# Co-design of Resource Limited Genetic Modules*

Carlos Eduardo Celeste Junior[1†], Ilaria Di Loreto[2†], Theodore Wu Grunberg[1], Maria Domenica Di Benedetto[2], Alessandro Borri[3] and Domitilla Del Vecchio[1]

***Abstract*— Modular composition of systems through defined input/output interfaces is a wide-spread engineering approach that allows to make the design of complicated systems tractable. Although this approach to design has percolated to the design of synthetic genetic circuits, it has proved challenging to obtain predictable design outcomes. In particular, context-dependence due to sharing a limited pool of cellular resources is a major factor that confounds modular composition of genetic modules. Here, we propose to use a systems framework in which resource sharing among different subsystems is explicitly modeled through disturbance inputs and disturbance outputs. Within this system description, resource sharing results in undesired connectivity among subsystems, which yet is explicitly accounted for design. Accordingly, we propose to use this system framework to co-design systems based on specifications that each subsystem should satisfy. To this end, we provide sufficient conditions on the system parameters such that the output of each subsystem in the network remains in a small interval around a desired value, as well as an algorithmic procedure to compute the feasible region for these parameters. In general, this framework can be used to design subsystems to satisfy a specification, while explicitly accounting for context-dependence.**

## I. INTRODUCTION

In traditional engineering disciplines, such as electrical engineering, mechanical engineering, and computer systems, modular composition is a standard approach to design complicated systems. The basic assumption is that subsystems can be characterized by their input/output behavior and that this behavior is not changing when systems are composed together. It has been long recognized that modular composition is challenging when engineering biology because, in the cell, there are many interactions among subsystems, which go beyond what we regard as the regulatory inputs and outputs that we use for connecting systems to one another [1], [2], [3]. There are many reasons for the failure of modularity in biological circuits, such as the effect of loads (retroactivity) on a system output caused by downstream circuits [4], [5], [6], unknown interactions between adjacent genetic sequences and factors [7], [8], [9], [10], as well as resource competition between systems [3], [2], [11], [12], [13], [14].

In this work, we specifically focus on lack of modularity due to resource sharing. Prior work on this topic experimentally demonstrated how two genetic modules become coupled when they become activated concurrently in the cell even when they are not connected through regulatory links [13]. Related work has further shown that this is the case even if one of the genetic modules is placed on the chromosome [14], highlighting even more this problem as a global perturbation to all genes in the cell. For genetic circuits, wherein more genetic modules are connected to each other through regulatory links, competition for resources among the modules leads to surprising emergent circuit behavior and mathematical models were introduced that well predict experimental outcomes [12]. These experimentally validated models were later adopted in a theoretical study aiming at designing local feedback controllers to insulate genetic modules from one another [15]. This line of work followed the general idea of capturing resource transactions through disturbance inputs to each genetic module and to solve a disturbance attenuation problem [16], [17].

In this paper, we rely on the mathematical models of genetic modules with resource competition developed in [12] and, in particular, on their reformulation introduced in [15]. In this reformulation, the load on resources that a module exerts appears as a disturbance output of the module. In turn, each module takes as disturbance input the ensemble load on the resources that all other modules exert. Based on this model, we determine conditions under which a given input/output specification for each of the genetic modules in the system can be met. To this end, we formulate a specification feasibility problem and provide a solution along with an algorithm to determine the feasible region in the design parameter space.

This paper is organized as follows. In Section II, we introduce the new system formulation and state the feasibility problem. In Section III, we provide the main theorem, which gives sufficient conditions for the satisfaction of the specifications, as well as the algorithm to compute the feasible parameter region. In Section IV, we provide an application example.

*This work was supported by AFOSR Grant # FA9550-22-1-0316 and partially supported by Center of Excellence DEWS

† These authors contributed equally to this project

[1] Corresponding Author: Carlos Eduardo Celeste Junior and Domitilla Del Vecchio, with the Department of Mechanical Engineering and Theodore Wu Grunberg with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, MA 02139, USA celeste8@mit.edu, ddv@mit.edu tgrunber@mit.edu

[2] Ilaria Di Loreto and Maria Domenica Di Benedetto, with the Department of Electrical and Information Engineering, University of L'Aquila, Palazzo Camponeschi, piazza Santa Margherita 2, 67100 L'Aquila ilaria.diloreto@graduate.univaq.it, mariadomenica.dibenedetto@univaq.it

[3] Alessandro Borri with the Institute for Systems Analysis and Computer Science "A. Ruberti", National Research Council of Italy, Via dei Taurini 19, 00185 Rome, Italy alessandro.borri@iasi.cnr.it

## II. PROBLEM FORMULATION

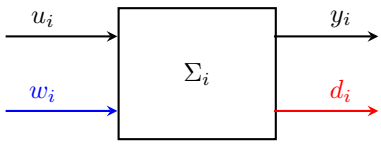


Fig. 1. Block diagram representation of subsystem $\Sigma_i$.

The system model we consider in this paper, for the process of gene expression [18], is depicted in Figure 1. This model describes the protein production process, while accounting for the fact that multiple such systems all share ribosomes required for gene expression [12], [13]. In what follows, we use the standard notation, in which for a species S we let *italics* $S$ denote its concentration.

