This article has been accepted for publication in IEEE Control Systems Letters. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/LCSYS.2022.3186840

Identifying competition phenotypes in synthetic biochemical circuits

M. Ali Al-Radhawi[1], Domitilla Del Vecchio[2], and Eduardo D. Sontag[1]

***Abstract*— Synthetic gene circuits require cellular resources, which are often limited. This leads to competition for resources by different genes, which alter a synthetic genetic circuit's behavior. However, the manner in which competition impacts behavior depends on the identity of the "bottleneck" resource, which might be difficult to discern from input-output data. In this paper, we aim at classifying the mathematical structures of resource competition in biochemical circuits. We find that some competition structures can be distinguished by their response to different competitors or resource levels. Specifically, we show that some response curves are always linear, convex, or concave. Furthermore, high levels of certain resources protect the behavior from low competition, while others do not. We also show that competition phenotypes respond differently to various interventions. Such differences can be used to eliminate candidate competition mechanisms when constructing models based on given data. On the other hand, we show that different networks can display mathematically equivalent competition phenotypes.**

Keywords: **Resource competition, model discrimination, synthetic biology, system identification.**

I. Introduction

A. Background

LIVING cells have the ability to perform sophisticated operations that include maintaining homeostasis against noise, responding appropriately to various input signals, constructing complex structures such as proteins, and adapting to novel environments. Reverse engineering the biochemical circuits responsible for implementing such operations has revealed various control mechanisms that include regulation of gene expression via transcription factors (TFs) and/or non-coding RNAs [1], [2]. This has inspired the development of engineering approaches that mimic natural circuits by inserting new synthetic circuits into cells to modify their behavior or create new functionalities. Applications are wide-ranging and include immunotherapy, programmed micro-organisms for diagnostics and therapy, biofuel production, and many others [3], [4], [5]. Despite the great promise, numerous challenges exist. In particular, genetic circuits utilize common resources for transcription and translation such as RNA polymerases (RNAPs), ribosomes, tRNAs, and others. Insertion of new circuits increases the load on the cell's reservoirs. This, in turn, can create indirect interactions [6] that impede the proper functioning of the circuit, retard cellular growth, or lead to premature apoptosis [7], [8]. Several approaches have been proposed to ameliorate this problem, including dynamic control [9], orthogonal ribosomes and RNAPs [10], [11], [12], forward engineering of the circuit to account for resource competition [13], and distributed computation [14], [15].

Several of the aforementioned approaches assume that it is possible to identify the mode of competition and the limited resources responsible for performance degradation. However, it is not always possible to infer the correct model of competition from the expression data of the circuit. Possible competition effects to account for include promoters competing for RNAPs, mRNAs competing for ribosomes, transcription factors competing for promoters, enzymes competing for substrates, substrates competing for enzymes, etc.

In this work, we identify *competition phenotypes*, i.e., features that could allow one to distinguish the scarce resource (or resources) responsible for performance deterioration. Are there *qualitatively different* types of competition effects? Are there *equivalent* effects that can be treated in a unified manner? Answers to these questions will help guide theoretical analysis as well as the design of targeted interventions that mitigate undesirable effects through, for example, the use of feedback control to regulate the level of a scarce resource, or the optimization of appropriate circuit parameters. We will ask how different aspects of gene expression are impacted by two factors: (1) the level of the resource being shared – such as an activating or repressing transcription factor (TF), RNAPs, or ribosomes– and (2) the level of competition from other biochemical species –such as other genes or mRNAs. We describe interventions that can be used to distinguish the different competition phenotypes. We also discover instances where, conversely, competition for different resources might result in mathematically equivalent competition prototypes.

B. Problem Setup

Notation: Chemical species are denoted by non-italic large caps, time-dependent state variables are denoted by small caps, and constants are denoted by large cap italics. For example, a protein is denoted by Y, its time-dependent concentration is denoted by $y(t)$ while its steady-state value is denoted by Y.

1) The system: Consider a synthetic circuit with an external input U and internal state vector x which includes the concentrations of promoters, mRNAs, proteins, etc. The output

This work was supported by AFOSR 22RT0159 and NSF DMS-2052455. M.A.R and E.D.S are with the Departments of Electrical & Computer Engineering, and Bioengineering, Northeastern University, Boston, MA. D.D.V is with the Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA. Emails: {malirdwi,e.sontag}@northeastern.edu, ddv@mit.edu.

This article has been accepted for publication in IEEE Control Systems Letters. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/LCSYS.2022.3186840

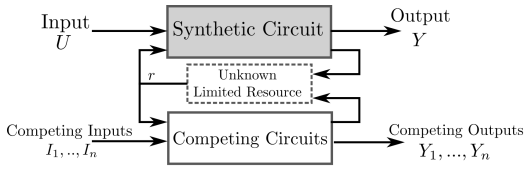


Fig. 1. Framework for studying the problem of competition model discrimination.

is denoted by y. The circuit utilizes the free limited resource r which is also utilized by other *competing* (or interfering) circuits with inputs $I_1,..,I_n$ and internal state vectors $z_1,..,z_n$. Figure 1 provides a pictorial representation of the system.

