Limitations and Trade-offs in Gene Expression due to Competition for Shared Cellular Resources

Andras Gyorgy and Domitilla Del Vecchio

***Abstract*— Gene circuits share transcriptional and translational resources in the cell. The fact that these common resources are available only in limited amounts leads to unexpected couplings in protein expressions. As a result, our predictive ability of describing the behavior of gene circuits is limited. In this paper, we consider the simultaneous expression of proteins and describe the coupling among protein concentrations due to competition for RNA polymerase and ribosomes. In particular, we identify the limitations and trade-offs in gene expression by characterizing the attainable combinations of protein concentrations. We further present two application examples of our results: we show that even in the absence of regulatory linkages, genes can seemingly behave as repressors, and surprisingly, as activators to each other, purely due to the limited availability of shared cellular resources.**

I. INTRODUCTION

One of the major bottlenecks in systems and synthetic biology is context-dependence [1], as it hinders our ability to accurately predict the behavior of complex systems from that of the composing modules [2]. This lack of modularity is particularly important when engineering biological systems using smaller components [3], as it often leads to a lengthy and *ad hoc* re-design process every time the context changes [4]. Context-dependence arises due to a number of different factors: unknown regulatory linkages; loading effects due to known regulatory interactions between components, a phenomenon known as retroactivity [5], [6]; metabolic burden [7]; cell growth [8]; and competition for shared cellular resources [9].

In this paper, we focus on the effects of competition for transcriptional and translational resources on gene expression. Since these resources are available only in limited amounts, they have to be reallocated every time new genes are introduced into the cell, or when the activity of already present genes changes. Due to the reallocation of these common resources, the over-expression of one gene can affect the growth rate of the cell [8], and it can decrease the expression of other genes [10]. As a result, the expression of different genes become coupled, even in the absence of regulatory linkages among them. To accurately predict and control the behavior of gene circuits, we must determine the distribution of shared resources, that is, the cellular economy of gene expression.

This work was supported by AFOSR grant FA9550-12-1-0129 and NIGMS grant P50 GM098792.

A. Gyorgy is with the Department of Electrical Engineering and Computer Science, MIT, Cambridge, MA 02139, USA gyorgy@mit.edu

D. Del Vecchio is with the Department of Mechanical Engineering, MIT, Cambridge, MA 02139, USA ddv@mit.edu

Here, we characterize how the expression of different genes become coupled due to competition for RNA polymerase (RNAP) and ribosomes. We focus on RNAP and ribosomes as their availability is considered to be the major limiting factor in transcription [11] and translation [10], respectively. We prove that due to the limited availability of these cellular resources, the attainable protein concentrations lie within the intersection of simplexes, and we show how these simplexes depend on various biochemical parameters, such as ribosome binding site (RBS) strength and DNA copy number. Building upon our results, we show that even in the absence of regulatory linkages, genes can seemingly repress and activate each other, as a result of the reallocation of limited resources. In particular, we first consider two genes and show that activating one decreases the expression of the other. We further demonstrate that this effect can be interpreted employing isocost lines, a tool originating from microeconomics to describe the affordable combinations of different products having only a limited budget. Second, in the case of three genes, we show that increasing the production of one protein can surprisingly increase the concentration of a second one by reallocating resources from the expression of the third protein.

Our work is closely related to recent efforts investigating the effects of shared cellular resources on gene circuits. In particular, in [12] and [13] the authors detail the effects of the limited availability of ribosomes causing translational crosstalk, a phenomenon verified experimentally in [14] in cell-free systems. A general framework for studying the effects of resource competition is presented in [15] using Metabolic Control Analysis [16], yielding response coefficients that describe local flux sensitivities in a gene network. Our work complements these results as we consider the role of both RNAP and ribosomes to characterize the global limitations and trade-offs in protein expression for n genes. Some of these results have been validated *in vivo* for two genes [17].

This paper is organized as follows. In Section II, the system of interest is introduced, together with the motivation and research question: *Having n genes, what are the limitations and trade-offs in gene expression due to competition for shared cellular resources?* In Section III, we determine the attainable protein concentrations and characterize how various biochemical parameters affect the interdependence in gene expression. In Section IV, we present two implications of the limited availability of RNAP and ribosomes on gene expression. Finally, we conclude our results and present future research directions in Section V.

II. SYSTEM MODEL AND PROBLEM FORMULATION

We consider a system in which n genes are expressed. In particular, each gene is first transcribed by RNAP to mRNA, then mRNA is translated by ribosomes to protein (Fig. 1A). Furthermore, we focus on the case when the transcription of each gene is regulated by a transcription factor (TF) as follows. In the case of gene i expressing protein p_i, TF u_i first binds to the empty promoter b_i^* forming the promoter complex b_i. Then, the binding of RNAP x to b_i can form the transriptionally active promoter complex c_i, resulting in the production of mRNA m_i at rate γ_i (encompassing the elongation reactions). Finally, mRNA decays at rate δ_i. Consequently, the reactions describing the transcriptional processes for gene i are as follows:

$$\mathrm{u}_i + \mathrm{b}_i^* \underset{\chi_i}{\overset{\zeta_i}{\rightleftharpoons}} \mathrm{b}_i, \quad \mathrm{b}_i + \mathrm{x} \underset{\kappa_i^-}{\overset{\kappa_i^+}{\rightleftharpoons}} \mathrm{c}_i, \quad \mathrm{c}_i \xrightarrow{\gamma_1} \mathrm{b}_i + \mathrm{x} + \mathrm{m}_i, \quad \mathrm{m}_i \xrightarrow{\delta_i} \emptyset.$$

Translation of m_i is initialized by the ribosome y binding to the RBS of the mRNA m_i, forming the translationally active complex d_i. The degradation of mRNA when bound to the ribosome occurs with rate constant $a_i\delta_i$ where $0 < a_i \leq 1$ ($a_i \to 0$ represents the case when the ribosome-bound mRNA is protected from degradation, whereas $a_i = 1$ models the scenario when ribosomes provide no protection against degradation, which is considered in what follows). Protein p_i is degraded at rate λ_i, whereas elongation and production are lumped together in one step with effective production rate constant π_i. Therefore, the reactions describing the translation processes for gene i are given by

$$\mathrm{m}_i + \mathrm{y} \underset{k_i^-}{\overset{k_i^+}{\rightleftharpoons}} \mathrm{d}_i \xrightarrow{a_i\delta_i} y, \qquad \mathrm{d}_i \xrightarrow{\pi_i} \mathrm{m}_i + \mathrm{y} + \mathrm{p}_i, \qquad \mathrm{p}_i \xrightarrow{\lambda_i} \emptyset.$$

Consequently, the corresponding differential equation model for $i = 1, 2, \ldots, n$ is given by

$$\begin{aligned}
\dot{b}_i &= (\zeta_i u_i b_i^* - \chi_i b_i) - \left(\kappa_i^+ x b_i - \kappa_i^- c_i\right) + \gamma_i c_i, \\
\dot{c}_i &= \left(\kappa_i^+ x b_i - \kappa_i^- c_i\right) - \gamma_i c_i, \\
\dot{m}_i &= \gamma_i c_i - \delta_i m_i - \left(k_i^+ m_i y - k_i^- d_i\right) + \pi_i d_i, \\
\dot{d}_i &= \left(k_i^+ m_i y - k_i^- d_i\right) - \pi_i d_i - a_i\delta_i d_i, \\
\dot{p}_i &= \pi_i d_i - \lambda_i p_i.
\end{aligned} \tag{1}$$

