Submitted, 2013 American Control Conference (ACC)
http://www.cds.caltech.edu/~murray/papers/gfm13-acc.html

Feedback architectures to regulate flux of components in artificial gene networks

Giulia Giordano, Elisa Franco and Richard M. Murray

***Abstract*— This paper focuses on RNA flux regulation for *in vitro* synthetic gene networks and considers architectures that can be scaled to an arbitrary number of species. Feedback loops are designed based on negative auto–regulation (which can minimize the potentially harmful amount of molecules not used to form useful products) and cross–activation (which can maximize the overall output flux): transcription rate matching can be achieved through proper feedback constants; negative feedback is faster and maintains stability. A possible experimental implementation of a three and four genes negative feedback architecture is also numerically studied.**

I. Introduction

Synthetic biology aims at designing from the bottom–up new biological circuits with specific functionalities, in order to devise innovative biotechnologies. Building new circuits also offers a powerful insight into the design principles present in nature and selected by evolution. Design and synthesis of artificial biotechnologies can be greatly streamlined by quantitative modeling, which provides rational explanations and quantitative assessment of the system performance; a control–theoretic approach is very powerful to investigate dynamic properties and robustness [1], [2].

In vitro synthetic gene networks have been recently proposed in [3], [4]: the activity of artificial, short DNA genes ("genelets") is regulated by their RNA outputs, through displacement of key activating strands bound to the genes. Thus, unlike *in vivo* transcriptional control, regulation is mediated by RNA species rather than by proteins. These networks are translation free (i.e. no proteins are produced) and are built with few biochemical components (DNA, RNA, two protein species off-the-shelf, and a well defined set of buffer reagents), but can exhibit by design complex behaviors such as bistability [3], [5] and oscillations [4], [6].

Binding of proteins and RNA underlies cell metabolism, gene expression and self–assembly phenomena. Synthetic biological systems also rely on the accuracy of programmed binding pathways among biological components. In many instances, binding of reagents has to occur with specific stoichiometric ratios: therefore, it is important to regulate production and degradation rates, i.e. the overall flux, of biochemical species, so that their concentrations fall within desired bounds. Since they can generate many complex behaviors, *in vitro* transcriptional circuits are a promising toolkit to control dynamics in molecular machines [7], [8], patterns [9] and computers [10]. Thus, we need scalable flux control architectures tailored to these synthetic gene networks. Feedback circuits regulating RNA transcription rates were first proposed in [11], [12], [2]. Suppose two RNA species bind in a 1:1 stoichiometry to form an output product: if their production/degradation rates are not matched, the reagent with the higher flux accumulates, creating a potentially harmful excess, and the flow of product is limited by the lower reagent flux. To equate the two production rates, negative feedback (self–repression) and positive feedback (cross–activation) architectures were proposed.

The objective of this paper is to systematize and scale up these two schemes, understanding their performance in a context where n genes operate together to produce outputs in a desired stoichiometry. Our main contributions are: 1) We categorize different interconnection and feedback architectures for the RNA outputs of the synthetic genes. This categorization is useful in a scenario where these RNA outputs will be employed as building blocks for complex nanostructures or as inputs for downstream circuits. 2) We numerically analyze the performance of these different architectures for schemes with 3 and 4 interconnected genelets, using plausible experimental parameters. Feedback occurs by stoichiometric interactions between RNA outputs and genelets. We find that negative feedback architectures are more scalable and respond with faster timescales. 3) We propose a viable implementation for a negative feedback architecture involving three and four genes.

Section II introduces artificial gene networks and our categorization of different interconnections; our proposed negative and positive feedback regulation is numerically analyzed. In Section III we describe a synthetic gene network implementation of a negative feedback scheme and numerically analyze the corresponding performance.

II. Feedback architectures to regulate production rates in synthetic gene networks

A. Artificial gene networks

Here we introduce artificial *in vitro* gene networks (transcriptional circuits) [3], with the support of Figure 1. Interactions among nucleic acids (DNA and RNA) can be programmed by choosing their sequences (ATCG); strand domains (subsequences of bases) having a particular function, Figure 1 (a), are identified with a specific color (once domain interactions are chosen, automated software tools

Giulia Giordano is with the University of Udine, Department of Mathematics and Computer Science, 33100 Udine, Italy. giordano.giulia@spes.uniud.it

Elisa Franco is with the University of California at Riverside, Department of Mechanical Engineering, Riverside, CA 92521, USA. efranco@engr.ucr.edu

Richard M. Murray is with the California Institute of Technology, Division of Engineering and Applied Sciences, Pasadena, CA 91125, USA. murray@cds.caltech.edu