The $i$-th subsystem is responsible for the expression of the $i$-th gene, where the mRNA $\mathrm{m_i}$ is transcribed at a rate $r_i$, which is then translated into protein $\mathrm{p_i}$. So, we define the $i$-th subsystem states $x_i = [m_i\ p_i]' \in \mathbb{R}^2_+$, with input $u_i = r_i \in \mathbb{R}_+$ and output $y_i = p_i \in \mathbb{R}_+$, as well as disturbance input $w_i \in \mathbb{R}_+$ and output $d_i \in \mathbb{R}_+$. With this, the subsystem dynamics are given by [15]

$$\begin{aligned} \dot{m}_i &= u_i - \delta_0 m_i \\ \dot{p}_i &= \alpha_i \frac{(m_i/k_i)}{1 + (m_i/k_i) + w_i} - \delta p_i \\ y_i &= p_i \\ d_i &= m_i/k_i, \end{aligned} \tag{1}$$

for $i = \{1, \dots, N\}$. Here, $\alpha_i$ is the translation rate constant, $k_i$ is the dissociation constant of mRNA binding with ribosome, $\delta$ is the decay rate constant of the protein, $\delta_0$ is the decay rate constant of mRNA. All parameters are strictly positive.

The disturbance input $w_i$ and disturbance output $d_i$ capture the unintended interactions among subsystems. Specifically, this model was derived in [15] and captures the fact that ribosomes are required in the translation step, where the mRNA binds to ribosomes to be translated to protein, which causes a "load" on the ribosome pool. In particular, the larger $m_i$ and the smaller $k_i$ (stronger ribosome binding site), the larger the load $d_i = m_i/k_i$ that subsystem $\Sigma_i$ applies to ribosomes. Because the decrease of translation rate that system $\Sigma_i$ experiences results from the overall load that all subsystems apply to ribosomes, we have that the disturbance input is given by

$$w_i = \sum_{j \neq i} d_j, \tag{2}$$

which represents the effect that load on ribosomes from all other subsystems has on the $i$-th subsystem. The full derivation of this model can be found in [15].

Since in this paper we are interested in guarantees on the steady state behavior of $N$ interconnected systems, we first prove uniqueness and stability of the equilibrium point.

**Lemma 1.** *System $\Sigma_i$ with interconnection (2) admits a unique equilibrium point. Furthermore, this equilibrium point is asymptotically stable for all parameter values.*

*Proof:* System $\Sigma_i$ equilibrium point is given by

$$m_i = u_i/\delta_0 \tag{3}$$

$$p_i = \frac{(\alpha_i/\delta)(u_i/k_i\delta_0)}{1 + \sum_{j=1}^{N}(u_j/k_j\delta_0)}. \tag{4}$$

For all parameters and inputs $u_i$, this equilibrium point is unique. Now to conclude about its stability we will use the quasi-steady state approximation for the mRNA, that is, since the mRNA dynamics are much faster than the protein dynamics [18], we can approximate the concentration of mRNA $m_i$ by its value at steady state given by (3). With this, we obtain the reduced order system

$$\dot{p}_i = \frac{1}{\alpha_i} \frac{(u_i/k_i\delta_0)}{1 + \sum_{j=1}^{N}(u_j/k_j\delta_0)} - \delta p_i, \tag{5}$$

which is a linear system with a constant input and Jacobian $J = \mathrm{diag}([-\delta \cdots -\delta])$. Since the eigenvalues have negative real part, more specifically $\lambda_i = -\delta$, for $i \in \{1, \dots, N\}$, the equilibrium point is asymptotically stable. Therefore, we can conclude that the system has a unique equilibrium point that is asymptotically stable for all parameter values. □

We are interested in steady state behavior, so we consider the following input/output steady state characteristic of system $\Sigma_i$:

$$y_i = \frac{\alpha_i}{\delta} \frac{(u_i/\delta_0 k_i)}{1 + (u_i/\delta_0 k_i) + w_i} \tag{6}$$

$$d_i = u_i/\delta_0 k_i, \tag{7}$$

and let $y_i^*$ be the output of the isolated system with $u_i = r_i^* > 0$, i.e., $y_i$ in (6) with $w_i = 0$, and nominal parameter values $\alpha_i = \alpha_i^*$, $\delta_i = \delta^*$, $\delta_0 = \delta_0^*$, and $k_i = k_i^* \in K_i$ with $K_i \subset \mathbb{R}^+$. Now let us define the steady state I/O maps $f_i : W_i \to D_i$ as

$$d_i = f_i(w_i) = \gamma_i w_i + \gamma_i, \quad \gamma_i = \frac{\delta y_i}{\alpha_i - \delta y_i}. \tag{8}$$

With this, our system specification is given as follows:

**Specification:** Given $u_i = r_i^*$, $y_i^*$, and fixed tolerances $\epsilon_i > 0$, $i = \{1, \dots, N\}$. The specifications on the connected systems given in (2), (6), (7) are given in the form

$$y_i \in [y_i^* - \epsilon_i, y_i^* + \epsilon_i],\ i \in \{1, ..., N\}. \tag{9}$$

Based on this specification, we seek to tackle two problems. First, we seek to determine sufficient conditions on the systems' parameters to satisfy this specification (Problem 1). The second problem is to design the systems such that the specification is met (Problem 2). For this problem, we regard the ribosome binding site strengths, captured by parameters $1/k_i$ (see [18]) as the design parameters since they are easily and quantitatively tunable.