A. RNAP and Ribosome Demand at the Steady State

Introduce the dissociation constants $\kappa_i = (\kappa_i^- + \gamma_i)/\kappa_i^+$ and $k_i = (k_i^- + \pi_i + \delta_i)/k_i^+$ for $i = 1, 2, \ldots, n$. Given that protein production and decay are much slower than binding and unbinding reactions [18], we have $\gamma_i \ll \kappa_i^-$ and $\pi_i, \delta_i \ll k_i^-$, so that $\kappa_i \approx \kappa_i^-/\kappa_i^+$ and $k_i \approx k_i^-/k_i^+$. The stronger the binding of RNAP to the promoter, the smaller κ_i, and similarly, the stronger the binding of ribosome to the RBS, the smaller k_i. Next, define

$$h_i = \frac{\gamma_i \eta_i}{\delta_i}, \qquad q_i = \frac{\pi_i}{\lambda_i} h_i \qquad \text{and} \qquad \mu_i = \frac{\chi_i}{\zeta_i}, \tag{2}$$

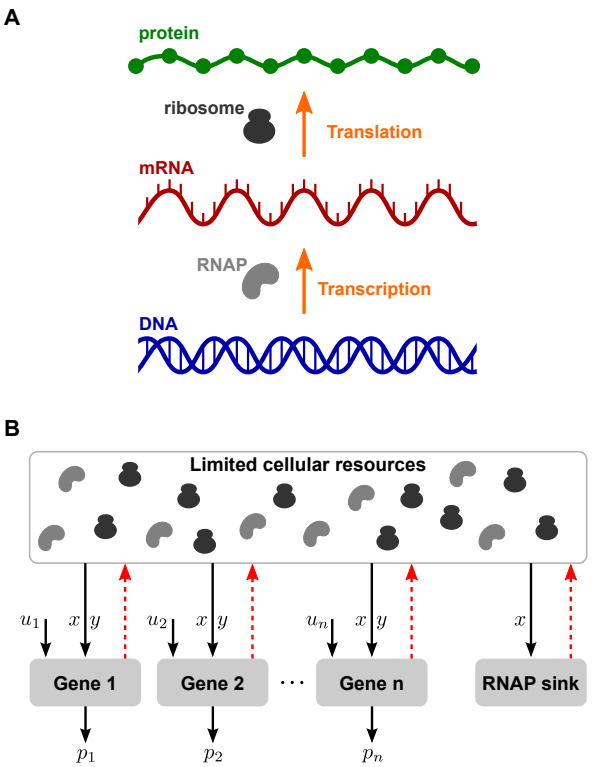

Fig. 1. Genes compete for transcriptional and translational resources. (A) DNA is transcribed into mRNA by RNAP, which is then translated to protein by ribosome. (B) The sharing of limited resources couples the expression of genes, even in the absence of regulatory linkages, as a result of loading (red) imposed by each gene on the pool of shared resources. The RNAP sink models the non-specific binding of RNAP (the non-specific binding of ribosomes could be treated similarly, neglected here to simplify exposition).

where μ_i is the dissociation constant of the TF u_i to the promoter of gene i. Furthermore let

$$\epsilon_i = \frac{\frac{u_i}{\mu_i}\left(1 + \frac{x}{\kappa_i}\right)}{1 + \frac{u_i}{\mu_i}\left(1 + \frac{x}{\kappa_i}\right)}, \qquad \text{for } i = 1, 2, \ldots, n. \tag{3}$$

Assuming that DNA concentration is constant [19], we have that $\eta_i = b_i^* + b_i + c_i$, where η_i is the total concentration of the promoter of gene i. We further have $\epsilon_i = (b_i + c_i)/\eta_i$, so that $\epsilon_i \in [0, 1)$ is the fraction of the promoter of gene i activated by u_i. At the steady state of (1), we have

$$c_i = \epsilon_i \eta_i \frac{x}{x + \kappa_i}, \qquad d_i = \epsilon_i h_i \frac{x}{x + \kappa_i} \frac{y}{y + k_i}, \tag{4}$$

whereas the concentration of protein p_i is

$$p_i = \epsilon_i q_i \frac{x}{x + \kappa_i} \frac{y}{y + k_i}, \qquad \text{for } i = 1, 2, \ldots, n. \tag{5}$$

We call c_i and d_i in (4) the *RNAP and ribosome demand* of gene i at the steady state, respectively, as they represent the concentration of RNAP and ribosomes bound to the promoter and mRNA, respectively. The protein concentrations p_i for $i = 1, 2, \ldots, n$ in (5) are implicitly coupled as the free concentration x and y of RNAP and ribosomes, respectively, depend on the demand by the genes (Fig. 1B), as we detail in the next section.

B. Modeling the Limited Availability of RNAP & Ribosomes

According to [20], RNAP can be divided into four main categories: immature RNAP, free RNAP, and RNAP bound specifically (and transcribing) and non-specifically to the chromosome. Based on [21], the cell has approximately 1500 RNAP molecules ($x_T = 1500$nM), among which about 200 are actively transcribing endogenous genes ($x_S = 200$nM) at low growth rate. Furthermore, [20] suggests that the ratio of immature RNAP is negligible, and the remaining 1300 molecules are partitioned as follows: 100 of them are free ($x = 100$nM), whereas 1200 are non-specifically bound ($x_N = 1200$nM). Consequently, the conservation law for RNAP without the additional genes of Fig. 1B is given by

$$x_T = x + x_S + x_N. \tag{6}$$

As for ribosomes, [21] reports that the number of ribosomes per cell is 6800 ($y_T = 6800$nM), 80% of which is active, that is, approximately 5500 ($y_S = 5500$nM) at low growth rate. According to [22], the concentration of free ribosomes is approximately 15%, so that the ratio of non-specifically bound ribosomes and immature ribosomes is about 5%. This is negligible compared to the fraction of active and free ribosomes, unlike in the case of RNAP. For simplicity, we treat this last 5% as if they belonged to the pool of free ribosomes (so that we slightly under-estimate the effect of competition for ribosomes). As a result, the conservation law for ribosomes without the additional genes of Fig. 1B is given by

$$y_T = y + y_S. \tag{7}$$

In [17], several exogenous genes are constitutively expressed with the simultaneous activation of an inducible gene. Since there is no appreciable change in growth rate even when using high copy number plasmids, we assume that x_S and y_S, representing the resources allocated to the gene expression of the host, are constant (however, if an overexpressed protein is toxic to the cell, the growth rate may decrease [8] and x_S and y_S might also be affected). Therefore, introduce $X = x_T - x_S$ and $Y = y_T - y_S$ denoting the concentration of available RNAP and ribosomes, respectively. To model the non-specific binding of RNAP, introduce the "RNAP sink" described by the reactions

$$\bar{\mathrm{b}} + \mathrm{x} \underset{\bar{\kappa}^-}{\overset{\bar{\kappa}^+}{\rightleftharpoons}} \bar{\mathrm{c}},$$

where $\bar{\eta} = \bar{b} + \bar{c}$ is the DNA concentration of this "RNAP sink". At the steady state we obtain that the concentration of RNAP sequestered by this sink is $\bar{c} = \bar{\eta}x/(x + \bar{\kappa})$ with $\bar{\kappa} = \bar{\kappa}^-/\bar{\kappa}^+$, and since the non-specific binding of RNAP is weak ($\bar{\kappa} \ll x$ by [23]), we obtain that $x_N = \bar{c} \approx x\bar{\eta}/\bar{\kappa}$.

