Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 52-1 (2019) 526–531

Analysis of Transcriptional Feedback Strategy for Reducing Interaction in Gene Expression Processes ⋆

Sebastián Nuñez * **Fabricio Garelli** * **Jesús Picó** ** **Hernán De Battista** *

* *Grupo de Control Aplicado (GCA), Instituto LEICI (UNLP-CONICET), Facultad de Ingeniería, Universidad Nacional de La Plata, La Plata (1900) Argentina. (e-mail: {sebastian.nuniez;fabricio;deba}@ing.unlp.edu.ar).*

** *Institut d'Automàtica i Informàtica Industrial, Universitat Politècnica de València, Valencia, España. (e-mail: jpico@ai2.upv.es)*

Abstract: Advances in genetic manipulation have allowed the overexpression of proteins and insertion of circuits in cells. However, the expected behaviour can be altered by the internal competition for the limited amount of cellular resources. In this work we analyse a feedback strategy based on transcription inhibition that aims to reduce the interaction in a two-protein expression system. The results allow interpreting the effects of negative feedback on the steady-state protein levels and how the realizable protein region is affected by the feedback loop.

Keywords: genetic circuits, protein production, feedback structure, dynamical systems

1. INTRODUCTION

One of the most important application of microbial cultures is the production of high-value proteins and other metabolites. To this end, cells are cultivated in controlled conditions where they uptake nutritional substances including carbon and nitrogen sources and minerals. The carbon and energy source is metabolised to obtain energy for growth and maintenance, and the building blocks required for cell division. The cell growth process includes the duplication of DNA and synthesis of a variety of proteins. The machinery involved in the gene expression process has a limited capacity [Gyorgy and Del Vecchio, 2014]. For instance, a growing *Escherichia coli* cell has a finite pool of RNA polymerases (RNAP) in the range of 1,000 - 10,000 units. There is also a finite pool of ribosomes for protein translation (e.g. about 14,000 at a doubling time of 1.0 h) [Milo et al., 2010]. Therefore, the expression of proteins occurs in an environment characterised by a finite pool of cellular resources. Then, the expression of heterologous DNA may lead to a significant reduction in the host cell resources.

It has been shown that the over-expression of the target metabolite may lead to cellular stress and undesired response of inserted genetic circuits [Borkowski et al., 2016, Qian et al., 2017]. The effects produced by limited cellular resources can be described with models that represent the gene expression process (i.e. the transcriptional and translational processes) [Hamadeh and Del Vecchio, 2014, Weiße et al., 2015, Carbonell-Ballestero et al., 2016]. The limited cell capacity can be represented with a finite quantity of the cellular resources involved in the protein formation process, i.e RNAPs for transcription and ribosomes for translation [Gyorgy and Del Vecchio, 2014, Hamadeh and Del Vecchio, 2014]. In [Gyorgy and Del Vecchio, 2014], the attainable region of steady-state protein concentrations was studied and a mathematical description of the realizable region was given. It was also found that the interaction produced by sharing finite cellular resources may result in an unexpected coupling between genetic circuits. In [Carbonell-Ballestero et al., 2016] it was shown that an increase in the demand of cellular resources leads to a reduction in gene expression levels and this can be represented with equations analogous to the ones that govern an electrical circuit with resistive loads. Particularly, the gene expression load associated with one gene was found analogous to one element of a series resistive load. Thus, the overall gene expression load can be represented as the sum of the individual loads.

Engineering may help to mitigate the detrimental effects associated to the metabolic overload. To this end, feedback structures have been proposed to operate in case of stressing conditions [Dragosits et al., 2012], and to reduce toxicity of intermediate metabolites [Liu et al., 2015]. In [Hamadeh and Del Vecchio, 2014] feedback strategies to mitigate the steady-state change in protein expression when a second protein is expressed were presented. Mathematical expressions exhibiting the improvements achieved via the feedback loops were found for the case of circuits with identical parameters. Recently, Shopera and co-workers provided experimental evidence on the improvements of a negative feedback loop for reducing the interaction between two genetic circuits [Shopera et al.,

⋆ This work was supported by funding projects from Agencia Nacional de Promoción Científica y Tecnológica (PICT 2014-2394, 2016-2258), CONICET (PIP 2015-0837), Universidad Nacional de La Plata (PPID-I008, I216) of Argentina; and from MINECO/AEI, EU (DPI2017-82896-C2-1-R).

10.1016/j.ifacol.2019.06.116

Table 1. Model reactions.