We model the entire system by a system of ordinary differential equations structured as follows:

$$\dot{x} = f(x, r, U), \tag{1}$$
$$\dot{z}_i = g_i(z_i, r, I_i),\ i = 1,..,n, \tag{2}$$
$$\dot{r} = \varphi(r, x, z_1,..,z_n), \tag{3}$$
$$y = h(x, r, U), \tag{4}$$

for some C^1 vector fields $f, g_1,..,g_n, \varphi$, and function h.

We assume that the limited resource is conserved. It partakes in the synthetic and competing circuits without being consumed or annihilated. Concretely, there exists nonnegative vectors $d_0,..d_n$ of compatible dimensions such that

$$r + d_0^T x + \textstyle\sum_{i=1}^n d_i^T z_i = R_T, \tag{5}$$

where R_T is the total resource, and r is the free (available) resource. In order for (1),(2),(3) to satisfy (5), the following relation is assumed to be satisfied: $\dot{r} + d_0^T \dot{x} + \sum_{i=1}^n d_i^T \dot{z}_i \equiv 0$.

In this paper, we perform our analysis at steady-state. We will assume that for each choice of the inputs $U, I_1,..,I_n$ and total resource R_T, there exists a steady-state $(X, Z_1,..,Z_n, R)$ that is globally asymptotically stable.

After eliminating the intermediate variables, the corresponding steady-state output Y can be written as:

$$Y = H(R_T, U, I_1,..,I_n), \tag{6}$$

for some function H.

2) Performance evaluation: The performance of a circuit with competition is compared to its performance with no competition. For this purpose, the *Competition-induced Performance Deterioration Ratio* (CPDR) ρ is defined as: $\rho := Y/Y|_{\text{competition}=0}$, where the total competition is $I := \sum_{i=1}^n I_i$.

3) Problem Formulation: The formulation in (1)-(4) implicitly assumes that r is the limited resource, while other resources are abundant and appear as kinetic constants in the functions $f, g_1,..,g_n, h$. Changing the identity of the limited resource will change the model (1)-(4). Our aim is to compare the qualitative differences in the steady-state input-output data that follow from the scarcity of different resources.

The paper is organized as follows. In section II, we discuss transcription/translation systems where the mRNA is protected by the ribosomes, while we drop that assumption in section III. We conclude with a discussion in section IV.

II. Transcription/Translation model: Ribosome protects mRNA

Consider a synthetic genetic circuit in which a promoter is inserted into a cell. The circuit should express a desired output protein Y. The input to the circuit is the total concentration of the input promoter U which is assumed to be constant; the unbound (free) promoter is denoted by U_f.

We first assume that mRNAs are protected from decay when bound to ribosomes, which is a valid assumption in many situations [16]. The transcription and translation reactions can be then written as follows:

$$\mathrm{U}_f + \mathrm{Q} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{E} \xrightarrow{\alpha_1} \mathrm{E} + \mathrm{M},\ \mathrm{M} \xrightarrow{\alpha_{-1}} \emptyset, \tag{7}$$

$$\mathrm{M} + \mathrm{R} \underset{w_{-1}}{\overset{w_1}{\rightleftharpoons}} \mathrm{F} \xrightarrow{\beta} \mathrm{F} + \mathrm{Y},\ \mathrm{Y} \xrightarrow{\beta_-} \emptyset, \tag{8}$$

where Q denotes RNAP, R denotes the ribosome, M denotes mRNA, E is the promoter-RNAP complex, and F is the mRNA-ribosome complex.

Free competing promoters $\mathrm{I}_{f,i}, i = 1,..,n$ can bind to RNAP and produce mRNAs M_i which can bind to the ribosome. This can be written as:

$$\mathrm{I}_{f,i} + \mathrm{Q} \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} \mathrm{E}_i \xrightarrow{\alpha_2} \mathrm{E}_i + \mathrm{M}_i,\ \mathrm{M}_i \xrightarrow{\alpha_{-2}} \emptyset, \tag{9}$$

$$\mathrm{M}_i + \mathrm{R} \underset{w_{-2}}{\overset{w_2}{\rightleftharpoons}} \mathrm{F}_i \xrightarrow{\beta} \mathrm{F}_i + \mathrm{Y}_i,\ \mathrm{Y}_i \xrightarrow{\beta_-} \emptyset, i = 1,..,n. \tag{10}$$

The promoters are conserved. Hence, we have:

$$U_f + E = U, \tag{11}$$
$$I_{f,i} + E_i = I_i, i = 1,..,n. \tag{12}$$

Let $K_j = k_{-j}/k_j, A_j = \alpha_j/\alpha_{-j}, W_j = w_{-j}/w_j, B = \beta/\beta_-,\ j = 1,2$ be RNAP Dissociation Ratio (DR), transcription ratio, ribosome DR, and protein expression ratio, respectively. By writing the irreversible reactions as $\emptyset \underset{\alpha_{-1}}{\overset{\alpha_1 E}{\rightleftharpoons}} \mathrm{M}$, $\emptyset \underset{\beta_-}{\overset{\beta F}{\rightleftharpoons}} \mathrm{Y}$, $\emptyset \underset{\alpha_{-2}}{\overset{\alpha_2 E_i}{\rightleftharpoons}} \mathrm{M}_i$, $\emptyset \underset{\beta_-}{\overset{\beta F_i}{\rightleftharpoons}} \mathrm{Y}_i$, the steady-states of the network (7)-(8),(9)-(10) can be computed by noting that the network is detailed-balanced, i.e. the forward and backward rates in each reaction are equal at steady state. Hence, the steady state values of the mRNAs M, M_i are determined by the equilibrium values $A_1 E, A_2 E_i, i = 1,..,n$, respectively, and are independent of the translation process. We get the following steady-state expressions:

$$E = \tfrac{QU}{K_1+Q}, E_i = \tfrac{QI_i}{K_2+Q}, M = A_1 E, M_i = A_2 E, \tag{13}$$
$$F = \tfrac{A_1}{W_1}\tfrac{RQU}{K_1+Q}, F_i = \tfrac{A_2}{W_2}\tfrac{RQI_i}{K_2+Q}, \tag{14}$$
$$Y = \tfrac{A_1 B}{W_1}\tfrac{RQU}{K_1+Q}, Y_i = \tfrac{A_2 B}{W_2}\tfrac{RQI_i}{K_2+Q}. \tag{15}$$

Note that Q, R are the levels of the *free* RNAP and ribosome, respectively.

In the formalism defined in §1.B and depending on the identity of the limited resource, Eq. (1) describes the dynamics of (7)-(8), while Eq. (2) describes the dynamics of (9)- (10). In order to write the output in the form (6), we will study several scenarios in which either the ribosome is limited, or RNAP is limited. We keep the other resource abundant to isolate the effects of the limitations in a single resource.

Remark 1: A slightly different mathematical model of translation (8) would have the mRNA/ribosome complex dis-

This article has been accepted for publication in IEEE Control Systems Letters. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/LCSYS.2022.3186840

sociating directly upon protein production, and would be as follows: $\mathrm{M}+\mathrm{R} \underset{\tilde{w}_-}{\overset{w}{\rightleftharpoons}} \mathrm{F} \xrightarrow{\beta} \mathrm{M}+\mathrm{R}+\mathrm{Y}$, $\mathrm{Y} \xrightarrow{\beta_-} \emptyset$. Nevertheless, the basic competition phenotype remains the same as can be seen by defining $w_- := \tilde{w}_- + \beta$ in (8).

Remark 2: An alternative way to regulate a target circuit is to use a small molecule such as AHL to activate the target promoter [7]. The RNAP binding reactions can be written as follows (compare to (7)): $\mathrm{U}_f^* \underset{\mu_-}{\overset{\mu}{\rightleftharpoons}} \mathrm{U}_f, \mathrm{U}_f + \mathrm{Q} \underset{\tilde{k}_{-1}}{\overset{\tilde{k}_1}{\rightleftharpoons}} \mathrm{E}$, where U_f^* is the free inactive promoter, and μ is proportional to the exogenous input. Our model encompasses this case also since it can be shown that the reparameterization $K_1 = (\tilde{k}_{-1}/\tilde{k}_1)(1+\mu_-/\mu)$ recovers the equations (13)-(15). In other words, the effective effect of the external input in our model is to regulate the RNAP dissociation ratio K_1. Note that AHL concentrations can be precisely regulated over four orders of magnitude [8]. Therefore, employing promoters inducible by AHL will not have a considerable effect on the dynamic range of K_1.

Parameter Ranges: In order to keep the numbers within biological ranges, we use the following parameters: $K_i \in [0.3, 10000]$ nM, $W_i \in [5, 2000]$ μM, $I \in [1, 1000]$ nM, $U \in [1, 1000]$ nM, $i = 1, 2$ [2], [7], [8]. Furthermore, whenever the following parameters are not treated as variables, we use the following numbers as in [7]: $Q_T = 500$ nM, $R_T = 1000$ nM, $A_i = 10$, $B = 300$, $i = 1, 2$.

A. Limited Ribosomes and Abundant RNAP (LRAP)

In this case, RNAP Q is unaffected by the circuit (7)-(8),(9)-(10), hence its level Q will be constant, $Q = Q_T$.

Let $F_I = \sum_{i=1}^{n} F_i$, then the conservation law (5) is

$$R + F + F_I = R_T. \tag{16}$$

Solving the resulting algebraic equations, the free resource at steady state is found to be:

$$R = \frac{R_T}{1+\gamma_1 U+\gamma_2 \sum_i I_i} = \frac{R_T}{1+\gamma_1 U+\gamma_2 I},$$

where $\gamma_j := (A_j Q_T)/(W_j(K_j + Q_T)), j = 1, 2$. The output is

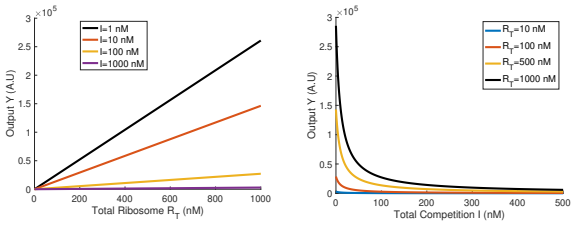
$$Y = \frac{RB}{W_1} M = B\gamma_1 \frac{UR_T}{1 + \gamma_1 U + \gamma_2 I}.$$

The output is linearly dependent on R_T, and takes an inhibiting Michaelis-Menten form with respect to the competition I. It also takes a Michaelis-Menten form with respect to U. The CPDR ρ is given as: $\rho = (1+\gamma_1 U)/(1+\gamma_1 U+\gamma_2 I)$, which is *independent* of the total resource R_T.