**Problem 1** (**Feasibility**). *Given $N$ subsystems of the form* (1) *and connection rule (2). Determine conditions on the systems' parameters such that the specification is met.*

The practical relevance of this problem stands in the fact that once multiple systems are concurrently operating in the cell, they may not be able to achieve their nominal outputs as they do in isolation because of decreased availability of gene expression resources to each of them. Therefore, we investigate to what extent it is still possible to meet the specifications as the number of subsystems increases and as the tolerance is changed. Indeed, it is reasonable to expect that with more systems, one may require a larger tolerance and hence a larger degradation of the system specification.

**Problem 2** (**Feasible Region**). *With all other parameters fixed, compute the region for the parameters $(k_1, \dots, k_N) \in K_1 \times \cdots \times K_N$ such that the specification is met.*

With this, we define the quantities

$$\tilde{\gamma}_i = \frac{\delta(y_i^* - \epsilon_i)}{\alpha_i - \delta(y_i^* - \epsilon_i)} \tag{10}$$

$$\hat{\gamma}_i = \frac{\delta(y_i^* + \epsilon_i)}{\alpha_i - \delta(y_i^* + \epsilon_i)}. \tag{11}$$

**Lemma 2.** *The following conditions*

$$\begin{cases} d_i \geq \tilde{\gamma}_i w_i + \tilde{\gamma}_i & (12) \\ d_i \leq \hat{\gamma}_i w_i + \hat{\gamma}_i, & (13) \end{cases}$$

*are equivalent to those in* (9).

*Proof:* We start by showing that (9) implies (12)-(13). The specifications given in (9) define lower and upper bounds on the output $y_i$, based on the tolerances $\epsilon_i$. With this, we can substitute these bounds on $\gamma_i$ as defined in (8), yielding

$$\tilde{\gamma}_i \leq \gamma_i \leq \hat{\gamma}_i, \tag{14}$$

with $\tilde{\gamma}_i$ as given in (10) and $\hat{\gamma}_i$ as given in (11). Now we substitute this into the I/O map given in equation (8), which results in

$$\begin{aligned} \tilde{\gamma}_i w_i + \tilde{\gamma}_i \leq & \gamma_i w_i + \gamma_i \leq \hat{\gamma}_i w_i + \hat{\gamma}_i \\ \Rightarrow \tilde{\gamma}_i w_i + \tilde{\gamma}_i \leq & \quad d_i \quad \leq \hat{\gamma}_i w_i + \hat{\gamma}_i, \end{aligned}$$

which are the conditions presented in (12)-(13).

Now we show that (12)-(13) implies (9). We start with (12), where we substitute $d_i$ as defined in (8):

$$\begin{aligned} d_i \geq \tilde{\gamma}_i w_i + \tilde{\gamma}_i \Rightarrow \gamma w_i + \gamma \geq \tilde{\gamma}_i w_i + \tilde{\gamma}_i \\ \Rightarrow \gamma(1 + w_i) \geq \tilde{\gamma}_i (1 + w_i) \Rightarrow \gamma \geq \tilde{\gamma}_i. \end{aligned}$$

Now we substitute the $\tilde{\gamma}_i$ as defined in (10) and $\gamma$ as defined in (8). This yields

$$\frac{\delta y_i}{\alpha_i - \delta y_i} \geq \frac{\delta(y_i^* - \epsilon_i)}{\alpha_i - \delta(y_i^* - \epsilon_i)} \Rightarrow \alpha_i \delta y_i \geq \alpha_i \delta(y_i^* - \epsilon_i).$$

With this, we obtain $y_i \geq (y_i^* - \epsilon_i)$. Now for (13), we substitute $d_i$ as defined in (8)

$$\begin{aligned} d_i \leq \hat{\gamma}_i w_i + \hat{\gamma}_i \Rightarrow \gamma w_i + \gamma \leq \hat{\gamma}_i w_i + \hat{\gamma}_i \\ \Rightarrow \gamma(1 + w_i) \leq \hat{\gamma}_i (1 + w_i) \Rightarrow \gamma \leq \hat{\gamma}_i. \end{aligned}$$

Now we substitute the $\hat{\gamma}_i$ as defined in (11) and $\gamma$ as defined in (8). This yields

$$\frac{\delta y_i}{\alpha_i - \delta y_i} \leq \frac{\delta(y_i^* + \epsilon_i)}{\alpha_i - \delta(y_i^* + \epsilon_i)} \Rightarrow \alpha_i \delta y_i \leq \alpha_i \delta(y_i^* + \epsilon_i).$$

With this, we obtain $y_i \leq (y_i^* + \epsilon_i)$. Thus, we have shown that the conditions given in (12)-(13) are equivalent to those in (9). □

## III. PROBLEM SOLUTION

We tackle Problem 1 first, that is, we want to determine if there exist parameters $(k_1, \dots, k_n)$ such that our steady state output $y_i$ stays in the prescribed region around $y_i^*$, with tolerances $\epsilon_i$.