#	Reaction	Description
R1	$\mathrm{DNA_i}^* + \mathrm{u_i} \underset{k_i\xi_i}{\overset{\xi_i}{\rightleftharpoons}} \mathrm{DNA_i}$	TF-DNA binding
R2	$\mathrm{DNA_i} + \mathrm{RNAP} \underset{\beta_i}{\overset{\alpha_i}{\rightleftharpoons}} \mathrm{c_i}$	RNAP-DNA binding
R3	$\mathrm{c_i} \xrightarrow{\tau_i} \mathrm{mRNA_i} + \mathrm{DNA_i} + \mathrm{RNAP}$	Transcription
R4	$\mathrm{mRNA_i} \xrightarrow{\delta_i} \varnothing$	mRNA degradation
R5	$\mathrm{mRNA_i} + \mathrm{Rib} \underset{\sigma_i}{\overset{\rho_i}{\rightleftharpoons}} \mathrm{cRib_i}$	Rib-mRNA binding
R6	$\mathrm{cRib_i} \xrightarrow{\pi_i} \mathrm{P_i} + \mathrm{mRNA_i} + \mathrm{Rib}$	Translation process
R7	$\mathrm{cRib_i} \xrightarrow{\overline{a}_i\delta_i} \mathrm{Rib}$	cRib degradation
R8	$\mathrm{P_i} \xrightarrow{\gamma_i} \varnothing$	Protein degradation

2017]. They implemented an indirect feedback strategy based on protein to protein interaction. All these results are promising for the development of robust systems that can operate despite cellular disturbances.

In this work, we analyse the transcription inhibition strategy proposed in [Hamadeh and Del Vecchio, 2014]. New results for obtaining the steady-state value of proteins are determined. Particularly, we focus on a two-circuits system with non-identical parameters that can be used to represent more realistic process conditions (for instance, circuits with different reaction constants, transcription factor level and feedback parameters). The results show how each parameter of the feedback loop affects the steady-state values of the system. This information could be useful for the design of the feedback system. The rest of the work is organised as follows. Section 2 describes the two protein gene expression model. Then, in Section 3 a feedback strategy based on transcription inhibition for reducing the coupling due to finite resources is analysed in detail. The results with the proposed expression are exemplified in Section 4 and finally, concluding remarks are given in Section 5.

2. PROCESS MODEL

2.1 Description

The process consists of two circuits, each producing a protein P_i. The list of model reactions is given in Table 1. The gene expression process begins with a transcription factor (TF) binding to the promoter DNA_i^* to form the promoter complex DNA_i (reaction R1). Binding of a free RNAP with DNA_i yields the complex c_i (reaction R2). Then, as a result of the transcription process, a messenger RNA mRNA_i is produced (reaction R3), whereas its degradation is considered in reaction R4. The binding of mRNA_i and a free ribosome gives the complex cRib_i

Table 2. List of variables.

Name	Description
DNA_i^*	gene i
u_i	transcription factor i
DNA_i	complex (DNA_i^* - u_i)
c_i	complex (RNAP - DNA_i)
mRNA_i	messenger RNA (protein i)
cRib_i	complex (mRNA_i - ribosome)
P_i	protein i
Rib	free ribosome
RNAP	free RNAP

Table 3. Model parameters for system (1)-(2) (from ref. [Hamadeh and Del Vecchio, 2014])

Parameter	Description	Value	Unit
ξ_i	Reaction constant (R1)	1	1/(nM h)
k_i		200	nM
α_i	Reaction constant (R2)	20	1/(nM h)
β_i		6000	1/h
τ_i	Reaction constant (R3)	250	1/h
δ_i	Degradation rate mRNA_i (R4)	10	1/h
ρ_i	Reaction constant (R5)	100	1/(nM h)
σ_i		10^6	1/h
π_i	Reaction constant (R6)	300	1/h
γ_i	Degradation rate P_i (R8)	1	1/h
RNAP_T	total RNAPs	27	nM
Rib_T	total ribosomes	13.5	nM
DNA_{iT}	gene concentration	500	nM

(reaction R5). Finally, after translation a protein P_i is produced and the mRNA and ribosomes are released (reaction R6). Degradation of cRib_i and P_i is considered in reactions R7 and R8, respectively. Under the assumption of mass-action kinetics for the reactions and first-order degradation for mRNAs and proteins, a set of differential equations can be obtained:

$$\dot{\mathrm{DNA}}_i^* = k_i\xi_i\mathrm{DNA}_i - \xi_i\mathrm{DNA}_i^* u_i, \tag{1a}$$

$$\dot{\mathrm{DNA}}_i = (\beta_i+\tau_i)c_i + \xi_i\mathrm{DNA}_i^* u_i - k_i\xi_i\mathrm{DNA}_i - \alpha_i\mathrm{DNA}_i\mathrm{RNAP}, \tag{1b}$$

$$\dot{c}_i = \alpha_i\mathrm{DNA}_i\mathrm{RNAP} - (\beta_i+\tau_i)c_i, \tag{1c}$$

$$\dot{\mathrm{mRNA}}_i = (\sigma_i+\pi_i)c\mathrm{Rib}_i + \tau_i c_i - \delta_i\mathrm{mRNA}_i - \rho_i\mathrm{mRNA}_i\,\mathrm{Rib}, \tag{1d}$$

$$\dot{c\mathrm{Rib}}_i = \rho_i\mathrm{mRNA}_i\,\mathrm{Rib} - (\sigma_i+\pi_i+\overline{a}_i\delta_i)c\mathrm{Rib}_i, \tag{1e}$$