The RNAP and ribosome demand of gene i is given by c_i and d_i in (4), respectively. Introduce $\bar{N} = 1 + \bar{\eta}/\bar{\kappa}$, so that upon addition of genes i ($i = 1, 2, \ldots, n$), (6)–(7) become

$$X = \bar{N}x + \sum_{i=1}^{n} \epsilon_i \eta_i \frac{x}{x + \kappa_i}, \tag{8}$$

$$Y = y + \sum_{i=1}^{n} \epsilon_i h_i \frac{x}{x + \kappa_i} \frac{y}{y + k_i}. \tag{9}$$

Let $\epsilon = (\epsilon_1, \epsilon_2, \ldots, \epsilon_n)^T$ and $u = (u_1, u_2, \ldots, u_n)^T$, and write (3) as

$$\epsilon = E(u, x), \tag{10}$$

so that (8) and (9) can be written with (10) as

$$X = F_\epsilon(\epsilon, x) \qquad \text{and} \qquad Y = G_\epsilon(\epsilon, x, y), \tag{11}$$

respectively, and (5) with $p = (p_1, p_2, \ldots, p_n)^T$ as

$$p = H_\epsilon(\epsilon, x, y). \tag{12}$$

C. Problem Formulation

Define

$$\begin{aligned} F(u, x) &= F_\epsilon(E(u, x), x), \\ G(u, x, y) &= G_\epsilon(E(u, x), x, y), \\ H(u, x, y) &= H_\epsilon(E(u, x), x, y), \end{aligned} \tag{13}$$

using (10)–(12), and introduce the sets $\mathcal{U} = [0, \infty)^n$ and

$$\begin{aligned} \mathcal{P} = \{p \mid p = H(u, x, y),\ X = F(u, x),\ Y = G(u, x, y), \\ x \in [0, X],\ y \in [0, Y],\ u \in \mathcal{U}\}, \end{aligned} \tag{14}$$

so that $\mathcal{P}$ is the set of protein concentrations attainable at the steady state. Therefore, we call $\mathcal{P}$ the *realizable region*. Here, we seek an explicit characterization of $\mathcal{P}$ solely in terms of p, instead of the definition in (14) involving u, x and y in the form of implicit constraints. As a result, we can answer the following questions. *How does the concentration of protein p_j change upon activation of gene i for $j \neq i$? To what extent is it possible to increase the concentration of p_i without affecting the concentration of p_j?* In other words, we seek to characterize the limitations and trade-offs in protein production due to the limited availability of RNAP and ribosomes.

III. REALIZABLE REGION

We characterize the realizable region $\mathcal{P}$ through a series of intermediate results. In particular, we first focus on the activation level ϵ_i of gene i for $i = 1, 2, \ldots, n$. Then, we consider a biologically reasonable approximation of (5) and (8)–(9) and characterize the corresponding set $\mathcal{S}$ of attainable protein concentrations. Finally, we prove that $\mathcal{P} \subseteq \mathcal{S}$.

A. Activation Level of Genes

Claim 1. *Take $F(u, x)$ and $G(u, x, y)$ defined in* (13)*. For $u \in \mathcal{U}$, there is a unique $(x, y) \in [0, X] \times [0, Y]$ such that $F(u, x) = X$ and $G(u, x, y) = Y$. As a result, there exist functions $f, g : \mathbb{R}^n \to \mathbb{R}$ such that $x = f(u)$ and $y = g(u)$.*

Proof: According to (13), we have

$$F(u, x) = \bar{N}x + \sum_{i=1}^{n} \frac{\frac{u_i}{\mu_i}\left(1 + \frac{x}{\kappa_i}\right)}{1 + \frac{u_i}{\mu_i}\left(1 + \frac{x}{\kappa_i}\right)} \eta_i \frac{x}{x + \kappa_i} - X. \tag{15}$$

Fix $u \in \mathcal{U}$. Since $F(u,x)$ is continuous and $F(u,0) = 0$ and $F(u,X) > X$ by (15), there is at least one $x \in [0,X]$ such that $F(u,x) = X$, according to the Intermediate Value Theorem [24]. Furthermore, since $F(u,x)$ in (15) is strictly increasing with x, there is exactly one $x \in [0,X]$ such that $F(u,x) = X$. Then, let $f : \mathbb{R}^n \to \mathbb{R}$ be the function that maps u to this unique x, that is, $F(u,f(u)) = X$. The proof for G can be constructed similarly. ■

With $H(u,x,y)$ defined in (13), introduce $A : \mathbb{R}^n \to \mathbb{R}^n$ as $A(u) = H(u,f(u),g(u))$, so that (14) can be written as

$$\mathcal{P} = \{p \mid p = A(u),\ u \in \mathcal{U}\}. \tag{16}$$

Claim 2. *Let $\mathcal{E} = [0,1)^n$. Take $F_\epsilon(\epsilon,x)$ and $G_\epsilon(\epsilon,x,y)$ from* (11)*. For $\epsilon \in \mathcal{E}$, there is a unique $(x,y) \in [0,X] \times [0,Y]$ such that $F_\epsilon(\epsilon,x) = X$ and $G_\epsilon(\epsilon,x,y) = Y$. As a result, there exist functions $f_\epsilon, g_\epsilon : \mathbb{R}^n \to \mathbb{R}$ such that $x = f_\epsilon(\epsilon)$ and $y = g_\epsilon(\epsilon)$.*

Proof: Similar to the proof of Claim 1. ■

Claim 3. *Take $u \in \mathcal{U}$, the functions f and g defined in Claim 1, together with f_ϵ and g_ϵ defined in Claim 2. Furthermore, consider $\epsilon = E(u,f(u))$ from* (10) *with $x = f(u)$. Then $f(u) = f_\epsilon(\epsilon)$ and $g(u) = g_\epsilon(\epsilon)$.*

Proof: By Claim 1, we have $X = F(u,(f(u)))$, yielding $X = F(u,(f(u))) = F_\epsilon(E(u,f(u)),f(u))$ from (13), and since $\epsilon = E(u,f(u))$ by assumption, we obtain $X = F_\epsilon(\epsilon,f(u))$. We further have $X = F_\epsilon(\epsilon,f_\epsilon(\epsilon))$ by Claim 2. As a result, we obtain that $x = f(u)$ and $x = f_\epsilon(\epsilon)$ are both solutions of $X = F_\epsilon(\epsilon,x)$, and since it has a unique solution by Claim 2, we conclude that $f(u) = f_\epsilon(\epsilon)$. The proof of $g(u) = g_\epsilon(\epsilon)$ can be constructed similarly. ■

With $H_\epsilon(\epsilon,x,y)$ defined in (12), introduce the function $A_\epsilon : \mathbb{R}^n \to \mathbb{R}^n$ as $A_\epsilon(\epsilon) = H_\epsilon(\epsilon,f_\epsilon(\epsilon),g_\epsilon(\epsilon))$ and the set

$$\mathcal{P}_\epsilon = \{p \mid p = A_\epsilon(\epsilon),\ \epsilon \in \mathcal{E}\}. \tag{17}$$

Lemma 1. *With $\mathcal{P}$ and $\mathcal{P}_\epsilon$ given in* (16) *and* (17)*, respectively, we obtain that $\mathcal{P} = \mathcal{P}_\epsilon$.*

Proof: Let $x = f(u)$ and $y = g(u)$ denote the unique solutions of $F(u,x) = X$ and $G(u,x,y) = Y$ with $(x,y) \in [0,X] \times [0,Y]$ for $u \in \mathcal{U}$, respectively (Claim 1). Referring to (11), let $x = f_\epsilon(\epsilon)$ and $y = g_\epsilon(\epsilon)$ denote the unique solutions of $F_\epsilon(\epsilon,x) = X$ and $G_\epsilon(\epsilon,x,y) = Y$ with $(x,y) \in [0,X] \times [0,Y]$ for $\epsilon \in \mathcal{E}$, respectively (Claim 2).