Let $w = (w_1, \dots, w_N)$ and $d = (d_1, \dots, d_N)$, then (2) implies that

$$w = Td,$$

with $T \in \mathbb{R}^{N \times N}$ the interconnection matrix defined as

$$\{T\}_{i,j} = \begin{cases} 0, & \text{if } i = j \\ 1, & \text{otherwise.} \end{cases} \tag{15}$$

In turn, (8) with $y_i = y_i^* - \epsilon_i$ can we rewritten in vector form as

$$d = \tilde{\gamma} w + \tilde{\gamma},$$

in which the gain matrix $\tilde{\gamma} \in \mathbb{R}^{N \times N}$ is defined as

$$\{\tilde{\gamma}\}_{i,j} = \begin{cases} \tilde{\gamma}_i = \frac{\delta(y_i^* - \epsilon_i)}{\alpha_i - \delta(y_i^* - \epsilon_i)}, & \text{if } i = j \\ 0, & \text{otherwise.} \end{cases} \tag{16}$$

The following Theorem provides a sufficient conditions to solve Problem 1. For a matrix $A$, we let $\rho(A)$ denote the spectral radius of $A$.

**Theorem 1.** *Let $\tilde{\gamma}$ be the gain matrix defined in* (16)*, and let $T$ be the interconnection matrix defined in* (15)*. If $\rho(\tilde{\gamma}T) < 1$, then Problem 1 has a solution.*

*Proof:* By Lemma 2, satisfaction of the specification is equivalent to (12)-(13) with $d_i \geq 0$. We then focus on providing sufficient conditions for (12)-(13) to be satisfied.

Let us consider just the constraints of the form (12), which, given the matrices $T$ and $\tilde{\gamma}$, defined in (15) and (16), can be rewritten as

$$d \geq \tilde{\gamma} T d + \underline{\tilde{\gamma}} \Longleftrightarrow (I - \tilde{\gamma} T) d \geq \underline{\tilde{\gamma}},$$

where $d = (d_1, \dots, d_N)$ and $\underline{\tilde{\gamma}} = (\tilde{\gamma}_1, \dots, \tilde{\gamma}_N)$. Since $\underline{\tilde{\gamma}} \geq 0$ and $d_i \geq 0$, for all $i$, must hold from the definition of the models, we have that if $(I - \tilde{\gamma} T)^{-1} \geq 0$, i.e., $(I - \tilde{\gamma} T)^{-1}$ has non-negative entries, then

$$d \geq (I - \tilde{\gamma} T)^{-1} \underline{\tilde{\gamma}} \geq 0 \Rightarrow d \geq 0. \tag{17}$$

A sufficient condition to prove that $(I-\tilde{\gamma}T)^{-1} \geq 0$ is given by checking that $(I-\tilde{\gamma}T)^{-1}$ is an $M$-matrix ([19]). Now, given that $M=(I-\tilde{\gamma}T)$ is such that $\{M\}_{i,j} \leq 0$ for all $i \neq j$, and given that $\tilde{\gamma}T \geq 0$, we can exploit the result stated in Lemma 2.5.2.1 in [19], picking $\alpha = 1$, and finally obtaining that

$$(I-\tilde{\gamma}T) \quad \text{is an } M\text{-matrix} \quad \iff 1 > \rho(\tilde{\gamma}T).$$

We can conclude that if $1 > \rho(\tilde{\gamma}T)$, then $(I-\tilde{\gamma}T)^{-1} \geq 0$ and hence (12) is satisfied for $d_i \geq 0$. Now, we prove the satisfaction of the conditions in (13). Let us consider values for $d_i$, such that, $d_i = \tilde{\gamma}_i w_i + \tilde{\gamma}_i$. By plugging these expressions for $d_i$ into (13), we obtain

$$w_i(\hat{\gamma}_i - \tilde{\gamma}_i) + (\hat{\gamma}_i - \tilde{\gamma}_i) \geq 0,$$

which is always true for $(k_1, \ldots, k_N)$ when (17) is satisfied, given that $\tilde{\gamma}_i \leq \hat{\gamma}_i$. To conclude, we have shown that if $\rho(\gamma T) < 1$, then (9) is satisfied. Therefore, Problem 1 has a solution. □

**Remark 1.** *Notice that the systems gains $\gamma_i$ are monotonically increasing with respect to $y_i$, meaning that the steady state I/O maps $d_i = f_i(w_i)$ are monotonically increasing with respect to $y_i$. In fact, if $y_i^* - \epsilon_i \leq y_i^* + \epsilon_i$, then $\tilde{\gamma}_i \leq \hat{\gamma}_i$ and $f_i|_{(y_i^*-\epsilon_i)} \leq f_i|_{(y_i^*+\epsilon_i)}$.*

With this we can move on to Problem 2, where we want to find the feasible region for the systems parameters $(k_1, \ldots, k_N)$. We consider first $N = 2$ as an illustrative example and then propose a general algorithm for arbitrary N.

### A. Illustrative Example

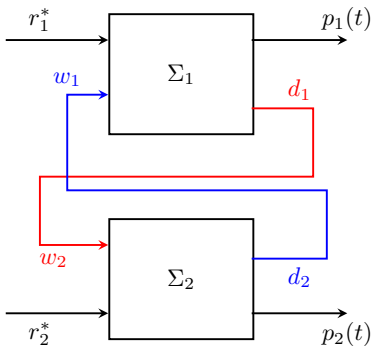


Fig. 2. Example $N = 2$ subsystem network block diagram.