$$\dot{P}_i = \pi_i\, c\mathrm{Rib}_i - \gamma_i P_i, \tag{1f}$$

with $i = 1, 2$. The variables and process parameters are described in Tables 2-3. The finite amount of resources is considered with the following algebraic equations:

$$\mathrm{DNA}_i^* + \mathrm{DNA}_i + c_i = \mathrm{DNA}_{iT}, \tag{2a}$$

$$\mathrm{RNAP} + c_1 + c_2 = \mathrm{RNAP}_T, \tag{2b}$$

$$\mathrm{Rib} + \mathrm{cRib}_1 + \mathrm{cRib}_2 = \mathrm{Rib}_T, \tag{2c}$$

where RNAP_T and Rib_T stand for the resources available for the protein circuits. Notice that the time evolution of free RNAP and ribosomes can be obtained from eqs. (2b) and (2c), respectively.

The following dissociation constants are defined:

$$K_i := \frac{\beta_i+\tau_i}{\alpha_i}, \tag{3a}$$

$$Q_i := \frac{\sigma_i+\pi_i+\overline{a}_i\delta_i}{\rho_i}. \tag{3b}$$

If the binding of RNAP to DNA is strong, then K_i is small. Similarly, the greater the ribosome binding to mRNA, the smaller is Q_i. Typically, binding and unbinding reactions are faster than the transcriptional and translational reactions, then $\tau_i \ll \beta_i$, $\pi_i \ll \sigma_i$ and consequently, $K_i \approx \beta_i/\alpha_i$ and $Q_i \approx \sigma_i/\rho_i$ [Gyorgy and Del Vecchio, 2014].

In order to obtain analytical expressions, the following assumptions are introduced.

Assumption 1. The steady-state concentrations of free RNAPs and ribosomes satisfy $\mathrm{RNAP} \ll K_i$ and $\mathrm{Rib} \ll Q_i$.

Assumption 2. The mRNA-ribosome complex (cRib_i) has the same decay rate as the mRNA molecules, i.e. parameter $\overline{a}_i = 1$.

The assumption 1 is based on experimental data and typical parameter values. Assumption 2 states that no additional protection is given by the ribosomes to the mRNAs being translated.

2.2 Effects of limited resources

It is well-known that limited cellular resources may have an impact in circuit's behaviour, target metabolite yields and microbial growth rate. Particularly, for a two-protein gene expression system the constraints in eq. (2) affect the steady-state values of proteins following a linear relationship that can be expressed as [Gyorgy and Del Vecchio, 2014, Hamadeh and Del Vecchio, 2014, Carbonell-Ballestero et al., 2016]:

$$\mathrm{RNAP}_T\mathrm{Rib}_T = aP_1 + b(u_2)P_2, \tag{4}$$

where a is constant and $b(u_2)$ is function of the transcription factor u_2. This expression can be obtained from eqs. (1), (2c) and taking into account that the term $\mathrm{RNAP}_T \cdot \mathrm{Rib}$ can be written as a linear function of P_1 and P_2 (appendix (A)). According to eq. (4), the expression of a protein is affected by the variation of a second protein. This coupling appears due to the limited cellular resources, i.e. the shared pool of RNAPs and ribosomes. Particularly, the slope of (4) is given by

$$\begin{aligned}\frac{dP_2}{dP_1} &= -\frac{a}{b(u_2)} \\ &= -\frac{\frac{\gamma_1}{\pi_1}\frac{\delta_1 Q_1}{\tau_1} + \frac{\gamma_1}{\pi_1}\mathrm{RNAP}_T}{\frac{\gamma_2}{\pi_2}\frac{\delta_2 Q_2}{\tau_2} + \frac{\gamma_2}{\pi_2}\mathrm{RNAP}_T + \frac{\gamma_2}{\pi_2}\frac{\delta_2 Q_2}{\tau_2}\frac{K_2}{\mathrm{DNA}_{2T}}(1+\frac{k_2}{u_2})}.\end{aligned} \tag{5}$$

One problem of interest is to reduce the degree of interaction between the two gene expression circuits (i.e. to reduce the slope in (4)). This would lead to increase modularity of circuits and reduce the probability of losing the intended circuit's behaviour.

3. ANALYSIS OF A FEEDBACK STRATEGY TO MITIGATE EFFECTS OF RESOURCE LIMITATION

With the aim of reducing the degree of coupling between gene expression circuits a variety of feedback structures have been proposed [Hamadeh and Del Vecchio, 2014, Shopera et al., 2017]. In this section we focus on a feedback loop presented in Hamadeh and Del Vecchio [2014], where each protein inhibits its own transcription. Although it has been shown the improvement of the feedback for reducing the slope in (5), the analysis was presented for circuits with identical parameters. Here, we obtain expressions to deal with the general case. This allows studying more realistic situations, for instance circuits with different parameters values, TFs concentration, or even the case in which only one protein inhibits its own transcription.

