant 0$.

B. Fixed Downstream Parameters

Consider first the case when the load parameters are fixed at $\eta_L = 10$ and $k_L = 1$. Our goal is to find V which satisfies the dissipativity inequality $\dot{V} + (x - \tilde{x})^2 \leqslant 0$. We introduce the auxiliary SOS decision variables $\sigma_1(x, \tilde{x}) \in \Sigma$ and $\sigma_2(x, \tilde{x}) \in \Sigma$. With these variables, the SOS constraint

$$-d\left[\dot{V} + (x - \tilde{x})^2\right] + \sum_{i=1}^{2} \sigma_i g_i \in \Sigma$$

ensures the desired inequality locally to $\mathcal{G}$, since both $-d\left[\dot{V} + (x - \tilde{x})^2\right] \geqslant LHS \geqslant 0$ and $d(x, \tilde{x}) > 0$ are true for all $(x, \tilde{x})^T \in \mathcal{G}$.

Implementing the constraint above and minimizing the value of the upper bound $V(-2, -2)$ with MATLAB using SOSTOOLS yields

$$\|x - \tilde{x}\|^2 \leqslant \begin{cases} 0.3044 & \deg V = 2, \\ 0.1216 & \deg V = 4, \\ 0.1214 & \deg V = 6, 8, \ldots \end{cases} \quad (13)$$

By increasing $\deg V$, a larger subspace of possible positive semi-definite storage functions is searched, at the expense of computational cost (7s, 14s, 62s, and 532s for $\deg V = 2, 4, 6, 8$, respectively).

From (13) it follows that the upper bound of the difference $\|x - \tilde{x}\|^2$ between the trajectories of the isolated and connected upstream module is no more than 0.1214, assuming the initial conditions $x(0) = \tilde{x}(0) = -2$. This bound turns out to be a very close estimate of the actual loading effect: by simulation we estimate a value of 0.1211.

C. One Uncertain Downstream Parameter

We now fix the initial conditions of the loaded and unloaded system to be $x(0) = \tilde{x}(0) = -2$, and consider two cases: first, $k_L = 1$ is fixed and $\eta_L \in [0, 10]$ is uncertain but bounded, and second, $\eta_L = 10$ is fixed and $k_L \in [1, 10]$ is uncertain but bounded.

As described in Section III-D, we introduce the auxiliary variable H that corresponds to the uncertain parameter η_L. The loaded dynamics are expressed in terms of the new parameter by re-writing the retroactivity term in (11) as

$$\bar{S}(\tilde{x}, H) = \frac{H/k_L}{(1 + \tilde{x}/k_L)^2},$$

and substituting into (10).

To bound the values of H we construct the polynomial $g_3 = H(H - 10)$, which is non-positive for $H \in [0, 10]$, and define a higher-dimensional $\mathcal{G} = \{(x, \tilde{x}, H) \mid g_i \leqslant 0 \text{ for } i = 1, 2, 3\}$, where $g_1 = x(x+2)$ and $g_2 = \tilde{x}(\tilde{x}+2)$ were defined previously. The resulting SOS constraint we must satisfy is

$$-d\left[\dot{V} + (x - \tilde{x})^2\right] + \sum_{i=1}^{3} \sigma_i g_i \in \Sigma,$$

with objective to minimize $V(-2, -2)$. This SOS program returns the upper bound of 0.1218 with $\deg(V) = 6$ (in 67s time), which is very close to the bound 0.1211 found in the nominal case of $\eta_L = 10$. The slight increase in the upper bound from 0.1214 is due to the uncertainty in η_L requiring an additional slack term $\sigma_3 g_3$ in the SOS constraint, which very slightly increases the conservatism of the resulting bound.

