Resource Sensor Design for Quantifying Resource Competition in Genetic Circuits

Cameron McBride and Domitilla Del Vecchio

***Abstract*— Sharing of cellular resources in genetic circuits negatively affects performance and often leads to unpredictable behavior. Measuring key metrics from experimental data that quantify resource sharing and its effect on a system's output would be highly useful. In this paper, we propose two metrics, Q and S, representing the quantity of resources used by a genetic circuit module and the sensitivity of the output of a module to resource disturbances, respectively. Together, Q and S may be used to estimate the change in the output of a module in response to the disturbances in the availability of resources. We cast the problem of finding these metrics as a parameter estimation problem and outline a simple procedure to estimate these metrics from data. Knowledge of Q and S for a circuit module enables prediction of the effects of resource sharing and allows for resource-aware design of genetic circuits.**

I. Introduction

As the field of synthetic biology has matured in recent years, it has become increasingly important to use modular design approaches to create sophisticated systems [1, 2]. Genetic circuits may be used for a variety of exciting applications in fields such as medicine, energy, or the environment [3–5]. However, one of the main barriers to predictable, modular design of genetic circuits is resource sharing [6–10], in which the sequestration of a finite pool of cellular resources, e.g. ribosomes, RNAP, or amino acids, causes unwanted coupling between otherwise uncoupled genes [8].

Much work has been done in characterizing resource sharing effects and in creating predictive models [8, 9, 11–17]. Most notably, a global resource monitor which characterizes the resource usage of a genetic circuit was developed in [18, 19]; however, this method is unable to predict changes in circuit behavior due resource sharing effects since it lacks a characterization of resource sharing effects on the genetic circuit of interest.

A method to allocate cellular resources in a synthetic circuit context has not been possible thus far. Furthermore, computational simulation of the circuits of interest and of their interaction with the host resources suffers from the problem that parameters are largely unknown. Hence, quantitative predictions are often difficult. Our goal is to develop a method to predict the effects of resource sharing on the behavior of a genetic circuit of interest based on a few standard experiments, with the idea that this information could be included in automated circuit optimization algorithms [20–22].

This work was supported by the National Science Foundation Award Number 1521925

C. McBride and D. Del Vecchio are with the Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, cmcbride@mit.edu, ddv@mit.edu

In particular, we consider measuring the resource sharing properties of a genetic circuit module. We propose two metrics to characterize resource sharing in a genetic circuit module i: Q_i and S_i, representing the resources used by the module and the sensitivity of the module to changes in resource availability, respectively. We then demonstrate that these metrics may be employed to quantitatively predict resource sharing effects and propose an experimental method to estimate Q_i and S_i. In Theorem 1, we find estimators $\hat{Q}_i$ and $\hat{S}_i$ for Q_i and S_i using data obtained through the proposed experimental method and quantify the error introduced by the estimators. In Corollary 1, we show that these estimators may be used to predict the output of a genetic circuit with resource sharing and quantify the estimation error. We show through simulation that we are able to accurately predict how the output of one module changes when other modules are included in the cellular environment. Overall, the metrics Q_i and S_i may be appended to a module's description along with the experimentally measured input/output characteristic and used to predict the quantitative output of an arbitrary combination of modules working together in the cellular environment.

This paper is organized as follows. In Section II, we introduce the general framework, define the metrics Q_i and S_i, and demonstrate that they may be used to quantitatively predict resource sharing effects. In Section III, we propose an experimental method that may be used to estimate Q_i and S_i. In Section IV, we present Theorem 1 and Corollary 1, finding estimators $\hat{Q}_i$ and $\hat{S}_i$ and an estimate on the output of the module under resource sharing. In Section V, we demonstrate the experimental method and find the estimates $\hat{Q}_i$ and $\hat{S}_i$ through simulation.

II. Problem Formulation

To design a genetic circuit, the design process is typically broken up into the design of separate modules. These modules are then combined to perform the desired function under the implicit assumption that their input/output behavior does not depend on the presence of other modules. Unfortuantely, this assumption is not valid in synthetic genetic circuits due to resource sharing [8, 13].

We begin by giving background on synthetic circuits and defining the problem setup. As shown in Figure 1a, a node in a genetic circuit is a dynamic process that takes a protein, known as a transcription factor, as input and produces a protein as output. A transcription factor input may regulate the output of the node either positively or negatively, with positive regulations being represented as "$\rightarrow$" and negative

Preprint submitted to 57th IEEE Conference on Decision and Control.
Received March 15, 2018.

regulations as "⊣". Then, a genetic circuit module is a collection of nodes with an input that may be controlled and an output that may be measured.

We consider the input/output behavior of a genetic circuit module i shown in Figure 1b. The input is u_i and the output is y_i, which we assume is some fluorescent reporter protein that may be measured. This gives the measurement of the output for module i as y_{i_0} for each input, u_i, when the module is in isolation, that is, it is the only synthetic genetic circuit in the cell. The module uses resources from an available pool that the cell has in limited quantity. Now, if module i and module j are placed within the same cell, as shown in Figure 1c, due to the interaction of modules i and j with the pool of resources, the output of the modules, y_i and y_j cannot be predicted solely from the measurement of input/output responses of the modules in isolation.

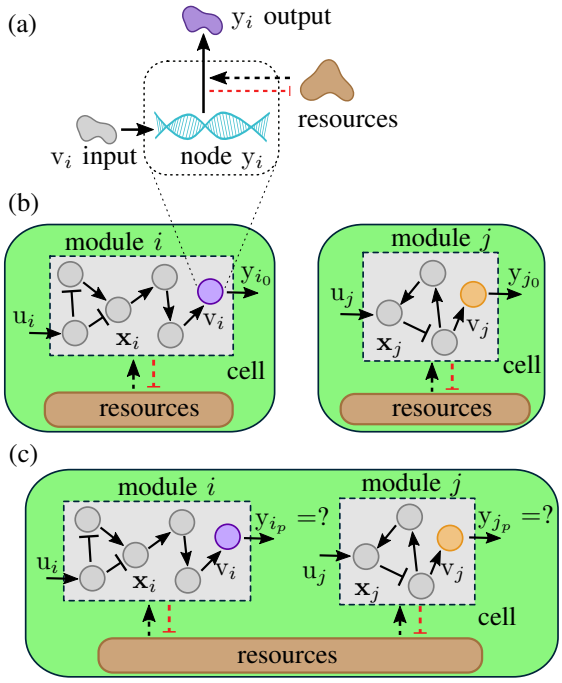

Fig. 1. (a) Representation of a node in a genetic circuit. The node takes the transcription factor protein v_i as input and produces the protein y_i as output. The node uses cellular resources for production of the protein y_i, represented by the black dashed arrow. Through using these resources, the node also sequesters these resources, represented by the red dashed arrow. (b) The input/output behavior of module i and module j are measured in isolation. All nodes interact with the shared cellular resources, shown by dashed arrows. Red dashed arrows indicate sequestration of resources. (c) Due to the sharing of cellular resources, when module i and module j are placed together in a cell, the output cannot be predicted solely by the input/output behavior of the modules done in isolation.