In order to obtain analytical expressions, the following variables are introduced:

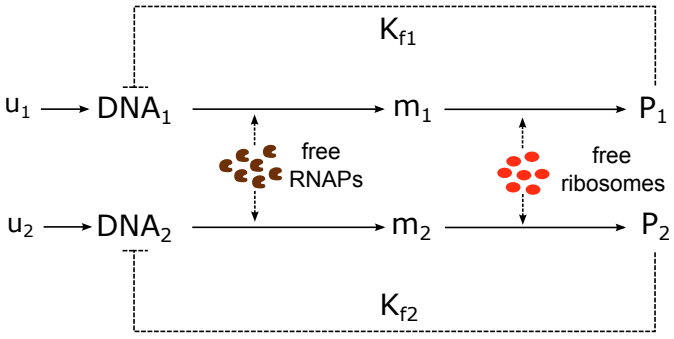


Fig. 1. Feedback structure with each protein inhibiting its own transcription process. The inhibition is implemented with reaction in (7).

$$a_i = \frac{\gamma_i}{\pi_i}\frac{Q_i\delta_i}{\tau_i} + \frac{\gamma_i}{\pi_i}\mathrm{RNAP}_T, \tag{6a}$$
$$b(u_i) = a_i + e_i d_i(u_i), \tag{6b}$$
$$d_i(u_i) = 1 + k_i/u_i, \tag{6c}$$
$$e_i = \frac{\gamma_i}{\pi_i}\frac{Q_i\delta_i}{\tau_i}\frac{K_i}{\mathrm{DNA}_{iT}}, \tag{6d}$$

where $u_i > 0$ is considered.

A schematic representation of the feedback structure is shown in Fig. 1. To consider each protein inhibiting its own transcription, the following reaction can be incorporated to the system described in Table 1

$$\mathrm{DNA_i} + \mathrm{P_i} \underset{\lambda_{ir}}{\overset{\lambda_{if}}{\rightleftharpoons}} \mathrm{I_{pi}}, \tag{7}$$

where I_{pi} is an intermediate product of the binding process. This strategy reduces the available DNA_i for binding with the free RNAP (reaction R2). Now the constraint for DNA conservation (eq. (2a)) is modified with the term $+I_{pi}$ to take into account the DNA_i-P_i binding. From (7), the dynamics of I_{pi} is given by:

$$\dot{I}_{pi} = \lambda_{if}\mathrm{DNA}_i P_i - \lambda_{ir}\mathrm{I}_{pi}, \tag{8}$$

and the dynamics of DNA_i and P_i (eqs. (1b) and (1f)) are modified accordingly with the terms $-\lambda_{if}\mathrm{DNA}_i P_i + \lambda_{ir}\mathrm{I}_{pi}$. Then, the steady-state relationship given in (4) is modified and can be written as follows:

$$\mathrm{RNAP}_T\mathrm{Rib}_T = a_1P_1 + b(u_2)P_2 + e_2K_{f2}P_2^2, \tag{9}$$

with the feedback parameter defined as $K_{fi} = \lambda_{if}/\lambda_{ir}$ (nM^{-1}) [Hamadeh and Del Vecchio, 2014].

The degree of coupling between the circuits is reduced as the feedback is incremented. Taking derivatives in (9) it follows that

$$\frac{dP_2}{dP_1} = \frac{-a_1}{b(u_2) + 2e_2K_{f2}P_2}. \tag{10}$$

Differing from eq. (5) where the slope is constant, in eq. (10) the slope depends not only on u_2 but also on P_2.

3.1 Determination of steady-state protein concentrations

Proposition 1. Given $u_i > 0$, the steady-state values of proteins in (1) subject to constraints (2) with the transcription inhibition feedback given by (7) can be obtained from the following equations:

$$\mathrm{RNAP}_T\mathrm{Rib}_T = a_1P_1 + b(u_2)P_2 + e_2K_{f2}P_2^2, \tag{11a}$$
$$\mathrm{RNAP}_T\mathrm{Rib}_T = a_2P_2 + b(u_1)P_1 + e_1K_{f1}P_1^2. \tag{11b}$$

Then, the states of (1) can be calculated from:

$$\mathrm{DNA}_i \approx \frac{\mathrm{DNA}_{iT}\, v_i}{1 + K_{fi} P_i v_i}, \tag{12a}$$

$$\mathrm{DNA}_i^* = \frac{k_i}{u_i}\mathrm{DNA}_i, \tag{12b}$$

$$\mathrm{mRNA}_i \approx \frac{\tau_i c_i}{\delta_i}, \tag{12c}$$

$$c\mathrm{Rib}_i = \frac{\mathrm{mRNA}_i}{Q_i}\mathrm{Rib}, \tag{12d}$$

with $v_i = u_i/(u_i + k_i)$ and

$$\mathrm{RNAP} = \frac{\mathrm{RNAP}_T}{1 + \frac{\mathrm{DNA}_1}{K_1} + \frac{\mathrm{DNA}_2}{K_2}}, \tag{13a}$$

$$c_i = \frac{\mathrm{DNA}_i}{K_i}\mathrm{RNAP}, \tag{13b}$$

$$Rib = \frac{\mathrm{Rib}_T}{1 + \frac{\mathrm{mRNA}_1}{Q_1} + \frac{\mathrm{mRNA}_2}{Q_2}}, \tag{13c}$$