This program has now certified that, if the downstream parameter η_L varies in the range $[0, 10]$ then the worst-case load caused by retroactivity will be such that $\|x - \tilde{x}\|^2$ is no greater than 0.1218. We can simulate a finite number of possible values η_L in this range. As shown in Fig. 2A, the greatest simulated error is found at $\eta_L = 10$. This result can be interpreted as follows. The greater η_L, the more downstream binding sites there are sequestering x, thus slowing down its dynamics, represented by the fact that $\frac{\partial \bar{S}}{\partial \eta_L} > 0$ in (10). As a result, we can simulate (8) and (10) with $\eta_L = 10$ and $k_L = 1$ and simply integrate the error numerically to find an upper bound when η_L is uncertain. This is not the case when k_L is uncertain, however, as we next show.

When $\eta_L = 10$ and $k_L \in [1, 10]$, we can follow a similar path but now using the auxiliary variable K corresponding to the uncertain parameter k_L. Introduce $g_4 = (K-1)(K-10)$, which is non-positive for $K \in [1, 10]$. With this, we define

$\mathcal{G} = \{(x, \tilde{x}, K) \mid g_i \leqslant 0 \text{ for } i = 1, 2, 4\}$. The resulting SOS constraint we must satisfy is

$$-d\left[\dot{V} + (x - \tilde{x})^2\right] + \sum_{i=1,2,4} \sigma_i g_i \in \Sigma,$$

with objective to minimize $V(-2, -2)$. This SOS program returns the upper bound of 0.1926 with $\deg(V) = 6$ (in 163s time). This program has now certified that, if the downstream parameter k_L varies in the range $[1, 10]$ then the worst-case error caused by retroactivity will be such that $\|x - \tilde{x}\|^2$ is no greater than 0.1926. As before, we can simulate a finite number of possible values of k_L in this range. As shown in panel B of Fig. 2, the greatest simulated error is 0.1864, which is achieved when $k_L = 2.7$. That is, the greatest error does not occur when the affinity to the downstream sites is the greatest ($k_L = 1$). This seemingly counter-intuitive phenomenon can be interpreted as follows.

For simplicity, we consider the unstranslated coordinate system, so that $x, \tilde{x} \in [2, 4]$. The concentration of x bound to the downstream binding sites is $c_L'' = \eta_L x/(x + k_L)$, which means that binding sites available for further binding of x is given by $c_L' = \eta_L - c_L'' = \eta_L k_L/(x + k_L)$ [7]. For instance, in the limit $k_L \to 0$ we have that $c_L' \to 0$. That is, when the affinity to downstream sites is the greatest possible, there are no more free binding sites because they are already occupied at $t = 0$. As a result, the connected system behaves as if it were isolated: $\bar{S} = 0$ when $k_L = 0$ implies that $f(x) = \tilde{f}(x)$; since $x(0) = \tilde{x}(0)$ we have that $x(t) = \tilde{x}(t)$ for $t \geqslant 0$. Since $\frac{\partial \bar{S}}{\partial k_L} > 0$ if $k_L < 2$ and $\frac{\partial \bar{S}}{\partial k_L} < 0$ if $k_L > 4$, we conclude that the error increases in the first case and decreases in the second case (verified numerically in Fig. 2B) following a similar reasoning as above. However, we do not know what happens for $k_L \in [2, 4]$, so that we would need to run numerous simulations of (8) and (10) and numerically integrate the error to find an upper bound on $\|x - \tilde{x}\|^2$. Conversely, using SOSTOOLS, we have certified the upper bound on the worst-case retroactive effect solving a single optimization problem.