We next study the properties of the output as it depends on the input and the competition.

As a function of the resource R_T, the output is linear as noted above, but is concave with respect to U. As a function of the competition, it is convex. Note $d^2Y/dI^2 > 0$ for all $I>0$. Figure 2 depicts the typical behavior of the competition phenotype associated with this mode of competition.

Remark 3: The fact that the CPDR is independent of the total resource might lead one to think that the total resource is irrelevant for reducing competition. However, this depends on how we define competition reduction. If we consider strategies



(a) (b)

Fig. 2. **The limited ribosome case (LRAP)**. (a) The output versus total ribosome for various competition levels, (b) The output versus total competition for various total ribosome levels. For both panels we have $\boldsymbol{K_1 = K_2 = 1nM, W_1 = W_2 = 5\mu M, U = 10nM}$.

to reinstate the level of the output Y to its competition-free level $Y|_{I=0}$, then we can increase the total resource to compensate for the reduction in the output due to competition. In particular, we can write the required total resource as $R_T^* = \frac{1+\gamma_1 U+\gamma_2 I}{B\gamma_1 U} Y|_{I=0}$.

B. Limited RNAPs and abundant ribosomes (LPAR)

In this case, the ribosome R will be unaffected by the circuit (7)-(8),(9)-(10), hence its level R will be assumed to be constant. In other words, we have $R = R_T$. Therefore, we write the conservation law (5) as:

$$Q + E + \textstyle\sum_i E_i = Q_T. \tag{17}$$

In the most general case, solving (17) for Q requires solving a cubic equation. Therefore, in this subsection, we assume that $K_1 = K_2 = K$ to simplify the analysis. This can be justified in cases when the gene and its competitors are located on the same plasmid. For instance, the RNAP DRs are taken to be $K_1 = K_2 = 200$ nM when simulating the data in [7]. In this case, the free RNAP Q and the output Y are

$$Q = \tfrac{1}{2}(Q_T - I - U - K) + \tfrac{1}{2}\sqrt{(K + U + I - Q_T)^2 + 4KQ_T},$$

$$Y = \frac{2Q_T R_T U A_1 B/W_1}{K + Q_T + I + U + \sqrt{(Q_T - K - U - I)^2 + 4KQ_T}}.$$

As Q_T grows without bound, we have $\lim_{Q_T \to \infty} Y = \frac{A_1 B}{W_1} U R_T$, i.e., the protein is expressed at maximum capacity. If I grows without bound, then $\lim_{I \to \infty} Y = 0$. The competition reduction ratio is written as follows: $\rho = \frac{K+Q_T+U+\sqrt{(Q_T-K-U)^2+4KQ_T}}{K+Q_T+I+U+\sqrt{(Q_T-K-U-I)^2+4KQ_T}}$ which depends on the total resource unlike the case in §2.1.

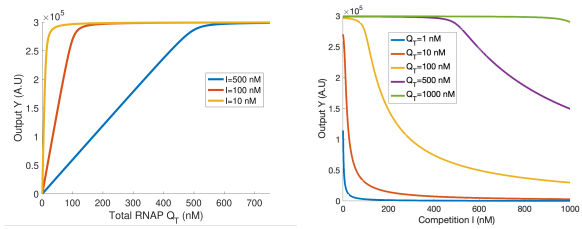
Convexity of the output as a function of the resource: The output is globally concave, because:

$$\frac{d^2Y}{dQ_T^2} = \frac{-2KA_1 BR_T U/W_1}{((Q_T - K - U - I)^2 + 4KQ_T)^{3/2}} < 0, \tag{18}$$

for all $Q_T \geq 0$. The typical competition phenotypes are plotted in Figure 3-a,b.

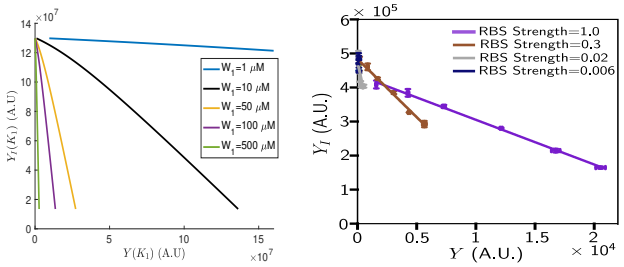
Convexity of the output as a function of the competition : Simulations show that the response is globally convex when Q_T is small. For larger Q_T, the response starts concave and then it has an inflection point. Using the same parameters above, when $Q_T = 1$ the response is convex at zero as verified by computing the second derivative of Y with respect to I at zero. Figure 3- b shows the transition from convexity to concavity with higher Q_T.

This article has been accepted for publication in IEEE Control Systems Letters. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/LCSYS.2022.3186840



(a) (b)

Fig. 3. **Limited RNAP case**. (a) The output versus total RNAP for various competition levels, (b) The output versus competition with various $\boldsymbol{Q_T}$. For both panels we have $\boldsymbol{K_1{=}K_2{=}1nM, W_1{=}W_2{=}10\mu M, U{=}10nM}$.