In the case in which $N = 2$, the system network takes the simple form shown in Figure 2. In this case, we have $w_1 = d_2$ and $w_2 = d_1$. The gains of the subsystems, for $y_1 = y_1^* - \epsilon_1$, $y_2 = y_2^* - \epsilon_2$ are given by $\tilde{\gamma}_1 = \delta(y_1^* - \epsilon_1)/(\alpha_1 - \delta(y_1^* - \epsilon_1))$ and $\tilde{\gamma}_2 = \delta(y_2^* - \epsilon_2)/(\alpha_2 - \delta(y_2^* - \epsilon_2))$. The gain matrix $\tilde{\gamma}$ and the interconnection matrix $T$ take the following forms:

$$\tilde{\gamma} = \begin{bmatrix} \tilde{\gamma}_1 & 0 \\ 0 & \tilde{\gamma}_2 \end{bmatrix}, \quad T = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

The eigenvalues of $\tilde{\gamma}T$ are given by $\lambda_1 = \sqrt{\tilde{\gamma}_1\tilde{\gamma}_2}$ and $\lambda_2 = -\sqrt{\tilde{\gamma}_1\tilde{\gamma}_2}$. Then, $\rho(\tilde{\gamma}T) = \sqrt{\tilde{\gamma}_1\tilde{\gamma}_2}$. As a consequence, for a solution to Problem 1 to exist, it is sufficient that $\sqrt{\tilde{\gamma}_1\tilde{\gamma}_2} < 1$. We next compute the region of $(1/k_1, 1/k_2)$ that ensures $\sqrt{\tilde{\gamma}_1\tilde{\gamma}_2} < 1$.

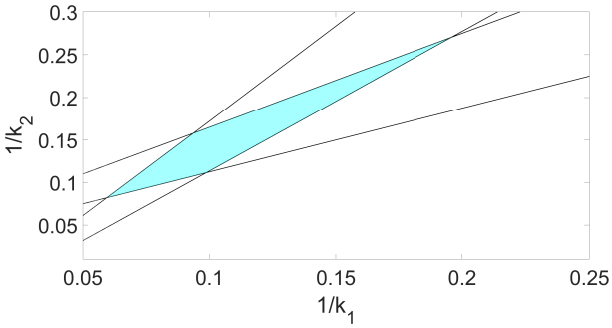


Fig. 3. Feasible region for $1/k_1$ and $1/k_2$, with $r_i^* = (1\ 1)$ $[nM]$, $y_i^* = (2\ 2)$ $[nM]$, $\epsilon_i = 0.1 y_i^* = (0.2\ 0.2)$ $[nM]$, $\alpha_i = (5.8\ 4.2)$ $[nM/hr]$, $\delta = 1$ $[1/hr]$, $\delta_0 = 0.05$ $[1/hr]$, which yields $\tilde{\gamma}_i = (0.45\ 0.75)$ and $\hat{\gamma}_i = (0.61\ 1.10)$.

To compute the feasible region, we first substitute (2) and (7) in (12)-(13), to obtain these inequalities in terms of $(1/k_1, 1/k_2)$

$$\tilde{\gamma}_1\left(\frac{u_2}{\delta_0}\cdot\frac{1}{k_2}\right) + \tilde{\gamma}_1 \leq \left(\frac{u_1}{\delta_0}\cdot\frac{1}{k_1}\right) \leq \hat{\gamma}_1\left(\frac{u_2}{\delta_0}\cdot\frac{1}{k_2}\right) + \hat{\gamma}_1 \tag{18}$$

$$\tilde{\gamma}_2\left(\frac{u_1}{\delta_0}\cdot\frac{1}{k_1}\right) + \tilde{\gamma}_2 \leq \left(\frac{u_2}{\delta_0}\cdot\frac{1}{k_2}\right) \leq \hat{\gamma}_2\left(\frac{u_1}{\delta_0}\cdot\frac{1}{k_1}\right) + \hat{\gamma}_2, \tag{19}$$

where $\tilde{\gamma}_i$ is as defined in (10) and $\hat{\gamma}_i$ is as defined in (11). Then, if $\tilde{\gamma}_1$ and $\tilde{\gamma}_2$ satisfy $\sqrt{\tilde{\gamma}_1\tilde{\gamma}_2} < 1$, we can compute the $(1/k_1, 1/k_2)$ feasible region directly from the inequalities (18)-(19), which is a linear program in the variables $(1/k_1, 1/k_2)$.

One possible solution is shown in Figure 3 in terms of $1/k_1$ and $1/k_2$, that is, the polygon in cyan contain all the points $(1/k_1, 1/k_2)$ for which the specification given in (9) holds. What we obtain is that inside the feasible region, we can decrease concurrently both $k_1$ and $k_2$, so that $d_1$ and $d_2$ also will increase. This, in turn, implies that also $p_1$, $p_2$ will increase, keeping on satisfying the specifications. On the other hand, on the boundaries of the feasible region, we can decrease either $k_1$ or $k_2$, in order to preserve the satisfaction of the specifications.

### B. General Solution to Problem 2

Now we consider the general case in which we have $N$ subsystems and provide an algorithm to determine the feasible region while allowing to change the tolerance $\epsilon_i$. Suppose we have a network composed of $N$ subsystems, with prescribed outputs $y_i^* = p_i^*$, with fixed input $u_i = r_i^* > 0$ and tolerances $\epsilon_i$, with $\tilde{\gamma}_i$ and $\hat{\gamma}_i$ defined in (10) and (11), respectively, for fixed parameter values $\alpha_i$, $\delta$ and $\delta_0$. Our goal is to find the feasible region for $1/k_i$, $i \in \{1, ..., N\}$.