$$I_{pi} = K_{fi}\mathrm{DNA}_i P_i. \tag{13d}$$

Proof 1. Eqs. (12)-(13) can be obtained from equating to zero the system dynamics. Particularly, in (12a) Assumption 1 is considered and (12c) follows from the dynamics of $\mathrm{mRNA}_i(t) + \mathrm{cRib}_i(t)$. Then, after algebraic manipulation, the ratio P_2/P_1 at steady-state can be written as

$$\frac{P_2}{P_1} = \frac{e_1}{e_2}\frac{v_2}{(1 + K_{f2} P_2 v_2)}\frac{(1 + K_{f1} P_1 v_1)}{v_1}. \tag{14}$$

By using $v_i = 1/d_i$ it follows that:

$$e_2 K_{f2} P_2^2 + e_2 d_2(u_2) P_2 = e_1 K_{f1} P_1^2 + e_1 d_1(u_1) P_1. \tag{15}$$

Finally, by replacing (11a) in (15), eq. (11b) is obtained. □

As a particular result of Proposition 1, when both circuits have the same set of parameters with $u_i \gg k_i$, then $d_i \approx 1$ and $e_1 = e_2$. Consequently, eq. (15) implies that $P_1 = P_2 = P$ and the protein values follow from the solution of eqs. (11a) and $P_1 = P_2$.

3.2 Description of the realizable protein region under negative feedback

The realizable region of proteins of system (1) subject to resource limitation can be bounded with the intersection of sets S_i [Gyorgy and Del Vecchio, 2014]:

$$S_i := \left\{ p \,|\, p \geq 0,\ \frac{p_i}{p_i^{max}} + \sum_{j=1 j\neq i}^{2} \frac{p_j}{p_j^{\infty}} < 1 \right\}, \tag{16}$$

where, for the constraints given in (2), the parameters of eq. (16) can be written as

$$p_i^{\infty} = \frac{\mathrm{RNAP}_T \mathrm{Rib}_T}{a_i}, \tag{17a}$$

$$p_i^{max} = \frac{\mathrm{RNAP}_T \mathrm{Rib}_T}{a_i + e_i}. \tag{17b}$$

Eq. (16) describes subsets of $\mathbb{R}^2$ with lines as upper bounds. Then, the intersection set is defined as

$$S = S_1 \cap S_2. \tag{18}$$

This set represents a bound for the realizable region for proteins P_1,P_2 in case of no feedback [Gyorgy and Del Vecchio, 2014]. Now, when the feedback strategy of Fig. 1 is implemented, for a given (u_1, u_2) the realizable region is contracted as K_{fi} is increased.

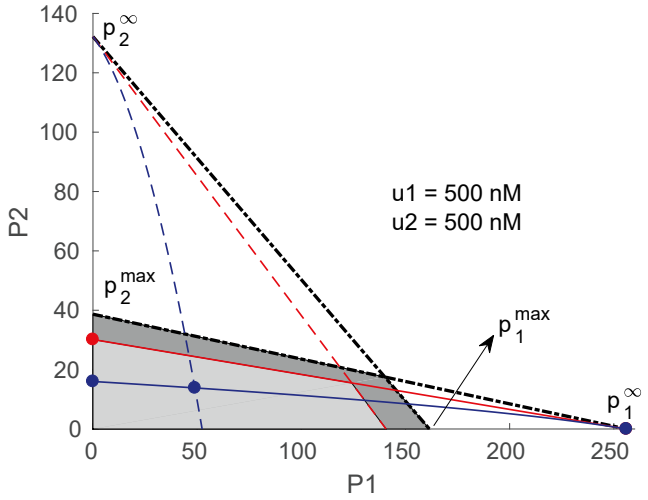


Fig. 2. Realizable protein region for system parameters in Table 3 with $\alpha_2 = \alpha_1/2$, $\rho_2 = \rho_1/2$, $\mathrm{DNA}_{2T} = \mathrm{DNA}_{1T}/2$ and $u_i = 500$ nM: $S_{(u_1,u_2,K_{f1},K_{f2})}$ with $K_{fi} = 0$ (grey), S (dark grey). The solution corresponding to $K_{fi} = 0.1$ nM^{-1} is the intersection of the curves drawn in blue color.

Proposition 2. The intersection of the areas under the curves (11a)-(11b) defines a convex set $S_{(u_1,u_2,K_{f1},K_{f2})}$ that fulfills $S_{(u_1,u_2,K_{f1},K_{f2})} \subseteq S$.