(a) (b)

Fig. 4. **Limited RNAP manifests as sensitivity to the ribosome DR $\boldsymbol{W_1}$ when parameterizing the outputs in terms of the RNAP DR $\boldsymbol{K_1}$.** (a) Simulated $(\boldsymbol{Y(K_1), Y_I(K_1)})$ parametric curves. The plots are computed by solving the cubic equation (17) numerically. The parameters are: $\boldsymbol{U = 500\ nM, I = 500\ nM, W_1 = 500\ \mu M, A_1 = A_2 = 10, K_2 = 10\ nM, W_2 = 10\mu M, K_1 \in [0.1, 10000]nM}$. (b) Data adapted from [7] where the curves are parameterized via the AHL concentration which corresponds to $\boldsymbol{K_1}$ in our model (see Remark 2), and by varying the RBS strengths which are inversely proportional to the phenomenological parameter $\boldsymbol{W_1}$ in our model.

C. Distinguishing the competition phenotypes

a) Convexity/concavity: We have been able to prove that, at steady state, limited ribosome and limited RNAP result in *qualitatively distinct competition phenotypes*. For the second, the output takes a Michaelis-Menten form and is globally convex (18) with respect to the resource (RNAP) level as shown in Figure 3-a, but, in contrast, for the first it is perfectly linear with respect to the resource (ribosome level) regardless if it is abundant or limited; see Fig. 2- a. Furthermore, a high level of RNAP provides *buffering* against low levels of competition in the second case as shown in Figure 3-b, while the output drops quickly even with low competition in the first case; see Fig. 2b. We are able to characterize this mathematically by proving that the output is initially concave (i.e., superlinear) with respect to competition mediated by limited RNAP, while it is globally convex (i.e., sublinear) with limited ribosomes. Thus, using *either* criterion (titrating resource, or titrating competitors), we can see a clear difference between these two types of context limitations. For instance, the level of competition can be controlled by adjusting the dosage of an inserted plasmid [7]. Hence, our criteria can be checked visually from the plot of the output versus the competitor dosage. From a different point of view, theory helps us *guess the source of competition* based upon experimental data.

b) Effect of the RNAP dissociation ratio: The effective RNAP DR K_1 can be modified by using an inducible promoter (see Remark 2). Keeping the remaining parameters fixed, we may think of the outputs Y and $Y_I := \sum_{i=1}^{n} Y_i$ as functions of K_1. Let us now analyze the trade-off between the two outputs when varying the parameter K_1 while keeping the resource levels constant. The resulting parametric curves are also known as *isocost curves* in the language of economics [7]. Such relations can be derived by solving Eq. (16) for K_1 in terms of Y, and then substituting it in the expression of Y_I.

In the case of LRAP (i.e., $Q = Q_T$), it can be shown that the relationship between the parameterized outputs is *linear* and is given by the following parametric equation:

$$Y_I(K_1) = \frac{BA_2IQ_T}{K_2W_2+A_2BIQ_T}(R_T - Y(K_1)). \qquad (19)$$

Note that (19) is *independent* of the ribosome DR W_1.

In comparison, the case of LPAR (i.e., $R = R_T$) is more complicated. The corresponding relationship between $Y_I(K_1)$ and $Y(K_1)$ is generally *nonlinear* and its computation requires solving a cubic equation, as noted before. In order to probe the effect of K_1, we use the simplifying assumption $Q \ll K_1, K_2$ which is often satisfied by practical systems [7]. Under this approximation, it can be shown again that the relationship between $Y(K_1)$ and $Y_I(K_1)$ is linear and is given by the parametric equation:

$$Y_I(K_1) = \frac{A_2I(R_TQ_TA_1B - W_1Y(K_1))}{A_1W_2(I+K_2)}. \qquad (20)$$

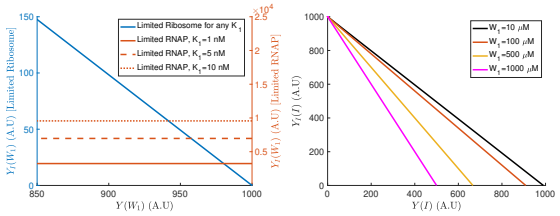
In this case, the relationship depends on W_1. Therefore, the two modes of competition (for RNAP or ribosomes) are in principle distinguishable by modifying W_1 via variable Ribosome Binding Site (RBS) strengths. Note that this conclusion stands even without the assumption made earlier.

Even though the linear approximation (20) may not hold in all situations, dependence on W_1 in the K_1-parameterized relationship between Y and Y_I indicates limited RNAP as seen in Figure 4-a) which depicts a simulation of the parameterized curves for various values of W_1. In comparison, Eq. (19) is linear and is independent of W_1. Figure 4-b) shows experimental data from [7] that depicts the same scenario. Our theoretical and computational prediction is consistent with the slope change noticed in the experimental data. We derived our conclusions assuming that RNAP is limited and that ribosomes are abundant. In practical situations, both resources might be in limited supplies [7].

c) Effect of the Ribosome Dissociation Ratio: Let us now consider modifying W_1, instead of K_1. We can again write the outputs as $Y(W_1), Y_I(W_1)$. In the case of LRAP, the relationship is again linear and it can be written in the same form as (19). It can be seen that it is independent of the RNAP DR K_1. In the case of LPAR, under the assumption $Q \ll K_1, K_2$, we get: $Y_I(W_1) = \frac{A_2BIK_1Q_TR_T}{W_2(K_2U+K_1I+K_1K_2)}$, which depends on the RNAP DR K_1 but is *constant*. This conclusion holds without the assumption above since the free RNAP Q which solves (17) is independent of ribosome DR W_1. Hence, the expressions $Y(W_1), Y_I(W_1)$ are not related.