To prove that $\mathcal{P} \subseteq \mathcal{P}_\epsilon$ we show that for every $u \in \mathcal{U}$ there is an $\epsilon \in \mathcal{E}$ such that $A(u) = A_\epsilon(\epsilon)$. First, consider $\epsilon = E(u,f(u))$, and given that $f(u) \in [0,X]$, we conclude that $\epsilon_i \in [0,1)$ by (3), so that $\epsilon \in \mathcal{E}$ by the definition of $\mathcal{E}$. Second, considering (13) implies $A(u) = H(u,f(u),g(u)) = H_\epsilon(E(u,f(u)),f(u),g(u))$, so that $\epsilon = E(u,f(u))$ together with $f(u) = f_\epsilon(\epsilon)$ and $g(u) = g_\epsilon(\epsilon)$ from Claim 3 yield $A(u) = H_\epsilon(\epsilon,f_\epsilon(\epsilon),g_\epsilon(\epsilon)) = A_\epsilon(\epsilon)$, where we used the definition of $A_\epsilon(\epsilon)$.

Similarly, to show that $\mathcal{P}_\epsilon \subseteq \mathcal{P}$ it is sufficient to prove that for every $\epsilon \in \mathcal{E}$ there is a $u \in \mathcal{U}$ such that $A(u) = A_\epsilon(\epsilon)$. Since (3) yields $u_i = \epsilon_i\mu_i\kappa_i/[(1-\epsilon_i)(\kappa_i+f_\epsilon(\epsilon))]$, and given that $\epsilon_i \in [0,1)$ as $\epsilon \in \mathcal{E}$, we obtain $u_i \in [0,\infty)$, so that $u \in \mathcal{U}$. The part $A(u) = A_\epsilon(\epsilon)$ can be showed as above. ■

By Lemma 1, in order to find the realizable region $\mathcal{P}$, it is sufficient to consider (5) with (8)–(9) for $\epsilon \in \mathcal{E}$, instead of considering (5) with (8)–(9) and with (3) for $u \in \mathcal{U}$.

B. Approximate Model & Approximate Realizable Region $\mathcal{S}$

As an intermediate step to characterize the realizable region $\mathcal{P}$, consider the (biologically reasonable, see Appendix) approximations $x \ll \kappa_i$ and $y \ll k_i$ for $i = 1,2,\ldots,n$, so that (8)–(9) and (5) take the form

$$X = \bar{N}x + \sum_{i=1}^{n} \epsilon_i \frac{\eta_i}{\kappa_i} x, \quad Y = y + \sum_{i=1}^{n} \epsilon_i \frac{h_i}{\kappa_i k_i} xy, \tag{18}$$

$$p_i = \epsilon_i \frac{q_i}{\kappa_i k_i} xy. \tag{19}$$

Expressing x and y from (18) as a function of ϵ and substituting these expressions into (19) yields

$$p_i = \frac{Q_i \epsilon_i}{\bar{N} + \sum_{i=1}^{n} H_i \epsilon_i}, \qquad \text{for } i = 1,2,\ldots,n \tag{20}$$

with

$$Q_i = \frac{q_i}{\kappa_i k_i} XY \qquad \text{and} \qquad H_i = \frac{h_i}{\kappa_i k_i} X + \frac{\eta_i}{\kappa_i}. \tag{21}$$

Furthermore, let $\hat{A} : \mathbb{R}^n \to \mathbb{R}^n$ be the function mapping ϵ to p according to (20), so that $p = \hat{A}(\epsilon)$. Next, define

$$p_i^{\max} = \frac{Q_i}{\bar{N} + H_i} \qquad \text{and} \qquad p_i^{\infty} = \frac{Q_i}{H_i}, \tag{22}$$

and introduce the simplex $\mathcal{S}_i$ for $i = 1,2,\ldots,n$ as

$$\mathcal{S}_i = \left\{ p \mid p \geq 0 \text{ and } \frac{p_i}{p_i^{\max}} + \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{p_j}{p_j^{\infty}} < 1 \right\}. \tag{23}$$

Lemma 2. *Let*

$$\mathcal{S} = \{p \mid p = \hat{A}(\epsilon),\ \epsilon \in \mathcal{E}\}. \tag{24}$$

Then, we obtain $\mathcal{S} = \cap_{i=1}^{n} \mathcal{S}_i$ where $\mathcal{S}_i$ is defined in (23)*.*

Proof: We first show $\mathcal{S} \subseteq \cap_{i=1}^{n} \mathcal{S}_i$ as follows. Introduce $\mathcal{E}_i = \{\epsilon \mid \epsilon_i \in [0,1) \text{ and } \epsilon_j \in [0,\infty) \text{ for } j \neq i\}$ and let $P_i = Q_i\epsilon_i/(\bar{N} + H_i\epsilon_i)$, so that we have $P_i < p_i^{\max}$ by (22). Furthermore, $p = \hat{A}(\epsilon)$ satisfies

$$\frac{p_i}{P_i} + \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{p_j}{p_j^{\infty}} = 1 \tag{25}$$

by substitution of (20) into (25). The fact that $\epsilon \in \mathcal{E}$ yields $p \geq 0$ by (19), and $P_i < p_i^{\max}$ with (25) result in

$$\frac{p_i}{p_i^{\max}} + \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{p_j}{p_j^{\infty}} < \frac{p_i}{P_i} + \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{p_j}{p_j^{\infty}} = 1,$$

so that $p \in \mathcal{S}_i$ by (23) for $\epsilon \in \mathcal{E}_i$. Combining this together with the fact that $\epsilon \in \mathcal{E} = \cap_{i=1}^{n} \mathcal{E}_i$ yields that $\mathcal{S} \subseteq \cap_{i=1}^{n} \mathcal{S}_i$.

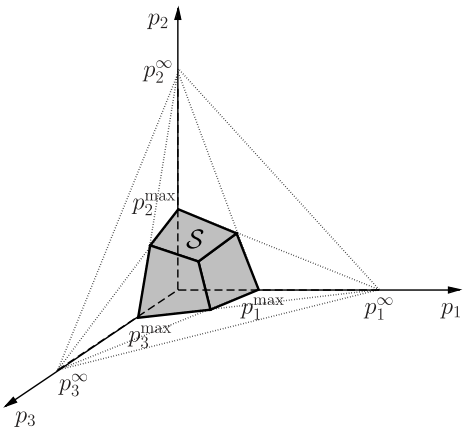

Fig. 2. In the case of $n = 3$, $\mathcal{S}$ in (24) is the intersection of the simplexes $\mathcal{S}_1$, $\mathcal{S}_2$ and $\mathcal{S}_3$ (Lemma 2), where $\mathcal{S}_i$ is defined in (23) as the simplex given by the origin, $p_i^{\max}$ on the p_i-axis and p_j^∞ on the p_j-axis for $j \neq i$.

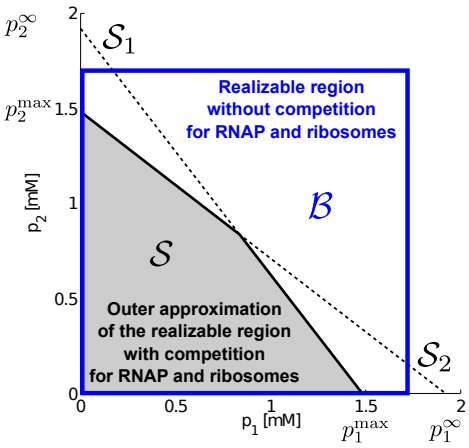

Fig. 3. The set of attainable protein concentrations without considering competition for RNAP and ribosomes is $\mathcal{B}$ given in (35). When considering the limited availability of RNAP and ribosomes given by (8)–(9), the set of attainable protein concentrations $\mathcal{P}$ in (16) lie inside the outer approximation $\mathcal{S}$ from (24). The set $\mathcal{S}$ is the intersection of the simplexes $\mathcal{S}_1$ and $\mathcal{S}_2$ in (23) (triangles with dashed border). Simulation parameters: $\eta_1 = \eta_2 = 100$nM, $\lambda_1 = \lambda_2 = 1\text{hr}^{-1}$, $\pi_1 = \pi_2 = 1500\text{hr}^{-1}$, $\delta_1 = \delta_2 = 10\text{hr}^{-1}$, $\gamma_1 = \gamma_2 = 500\text{hr}^{-1}$ $\kappa_1 = \kappa_2 = 150$nM, $k_1 = k_2 = 1000$nM, $\bar{N} = 13$, $X = Y = 1300$nM, $x_0 = 100$nM and $y_0 = 1300$nM.