Proof 2. As a function of P_1, eqs. (11a)-(11b) are concave functions. Then, the area below each curve (i.e the hypograph) is a convex set and its intersection with the first quadrant ($\mathbb{R}^2_+$) defines the convex set $S_{(u_1,u_2,K_{f1},K_{f2})}$. Besides, it can be shown that for (11a)-(11b) the parameters p_1^{∞}, p_2^{∞} are the same as in eq. (17a). However, the values of p_1^{max}, p_2^{max} are reduced as K_{f1}, K_{f2} are increased (due to concavity of the functions). Then, the intersection area below the curves is in S. □

If the TF concentration is not large (e.g. $u_i \approx k_i$) and $K_{fi} = 0$ is considered, then $S_{(u_1,u_2,0,0)} \subset S$. This is due to the fact that

$$p_{iU}^{max}(u_1, u_2) = \frac{\mathrm{RNAP}_T \mathrm{Rib}_T}{a_i + e_i d_i(u_i)} < p_i^{max}. \tag{19}$$

On the contrary, $p_{iU}^{max}(u_1, u_2) \to p_i^{max}$ when $u_i \gg k_i$ (i.e. in case of full induction). Then, the feedback system is realizable (i.e. $K_{fi} \geq 0$ can be selected) when the desired proteins values (p_1^*, p_2^*) are in S.

From (11a)-(11b) the feedback parameters K_{f1}, K_{f2} can be obtained for a realizable point $(p_1^*, p_2^*) \in S_{(\overline{u}_1,\overline{u}_2,0,0)}$ where $\overline{u}_i$ stands for the maximum TF value:

$$K_{f1} = \frac{\mathrm{RNAP}_T \mathrm{Rib}_T - (a_1 + e_1 d_1) p_1^* - a_2 p_2^*}{e_1 p_1^{*2}}, \tag{20a}$$

$$K_{f2} = \frac{\mathrm{RNAP}_T \mathrm{Rib}_T - (a_2 + e_2 d_2) p_2^* - a_1 p_1^*}{e_2 p_2^{*2}}. \tag{20b}$$

As stated above when large TF concentration is considered then $\overline{u}_i \gg k_i$ and $d_i(u_i) = 1$ can be replaced.

As an example, the set $S_{(u_1,u_2,0,0)}$ is depicted in Fig. 2 for the parameters of Table 3, where the coefficients α_2, ρ_2 and DNA_{2T} were modified (see caption of the figure). Notice

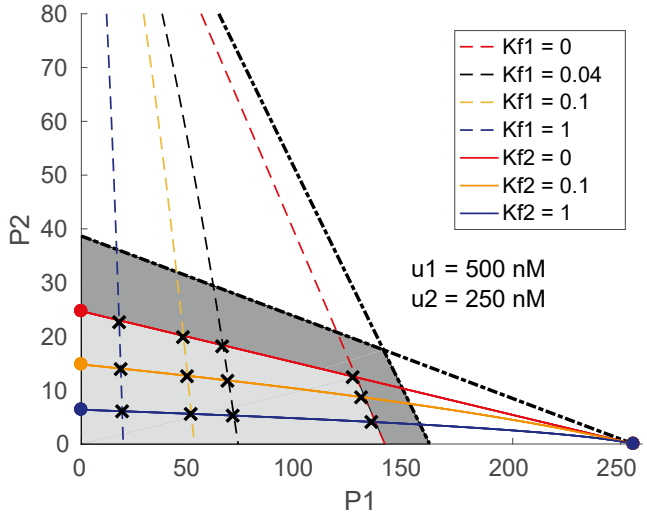


Fig. 3. Steady-state values of proteins estimated with the curves in Proposition 1 for different values of the feedback parameters K_{f1}, K_{f2}.

that for the illustrated case of $u_i = 500$ nM with $k_i = 200$ nM, it is observed that $S_{(u_1,u_2,0,0)} \subset S$. Also, the curves corresponding to the particular solution of $K_{fi} = 0.1$ nM^{-1} are shown in blue color.

3.3 Extension to a system with three proteins

Although a two-protein system has been studied, the previous results can be extended. For example, in a system with three genes the constraints in (2b) and (2c) can be written as $\text{RNAP}+\sum_i^3 c_i = \text{RNAP}_T$ and $\text{Rib}+\sum_i^3 \text{cRib}_i = \text{Rib}_T$, respectively. Then, new expressions of the form of eq. (11) can be obtained:

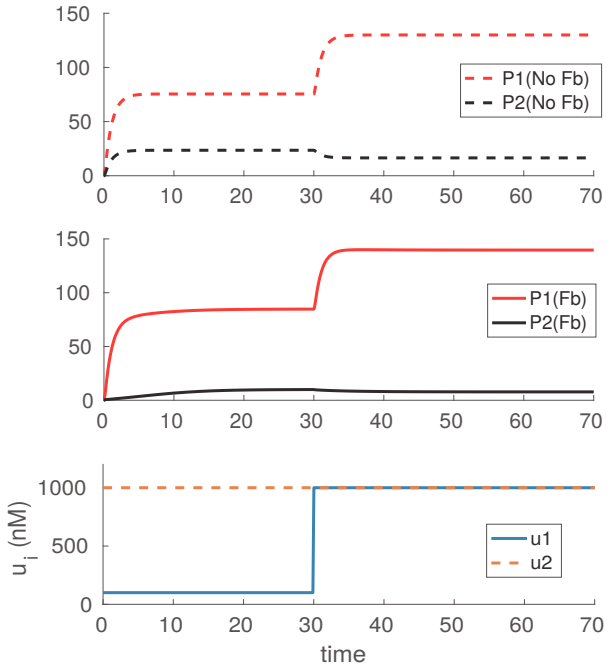
$$\text{RNAP}_T\text{Rib}_T = b(u_j)P_j + e_j K_{fj} P_j^2 + \sum_{i=1,i\neq j}^{3} a_i P_i, \quad (21)$$

for $j = 1, 2, 3$. Eq. (21) corresponds to surfaces in $\mathbb{R}^3$ and can be used to estimate the steady-state protein concentrations and to analyze the effects when a third protein is expressed.