D. Two Uncertain Downstream Parameters

Now suppose that both of the loading module's parameters are uncertain, taking values $k_L \in [1, 10]$ and $\eta_L \in [0, 10]$. We define the auxiliary variables K and H respectively corresponding to these two parameters, and include these as additional variables together with x and $\tilde{x}$. We now construct the higher-dimensional region $\mathcal{G}$ of $(x, \tilde{x}, H, K)$-space to be defined by $g_i \leqslant 0$ where $i = 1, 2, 3, 4$. To set up the SOS program, we define the decision variable $V(x, \tilde{x}) \in \Sigma$ and the auxiliary decision variables $\sigma_i(x, \tilde{x}, H, K) \in \Sigma$ for $i = 1, \ldots, 4$. As before, the time derivative of V is a rational function with a polynomial denominator $d(x, \tilde{x}, H, K)$. Thus, the SOS constraint we need to apply is

$$-d\left[\dot{V} + (x - \tilde{x})^2\right] + \sum_{i=1}^{4} \sigma_i g_i \in \Sigma,$$

with the objective of minimizing $V(-2, -2)$. By searching over $V \in \Sigma$ with $\deg(V) = 6$ we find a storage function

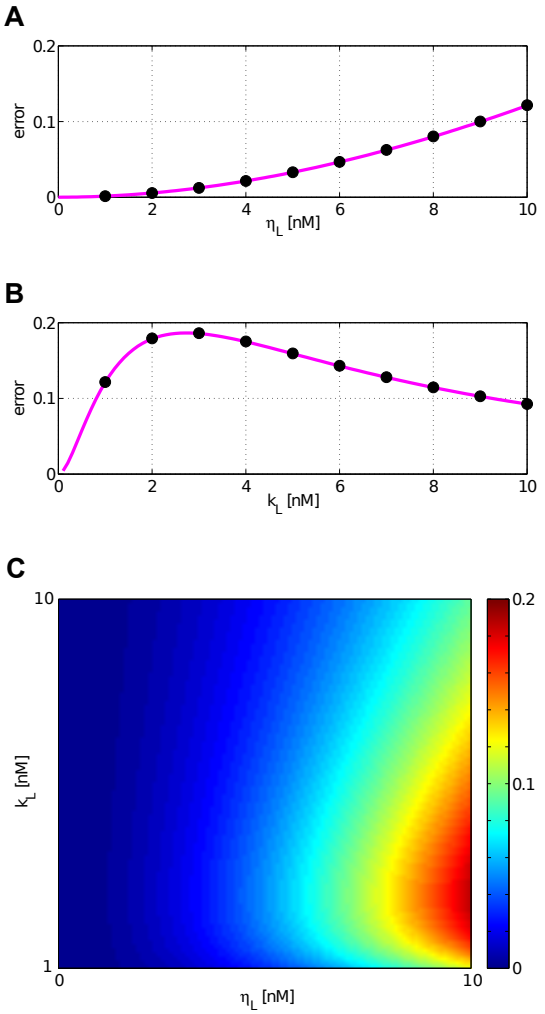

Fig. 2. Error $\|x - \tilde{x}\|^2$ for the isolated and connected modules given by (8) and (10), respectively. Magenta curve denotes the simulated error $\|x - \tilde{x}\|^2$, whereas black dots represent the upper bound obtained by using SOSTOOLS with deg(V) = 4. (A) The dissociation constant $k_L = 1$ is fixed, the DNA copy number $\eta_L \in [0, 10]$ is uncertain but bounded. (B) The DNA copy number $\eta_L = 10$ is fixed, the dissociation constant $k_L \in [1, 10]$ is uncertain but bounded. (C) Simulated error $\|x - \tilde{x}\|^2$ in the range $(k_L, \eta_L) \in [1, 10] \times [0, 10]$.

V (in 296s time) such that $V(-2, -2) = 0.2015$. This certifies that, for any value of the parameters $(\eta_L, k_L) \in [0, 10] \times [1, 10]$, the retroactive effect on the upstream module will be no greater than $\|x - \tilde{x}\|^2 \leqslant 0.2015$. By gridding parameter space and simulating the resulting system, Fig. 2C shows that the discovered upper bound is close to a resulting lower bound on the worst-case effect, simulated to be achieved for $\eta_L = 10$ and $k_L = 2.7$. It is important to note that if the parameter uncertainty increases, the time it takes to compute the difference between the trajectories of (8) and (10) with the same resolution increases. For instance, if the parameter uncertainty becomes $(\eta_L, k_L) \in [0, 100] \times [1, 100]$, the computational time increases 100-fold (approximately 0.01s for a given (η_L, k_L) pair). Conversely, the computational time required using our approach remains relatively unaltered (in fact, it decreases from 296s to 275s).