Second, we prove $\cap_{i=1}^n \mathcal{S}_i \subseteq \mathcal{S}$ by showing that for any $p \in \cap_{i=1}^n \mathcal{S}_i$ there exists an $\epsilon \in \mathcal{E}$ such that $p = \hat{A}(\epsilon)$. To this end, pick $p \in \cap_{i=1}^n \mathcal{S}_i$ and define

$$P_i = \frac{p_i}{1 - \sum_{\substack{j=1 \\ j\neq i}}^n \frac{p_j}{p_j^\infty}} \qquad \text{and} \qquad \epsilon_i = \frac{\bar{N} P_i}{Q_i - H_i P_i} \tag{26}$$

for $i = 1, 2, \ldots, n$. Substituting ϵ into (20) we obtain that $p = \hat{A}(\epsilon)$. Therefore, it is only left to show that $\epsilon \in \mathcal{E}$. Given that $p \in \cap_{i=1}^n \mathcal{S}_i$, we obtain by (23) that

$$0 \leq p_i < p_i^{\max} \qquad \text{and} \qquad 0 \leq \sum_{\substack{j=1 \\ j\neq i}}^n \frac{p_j}{p_j^\infty} < 1.$$

Combining this together with (26) yields that $P_i \in [0, p_i^{\max})$. Having $P_i = 0$ and $P_i = p_i^{\max}$ result in $\epsilon_i = 0$ and $\epsilon_i = 1$ in (26) by (22). Furthermore, as ϵ_i in (26) is a strictly increasing function of P_i for $P_i \in [0, p_i^{\max})$, we conclude that $\epsilon_i \in [0, 1)$ for $i = 1, 2, \ldots, n$, so that $\epsilon \in \mathcal{E}$. ■

The realizable region $\mathcal{S}$ of protein concentrations when $x \ll \kappa_i$ and $y \ll k_i$ is given as $\mathcal{S} = \cap_{i=1}^n \mathcal{S}_i$ by Lemma 2, where $\mathcal{S}_i$ is the n-dimensional simplex defined by the following $n+1$ vertices: the origin, $p_i^{\max}$ on the p_i-axis and p_j^∞ on the p_j-axis for $j \neq i$ (see Fig. 2). Furthermore, the dependence of $\mathcal{S}_i$ on the biochemical parameters is given by the expressions of $p_i^{\max}$ and p_i^∞ in (22). For instance, both $p_i^{\max}$ and p_i^∞ increase as k_i decreases (stronger RBS for gene i), and $p_i^{\max}$ increases while p_i^∞ remains unaffected as κ_i decreases (stronger promoter for gene i).

C. The Realizable Region $\mathcal{P}$ Lies Inside $\mathcal{S}$

We next show that even when the approximations $x \ll \kappa_i$ and $y \ll k_i$ do not hold, the set of attainable protein concentrations given by $\mathcal{P}$ in (16) lie within $\mathcal{S}$ in (24).

Theorem 1. *Considering $\mathcal{P}$ and $\mathcal{S}$ defined in* (16) *and* (24)*, respectively, we obtain that $\mathcal{P} \subseteq \mathcal{S}$.*

Proof: With $\mathcal{P}_\epsilon$ defined in (17), we have $\mathcal{P}_\epsilon = \mathcal{P}$ by Lemma 1, so that it is sufficient to show that $\mathcal{P}_\epsilon \subseteq \mathcal{S}$ to prove $\mathcal{P} \subseteq \mathcal{S}$. To this end, fix $\epsilon \in \mathcal{E}$ and let $p = A_\epsilon(\epsilon)$. If we can show that $p \in \mathcal{S}_i$ for $i = 1, 2, \ldots, n$, it implies that $p \in \mathcal{S}$ since $\mathcal{S} = \cap_{i=1}^n \mathcal{S}_i$ by Lemma 2, yielding $\mathcal{P}_\epsilon \subseteq \mathcal{S}$.

To show that $p \in \mathcal{S}_i$ for $i = 1, 2, \ldots, n$, define

$$\alpha_i = \frac{\kappa_i}{x + \kappa_i}, \qquad \beta_i = \frac{k_i}{y + k_i}, \qquad \epsilon_i' = \alpha_i \beta_i \epsilon_i, \tag{27}$$

so that (4)–(5) and (8)–(9) become

$$c_i = \frac{\epsilon_i'}{\beta_i} \frac{\eta_i}{\kappa_i} x, \quad d_i = \epsilon_i' \frac{h_i}{\kappa_i k_i} xy, \quad p_i = \epsilon_i' \frac{q_i}{\kappa_i k_i} xy, \tag{28}$$

$$x = \frac{X}{\bar{N} + \sum_{i=1}^n \frac{\eta_i}{\kappa_i} \epsilon_i'}, \qquad y = \frac{Y}{1 + \sum_{i=1}^n \frac{h_i}{\kappa_i k_i} \epsilon_i' x}. \tag{29}$$

As a result, with Q_i from (21) and with $H_i' = h_i X/(\kappa_i k_i) + \eta_i/(\beta_i \kappa_i)$, we can write p_i in (28) with (29) as

$$p_i = \frac{Q_i \epsilon_i'}{\bar{N} + \sum_{i=1}^n H_i' \epsilon_i'}. \tag{30}$$

Furthermore, introduce $\tilde{p}_i = (Q_i \epsilon_i')/(\bar{N} + \sum_{i=1}^n H_i \epsilon_i')$ and let $\hat{p} = (\hat{p}_1, \hat{p}_2, \ldots, \hat{p}_n)^T$ where $\hat{p}_i$ is given by (20).

The fact that $\alpha_i, \beta_i \in (0, 1)$ yields $\epsilon_i' \in [0, \epsilon_i)$ by (27) and $H_i \in (0, H_i')$ by (21). Since $\epsilon_i' \in [0, \epsilon_i)$ implies $\tilde{p}_i < \hat{p}_i$, and similarly, $H_i \in (0, H_i')$ yields $p_i < \tilde{p}_i$, we obtain

$$0 \leq p_i < \tilde{p}_i < \hat{p}_i. \tag{31}$$

Furthermore, from Lemma 2 we have

$$\frac{\hat{p}_i}{p_i^{\max}} + \sum_{\substack{j=1 \\ j\neq i}}^n \frac{\hat{p}_j}{p_j^\infty} - 1 < 0, \tag{32}$$

and combining (31)–(32) yields

$$\frac{p_i}{p_i^{\max}} + \sum_{\substack{j=1 \\ j\neq i}}^n \frac{p_j}{p_j^\infty} < \frac{\hat{p}_i}{p_i^{\max}} + \sum_{\substack{j=1 \\ j\neq i}}^n \frac{\hat{p}_j}{p_j^\infty} < 1. \tag{33}$$

We have $p_i \geq 0$ by (31). Together with (33) this implies that $p \in \mathcal{S}_i$ for $i = 1, 2, \ldots, n$ by (23), concluding the proof. ■

Introduce x_0 and y_0 such that

$$X = F_\epsilon(0, x_0) \qquad \text{and} \qquad Y = G_\epsilon(0, x_0, y_0), \tag{34}$$

that is, x_0 and y_0 denote the concentration of free RNAP and ribosomes, respectively, when none of the genes in Fig. 1B are activated ($\epsilon_i = 0$ for $i = 1, 2, \ldots, n$). Next, define

$$\mathcal{B} = \{p \mid p = H(u, x_0, y_0),\ u \in \mathcal{U}\}, \tag{35}$$

representing the set of attainable protein concentrations without considering competition for RNAP and ribosomes (so that $x = x_0$ and $y = y_0$), see Fig. 3 for a particular example when $n = 2$. Since $\mathcal{S} \subset \mathcal{B}$ in Fig. 3, if $(p_1, p_2) \in \mathcal{B} \setminus \mathcal{S}$ then $(p_1, p_2) \notin \mathcal{P}$. As a result, without considering competition for RNAP and ribosomes we would erroneously conclude that the protein concentrations (p_1, p_2) are attainable.