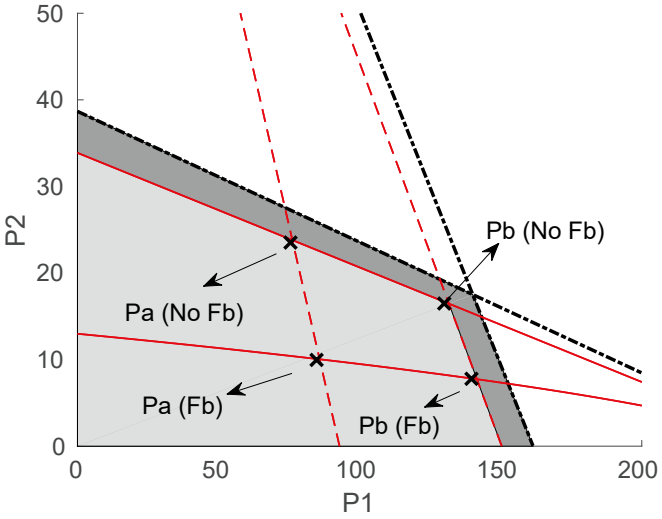
4. EXAMPLES

Numerical results of the eqs. obtained in Section 3 are presented. In order to consider different parameter values, the second circuit has the following modification $\alpha_2 = \alpha_1/2$, $\rho_2 = \rho_1/2$, $\text{DNA}_{2T} = \text{DNA}_{1T}/2$. This results in a circuit 2 with the following dissociation constants $K_2 = 2K_1$, $Q_2 = 2Q_1$.

Fig. 3 presents a comparison between the steady-state protein values of system (1)-(2) (cross marks) and the estimated values with Proposition 1 for the TFs concentrations $u_1 = 500$ nM, $u_2 = 250$ nM. The sets $S_{(u_1,u_2,K_{f1},K_{f2})}$ and S are filled in light-grey and dark-grey color, respectively. A good agreement is observed between the proteins concentrations and the values obtained with the proposed curves. This figure can be used to explain the effect of the feedback parameters for constant concentrations of TFs. In case of $K_{fi} = 0$ curves (11a)-(11b) are straight lines that join the points p_i^{∞} and $p_{iU}^{max}(u_1, u_2)$. If K_{f1} is increased,



(a) Time evolution of proteins and steps applied in u_i with feedback (Fb) and without feedback (No Fb).



(b) Curves (11a)-(11b) for fixed feedback parameters and different TFs concentration. Protein concentrations before the step change (Pa) and at t = 70 h (Pb) are represented with cross marks.

Fig. 4. Example of the two protein system with feedback parameters $K_{f1} = 0$ and $K_{f2} = 0.2$ nM^{-1}.

the new steady-state point moves to the left following the corresponding curve for $K_{f1} > 0$ (dotted lines). Since the resources for the second circuit are increased, the value of P_2 also increases following the curve corresponding to K_{f2} (solid lines). Similarly, when the parameter K_{f2} is increased, then P_2 is reduced while P_1 increases following the curve of K_{f1}. Thus, these tuning parameters affect the protein values and the coupling that exists between the circuits due to the shared resources (RNAPs, ribosomes).

Additionally, the range of proteins obtained with feedback parameters K_{f1}, K_{f2} can be analysed in case of variations in the TFs (u_1, u_2). Fig. 4 compares a system without feed-

back with one where only the second expression circuit has the feedback loop described in eq. (7). In this case, P_2 is the protein to be regulated despite variations of P_1. The time evolution of the input and output variables is presented in Fig. 4(a). A step change from $(u_1, u_2) = (100, 1000)$ nM to $(u_1, u_2) = (1000, 1000)$ nM was applied at t = 30 h. As expected, the variation of P_2 in the fedback system (black-solid line) was smaller when compared with the system without feedback (black-dashed line), despite the variation in P_1 is similar in both cases (red lines). This situation is also represented in the P_1, P_2 coordinates (Fig. 4(b)). In this plot, P_a stands for the proteins concentrations before the step change and P_b represents the values at t = 70 h. The curves of Prop. 1 are plotted for constant values of K_{fi} and both levels of (u_1, u_2).