V. DISCUSSION

In this paper, we presented a method to quantify how downstream modules affect the behavior of upstream modules in gene networks. In particular, we showed that the effect of retroactivity from a downstream module on the trajectory of an upstream module can be upper bounded via dissipativity analysis using SOSTOOLS. We also demonstrated how this method can be applied in the presence of parameter uncertainties, which is a common situation in biological systems as parameters are usually hard to measure precisely.

Quantifying the effect of interactions in synthetic gene networks is increasingly important as more functional modules become available, but their behavior once interconnected is not reliably predictable [22]. Our method could aid the design of multi-module systems [23] by accounting for the effects of retroactivity arising from the interactions. This way the lengthy and *ad hoc* process of tuning parameters experimentally could be significantly reduced.

One of the main advantages of the method presented in this paper is the fact that we can provide a guaranteed upper bound on the error in the presence of parameter uncertainties. Instead of running simulations for every possible combination of different parameter values and initial conditions by gridding the space of interest, only one optimization problem is required to be solved. Furthermore, our method scales favorably with the volume of parameter uncertainty, unlike the simulation-based approach. Additionally, via gridding the parameter space we could miss parameter values yielding extreme error values (for instance, leading to resonance), so that the upper bound obtained this way is not guaranteed to be valid, unlike in the case of the presented method.

If the loaded and unloaded modules have different steady state values, or approach different limit cycles, alternative measures of the trajectory difference may become important. If the SOS method above returns a feasible storage function, then we have proved a finite L_2 norm to the trajectory difference and hence can guarantee an equal steady state value between the loaded and unloaded modules. Nevertheless, further work would be required in order to apply optimization methods to other measures of trajectory perturbation.

Unfortunately, the size of the required SOS program increases with the problem dimension, depending on both the numbers of species and of uncertain parameters. Consequently, the resulting optimization problem may not be solved due to lack of computational capacity. As a result, we are currently investigating how to partition the high dimensional global optimization problem into several smaller dimensional local optimization problems [21]. In particular, we seek local storage functions first as a basis for a global storage function [24]. For instance, if two downstream modules share a common upstream module, we may re-use the storage functions obtained by considering only one load to bound the error when both loads are present.

ACKNOWLEDGMENT

The authors thank Antonis Papachristodoulou and Domitilla Del Vecchio for comments on the manuscript.

REFERENCES

[1] D.A. Lauffenburger "Cell signaling pathways as control modules: complexity for simplicity?" PNAS, vol 97, pp. 5031-5033, 2000.

[2] S. Cardinale and A.P. Arkin, "Contextualizing context for synthetic biology - identifying causes of failure of synthetic biological systems," Biotech Journ, vol. 7, pp. 856-866, 2012.

[3] F.K. Balagadde, L. You, C.L. Hansen, F.H. Arnold and S.R. Quake, "Long-term monitoring of bacteria undergoing programmed population control in a microchemostat," Science, vol. 309, pp. 137-140, 2005.

[4] M. Scott, C. Gunderson, E. Mateescu, Z. Zhang and T. Hwa, "Interde pendence of cell growth and gene expression: origins and consequences," Science, vol. 330, pp. 1099-1102, 2010.

[5] J. Neupert, D. Karcher and R. Bock, "Design of simple synthetic RNA thermometers for temperature-controlled gene expression in *Escherichia coli*," Nucleic Acids Res, vol. 36, pp. e124, 2008.

[6] D. Del Vecchio, A.J. Ninfa and E.D. Sontag, "Modular cell biology: retroactivity and insulation," Molecular Sys Biol, vol. 4, 2008.