Figure 5-a) shows that the case of limited ribosomes manifests as a decreasing linear relationship when parameterized by W_1 (which can be experimentally controlled by RBS strengths), but it is independent of K_1 (which can be controlled by utilizing an inducible promoter). On the other hand, the

This article has been accepted for publication in IEEE Control Systems Letters. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/LCSYS.2022.3186840



(a) (b)

Fig. 5. **Various interventions can be used to distinguish between limited RNAP versus limited Ribosome cases** (a) Simulated $(\boldsymbol{Y(W_1), Y_I(W_1)})$ curves for various $\boldsymbol{K_1}$ levels. The parameters are: $\boldsymbol{U = 500\ nM, I = 1\ nM, K_2 = 10\ nM, W_2 = 2000\mu M, W_1 \in [5, 3000]\mu M}$. The curves for the limited RNAP case are computed by solving the cubic equation (17) numerically. (b) Simulated $(\boldsymbol{Y(I), Y_I(I)})$ curves for various $\boldsymbol{W_1}$ levels for the case of LRAP. The parameters are $\boldsymbol{U = 1\ nM, K_1 = K_2 = 1000nM, W_2 = 10\mu M}$.

	LRAP	LPAR
Resource Convexity	globally linear	globally concave
Competition Convexity	globally convex	globally convex (low resource) concave then convex (high resource)
Y_I vs Y parameterized by RNAP DR (K_1)	Linear. Independent of W_1	Nonlinear. Depends on W_1
Y_I vs Y parameterized by Ribosome DR (W_1)	Linear. Independent of K_1	Constant. Depends on K_1
Y_I vs Y parameterized by competitor copy number (I)	Linear. Slope depends on W_1	Nonlinear. Both slope & intercept depend on W_1

Table I: Comparison of the competition phenotypes discussed in §2.

same relationship is constant in the case of limited RNAP, but the level is dependent on K_1. This manifests by examining the Y_I-axis intercepts, which depend on K_1 in the case of LPAR, but are independent of it in the case of LRAP.

d) Effect of the total copy number of the competitor: We consider next modifying I while keeping U fixed. We write the outputs as $Y(I), Y_I(I)$.

In the case of LRAP, we get the *linear* relationship: $Y_I(I) = R_T - \left(\frac{W_1(K_1+Q_T)}{A_1BQ_TU} + 1\right) Y(I)$, where the Y_I-axis intercept depends only on R_T.

In the case of LPAR, the relationship is generally nonlinear, but under the assumption $Q \ll K_1, K_2$ we get the following linear relationship: $Y_I(I) = \frac{1}{W_1} A_2BQ_TR_T - \left(\frac{A_2W_1(K_1+U)}{A_1UW_2}\right) Y(I)$, where the Y_I-axis intercept depends on the ribosome DR W_1. Figure 5-b) depicts the parametric curves corresponding to LRAP case where it can be seen that the Y_I-axis intercept is constant for various W_1.

e) Summary: The results are summarized in Table I.

III. Transcription/Translation model: Ribosome does not protect mRNA

In many situations, ribosomes do not protect mRNAs completely from decay [16]. To model this, we modify (7)-(8),(9)-(10) by adding the following reactions which describe the decay of mRNA while bound to the ribosome:

$$\mathrm{F} \xrightarrow{\alpha^*_{-1}} \mathrm{R},\ \mathrm{F}_i \xrightarrow{\alpha^*_{-2}} \mathrm{R}, i = 1, .., n. \tag{21}$$

The network is no longer detailed balanced after the inclusion of decay of the bound mRNA-ribosome complexes. The steady-values of M, M_i are now dependent on the resource R. But we still have the promoter-RNAP complex expressed at steady-state as:

$$E = \frac{QU}{K_1+Q}, E_i = \frac{QI_i}{K_2+Q}. \tag{22}$$

Let $a_j = \alpha^*_{-j}/\alpha_{-j}$ be the ratio of the decay rates of the bound and unbound mRNA. The steady-state values of the mRNAs are given as: $M = A_1E/(1 + Ra_1/W_1), M_i = A_2E_i/(1 + Ra_2/W_2)$, where $W_j = (\alpha^*_{-j} + w_{-j})/w_j, j = 1, 2$. The steady-values of the mRNA-ribosome complexes and the output in terms of the free resource R are given as: $F = \frac{A_1RQE}{W_1+a_1R}, F_i = \frac{A_2RQE_i}{W_2+a_2R}, Y = \frac{AB}{W}\frac{RQU}{K_1+Q}\left(\frac{1}{1+Ra_1/W_1}\right)$. Observe that when $\alpha^*_{-1} = \alpha^*_{-2} = 0$, we recover the case discussed in §2. We study different scenarios next.