IV. Practical Implications of the Limited Availability of RNAP and Ribosomes

Here, we present two examples of how protein concentrations become coupled due to competition for shared resources, and we detail the resulting limitations and trade-offs building upon our results from the previous section. In particular, we show that proteins can seemingly behave both as repressors and as activators, purely as an effect of the limited availability of RNAP and ribosomes.

A. Lateral Inhibition with Two Genes

Consider two genes, and for simplicity, focus on the biologically reasonable approximations (see Appendix) $x \ll \kappa_i$ and $y \ll k_i$ for $i = 1, 2$, so that the realizable set of protein concentrations $\mathcal{P}$ is equal to $\mathcal{S}$ by Lemma 2. We investigate the limitations and trade-offs in protein expression, i.e., how p_2 changes as the expression of p_1 increases.

Fix the activation ϵ_2 of gene 2 ($\epsilon_2 = \epsilon_2^*$), while increasing the activation ϵ_1 of gene 1. Without considering competition for the shared resources, the set of attainable protein concentrations is $\mathcal{B}$ given in (35), see Fig. 4A. In this case, the concentration x and y of free RNAP and ribosomes, respectively, are independent of the value of ϵ_1 and ϵ_2 ($x = x_0$ and $y = y_0$, see (34)). As a result, p_1 increases while p_2 remains unaffected when increasing the activation ϵ_1 of gene 1 by (16). That is, the attainable pairs (p_1, p_2) lie along a horizontal line (Fig. 4A). However, due to the limited availability of resources, p_2 decreases by (20) as the activation ϵ_1 of gene 1 increases, since some of the resources have to be reallocated from gene 2 to gene 1. Referring to (20), the pair (p_1, p_2) satisfies the linear constraint

$$\underbrace{\left(A_1 + \frac{B_1}{X}\right)}_{\alpha} p_1 + \underbrace{\left(A_2 + \frac{B_2 + C/\epsilon_2^*}{X}\right)}_{\beta} p_2 = Y \tag{36}$$

with $A_i = \lambda_i/\pi_i$, $B_i = \delta_i k_i A_i/\gamma_i$ and $C = B_2\kappa_2 k_2 \bar{N}/\eta_2$. Since $\frac{\partial p_1}{\partial \epsilon_1} > 0$ and $\frac{\partial p_2}{\partial \epsilon_1} < 0$ from (20), the pair (p_1, p_2) moves along the line (36) from left to right by increasing ϵ_1 (red line in Fig. 4B). The top boundary of $\mathcal{S}$ in Fig. 4B is given by (36) when $\epsilon_2 = 1$ (gene 2 is fully activated), and similarly, the right boundary of $\mathcal{S}$ corresponds to the case when $\epsilon_1 = 1$ (gene 1 is fully activated).

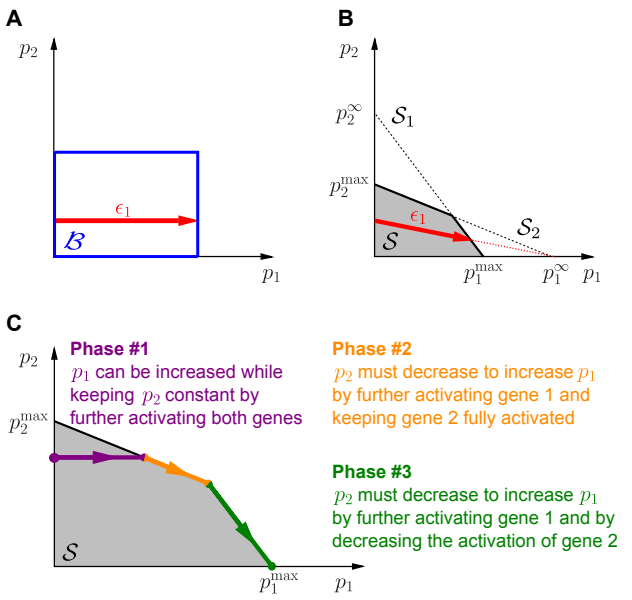

Fig. 4. Limitations and trade-offs in protein concentrations in the case of two genes. (A) Without competition for RNAP and ribosomes, the concentration of p_2 would not decrease when activating gene 1 (red). (B) When the genes are competing for RNAP and ribosomes, resources need to be reallocated from p_2 for the production of p_1. As a result, the concentration of p_2 decreases upon activation of gene 1 (red). (C) We can first increase the expression of gene 1 without decreasing p_2 (purple), but once reaching the boundary of the realizable region $\mathcal{S}$, the expression of gene 2 must decrease to further increase p_1 (orange and green).

The linear constraint in (36) can be interpreted as an isocost line [25], a concept introduced in microeconomics to describe what combinations of two products can be purchased with a limited budget. Here, the products are p_1 and p_2 with prices α and β, respectively, whereas the budget Y is the concentration of available ribosomes. Increasing the availability of resources (RNAP and ribosomes) allows for purchasing more products: the value of p_1 and p_2 can be increased simultaneously. In particular, increasing the concentration X of available RNAP molecules decreases the prices α and β, whereas increasing the concentration of available ribosomes Y increases the budget. Furthermore, the isocost line describes how changing the biochemical parameters of a gene affects the extent of competition due to the limited availability of resources. In particular, the slope of the isocost line is $-\alpha/\beta$ by (36), so that producing an extra p_1 decreases the concentration of p_2 by α/β. The "more expensive" p_1 compared to p_2 (the greater α/β), the more p_2 have to be sacrificed in order to purchase an additional unit of p_1. For instance, α decreases with the dissociation constant k_1, so that stronger RBS for gene 1 makes the isocost line flatter by (36), verified *in vivo* in [17].