[7] A. Gyorgy and D. Del Vecchio, "Modular composition of gene transcription networks," PLoS Comput Biol 10(3): e1003486, 2014.

[8] S. Jayanthi, K.S. Nilgiriwala and D. Del Vecchio, "Retroactivity controls the temporal dynamics of gene transcription," ACS Synt Biol, vol. 2, pp. 431-441, 2013.

[9] E. Franco, E. Friedrichs, J. Kim, R. Jungmann, R.M. Murray, E. Winfree and F. C. Simmel, "Timing molecular motion and production with a synthetic transcriptional clock," PNAS, vol. 108, pp. E787, 2011.

[10] P. Jiang, A.C. Ventura, E.D. Sontag, S.D. Merajver, A.J. Ninfa and D. Del Vecchio, "Load-induced modulation of signal transduction networks," Science Signaling, vol. 4, pp. ra67, 2011.

[11] Y. Kim, Z. Paroush, K. Nairz, E. Hafen, G. Jimnez and S.Y. Shvartsman, "Substrate-dependent control of MAPK phosphorylation *in vivo*," Mol Sys Biol, vol. 7, no. 467, 2011.

[12] W.M. Haddad and V. Chellaboina, Nonlinear Dynamical Systems and Control: A Lyapunov-Based Approach. Princeton University Press, 2008

[13] P.A. Parrilo, "Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization," PhD thesis, Caltech, Pasadena CA, 2000.

[14] A. Papachristodoulou, J. Anderson, G. Valmorbida, S. Prajna, P. Seiler and P.A. Parrilo, "SOSTOOLS: Sum of Squares optimization toolbox for MATLAB," 2013. Available from `http://www.eng.ox.ac.uk/control/sostools`.

[15] A. Papachristodoulou and S. Prajna, "On the construction of Lyapunov functions using the Sum of Squares decomposition," in Proc of IEEE Conference on Decision and Control, 2002

[16] N. Cookson, W. Mather, T. Danino, O. Mondragon-Palomino, R. Williams, L. Tsimring and J. Hasty, "Queueing up for enzymatic processing: correlated signaling through coupled degradation," Mol Sys Biol, vol. 7, 2011.

[17] D. Siegal-Gaskins, Z.A. Tuza, J. Kim, V. Noireaux and R.M. Murray, "Gene circuit performance characterization and resource usage in a cell-free 'breadboard'," ACS Synth Biol, vol. 3, pp. 416–425, 2014.

[18] M. Scott, C.W. Gunderson, E.M. Mateescu, Z. Zhang and T. Hwa, "Interdependence of cell growth and gene expression: origins and consequences," Science, vol. 330, 2010.

[19] D. Mishra, P.M. Rivera-Ortiz, A. Lin, D. Del Vecchio and R. Weiss, "A load driver device for engineering modularity in biological networks," Nat Biotech, Accepted, 2014.

[20] T.P. Prescott and A. Papachristodoulou, "Guaranteed error bounds for structured complexity reduction of biochemical networks," J Theor Biol, vol. 304, pp. 172–182, 2012

[21] T.P. Prescott and A. Papachristodoulou, "Structured storage functions for cascaded systems," in Proc of IEEE Conference on Decision and Control, 2014

[22] P.E.M. Purnick and R. Weiss, "The second wave of synthetic biology: from modules to systems," Nature Reviews Molecular Cell Biology vol. 10, pp. 410–422, 2009.

[23] M. Miller, M. Hafner, E.D. Sontag, N. Davidsohn, S. Subramanian, P.E.M. Purnick, D. Lauffenburger and R. Weiss, "Modular design of artificial tissue homeostasis: robust control through synthetic cellular heterogeneity," PLoS Comput Biol vol. 8, pp. e1002579, 2012.

[24] S. Coogan and M. Arcak, "Verifying safety of interconnected passive systems using SOS programming," in Proc of IEEE Conference on Decision and Control, 2013.