In order to achieve this, we consider the inequalities in (12)-(13), as they describe the feasible region. These inequalities are linear with respect to $d_i$ since $w_i = \sum_{j \neq i} d_j$, so we will first compute the polygon that describes the feasible region for $d_i$, by computing its vertices, then we use the linear relationship between $d_i$ and $1/k_i$ given in (7) to obtain the vertices for the polygon that describes the $1/k_i$ feasible region.

To do this, we solve the following linear system of equations

$$d = \beta T d + \underline{\beta} \Longleftrightarrow d = (I - \beta T)^{-1} \underline{\beta}, \tag{20}$$

where $\underline{\beta} = (\beta_1, \ldots, \beta_N)$, $d = (d_1, \ldots, d_N)$, $\beta = \text{diag}(\underline{\beta})$ and $T$ is as defined in (15). It is important to note that the conditions from Theorem 1 guarantee that the matrix $(I - \beta T)$ is invertible. We define $\beta_i$ as having two possible values, $\tilde{\gamma}_i$, as given in (10), or $\hat{\gamma}_i$, as given in (11) for $i \in \{1, \ldots, N\}$. We then solve (20) for all possible $(\beta_1, ..., \beta_N)$ tuples such that $\beta_i = \tilde{\gamma}_i$ or $\beta_i = \hat{\gamma}_i$. Then, to find the vertices for the $1/k_i$ feasible region we use the relationship $1/k_i = \delta_0 d_i / u_i$ that comes from (7).

Next, to aid in the choice of the tolerance $\epsilon_i$, we introduce a minimization problem that returns suitable values $\epsilon_i$, $\tilde{\gamma}_i$ and $\hat{\gamma}_i$, , for fixed parameters $\alpha_i$, $\delta$ and $\delta_0$.

**Tolerance Minimization Problem**

$$\begin{aligned} \min \quad & \sum_{i=1}^{N} \epsilon_i \\ \text{s.t.} \quad & \\ & \epsilon_{min} \leq \epsilon_i \leq \epsilon_{max} \\ & \tilde{\gamma}_i = \frac{\delta(p_i^* - \epsilon_i)}{\alpha_i - \delta(p_i^* - \epsilon_i)} \\ & \rho(\text{diag}(\tilde{\gamma}_i) T) < 1 \end{aligned}$$

To solve this minimization problem we use the YALMIP toolbox for MATLAB [20]. The bounds on the tolerance $\epsilon_{min}$ and $\epsilon_{max}$ affect the size of the feasible region, which is useful in practice as it is challenging to experimentally set the values of $k_i$ with precision. So, with this in mind, we have introduced a lower bound on the tolerance $\epsilon_i$, which, in turn, makes the feasible region larger, i.e., provides a trade-off between performance and implementability of the design.

## IV. APPLICATION EXAMPLES

Let us consider an example scenario, in which we have a network composed of $N = 2$ subsystems and show the effect of the minimum tolerance $\epsilon_{min}$ on the feasible region. For this, we will use the following parameters, the subsystem input and output $r^* = p^* = (9,\ 1)$ $[nM]$, the translation rate constant $\alpha_i = (2,\ 0.5)$ $[nM/hr]$, the decay rate constant for the protein is $\delta = 0.0770$ $[1/hr]$ and for the mRNA $\delta_0 = 0.0693$ $[1/hr]$. Moreover, we set the maximum tolerance $\epsilon_{max} = 0.3p^* = (2.7,\ 0.3)$ $[nM]$ and for the minimum tolerance we test two different values, the first $\epsilon_{min} = 0.1p^* = (0.9,\ 0.1)$ $[nM]$ and the second $\epsilon_{min} = 0.02p^* = (0.18,\ 0.02)$ $[nM]$.

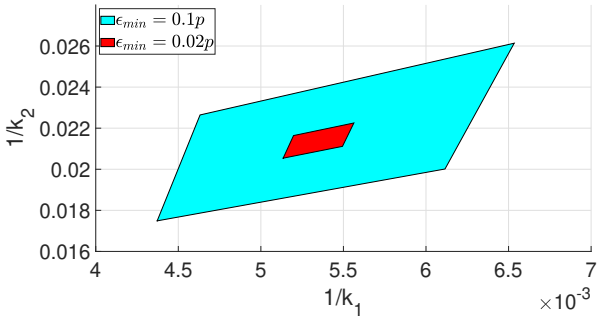


Fig. 4. Feasible region for $1/k_1$ and $1/k_2$ with different values of $\epsilon_{min}$.

Figure 4 presents the $(1/k_1, 1/k_2)$ feasible region for the two values of $\epsilon_{min}$. From the figure, we see that changing this variable affects the size of the feasible region, but not its shape. This occurs because as we increase $\epsilon_i$ we also decrease $\tilde{\gamma}_i$ and increase $\hat{\gamma}_i$. Decreasing $\tilde{\gamma}_i$ will make the $1/k_i$ coordinates of some of the vertices smaller (the ones closest to the origin in the $1/k_i$ axis). Increasing $\hat{\gamma}_i$ will make the $1/k_i$ coordinates of the remaining vertices larger (the ones furthest from the origin in the $1/k_i$ axis). Taken together, these result into the observed increase in the size of the feasible region. We conclude that $\epsilon_{min}$ is the parameter to be adjusted if the feasible region is too small.