We begin with the ordinary differential equation (ODE) model for protein production of the output node of module i, y_i, taking into account ribosome and RNAP sharing [23] and experimentally validated in [13], given as

$$\frac{dy_i}{dt} = \frac{\alpha_{y_i} F_{y_i}(v_i)}{1 + \sum_{j\in\mathcal{I}_i} w_j F_j(\mathbf{x}_i) + \sum_{k\notin\mathcal{I}_i} w_k F_k(\mathbf{z})} - \delta y_i. \quad (1)$$

Here, y_i is the concentration of the protein y_i, α_{y_i} is a scaling factor on the rate of production of y_i, $F_{y_i}(v_i)$ is a normalized Hill function [23, 24] with input v_i and has the form of $F_{y_i}(v_i) = \frac{1+a_i {v_i}^n}{1+b_i {v_i}^n}$ for one input with a_i and b_i as constants, $\mathbf{x}_i$ is a vector of the concentrations of proteins inside the module, $\mathcal{I}_i$ is the set of indices for all proteins within module i including y_i, and $\sum_{j\in\mathcal{I}_i} w_j F_j(\mathbf{x}_i)$ is proportional to the quantity of resources used by proteins within the module, $\mathbf{z}$ is a vector of the concentrations of proteins external to module i, and δ represents dilution and degradation of proteins due to cell growth. The term $\sum_{k\notin\mathcal{I}_i} w_k F_k(\mathbf{z})$ quantifies the resources used by nodes external to module i (e.g. the nodes of module j in Figure 1c) and is considered as an external disturbance. For notational simplicity, we define $w(\mathbf{z}) \triangleq \sum_{k\notin\mathcal{I}_i} w_k F_k(\mathbf{z})$ and drop the dependency of $w(\mathbf{z})$ on $\mathbf{z}$.

We assume there are no connections between the module and the nodes external to the module. We assume that y_i is a fluorescent reporter and is included as part of module i. We denote the steady state concentration of y_i as $y_i^{ss}(\mathbf{x}_i)$ and let $v_i, y_i \in \mathbf{x}_i$ for every module i. It has been experimentally shown that the resources in synthetic circuits are not shared with the resources for the cell's metabolism [25], so we do not consider resource sharing with the cell's metabolism in the model.

We now present the definitions of two metrics for a genetic circuit module i: Q_i and S_i in Definitions 1 and 2, respectively.

Definition 1: The quantity of resources used by module i, or resource usage, is defined as

$$Q_i(\mathbf{x}_i) = \sum_{j\in\mathcal{I}_i} w_j F_j(\mathbf{x}_i) \quad (2)$$

Definition 2: The sensitivity of the output of the module with respect to changes in resource availability is defined as

$$S_i(\mathbf{x}_i) = \frac{dy_i^{ss}(\mathbf{x}_i)}{dw} \frac{1}{y_i^{ss}(\mathbf{x}_i)|_w} \quad (3)$$

Both $Q_i(\mathbf{x}_i)$ and $S_i(\mathbf{x}_i)$ are nondimensional metrics that depend on the states of module i. $Q_i(\mathbf{x}_i)$ is a positive real number and $S_i(\mathbf{x}_i)$ may be any real number, but is typically negative since concentrations tend to decrease as resource competition is increased. For brevity in notation, we drop the explicit dependence of both $Q_i(\mathbf{x}_i)$ and $S_i(\mathbf{x}_i)$ on $\mathbf{x}_i$ for the remainder of this paper. We now show some useful properties of the metrics Q_i and S_i.

Suppose we have multiple genetic circuit modules i for $i \in \mathcal{A}$, and we wish to find the net resource usage, Q_i, of all modules combined. We define a module A as a module containing all nodes in modules i for $i \in \mathcal{A}$. Then, the total resource usage of the module $Q_\mathcal{A}$ is given as

$$Q_\mathcal{A} = \sum_{i\in\mathcal{A}} \left(\sum_{j\in\mathcal{I}_i} w_j F_j(\mathbf{x}_i) \right) = \sum_{i\in\mathcal{A}} Q_i, \quad (4)$$

following from Definition 1. Then, the resource usage for a module is the sum of the resource usages of all submodules.

Preprint submitted to 57th IEEE Conference on Decision and Control.
Received March 15, 2018.

A. Using Q_i and S_i for Prediction

We now consider the situation where two or more modules are present in the cellular context (Figure 1(c)), and seek to quantify the module's output. We show that prediction of the output of each module is enabled by knowledge of the input/output relation of the module in isolation, Q_i, and S_i for each module i.

The total resource usage from all modules is the sum of the resource usage due to each individual module within the host cell. Therefore, the effective resource disturbance, w, on module i from all other modules in a system is $w = \sum_{j \neq i} Q_j$ where Q_j is the resource usage of module j. For analysis purposes, we assume that w is sufficiently small such that we may use a Taylor approximation for a few key functions. However, as we show in Section V, our predictions still perform well went w is not small, and, in particular, always better than predictions made neglecting resource sharing. Using Definition 2, we find:

Claim 1: The output of module i, y_{i_p}, when sharing the resource pool with other genetic circuit modules $j = 1, \ldots, n$ is

$$y_{i_p}^{ss}(\mathbf{x}_i) = \frac{y_{i_0}^{ss}(\mathbf{x}_i)}{1 - S_i \left(\sum_{j \neq i} Q_j \right)} + \mathcal{O}\left(\frac{d^2 y_i^{ss}}{dw^2} w^2 \right). \quad (5)$$

where $y_{i_0}(\mathbf{x}_i)$ is the output of module i measured in isolation, S_i is the sensitivity of module i, and $w = \sum_{j \neq i} Q_j$ is the resource usage of all modules sharing the cellular resources except for module i.