According to (10), the interaction is reduced as K_{f2} is incremented. Also, the improvement increases with the dissociation constant K_2. Then, it is expected that the uncoupling will be more effective in scenarios where binding between RNAPs and gene promoters is not so strong (high K_2). The proposed equations have shown the variables involved in the gene expression process and could be used for designing the feedback strategy in cases with non-identical model parameters. From eq. (20), in case of full induction only one set of K_{f1}, K_{f2} can be found for a given point (p_1^*, p_2^*). Then, in order to design the feedback parameters for a given (p_1^*, p_2^*) with a desired degree of interaction (eq. (10)), a design involving not only the K_{fi} but also the dissociation constants can be considered.

5. CONCLUSION

In this work we proposed new expressions to determine the steady-state protein values for a feedback gene expression system. The proposal focused on the general case of non-identical circuit parameters. Since each curve found in Prop. 1 depends only on one feedback parameter, it can be used to explain how the feedback affect the realizable region of proteins.

Future work will address a detailed analysis of systems with an arbitrary number of genes and the simultaneous design of the feedback and process parameters for obtaining a desired steady-state value with a predefined coupling.

Appendix A. AUXILIARY CALCULATIONS

The steady-state value of the product $\mathrm{RNAP}_T \cdot \mathrm{Rib}$ depends linearly on the values of P_1 and P_2. This follows from eq. (2b):

$$\mathrm{RNAP}_T = c_1 + c_2 + \mathrm{RNAP}, \tag{A.1a}$$

$$\mathrm{RNAP}_T = \frac{\mathrm{mRNA}_1\delta_1}{\tau_1} + \frac{\mathrm{mRNA}_2\delta_2}{\tau_2} + \frac{c_2 K_2}{\mathrm{DNA}_2}, \tag{A.1b}$$

$$\mathrm{RNAP}_T = \frac{\mathrm{mRNA}_1\delta_1}{\tau_1} + \frac{\mathrm{mRNA}_2\delta_2}{\tau_2}\left(1 + \frac{K_2 d_2}{\mathrm{DNA}_{2T}}\right), \tag{A.1c}$$

$$\mathrm{RNAP}_T \cdot \mathrm{Rib} = \frac{\gamma_1 P_1}{\pi_1}\frac{Q_1\delta_1}{\tau_1} + \frac{\gamma_2 P_2}{\pi_2}\frac{Q_2\delta_2}{\tau_2}\left(1 + \frac{K_2 d_2}{\mathrm{DNA}_{2T}}\right). \tag{A.1d}$$

The last equality was obtained by using $\mathrm{mRNA}_i\mathrm{Rib} = \mathrm{cRib}_i Q_i$ and $\mathrm{cRib}_i = \gamma_i P_i/\pi_i$.

REFERENCES

Olivier Borkowski, Francesca Ceroni, Guy-Bart Stan, and Tom Ellis. Overloaded and stressed: whole-cell considerations for bacterial synthetic biology. *Current Opinion in Microbiology*, 33:123 – 130, 2016. ISSN 1369-5274.

M. Carbonell-Ballestero, E. Garcia-Ramallo, R. Montañez, C. Rodriguez-Caso, and J. Macía. Dealing with the genetic load in bacterial synthetic biology circuits: convergences with the Ohm's law. *Nucleic Acids Res*, 44 (1):496–507, 2016. ISSN 0305-1048.

Martin Dragosits, Daniel Nicklas, and Ilias Tagkopoulos. A synthetic biology approach to self-regulatory recombinant protein production in *Escherichia coli*. *Journal of Biological Engineering*, 6(1):2, 2012. ISSN 1754-1611.

A. Gyorgy and D. Del Vecchio. Limitations and trade-offs in gene expression due to competition for shared cellular resources. In *53rd IEEE Conference on Decision and Control*, pages 5431–5436, 2014.

A. Hamadeh and D. Del Vecchio. Mitigation of resource competition in synthetic genetic circuits through feedback regulation. In *53rd IEEE Conference on Decision and Control*, pages 3829–3834, Dec 2014.

Di Liu, Yi Xiao, Bradley S. Evans, and Fuzhong Zhang. Negative feedback regulation of fatty acid production based on a malonyl-coa sensor actuator. *ACS Synthetic Biology*, 4(2):132 – 140, 2015. PMID: 24377365.

Ron Milo, Paul Jorgensen, Uri Moran, Griffin Weber, and Michael Springer. Bionumbers - the database of key numbers in molecular and cell biology. *Nucleic Acids Research*, 38(suppl 1):D750–D753, 2010.

Yili Qian, Hsin-Ho Huang, Jose Jimenez, and Domitilla Del Vecchio. Resource competition shapes the response of genetic circuits. *ACS Synthetic Biology*, 6(7):1263–1272, 2017. PMID: 28350160.

Tatenda Shopera, Lian He, Tolutola Oyetunde, Yinjie J. Tang, and Tae Seok Moon. Decoupling resource-coupled gene expression in living cells. *ACS Synthetic Biology*, 6(8):1596–1604, 2017. PMID: 28459541.

Andrea Y. Weiße, Diego A. Oyarzún, Vincent Danos, and Peter S. Swain. Mechanistic links between cellular trade-offs, gene expression, and growth. *Proceedings of the National Academy of Sciences*, 112(9):E1038–E1047, 2015.