A. Limited ribosomes and abundant RNAP

As before, we have abundant RNAP, hence $Q = Q_T$. The only conservation law is (16). Therefore, we need to solve for the free ribosome R. Let E be as given in (22), and let $E_I = \sum_i E_i$ where E_i is defined in (22). The conservation law (16) results in a cubic equation. Therefore, we assume that $A = A_1 = A_2, W = W_1 = W_2, K = K_1{=}K_2, a = a_1 = a_2$ in order to simplify the analysis. We get (23). The output can be written as in (24). Note that when $a_1 = a_2 = 0$, we get the case discussed in §2.1. Next, we show that the properties of the system above can be deduced by studying a different system that has been studied earlier.

B. Different systems have the same phenotype

It is perhaps surprising, that *two very different biological systems may lead to mathematically identical competition phenotypes*. To formalize this, let us write the (6) as: $Y = H(V; \Lambda)$ where the inputs are $V = [U, R_T, I_1, .., I_n]^T$, and Λ are a set of parameters (kinetic rates, for example) that appear in (1)-(4). Hence, we have a steady-state expression $H_1(V_1; \Lambda_1)$ that gives us the amount of output, as well as a second function $H_2(V_2; \Lambda_2)$ for a different system; equivalent phenotypes will have the property that there is a diffeomorphism $\psi : (V_1, \Lambda_1) \mapsto (V_2, \Lambda_2)$ so that every function $H_2(V_2; \Lambda_2)$ can be written as $H_1(\psi(V_1, \Lambda_1))$.

As a concrete example, let us revisit scenarios §2.2 and §3.1 discussed earlier with $A_1 = A_2 = A$. For the system in §3.1, we consider the case in which the mRNA decays at the same rate, whether it is bound to the ribosome or not, i.e., $a_1 = a_2 = 1$. One can prove that the two systems are equivalent, under a reparameterization given as follows: $k_-/k \mapsto (w_- + \alpha_-)/w$, $Q_T \mapsto R_T$, $U \mapsto \alpha UQ_T/(\alpha_-(K + Q_T))$, $BAUR_T/(W + R_T) \mapsto B$, $I \mapsto \alpha IQ_T/(\alpha_-(K + Q_T))$, where $I := \sum_i I_i$. Note that the underlying biochemical systems are very different, and the two systems of ODEs are distinct. In fact, they result in different transient behavior. However, the steady-states are the same, as shown theoretically and illustrated in Fig. 6.

C. Abundant ribosome and limited RNAP

Similar to the previous section, we let $R = R_T$. Hence, the only conservation law is (17). Since decay affects only the

This article has been accepted for publication in IEEE Control Systems Letters. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/LCSYS.2022.3186840

$$R = (2R_T W) \Big/ \left(W + A(E + E_I) - aR_T + \sqrt{(W + A(E + E_I) - aR_T)^2 + 4aR_T W} \right). \tag{23}$$

$$Y = (2R_T BAE) \Big/ \left(W + A(E + E_I) + (2 - a)R_T + \sqrt{(W + A(E + E_I) - aR_T)^2 + 4aR_T W} \right). \tag{24}$$

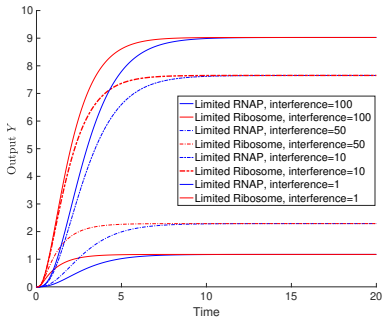


Fig. 6. Two different competition systems with different transient are identical at steady state.

ribosome-mRNA complex, we can see immediately that this case is very similar to the case discussed in §2.2 except for an additional factor. Hence, similar analysis can be replicated.

IV. Discussion

1) Generalization: The competition phenotypes studied in the previous sections can be generalized to other biological contexts by classifying them into two main categories: *externally-regulated targets*, and *conserved targets*. Ribosome competition studied in §2.1 can be studied under the first category. This is since the target which consumes the resource is the mRNA which is not conserved. Hence, the model can be essentially written as follows:

$$\text{(Circuit)} \quad \emptyset \underset{\alpha_-}{\overset{\alpha U}{\rightleftharpoons}} \mathrm{M}, \quad \mathrm{R} + \mathrm{M} \underset{k_-}{\overset{k}{\rightleftharpoons}} \mathrm{Y}$$

$$\text{(Competition)} \quad \emptyset \underset{\alpha_-}{\overset{\alpha I_i}{\rightleftharpoons}} \mathrm{M}_i, \ \mathrm{R} + \mathrm{M}_i \underset{k_-}{\overset{k}{\rightleftharpoons}} \mathrm{Y}_i, i = 1, .., n.$$

In addition to mRNAs, the above model can represent ligand-activated enzymes $\mathrm{M}, \mathrm{M}_i, i = 1, .., n$ competing for a substrate R, single-guide RNAs (sgRNAs) competing for a limited amount of dCas9 in CRISPRi [17], [18], or externally-activated TFs competing for a single promoter.

On the other hand, the case discussed in §2.2 has the RNAP as a limited resource. Hence, the target that consumes the resource is the promoter which is conserved. Hence, the model can be written essentially as follows:

$$\text{(Circuit)} \qquad \mathrm{R} + \mathrm{M} \underset{k_-}{\overset{k}{\rightleftharpoons}} \mathrm{Y},$$

$$\text{(Competition)} \qquad \mathrm{R} + \mathrm{M}_i \underset{k_-}{\overset{k}{\rightleftharpoons}} \mathrm{Y}_i, i = 1, .., n.$$

In addition to promoter, the above model can represent conserved enzymes $\mathrm{M}, \mathrm{M}_i, i = 1, .., n$ competing for a substrate R, or conserved TFs competing for a single promoter.