Proof: The external disturbance is the sum of the resource usage $w = \sum_{j \neq i} Q_j$. Then, by Definition 2, we have

$$S_i \cdot w = \frac{dy_i^{ss}(\mathbf{x}_i)}{dw} \frac{w}{y_i^{ss}(\mathbf{x}_i)}. \quad (6)$$

Substituting for $\frac{dy_i^{ss}}{dw}$ using the Taylor series for y_i with respect to w and simplifying, we have

$$S_i \left(\sum_{j \neq i} Q_j \right) = \frac{y_{i_p}^{ss}(\mathbf{x}_i) - y_{i_0}^{ss}(\mathbf{x}_i)}{y_{i_p}^{ss}(\mathbf{x}_i)} + \mathcal{O}\left(\frac{d^2 y_i^{ss}}{dw^2} w^2 \right), \quad (7)$$

where $y_{i_0}^{ss}(\mathbf{x}_i)$ is the output of module i measured in isolation for input u_i and $y_{i_p}^{ss}(\mathbf{x}_i)$ is the steady state output of module i for input to the module u_i while sharing resources with other modules. Solving for $y_{i_p}^{ss}(\mathbf{x}_i)$ in (7) gives (5). ■

From Claim 1, it can be seen that the behavior of any module when sharing the cellular resources with other modules may be predicted with knowledge of the input/output behavior measured in isolation, $y_{i_0}^{ss}(\mathbf{x}_i)$, S_i for module i, and Q_i for all other modules sharing the cellular resources. Therefore, we propose a procedure to estimate Q_i and S_i from data for a given module. Then, Claim 1 may be used to predict the output of each module when multiple modules are present in the cell and share the same resources.

III. Measurement Procedure

We now present a series of experiments used to estimate the metrics Q_i and S_i for module i by measuring the module with a *known* resource disturbance and observing the change on the input/output behavior.

These experiments are separated into two distinct steps. The first three experiments are used to characterize a disturbance to the pool of resources, w. The final two experiments then characterizes the module's properties with respect to resource sharing given that the resource disturbance from the first three experiments is well characterized.

The proposed measurement procedure consists of measurements of combinations of the module under investigation, a fluorescent reporter protein to measure the output of the module (green fluorescent protein, GFP), and an additional fluorescent reporter (red fluorescent protein, RFP) representing an external disturbance to the pool of resources, w. By observing the change in the output of the module (GFP) and the change in the RFP concentration, we are able to estimate $\hat{Q}_i$ and $\hat{S}_i$ for module i. Any reporter protein may be substituted, but for brevity we will refer to the output of the module as GFP and the external disturbance as RFP.

In Experiments 1, 2, and 3, we first use a constitutive fluorescent reporter protein (yellow fluorescent protein, YFP) to characterize the RFP disturbance. Then, in Experiments 4 and 5, we characterize the module. Figure 2 shows a diagram of each experiment for this procedure to estimate Q_i and S_i.

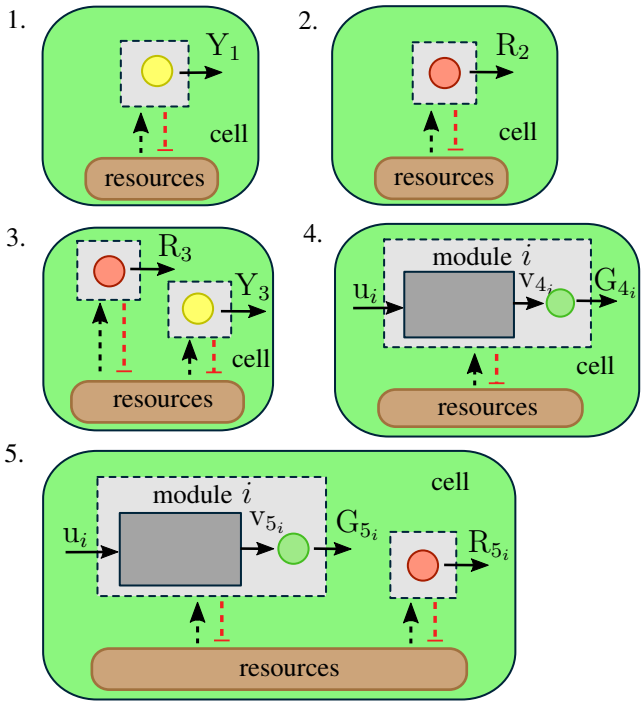

Fig. 2. Diagram for all experiments to run to estimate Q_i and S_i for module i. The steady state fluorescent concentration must be measured for each case and for each value of the input to the module, u_i.

We present a list of the steps in the proposed experimental procedure. Each step corresponds to a panel of Figure 2.

Preprint submitted to 57th IEEE Conference on Decision and Control.
Received March 15, 2018.

1) Measure the steady state output of a constitutive YFP reporter in isolation as shown in Figure 2-1. Solving (1) for the steady state of the output gives the steady state concentration of YFP, Y_1^{ss}, as

$$Y_1^{ss} = \frac{\alpha_Y/\delta}{1+w_Y}, \tag{8}$$

where α_Y is the scaling factor for protein production, δ is the dilution due to cell growth, and w_Y is the resource usage of the YFP node (compare with $\sum_{j\in\mathcal{I}_i} w_j F_j(\mathbf{x})$ in (1)). The external disturbance in (1) is $w(\mathbf{z}) = 0$ in this case since there are no modules present external to the YFP node.

2) Measure the steady state output of a constitutive RFP reporter in isolation as shown in Figure 2-2. Solving (1) for the steady state of the output gives the steady state concentration of RFP, R_2^{ss}, as

$$R_2^{ss} = \frac{\alpha_R/\delta}{1+w_R}, \tag{9}$$

where α_R is the scaling factor for protein production, and w_R is the resource usage of the RFP node (compare with $\sum_{j\in\mathcal{I}_i} w_j F_j(\mathbf{x})$ in (1)). The external disturbance in (1) is $w(\mathbf{z}) = 0$ in this case since there are no modules present external to the RFP node.

3) Measure the steady state output of the RFP reporter and GFP module, both constitutively expressed in the same host cell as shown in Figure 2-3. Solving (1) for the steady state of the output gives the steady state concentration of YFP and RFP, Y_3^{ss} and R_3^{ss}, as

$$Y_3^{ss} = \frac{\alpha_Y/\delta}{1+w_R+w_Y} \quad R_3^{ss} = \frac{\alpha_R/\delta}{1+w_R+w_Y}. \tag{10}$$

Comparing with (1), we consider the RFP and YFP nodes to be separate modules, so w_Y is the internal resource usage and w_R is the external resource usage for the YFP node, while for the RFP node, w_R is the internal resource usage and w_Y is the external resource usage.