Now we consider another example scenario, where we have a network composed of $N = 3$ subsystems and we wish to maintain the outputs of all subsystems around the same value of $p^* = (100,\ 100,\ 100)$ $[nM]$ with a minimum tolerance of $\epsilon_{min,i} = 0.2p_i^* = 20$ $[nM]$ and a maximum tolerance of $\epsilon_{max,i} = 0.3p_i^* = 30$ $[nM]$. Moreover, the inputs $r^* = (100,\ 100,\ 100)$ $[nM]$, the translation rate constant $\alpha_i = (43,\ 89,\ 62)$ $[nM/hr]$, the decay rate constant for the protein is $\delta = 0.0770$ $[1/hr]$ and for the mRNA $\delta_0 = 0.0693$ $[1/hr]$.

Solving the minimization problem, we obtain values for the tolerance $\epsilon_i$, and the gains $\tilde{\gamma}_i$ and $\hat{\gamma}_i$. Table I presents the values for these variables. Note that in this case the tolerance is the same as the minimum tolerance specified, that is, the feasible region we will obtain can be made smaller if the designer wishes and is able to implement the $k_i$ with greater precision.

| $i$ | 1 | 2 | 3 |
|---|---|---|---|
| $\epsilon_i$ $[nM]$ | 20 | 20 | 20 |
| $\tilde{\gamma}_i$ | 0.1673 | 0.0744 | 0.1103 |
| $\hat{\gamma}_i$ | 0.2738 | 0.1159 | 0.1752 |

TABLE I

$\epsilon_i$ TOLERANCES, $\tilde{\gamma}_i$ AND $\hat{\gamma}_i$ GAINS FOR THE CASE $N = 3$ SUBSYSTEMS EXAMPLE.

Furthermore, using (20) we can find the vertices of the $(1/k_1, 1/k_2, 1/k_3)$ feasible region shown in Table II and in Figure 5, where a plot of the $(1/k_1, 1/k_2, 1/k_3)$ feasible

region is displayed.

| $1/k_1$ | $1/k_2$ | $1/k_3$ |
|---|---|---|
| $1.44 \cdot 10^{-4}$ | $0.70 \cdot 10^{-4}$ | $1.00 \cdot 10^{-4}$ |
| $1.56 \cdot 10^{-4}$ | $0.75 \cdot 10^{-4}$ | $1.62 \cdot 10^{-4}$ |
| $1.52 \cdot 10^{-4}$ | $1.10 \cdot 10^{-4}$ | $1.05 \cdot 10^{-4}$ |
| $1.64 \cdot 10^{-4}$ | $1.19 \cdot 10^{-4}$ | $1.71 \cdot 10^{-4}$ |
| $2.42 \cdot 10^{-4}$ | $0.78 \cdot 10^{-4}$ | $1.12 \cdot 10^{-4}$ |
| $2.63 \cdot 10^{-4}$ | $0.85 \cdot 10^{-4}$ | $1.82 \cdot 10^{-4}$ |
| $2.56 \cdot 10^{-4}$ | $1.24 \cdot 10^{-4}$ | $1.18 \cdot 10^{-4}$ |
| $2.80 \cdot 10^{-4}$ | $1.35 \cdot 10^{-4}$ | $1.94 \cdot 10^{-4}$ |

TABLE II

$1/k$ FEASIBLE REGION VERTICES COORDINATES IN THE $N = 3$ EXAMPLE.

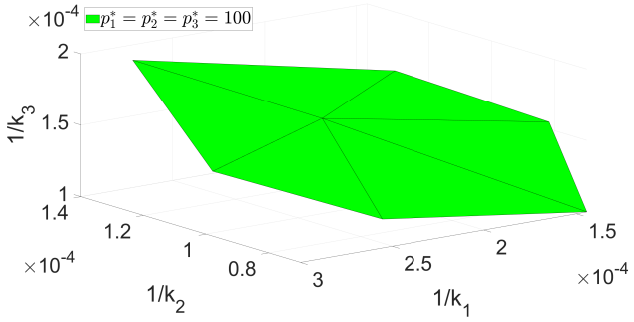


Fig. 5. Feasible region for $1/k_1$, $1/k_2$ and $1/k_3$.

Now we consider two modifications to this scenario, in the first one we want $\Sigma_3$ to increase its production to $p_3^* = 300$ $[nM]$ and in addition to this, in the second modification, we want $\Sigma_2$ to also increase its production to $p_2^* = 275$ $[nM]$. Moreover, we perform these modifications while maintaining all other parameters at their nominal values, but the tolerances $\epsilon_{min}$ and $\epsilon_{max}$ depend on the desired output values $p_i^*$, so the relationships remain the same, but the actual values change.

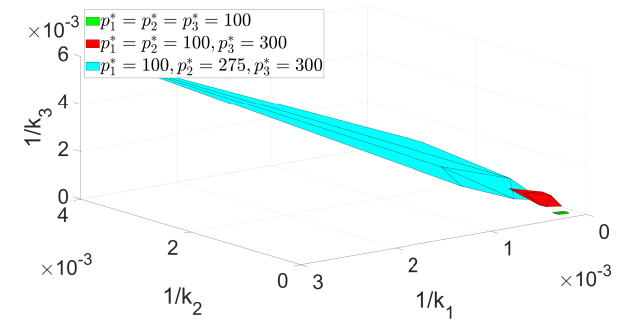


Fig. 6. Feasible region for $1/k_1$, $1/k_2$ and $1/k_3$ for different desired output values $p^*$.