2) Limitations: Our framework has included multiple simplifications to allow analytical derivations and facilitate clear interpretations. For instance, we assumed that the competitors behave similarly to each other. This can be justified when the target gene and its competitors are co-located on the same plasmid with high copy numbers, and with their own orthogonal RNAPs and ribosomes to minimize interactions with the rest of the genome [10], [7], [11]. Needless to say, this does not always hold. In addition, we have assumed a fixed number of competitors and a constant amount of allocated resources. However, the number of active pathways and the amount of allocated resources in a cell change dynamically depending on stress and growth conditions [19]. Studying such scenarios is subject to future work.

References

[1] U. Alon, *An introduction to systems biology: design principles of biological circuits*. Chapman and Hall/CRC, 2006.
[2] D. Del Vecchio and R. M. Murray, *Biomolecular feedback systems*. Princeton University Press, 2014.
[3] A. S. Khalil and J. J. Collins, "Synthetic biology: applications come of age," *Nature Reviews Genetics*, vol. 11, no. 5, pp. 367–379, 2010.
[4] P.-F. Xia, H. Ling, J. L. Foo, and M. W. Chang, "Synthetic genetic circuits for programmable biological functionalities," *Biotechnology Advances*, vol. 37, no. 6, p. 107393, 2019.
[5] X. Tan, J. H. Letendre, J. J. Collins, and W. W. Wong, "Synthetic biology in the clinic: engineering vaccines, diagnostics, and therapeutics," *Cell*, vol. 184, no. 4, pp. 881–898, 2021.
[6] S. Prabakaran, J. Gunawardena, and E. Sontag, "Paradoxical results in perturbation-based signaling network reconstruction," *Biophysical Journal*, vol. 106, pp. 2720–2728, 2014.
[7] A. Gyorgy, J. I. Jiménez, J. Yazbek, H.-H. Huang, H. Chung, R. Weiss, and D. Del Vecchio, "Isocost lines describe the cellular economy of genetic circuits," *Biophysical Journal*, vol. 109, no. 3, pp. 639–646, 2015.
[8] Y. Qian, H.-H. Huang, J. I. Jiménez, and D. Del Vecchio, "Resource competition shapes the response of genetic circuits," *ACS Synthetic Biology*, vol. 6, no. 7, pp. 1263–1272, 2017.
[9] H.-H. Huang, Y. Qian, and D. Del Vecchio, "A quasi-integral controller for adaptation of genetic modules to variable ribosome demand," *Nature communications*, vol. 9, no. 1, pp. 1–12, 2018.
[10] V. Tunitskaya and S. Kochetkov, "Structural–functional analysis of bacteriophage T7 RNA polymerase," *Biochemistry (Moscow)*, vol. 67, no. 10, pp. 1124–1135, 2002.
[11] A. P. Darlington, J. Kim, J. I. Jiménez, and D. G. Bates, "Dynamic allocation of orthogonal ribosomes facilitates uncoupling of co-expressed genes," *Nature Communications*, vol. 9, no. 1, pp. 1–12, 2018.
[12] J. Miller, M. A. Al-Radhawi, and E. D. Sontag, "Mediating ribosomal competition by splitting pools," *IEEE Control Systems Letters*, vol. 5, no. 5, pp. 1555–1560, 2021.
[13] C. D. McBride and D. Del Vecchio, "Predicting composition of genetic circuits with resource competition: demand and sensitivity," *ACS Synthetic Biology*, vol. 10, no. 12, pp. 3330–3342, 2021.
[14] M. A. Al-Radhawi, A. P. Tran, E. A. Ernst, T. Chen, C. A. Voigt, and E. D. Sontag, "Distributed implementation of boolean functions by transcriptional synthetic circuits," *ACS Synthetic Biology*, vol. 9, no. 8, pp. 2172–2187, 2020.
[15] T. Chen, M. A. Al-Radhawi, C. A. Voigt, and E. D. Sontag, "A synthetic distributed genetic multi-bit counter," *IScience*, vol. 24, no. 12, p. 103526, 2021.
[16] A. Deana and J. G. Belasco, "Lost in translation: the influence of ribosomes on bacterial mRNA decay," *Genes & Development*, vol. 19, no. 21, pp. 2526–2533, 2005.
[17] S. Zhang and C. A. Voigt, "Engineered dCas9 with reduced toxicity in bacteria: implications for genetic circuit design," *Nucleic Acids Research*, vol. 46, no. 20, pp. 11 115–11 125, 2018.
[18] P.-Y. Chen, Y. Qian, and D. Del Vecchio, "A model for resource competition in CRISPR-mediated gene repression," in *2018 IEEE Conference on Decision and Control (CDC)*, 2018, pp. 4333–4338.
[19] R. Balakrishnan, R. T. de Silva, T. Hwa, and J. Cremer, "Suboptimal resource allocation in changing environments constrains response and growth in bacteria," *Molecular systems biology*, vol. 17, no. 12, p. e10597, 2021.