4) Measure the output (GFP) of module i at steady state for various levels of the input to the module, u_i, as shown in Figure 2-4. Solving (1) for the steady state of the output gives the steady state concentration of GFP, G_4^{ss}, as

$$G_{4_i}^{ss} = \frac{\alpha_{G_i} F_{G_i}(v_{4_i})/\delta}{1+\sum_{j\in\mathcal{I}_i} w_j F_j(\mathbf{x}_{4_i})}, \tag{11}$$

where α_{G_i} is the scaling factor for GFP production, $F_{G_i}(v_{4_i})$ is the Hill function relating the production rate of GFP to the input to the GFP node, v_{4_i}, and $\sum_{j\in\mathcal{I}_i} w_j F_j(\mathbf{x}_{4_i})$ represents the resources usage by module i and the GFP reporter node.

5) Measure the output of module i (GFP) at steady state for each input value to the module, u, with the RFP resource disturbance used in Experiments 2 and 3 in the same host cell, as shown in Figure 2-5. Solving (1) for the steady state of the output gives the steady state concentration of GFP and RFP, G_5^{ss} and R_5^{ss}, as

$$G_{5_i}^{ss} = \frac{\alpha_{G_i} F_{G_i}(v_{5_i})/\delta}{1+w_R+\sum_{j\in\mathcal{I}_i} w_j F_j(\mathbf{x}_{5_i})} \tag{12a}$$

$$R_{5_i}^{ss} = \frac{\alpha_R/\delta}{1+w_R+\sum_{j\in\mathcal{I}_i} w_j F_j(\mathbf{x}_{5_i})}, \tag{12b}$$

where α_{G_i} is the scaling factor for GFP production, $F_{G_i}(v_{5_i})$ is the Hill function relating the production rate of GFP to the input to the GFP node, v_{5_i}, and $\sum_{j\in\mathcal{I}_i} w_j F_j(\mathbf{x}_{5_i})$ represents the resources usage by module i and the GFP reporter node. For the GFP output, we consider the RFP node to be external to module i, so w_R is compared to $\sum_{k\notin\mathcal{I}_i} w_k F_k(\mathbf{z})$ in (1). Conversely, the module is considered to be external to module i for the RFP node, so w_R represents the internal resource usage and $\sum_{j\in\mathcal{I}_i} w_j F_j(\mathbf{x}_{5_i})$ represents the external resource usage.

With these five experiments, it is possible to solve for Q_i and S_i using the equations for the steady state of each measurement. For the remainder of the analysis, we only consider steady state concentrations, so the ss will be dropped from each measured quantity for brevity. If multiple modules need to be characterized, only Experiments 4 and 5 need to be repeated for each module.

A. Assumptions

Standard assumptions to model genetic circuits using deterministic ODE models are listed in Appendix A. All additional assumptions are listed here.

1) The presence of the synthetic circuit does not affect the growth of the cell [26] and the growth conditions are the same for all experiments. Dilution and degradation of proteins does not require resources and is the same for all measured proteins in all experiments, i.e δ is the same for all experiments.
2) Retroactivity [27] does not affect the steady state behavior.
3) The resource usage of the module is not significantly affected by the disturbance, w, i.e. $\left(\frac{\partial Q_i}{\partial \mathbf{x}_i}\cdot\frac{d\mathbf{x}_i}{dw}\right) \ll 1$.
4) The total second derivative of $F_{G_i}(v_i)$ with respect to the disturbance is small, i.e $\left(\frac{\partial^2 F_{G_i}}{\partial v_i^2}\cdot\frac{d^2 v_i}{dw^2}\right) \ll 1$.

We discuss the validity of these assumptions further in Section IV.

IV. Estimating Metrics from Measurements

We now use the values for the steady states Y_1, R_2, Y_3, R_3, G_{4_i}, G_{5_i}, and R_{5_i} from Experiments 1–5 to find estimators for Q_i and S_i.

Preprint submitted to 57th IEEE Conference on Decision and Control.
Received March 15, 2018.

A. Characterization of Resource Disturbance

We use the steady state concentrations of YFP and RFP (Y_1, R_2, Y_3, R_3) from Experiments 1, 2, and 3 and solve for w_R and w_Y. Knowledge of w_R will then be used to estimate Q_i and S_i for the module. We define the dimensionless quantities

$$\hat{Y}_0 \triangleq \frac{Y_3}{Y_1} \quad \hat{R}_0 \triangleq \frac{R_3}{R_2}. \tag{13}$$

Lemma 1: The resource usage for both constitutive YFP, w_Y, and constitutive RFP, w_R, are given as

$$w_Y = \frac{1-\hat{R}_0}{\hat{R}_0+\hat{Y}_0-1} \tag{14a}$$

$$w_R = \frac{1-\hat{Y}_0}{\hat{R}_0+\hat{Y}_0-1}. \tag{14b}$$

Proof: Proof follows by substituting (8), (9), and (10) into (13) and solving for w_R and w_Y. ■

B. Characterization of Module

We now find estimators for the metrics Q_i and S_i in Theorem 1 using the information obtained in Experiments 4 and 5. Then, Corollary 1 quantifies the error introduced due to the estimation of these metrics in predicting the input/output behavior of module i when in the same context as other genetic modules. We begin by defining the quantities for each module i

$$\hat{G}_i = \frac{G_{5_i}}{G_{4_i}} \quad \hat{R}_i = \frac{R_{5_i}}{R_2} \quad \hat{F}_i = \frac{F_{G_i}(v_{5_i})}{F_{G_i}(v_{4_i})}. \tag{15}$$

Furthermore, from Definition 1, the total derivative of Q_i with respect to w is given as

$$\frac{dQ_i}{dw} \triangleq \frac{d}{dw_R}\left(\sum_{j\in\mathcal{I}_i} w_j F_j(\mathbf{x}_{4_i})\right) = \sum_{j\in\mathcal{I}_i} w_j \frac{\partial F_j(\mathbf{x}_{4_i})}{\partial \mathbf{x}_{4_i}} \frac{d\mathbf{x}_{4_i}}{dw}. \tag{16}$$

and the second total derivative of the function $F_{G_i}(v_i)$ with respect to w is

$$\frac{d^2F_{G_i}(v_i)}{dw^2} \triangleq \frac{\partial^2 F_{G_i}(v_i)}{\partial v_i^2}\frac{dv_i^2}{dw_R^2}. \tag{17}$$