Figure 6 shows the effects of the modifications to the desired output levels in the feasible region, where we can see that as we demand more protein production from the systems we stretch the feasible region. This is due to the increase in $p_i^*$, which makes $\tilde{\gamma}_i$ and $\hat{\gamma}_i$ also increase, causing an increase in the $1/k_i$ coordinates of the vertices. This is especially true for $\hat{\gamma}_i$, which sees the largest increase due to the fact that it depends on the sum of two variables that have increased in value $p_i^*$ and $\epsilon_i$.

## V. CONCLUSION AND FUTURE WORK

In this work, we introduced a systems composition formulation, in which resource sharing of genetic modules is explicitly captured through disturbance inputs and disturbance outputs. We used this framework to provide sufficient conditions to achieve desired subsystems' specifications despite resource sharing. Our approach is based on Theorem 1, which gives sufficient conditions for the existence of a feasible region for the dissociation constants $k_i$, as well as a procedure to find suitable tolerances $\epsilon_i$ for the output specification and the corresponding feasible region. An illustrative example and the analysis of a 3 subsystem case demonstrate how to tackle this problem with our approach.

Future work, will extend the approach to handle time-varying specifications and also systems interconnected by regulatory interactions. Additionally, we seek to incorporate other resource sharing effects, for example, due to sharing of degradation resources (proteases and micro RNAs) and to obtain similar conditions as those in Theorem 1.

## REFERENCES


[1] Cardinale, S., and Arkin, A. P., "Contextualizing context for synthetic biology–identifying causes of failure of synthetic biological systems", *Biotechnology journal*, 2012.

[2] Grunberg,T. W., and Del Vecchio, D., "Modular Analysis and Design of Biological Circuits". *Current Opinion in Biotechnology*, 2019.

[3] McBride, C. D., Grunberg, T. W., and Del Vecchio D., "Design of Genetic Circuits that are Robust to Resource Competition", *Current Opinion in Systems Biology*, 2021.

[4] Del Vecchio, D., Ninfa, A. J., and Sontag, E. D., "Modular cell biology: retroactivity and insulation", *Molecular systems biology*, 2008

[5] Jayanthi, S., Nilgiriwala, K. S., and Del Vecchio, D., "Retroactivity controls the temporal dynamics of gene transcription", *ACS synthetic biology*, 2013.

[6] Mishra, D., Rivera, P. M., Lin, A., Del Vecchio, D., and Weiss, R., "A load driver device for engineering modularity in biological networks", *Nature biotechnology*, 2014.

[7] Del Vecchio, D., "Modularity, context-dependence, and insulation in engineered biological circuits", *Trends in biotechnology*, 2015.

[8] Macdonald, L. E., Durbin, R. K., Dunn, J. J., and McAllister, W. T., "Characterization of two types of termination signal for bacteriophage T7 RNA polymerase" *Journal of molecular biology*, 1994.

[9] Rhodius, V. A., Mutalik, V. K., and Gross, C. A., "Predicting the strength of UP-elements and full-length E. coli $\sigma$E promoters", *Nucleic acids research*, 2012.

[10] Kosuri, S., Goodman, D. B., Cambray, G., Mutalik, V. K., Gao, Y., Arkin, A. P., Endy, D., and Church, G. M., "Composability of regulatory sequences controlling transcription and translation in Escherichia coli." *Proceedings of the National Academy of Sciences of the United States of America*, 2013.

[11] Del Vecchio, D., Qian, Y., Murray, R. M., and Sontag, E. D., "Future systems and control research in synthetic biology", *Annual Reviews in Control*, 2018.

[12] Qian, Y., Huang, H., Jiménez, J. I., and Del Vecchio, D., "Resource Competition Shapes the Response of Genetic Circuits", *ACS Synthetic Biology*, 2017.

[13] Gyorgy, A., Jiménez, J. I., Yazbek, J., Huang, H. H., Chung, H., Weiss, R., and Del Vecchio, D., "Isocost Lines Describe the Cellular Economy of Genetic Circuits", *Biophysical Journal*, 2015.

[14] Ceroni, F., Algar, R., Stan, G. B., and Ellis, T., "Quantifying cellular capacity identifies gene expression designs with reduced burden", *Nature Methods*, 2015.

[15] Qian, Y., and Del Vecchio, D., "Robustness of Networked Systems to Unintended Interactions With Application to Engineered Genetic Circuits", *IEEE Transactions on Control of Network Systems*, 2021.

[16] Huang, H., Qian, Y., and Del Vecchio, D., "A quasi-integral controller for adaptation of genetic modules to variable ribosome demand", *Nature Communications*, 2018.

[17] Shopera, T., He, L., Oyetunde, T., Tang, Y. J., and Moon, T. S., "Decoupling Resource-Coupled Gene Expression in Living Cells", *ACS synthetic biology*, 2017.

[18] Del Vecchio, D., and Murray, R. M., "Biomolecular Feedback Systems", *Princeton University Press*, 2014.

[19] Horn, R. A., and Johnson, C. R., "Topics in Matrix Analysis", *Cambridge University Press*, 1991.

[20] Löfberg, J., "YALMIP : A Toolbox for Modeling and Optimization in MATLAB", *In Proceedings of the CACSD Conference*, 2004.