Without competition for shared resources, increasing the activation ϵ_1 of gene 1 does not affect p_2 (Fig. 4A). However, due to the limited availability of RNAP and ribosomes, the

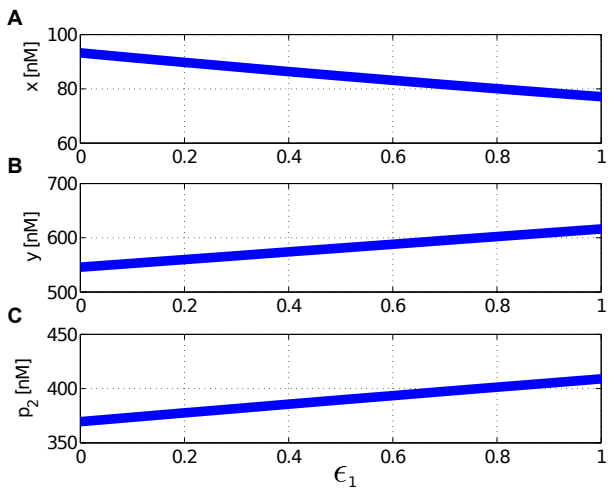

Fig. 5. Lateral activation in the case of $n = 3$ genes upon activation of gene 1. (A) The free concentration x of RNAP decreases (however, the RNAPs allocated to gene 2 are not sequestered as $\kappa_2 \ll x$). (B) The free concentration y of ribosomes increases as the activation level of gene 1 increases, since they are freed up from gene 3. (C) The concentration p_2 of p_2 increases as the activation level of gene 1 increases. Simulation parameters: $X = Y = 1300\text{nM}$, $\bar{N} = 13$, $\gamma_1 = \gamma_2 = \gamma_3 = 500\text{hr}^{-1}$, $\pi_1 = \pi_2 = \pi_3 = 1500\text{hr}^{-1}$, $\delta_1 = \delta_2 = \delta_3 = 10\text{hr}^{-1}$, $\lambda_1 = \lambda_2 = \lambda_3 = 1\text{hr}^{-1}$, $\eta_1 = 100\text{nM}$, $\eta_2 = 10\text{nM}$, $\eta_3 = 100\text{nM}$, $\kappa_1 = 100\text{nM}$, $\kappa_2 = 10\text{nM}$, $\kappa_3 = 1000\text{nM}$, $k_1 = 10\text{M}$, $k_2 = 1\text{mM}$ and $k_3 = 1\mu\text{M}$.

expression of gene 2 decreases when activating p_1 according to (36), as $\frac{\partial p_2}{\partial \epsilon_1} < 0 < \frac{\partial p_1}{\partial \epsilon_1}$ by (20). As a result, to increase p_2 and keep p_1 unaffected (phase #1 in Fig. 4C), we must increase the activation of both genes: by (20), increasing ϵ_1 yields greater expression of p_1, and the resulting decrease in p_2 can be compensated by increasing ϵ_2. However, when gene 2 becomes fully activated ($\epsilon_2 \to 1$), compensation is no longer possible, so that further increasing the activation ϵ_1 of gene 1 decreases p_1 (phase #2 in Fig. 4C). Finally, when gene 1 becomes fully activated ($\epsilon_1 \to 1$), the concentration of p_1 cannot be further increased while keeping gene 2 fully activated. Instead, we must decrease ϵ_2 so that resources can be reallocated to the expression of p_1 (phase #3 in Fig. 4C).

B. Lateral Activation with Three Genes

Since genes compete for the shared resources, one would expect that activating one gene decreases the expression of a different one. Here, we show that this is not always the case, and that counter-intuitively, unconnected genes can behave as activators to each other due to the limited availability of resources. To this end, consider $\kappa_2 \ll x \ll \kappa_1, \kappa_3$ and $y \ll k_1, k_2, k_3$, so that the promoter of gene 2 is saturated with RNAP. Furthermore, we focus on the case when genes 2 and 3 are fully activated ($\epsilon_2, \epsilon_3 \to 1$). Considering (8)–(9) when $\kappa_2 \ll x \ll \kappa_1, \kappa_3$ and $y \ll k_1, k_2, k_3$, and taking the derivative of p_2 in (5) with respect to ϵ_1 yields

$$\operatorname{sgn}\left(\frac{\mathrm{d}p_2}{\mathrm{d}\epsilon_1}\right) = \operatorname{sgn}\left(\frac{\eta_1}{\kappa_1}\frac{h_3}{\kappa_3 k_3} - \left(\bar{N} + \frac{\eta_3}{\kappa_3}\right)\frac{h_1}{\kappa_1 k_1}\right).$$

As a result, $\frac{\mathrm{d}p_2}{\mathrm{d}\epsilon_1} > 0$ if, for instance, k_1 is sufficiently large (the RBS of the mRNA encoding p_1 is sufficiently weak). In this case, activating gene 1 increases the concentration of p_2, despite gene 2 being already fully activated ($\epsilon_2 \to 1$).

This seemingly counter-intuitive result can be explained as follows. Activating gene 1 results in an increased demand for RNAP, consequently, less RNAP is available for the other two genes by (8). However, since the promoter of gene 2 is particularly strong ($\kappa_2 \ll x$), it stays saturated with RNAP, that is, the concentration of mRNA encoding p_2 remains about the same. By contrast, the promoter of gene 3 is weak ($\kappa_3 \gg x$), so that less mRNA encoding p_3 is produced, i.e., the demand for ribosomes by gene 3 decreases by (4). In the meantime, if the RBS of the mRNA encoding p_1 is weak (k_1 is sufficiently large), the demand for ribosomes by gene 1 is negligible by (4). Consequently, the ribosomes not used by gene 3 can be used by gene 2. In summary, the key features for obtaining the phenomenon are weak RBS in gene 1, strong promoter in gene 2 and weak promoter in gene 3. This lateral activation phenomenon is demonstrated in Fig. 5 (without using the approximations $\kappa_2 \ll x \ll \kappa_1, \kappa_3$ and $y \ll k_1, k_2, k_3$).

V. DISCUSSION

In this paper, we have characterized how the concentration of proteins become coupled due to competition for shared cellular resources, even in the absence of regulatory linkages. In particular, we showed that the realizable region $\mathcal{P}$ of protein concentrations lies within $\mathcal{S}$, which is a biologically reasonable outer approximation (see Appendix) easily calculated from (23)–(24). Building on this result, we determined the limitations and trade-offs in gene expression due to the limited availability of RNAP and ribosomes, and how they depend on various biochemical parameters. Furthermore, we demonstrated that the coupling in protein concentrations due to competition for RNAP and ribosomes can be interpreted using isocost lines, a concept introduced in microeconomics to describe the attainable combinations of products having a limited budget. Finally, we presented the counter-intuitive phenomenon of lateral activation, in which inducing the expression of one protein can increase the production of a second one, by reallocating resources from a third, serving as a buffer for shared resources.

A natural extension of the results presented here is considering regulatory linkages among genes, thus enabling the description of how the limited availability of resources couples the expression of different proteins in arbitrary gene networks. We are further working on the extension of the presented framework to describe the dynamic behavior of gene circuits. A particularly interesting research direction is combining the results of [6], describing the effects of sharing transcription factors on the dynamics of modules, and the result presented here, characterizing the stationary effects of the limited availability of transcriptional and translational machinery. As a result, one could account for two of the major causes of context-dependence in systems and synthetic biology in a unified mathematical framework, allowing a more detailed understanding of natural systems, and the design of multi-module systems with predictable behavior.

TABLE I

TYPICAL VALUES OF BIOCHEMICAL PARAMETERS

Parameter	Value	Unit	References
X	1300	nM	[20]
Y	1300	nM	[21], [22]
$\bar{N}$	13	-	[20]
κ_i	1000	nM	[26]
δ_i	10	hr^{-1}	[29]
γ_i	500	hr^{-1}	[21], [30], [27]
π_i	1500	hr^{-1}	[21], [30]
λ_i	1	hr^{-1}	[31]

VI. APPENDIX

The dissociation constant of the T7 RNAP to its promoter is approximately 200nM [26], and since this binding is considerably stronger than that of bacterial RNAP, we conclude that $\kappa_i \gg 200\text{nM}$, suggesting $x \ll \kappa_i$ as $x \approx 100\text{nM}$.

According to [27], as many as 20 RNAP molecules can simultaneously transcribe a gene. Instead of having one gene recruiting a maximum of ω RNAP molecules, we consider ω genes allowed to recruit at most one RNAP at a time, as if the DNA copy number was $\omega\eta$ instead of η (we use a low-range value of $\omega = 5$ denoting the number of RNAP molecules simultaneously transcribing a gene). Similarly, according to [21], several ribosomes can simultaneously translate each mRNA, up to a few dozen depending on the growth rate. Instead of having m mRNA molecules, each of which can be bound to ϕ ribosomes at any given time, we consider ϕm mRNA molecules allowed to be bound to a single ribosome. This can be achieved by considering the effective production rate $\phi\gamma$ instead of γ (we use a low-range value of $\phi = 5$ denoting the number of translations per mRNA).