Theorem 1: Let the estimator $\hat{Q}_i$ for Q_i in Definition 1 be

$$\hat{Q}_i \triangleq \left(\frac{1}{\hat{R}_i}-1\right)(1+w_R) \tag{18}$$

and let the estimator $\hat{S}_i$ for S_i in Definition 2 be

$$\hat{S}_i \triangleq \frac{\hat{G}_i-1}{\hat{G}_i}\left(\frac{1+w_R(1-\hat{R}_i)}{w_R(1+w_R)}\right). \tag{19}$$

Then

$$\left|Q_i-\hat{Q}_i\right| = \mathcal{O}\left(\frac{dQ_i}{dw}w_R\right), \tag{20}$$

and

$$\left|S_i-\hat{S}_i\right| = \mathcal{O}\left(\frac{dQ_i}{dw}w_R\right) + \mathcal{O}\left(\frac{d^2F_{G_i}}{dw^2}w_R^2\right). \tag{21}$$

Proof: We combine (9), (11), and (12) with (15), which gives

$$\hat{G}_i = \frac{F_{G_i}(v_{5_i})}{F_{G_i}(v_{4_i})}\left(\frac{1+\sum_{j\in\mathcal{I}_i} w_j F_j(\mathbf{x}_{4_i})}{1+w_R+\sum_{j\in\mathcal{I}_i} w_j F_j(\mathbf{x}_{5_i})}\right) \tag{22a}$$

$$\hat{R}_i = \frac{1+w_R}{1+w_R+\sum_{j\in\mathcal{I}_i} w_j F_j(\mathbf{x}_{5_i})}. \tag{22b}$$

Using the Taylor approximation for $\sum_{j\in\mathcal{I}_i} w_j F_j(\mathbf{x}_{5_i})$ with respect to the disturbance, w_R, we have $\sum_{j\in\mathcal{I}_i} w_j F_j(\mathbf{x}_{5_i}) = \sum_{j\in\mathcal{I}_i} w_j F_j(\mathbf{x}_{4_i}) + \mathcal{O}\left(\frac{dQ_i}{dw}w_R\right)$. Then (22) becomes

$$\hat{G}_i = \hat{F}_i\left(\frac{1+\sum_{j\in\mathcal{I}_i} w_j F_j(\mathbf{x}_{4_i})}{1+w_R+\sum_{j\in\mathcal{I}_i} w_j F_j(\mathbf{x}_{4_i})+\mathcal{O}\left(\frac{dQ_i}{dw}w_R\right)}\right) \tag{23a}$$

$$\hat{R}_i = \frac{1+w_R}{1+w_R+\sum_{j\in\mathcal{I}_i} w_j F_j(\mathbf{x}_{4_i})+\mathcal{O}\left(\frac{dQ_i}{dw}w_R\right)}. \tag{23b}$$

We solve for $\sum_{j\in\mathcal{I}_i} w_j F_j(\mathbf{x}_{4_i})$ and $\hat{F}_i$ in (23) simultaneously to find

$$\sum_{j\in\mathcal{I}_i} w_j F_j(\mathbf{x}_{4_i}) = \left(\frac{1}{\hat{R}}-1\right)(1+w_R)+\mathcal{O}\left(\frac{dQ_i}{dw}w_R\right) \tag{24a}$$

$$\hat{F}_i = \frac{\hat{G}_i(1+w_R)}{1+w_R(1-\hat{R}_i)}+\mathcal{O}\left(\frac{dQ_i}{dw}w_R\right). \tag{24b}$$

Then, by (24a), the estimator for Q_i is given as

$$Q_i = \left(\frac{1}{\hat{R}_i}-1\right)(1+w_R)+\mathcal{O}\left(\frac{dQ_i}{dw}w_R\right). \tag{25}$$

We now use Definition 2 with (12) to find an estimator of S_i as

$$\hat{S}_i = \frac{dG_{5_i}}{dw_R}\frac{1}{G_{5_i}} \tag{26a}$$

$$\hat{S}_i = \frac{1}{F_{G_i}(v_{5_i})}\frac{dF_{G_i}(v_{5_i})}{dw_R} - \frac{1+\frac{dQ_i}{dw_R}}{1+\sum_{j\in\mathcal{I}_i} w_j F_j(\mathbf{x}_{5_i})+w_R}. \tag{26b}$$

We again use the Taylor approximation $\sum_{j\in\mathcal{I}_i} w_j F_j(\mathbf{x}_{5_i}) = \sum_{j\in\mathcal{I}_i} w_j F_j(\mathbf{x}_{4_i}) + \mathcal{O}\left(\frac{dQ_i}{dw}\right)$ and approximate the derivative $\frac{dF_{G_i}(v_5)}{dw_R} = \frac{F_{G_i}(v_5)-F_{G_i}(v_{4_i})}{w_R}+\mathcal{O}\left(\frac{d^2F_{G_i}}{dw^2}w_R^2\right)$. Then we have

$$\hat{S}_i = \frac{1-1/\hat{F}_i}{w_R} - \frac{1}{1+w_R+\sum_{j\in\mathcal{I}_i} w_j F_j(\mathbf{x}_{4_i})} + \mathcal{O}\left(\frac{dQ_i}{dw}w_R\right)+\mathcal{O}\left(\frac{d^2F_{G_i}}{dw^2}w_R^2\right). \tag{27}$$

Preprint submitted to 57th IEEE Conference on Decision and Control.
Received March 15, 2018.

Substituting into (27) for $\sum_{j\in\mathcal{I}_i} w_j F_j(\mathbf{x}_{4_i})$ from (24a) and for $\hat{F}_i$ from (24b) and simplifying, we obtain

$$S_i = \frac{\hat{G}_i - 1}{\hat{G}_i}\left(\frac{1+w_R(1-\hat{R}_i)}{w_R(1+w_R)}\right) + \mathcal{O}\left(\frac{dQ_i}{dw}w_R\right) + \mathcal{O}\left(\frac{d^2F_{G_i}}{dw^2}w_R^2\right). \quad (28)$$

The error quantification of $\hat{S}_i$ follows directly from (28). ■

Note that the estimation error for both $\hat{Q}_i$ and $\hat{S}_i$ are small if Assumptions 3 and 4 are satisfied. Furthermore, as the disturbance $w_R \to 0$, $w_R > 0$, then $\hat{Q}_i \to Q_i$ and $\hat{S}_i \to S_i$. This is due to the fact that the resource usage $\sum_{j\in\mathcal{I}_i} w_j F_j(\mathbf{x}_4)$ in Experiment 4 and the resource usage, $\sum_{j\in\mathcal{I}_i} w_j F_j(\mathbf{x}_5)$ in Experiment 5 becomes closer to each other for smaller w_R. Additionally, the approximation of the derivative, $\frac{dy_i^{ss}(\mathbf{x})}{dw}$, using the Taylor series expansion in Claim 1 and Theorem 1 becomes better as w_R decreases. We discuss the range of validity of this assumption and the resultant errors in Section V.

With Theorem 1, we can find estimators $\hat{Q}_i$ and $\hat{S}_i$ for every module and use Claim 1 with these estimators to find a prediction for the output of module i when perturbed by the other modules with total resource disturbance $w = \sum_{j\neq i} \hat{Q}_j$. This is stated formally in Corollary 1.

Corollary 1: Let the estimate for the output of module i, $\hat{y}_{i_p}(\mathbf{x}_i)$, be

$$\hat{y}_{i_p}(\mathbf{x}_i) = \frac{y_{i_0}(\mathbf{x}_i)}{1 - \hat{S}_i\left(\sum_{j\neq i}\hat{Q}_j\right)}, \quad (29)$$

where $y_{i_0}(\mathbf{x}_i)$ is the output of module i measured in isolation for the input to the module u_i. Then the estimation error is given as

$$|\hat{y}_{i_p}(\mathbf{x}_i) - y_{i_p}(\mathbf{x}_i)| = \mathcal{O}\left(\frac{d^2y_i}{dw^2}w_R^2\right) + \mathcal{O}\left(\sum_{j\neq i}\frac{d\hat{Q}_j}{dw}w_R\right) + \mathcal{O}\left(\frac{d^2F_{G_i}}{dw^2}w_R^2\right). \quad (30)$$

Proof: The proof follows directly by applying the estimators $\hat{S}_i$ and $\hat{Q}_i$ in Theorem 1 to Claim 1. ■

V. Simulation Example

To validate the performance of our procedure at predicting changes is genetic circuit behavior when combined with other modules in a host cell we simulate three different modules, A, B, and C in isolation using the reduced model in (1) (shown in Figures 3, 4, and 5) and perform Experiments 1–5 for each module to find $\hat{Q}_i$ and $\hat{S}_i$ for $i \in \{A, B, C\}$. We then simulated the three modules together in all combinations of pairs (A with B, B with C, and A with C) and compared the output for each module predicted using $\hat{Q}_i$ and $\hat{S}_i$ with Corollary 1 to the actual output for each.

We simulated all combinations of pairs of modules and compared the predictions made using the estimators $\hat{Q}_i$ and $\hat{S}_i$ for each module with the actual input/output behavior. Figures 6, 7, and 8 show the input/output behavior of each module simulated in isolation and with other modules. The output of each module when combined with other modules is compared with the predicted output using $\hat{Q}_i$ and $\hat{S}_i$ in (5). For each of the modules considered, all predictions are accurate to within 2% of the actual behavior simulated, while the error introduced by neglecting all resource sharing ranges from 2.5% to 68%. The estimates using $\hat{Q}_i$ and $\hat{S}_i$ are *uniformly better* than predictions made neglecting resource sharing. All parameters used in the simulation are given in Table I.

Remark 1: In our derivation, we made the assumption that w is small in order to use the Taylor approximation for the output y_i. Figures 6, 7, and 8 show that this assumption is valid and gives good prediction results with about 2% error for $w = \sum_{j\neq i} Q_j$ up to approximately 0.7. When $w > 1$, this assumption begins to break down and the results become worse, giving approximately 15–25% error. However, when $w > 1$, our approach still gives uniformly better predictions than an approach that neglects resource sharing which gives errors on the order of 50–80%.

VI. Conclusions

In this paper, we introduced two metrics for a genetic circuit module i: Q_i, the quantity of cellular resources used by a genetic circuit, and S_i, the sensitivity of the output of that circuit to changes in resource availability. We showed that Q_i and S_i may be used to predict the changes in output of a genetic circuit due to changes in the availability of cellular resources. We then showed that these metrics may be estimated through a five-step experimental procedure and found estimators $\hat{Q}_i$ and $\hat{S}_i$ for Q_i and S_i in Theorem 1 from information obtained through the procedure. In Corollary 1, we demonstrated that $\hat{Q}_i$ and $\hat{S}_i$ may be used to estimate

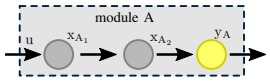

Fig. 3. Module A–a genetic three-protein activation cascade. The input v activates the protein $\mathrm{x_{A_1}}$ which activates $\mathrm{x_{A_2}}$ which then activates the output, $\mathrm{y_A}$.

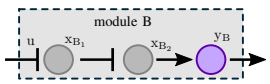

Fig. 4. Module B–a genetic three-protein repression cascade. The input v represses the protein $\mathrm{x_{B_1}}$, which represses the production of $\mathrm{x_{B_2}}$, which then activates the output, $\mathrm{y_B}$.

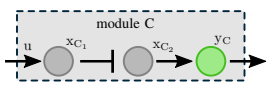

Fig. 5. Module C–a genetic three-protein activation cascade. The input v activates the production of the protein $\mathrm{x_{C_1}}$, which represses $\mathrm{x_{C_2}}$, which then activates the output, $\mathrm{y_C}$.

Preprint submitted to 57th IEEE Conference on Decision and Control.
Received March 15, 2018.

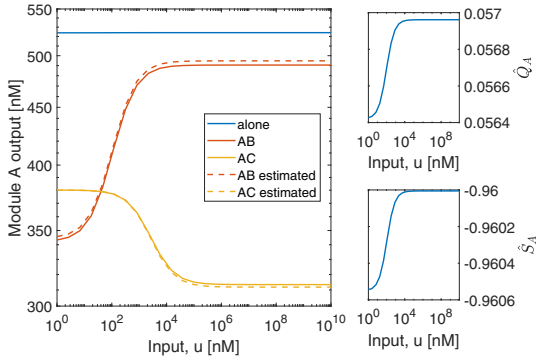

Fig. 6. (left) The input/output relation for various levels of input, u, of module A is given in blue while the input/output relation for module A while sharing the resource pool with modules B and C are given in red and yellow, respectively. The output of module A in isolation increases as the input is increased, but this is not easily seen due to the scaling of the figure. The estimated output using $\hat{Q}_A$ and $\hat{S}_A$ for both cases is given by the dashed red and yellow lines. (upper right) Resource usage estimator $\hat{Q}_A$ for module A as a function of the input. (lower right) Resource sensitivity estimator $\hat{S}_A$ as a function of the input.