Considering the typical value of biochemical parameters given in Tab. I with $k = 1000\text{nM}$, we obtain $p \approx 10\mu\text{M}$, which is comparable to the concentration of one of the most abundant proteins in *E. coli* [28]. Therefore, we approximate the binding of ribosomes to the RBS of the mRNA to be significantly weaker than 1000nM, so that $k_i \gg 1000\text{nM}$. Combining this with the fact that the concentration of free ribosomes is $y = 1300\text{nM}$ suggests that $y \ll k_i$.

REFERENCES

[1] S. Cardinale and A.P. Arkin, Contextualizing context for synthetic biology - identifying causes of failure of synthetic biological systems, *Biotechnol. J.*, vol. 7, no. 7, 2012, pp. 856-66.

[2] L.H. Hartwell, J.J. Hopfield, S. Leibler and A.W. Murray, From molecular to modular cell biology, *Nature*, vol. 402, no. 6761 Suppl, 1999, pp. C47-52.

[3] P.E.M. Purnick and R. Weiss, The second wave of synthetic biology: from modules to systems, *Nature Reviews Molecular Cell Biology*, vol. 10, no. 6, 2009, pp. 410-22.

[4] A.L. Slusarczyk, A. Lin and R. Weiss, Foundations for the design and implementation of synthetic genetic circuits, *Nature Reviews Genetics*, vol. 13, no. 6, 2012, pp. 406-20.

[5] D. Del Vecchio, A.J. Ninfa and E.D. Sontag, Modular cell biology: retroactivity and insulation, *Nature/EMBO Molecular Systems Biology*, vol. 4, no. 161, 2008.

[6] A. Gyorgy and D. Del Vecchio, Modular composition of complex gene transcription networks, *PLoS Computational Biology*, vol. 10, no. 3, 2014, e1003486

[7] W.E. Bentley , N. Mirjalili, D.C. Andersen, R.H. Davis and D.S. Kompal, Plasmid-encoded protein: the principal factor in the "metabolic burden" associated with recombinant bacteria, *Biotechnology and Bioengineering*, vol. 35 no. 7, 1990, pp. 668-81.

[8] M. Scott, C.W. Gunderson, E.M. Mateescu, Z. Zhang and T. Hwa, Interdependence of cell growth and gene expression: origins and consequences, *Science*, vol. 330, no. 6007, 2010, pp. 1099-1102.

[9] N.A. Cookson, W.H. Mather, T. Danino, O. Mondragon-Palomino, R.J. Williams, L.S. Tsimring and J. Hasty, Queueing up for enzymatic processing: correlated signaling through coupled degradation, *Molecular Systems Biology*, vol. 7, 2011, 561.

[10] J. Vind, M.A. Sorensen, M.D. Rasmussen and S Pedersen, Synthesis of proteins in *Escherichia coli* is limited by the concentration of free ribosomes: expression from reporter genes does not always reflect functional mRNA levels, *J Mol Biol*, vol. 231, no. 3, 1993, pp. 678-88.

[11] G. Churchward, H. Bremer and R. Young, Transcription in bacteria at different DNA concentrations, *J. Bacteriol.*, vol. 150, 1982, pp. 572-81.

[12] W.H. Mather, J. Hasty, L.S. Tsimring and R.J. Williams, Translational cross talk in gene networks, *Biophysical Journal*, vol. 104. no. 11, 2013, pp. 2564-72.

[13] R. Algar, T. Ellis and G.-B. Stan, Modelling the burden caused by gene expression: an *in silico* investigation into the interactions between synthetic gene circuits and their chassis cell, *arXiv:1309.7798*, 2013.

[14] D. Siegal-Gaskins, V. Noireaux and R.M. Murray, Biomolecular resource utilization in elementary cell-free gene circuits, *In Proc of the American Control Conference*, 2013.

[15] D. De Vos, F.J. Bruggeman, H.V. Westerhoff and B.M. Bakker, How molecular competition influences fluxes in gene expression networks, *PLoS One*, vol. 6, 2011, e28494.

[16] M.A. Savageau, *Biochemical systems analysis. A study of function and design in molecular biology*, Addison-Wesley Publishing Co.; 1976.

[17] J.I. Jimenez, A. Gyorgy, H. Chung, R. Weiss and D. Del Vecchio, Isocost lines describe the cellular economy of genetic circuits, *Under review*

[18] U. Alon, *An Introduction to systems biology - Design principles of biological circuits*, Chapman and Hall; 2006.

[19] T. Akerlund, K. Nordstrom and R. Bernander, Analysis of cell size and DNA content in exponentially growing and stationary-phase batch cultures of Escherichia coli, *J Bacteriol*, vol. 177, 1995, pp. 6791-97.

[20] S. Klumpp and T. Hwa, Growth-rate-dependent partitioning of RNA polymerases in bacteria, *PNAS*, vol. 105, 2008, pp. 20245-50.

[21] H. Bremer and P.P. Dennis, Modulation of chemical composition and other parameters of the cell by growth rate, *Escherichia coli and Salmonella: cellular and molecular biology*, vol. 2, 1996, pp. 1553-69.

[22] A.D. Tadmor and T. Tlusty, A coarse-grained biophysical model of *E. coli* and its application to perturbation of the rRNA operon copy number, *PLOS Computational Biology*, vo. 4, 2008.

[23] I.L. Grigorova, N.J. Phleger, V.K. Mutalik and C.A. Gross, Insights into transcriptional regulation and σ competition from an equilibrium model of RNA polymerase binding to DNA, *PNAS*, vol. 103, no. 14, 2006, pp. 5332-37.

[24] J.E. Marsden, *Elementary classical analysis*. W.H. Freeman and Company, San Francisco; 1974.

[25] R.S. Pindyck and D.L. Rubinfeld, *Microeconomics*. Pearson Practice Hall, NJ; 2005.

[26] V.L. Tunitskaya and S.N. Kochetkov, Structural-functional analysis of bacteriophage T7 RNA polymerase, *Biochemistry (Moscow)*, vol. 67, no. 10, 2002, pp. 1124-35.

[27] H. Bremer, P.P. Dennis and M. Ehrenberg, Free RNA polymerase and modeling global transcription in *Escherichia coli*", *Biochimie*, col. 85, no. 6, 2003, pp. 597-609.

[28] T. Van den Boom and J.E. Cronan Jr., Genetics and regulation of bacterial lipid metabolism, *Annu Rev Microbiol*, vol. 43, no. 1, 1989, pp. 317-43.

[29] J.A. Bernstein, A.B. Khodursky, P.-H. Lin, S. Lin-Chao and S.N. Cohen, Global analysis of mRNA decay and abundance in *Escherichia coli* at single-gene resolution using two-color fluorescent DNA microarrays, *PNAS*, vol. 99, no. 15, 2002, pp. 9697-9702.

[30] R.Y. Tsien, The green fluorescent protein, *Annual review of biochemistry*, vol. 67, no. 1, 1998, pp. 509-44.

[31] G. Kramer, R.R. Sprenger, M.A. Nessen, W. Roseboom, D. Speijer, L. de Jong, M.J.T. de Mattos, J.W. Back and C.G. de Koster, Proteome-wide alterations in *Escherichia coli* translation rates upon anaerobiosis, *Molecular & Cellular Proteomics*, vol. 9, no. 11, 2010, pp. 2508-16.