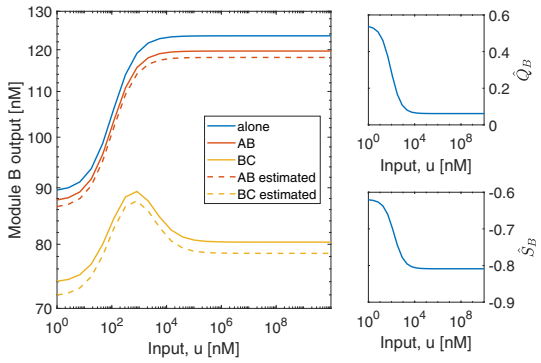

Fig. 7. (left) The input/output relation for various levels of input, u, of module B is given in blue while the input/output relation for module B while sharing the resource pool with modules A and C are given in red and yellow, respectively. The estimated output using $\hat{Q}_B$ and $\hat{S}_B$ for both cases is given by the dashed red and yellow lines. (upper right) Resource usage estimator $\hat{Q}_B$ for module B as a function of the input. (lower right) Resource sensitivity estimator $\hat{S}_B$ as a function of the input.

the output of module i in the context of other genetic circuit modules sharing cellular resources. By measuring $\hat{Q}_i$ and $\hat{S}_i$ for every module, we are able to predict the output of each module. This allows genetic circuit modules to be characterized individually in isolation, then, using $\hat{Q}_i$ and $\hat{S}_i$, predict the output of the module when combined with other modules, helping to avoid lengthy, ad hoc genetic circuit design and may inform design decisions. We performed the measurement procedure for sample genetic circuits in simulation to verify our results. The simulations show that using information about $\hat{Q}_i$ and $\hat{S}_i$ improves the predicted error in output from up to 67% when resource sharing is neglected to less then 2% using $\hat{Q}_i$ and $\hat{S}_i$ for prediction for all input levels. In the future, we wish to experimentally verify these metrics and analyze their performance with noisy measurements.

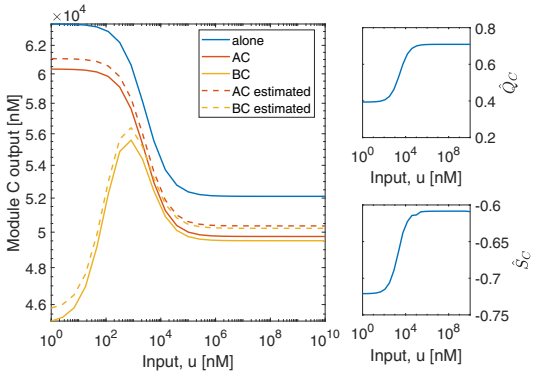

Fig. 8. (left) The input/output relation for various levels of input, u, of module C is given in blue while the input/output relation for module C while sharing the resource pool with modules A and B are given in red and yellow, respectively. The estimated output using $\hat{Q}_C$ and $\hat{S}_C$ for both cases is given by the dashed red and yellow lines. (upper right) Resource usage estimator $\hat{Q}_C$ for module C as a function of the input. (lower right) Resource sensitivity estimator $\hat{S}_C$ as a function of the input.

TABLE I

VALUES USED FOR SIMULATIONS OF RESOURCE SENSOR

description	variable	value
DNA copy numbers module A	DNA_A	134, 324, 503 nM
DNA copy numbers module B	DNA_B	280, 543, 196 nM
DNA copy numbers module C	DNA_C	684, 226, 505 nM
RFP copy number	DNA_R	50 nM
DNA transcription rate	k_1	$437\,\mathrm{h}^{-1}$
mRNA translation rate	k_2	$551\,\mathrm{h}^{-1}$
mRNA degradation rate	δ_1	$7.26\,\mathrm{h}^{-1}$
protein degradation rate	δ_2	$1.15\,\mathrm{h}^{-1}$
RNAP total concentration	$RNAP_{tot}$	491 nM
ribosome total concentration	$Ribo_{tot}$	594 nM
protein–DNA binding constant module A	K_{0A}	98, 722, 121 nM
protein–DNA binding constant module B	K_{0B}	70, 49, 1.36×10^3 nM
protein–DNA binding constant module C	K_{0C}	3.24×10^3, 26.0, 582 nM
leaky RNAP–DNA binding constant module A	K'_A	1.21×10^5, 8.28×10^5, 9.62×10^3 nM
leaky RNAP–DNA binding constant module B	K'_B	1.39×10^3, 1.11×10^4, 2.26×10^4 nM
leaky RNAP–DNA binding constant module C	K'_C	2.94×10^4, 4.28×10^5, 1.15×10^4 nM
RNAP–DNA binding constant module A	K_{1A}	1.06×10^5, 4.52×10^3, 9.32×10^3 nM
RNAP–DNA binding constant module B	K_{1B}	3.15×10^4, 2.30×10^4, 4.38×10^4 nM
RNAP–DNA binding constant module C	K_{1C}	2.92×10^3, 1.69×10^5, 1.04×10^4 nM
mRNA–ribosome binding constant module A	K_{2A}	3.93×10^5, 2.23×10^5, 7.68×10^5 nM
mRNA–ribosome binding constant module B	K_{2B}	1.27×10^4, 1.94×10^5, 4.37×10^5 nM
mRNA–ribosome binding constant module C	K_{2C}	4.20×10^4, 8.11×10^5, 4.03×10^3 nM

VII. ACKNOWLEDGMENTS

The authors would like to thank Theodore Grunberg for his helpful discussions. We would also like to thank Yili Qian for his helpful discussions and helping to proof read the paper.

REFERENCES

[1] D. E. Cameron, C. J. Bashor, and J. J. Collins, "A brief history of synthetic biology," *Nat Rev Micro*, vol. 12, pp. 381–390, May 2014.

[2] D. Del Vecchio, "Modularity, context-dependence, and insulation in engineered biological circuits," *Trends Biotechnol.*, vol. 33, pp. 111–119, Feb. 2015.

[3] C. T. Y. Chan, J. W. Lee, D. E. Cameron, C. J. Bashor, and J. J. Collins, "'Deadman' and 'Passcode' microbial kill switches for bacterial containment," *Nat Chem Biol*, vol. 12, pp. 82–86, Feb. 2016.

[4] Y. Qian, C. McBride, and D. D. Vecchio, "Programming Cells to Work for Us," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 1, no. 1, p. null, 2018.

[5] O. Wright, G.-B. Stan, and T. Ellis, "Building-in biosafety for synthetic biology," *Microbiology*, vol. 159, no. 7, pp. 1221–1235, 2013.
[6] P. E. M. Purnick and R. Weiss, "The second wave of synthetic biology: from modules to systems," *Nat Rev Mol Cell Biol*, vol. 10, pp. 410–422, June 2009.
[7] J. J. Tyson, K. C. Chen, and B. Novak, "Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell," *Current Opinion in Cell Biology*, vol. 15, pp. 221–231, Apr. 2003.
[8] A. Gyorgy, J. Jiménez, J. Yazbek, H.-H. Huang, H. Chung, R. Weiss, and D. DelVecchio, "Isocost Lines Describe the Cellular Economy of Genetic Circuits," *Biophysical Journal*, vol. 109, pp. 639–646, Aug. 2015.
[9] S. Cardinale and A. P. Arkin, "Contextualizing context for synthetic biology–identifying causes of failure of synthetic biological systems," *Biotechnol J*, vol. 7, pp. 856–866, July 2012.
[10] J. A. J. Arpino, E. J. Hancock, J. Anderson, M. Barahona, G.-B. V. Stan, A. Papachristodoulou, and K. Polizzi, "Tuning the dials of Synthetic Biology," *Microbiology*, vol. 159, no. 7, pp. 1236–1253, 2013.
[11] O. Borkowski, F. Ceroni, G.-B. Stan, and T. Ellis, "Overloaded and stressed: whole-cell considerations for bacterial synthetic biology," *Current Opinion in Microbiology*, vol. 33, pp. 123–130, Oct. 2016.
[12] M. Jens and N. Rajewsky, "Competition between target sites of regulators shapes post-transcriptional gene regulation," *Nat Rev Genet*, vol. 16, pp. 113–126, Feb. 2015.
[13] Y. Qian, H.-H. Huang, J. I. Jiménez, and D. Del Vecchio, "Resource Competition Shapes the Response of Genetic Circuits," *ACS Synth. Biol.*, vol. 6, pp. 1263–1272, July 2017.
[14] S. Klumpp and T. Hwa, "Bacterial growth: global effects on gene expression, growth feedback and proteome partition," *Curr. Opin. Biotechnol.*, vol. 28, pp. 96–102, Aug. 2014.
[15] S. Klumpp, M. Scott, S. Pedersen, and T. Hwa, "Molecular crowding limits translation and cell growth," *PNAS*, vol. 110, pp. 16754–16759, Oct. 2013.
[16] S. Klumpp, J. Dong, and T. Hwa, "On Ribosome Load, Codon Bias and Protein Abundance," *PLOS ONE*, vol. 7, p. e48542, Nov. 2012.
[17] T. E. Gorochowski, I. Avcilar-Kucukgoze, R. A. L. Bovenberg, J. A. Roubos, and Z. Ignatova, "A Minimal Model of Ribosome Allocation Dynamics Captures Trade-offs in Expression between Endogenous and Synthetic Genes," *ACS Synth. Biol.*, vol. 5, pp. 710–720, July 2016.
[18] F. Ceroni, R. Algar, G.-B. Stan, and T. Ellis, "Quantifying cellular capacity identifies gene expression designs with reduced burden," *Nature Methods*, vol. 12, p. 415, Apr. 2015.
[19] F. Ceroni, B. A. Blount, and T. Ellis, "Sensing the Right Time to Be Productive," *Cell Systems*, vol. 3, pp. 116–117, Aug. 2016.
[20] P. Mohammadi, N. Beerenwinkel, and Y. Benenson, "Automated Design of Synthetic Cell Classifier Circuits Using a Two-Step Optimization Strategy," *Cell Systems*, vol. 4, pp. 207–218.e14, Feb. 2017.
[21] A. A. K. Nielsen, B. S. Der, J. Shin, P. Vaidyanathan, V. Paralanov, E. A. Strychalski, D. Ross, D. Densmore, and C. A. Voigt, "Genetic circuit design automation," *Science*, vol. 352, p. aac7341, Apr. 2016.
[22] J. MacDonald, C. Barnes, R. Kitney, P. Freemont, and G.-B. Stan, "Computational design approaches and tools for synthetic biology," *Integrative Biology*, vol. 3, no. 2, pp. 97–108, 2011.
[23] Y. Qian and D. Del Vecchio, "Effective interaction graphs arising from resource limitations in gene networks," in *2015 American Control Conference (ACC)*, pp. 4417–4423, July 2015.
[24] A. Hill, "The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves," *J Physiol*, vol. 40, pp. iv–vii, 1910.
[25] A. Y. Weiße, D. A. Oyarzún, V. Danos, and P. S. Swain, "Mechanistic links between cellular trade-offs, gene expression, and growth," *Proc Natl Acad Sci U S A*, vol. 112, pp. E1038–E1047, Mar. 2015.
[26] M. S. Bienick, K. W. Young, J. R. Klesmith, E. E. Detwiler, K. J. Tomek, and T. A. Whitehead, "The Interrelationship between Promoter Strength, Gene Expression, and Growth Rate," *PLoS One*, vol. 9, Oct. 2014.
[27] D. Del Vecchio, A. J. Ninfa, and E. D. Sontag, "Modular cell biology: retroactivity and insulation," *Mol Syst Biol*, vol. 4, p. 161, Feb. 2008.

APPENDIX

A. Assumptions

(a) The dynamics of protein production may be modeled using a deterministic ODE model. This requires that the cell volume be large enough and protein counts be high enough (typically larger than 1000).
(b) The complexes formed all reactions reaches the quasi-steady state significantly faster than the dynamics of protein production and degradation. Additionally, mRNA dynamics are much faster than protein dynamics.
(c) The system can be modeled in the form of (1). See [23] for a derivation of this model.
(d) The cell is well stirred so spacial effects are negligible.

Preprint submitted to 57th IEEE Conference on Decision and Control.
Received March 15, 2018.