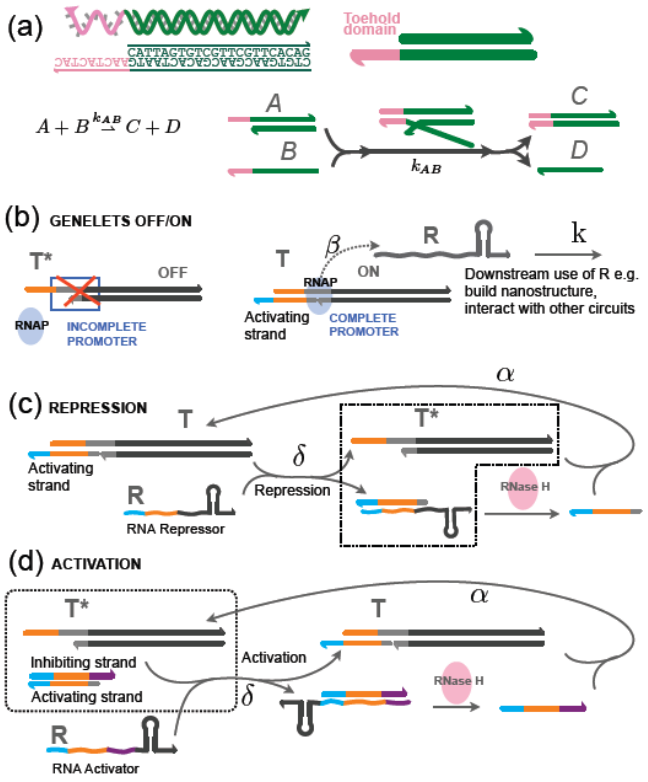


Fig. 1: (a) Domain representation of nucleic acids and branch migration (b) On/off state of genelets (c) RNA-mediated repression (d) RNA-mediated activation.

can be used to find optimal sequences [10]). The arrow on every strand represented in Figure 1 represents the $5'$ end. Reactions in nucleic acid systems occur by hybridization (two single stranded, complementary nucleic acids bind to form a double stranded complex) and by toehold-mediated branch migration [13], exemplified in Figure 1 (a): species A and B interact through the exposed pink domain and switch to a new, thermodynamically more favorable configuration, creating species D and C. The reaction speed is determined by the length of the toehold domain and is typically tunable in the range 10–10^6/M/s for 1–8 bases toeholds.

Figure 1 (b) introduces artificial genelets; synthetic DNA templates are copied (transcribed) into RNA using T7 RNA polymerase (RNAP); if the RNAP binding region, called promoter, is incomplete (partially single stranded), the genelet is off (T^*). When the double stranded region is reformed by the appropriate DNA activating strand, the genelet is on (T) and RNA output R is produced. The total amount of a genelet is constant, i.e. $[\mathrm{T}] + [\mathrm{T}^*] = [\mathrm{T}^{\mathrm{tot}}]$. We assume the enzyme operates in a linear regime, thus R is produced at rate β: $\mathrm{T} \xrightarrow{\beta} \mathrm{R} + \mathrm{T}$. Output R can be used downstream at a rate k, for instance to interact with other RNA species in circuit dynamics or to form nanostructures [7].

In the next paragraphs, we describe repression and activation of genelets with a set of aggregate reactions; we obtain intuitive models that bear relevance to general molecular networks. Detailed models are in Section III.

In Figure 1 (c) we show how a genelet can be repressed by an RNA species R. This pathway is at the basis of our negative feedback circuits. By design, R displaces part of the promoter in the activating strand through toehold-mediated branch migration: $\mathrm{R} + \mathrm{T} \xrightarrow{\delta} \mathrm{T}^*$, where δ is proportional to the length of the toehold domain (cyan domain). We lump the species in the dashed box into species T^*. The inactive gene T^* reverts at rate α to its active form T thanks to the action of RNase H, an enzyme which degrades RNA in RNA/DNA duplexes; R is degraded and the activating strand binds again to the template forming T^*.

In Figure 1 (d) we show how a genelet can be activated by an RNA species R. This pathway is at the basis of our positive feedback designs. In this case, species T^* (dashed box) is comprised of the inactive genelet and a DNA inhibitor-activator complex (the activating strand is sequestered by design). Again, by suitable domain design, R releases the activating–strand through toehold (violet domain) mediated displacement: thus, T^* is converted into T with rate δ (proportional to the toehold length). Now two species coexist: active template T and the complex $\mathrm{R}\cdot$(DNA inhibiting strand). Again, the active gene T reverts to its inactive form T^* at rate α, thanks to the action of RNase H, which releases the inactivating DNA strand.

B. *Output interconnection for n genes*

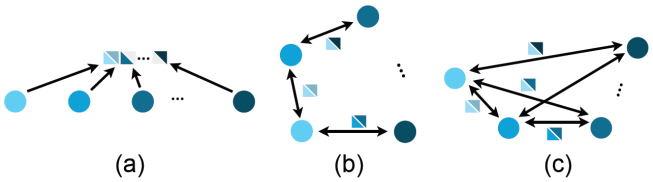


Fig. 2: Output connection schemes with n genes: circles represent genes, triangles represent their RNA outputs. (a) single product, (b) neighbor, (c) handshake.

We consider a set of n genelets and different interaction scenarios. We say that genelets are interconnected when their RNA outputs bind to form one or more products.

A *single product* interconnection occurs when a single RNA complex (for instance, a large nanostructure) is formed from the simultaneous interaction of all RNA species, as shown in the scheme of Figure 2 (a).

A network of genelets may be designed to produce different subcomponents, that may later assemble into a larger product. In this scenario, we can take two extreme cases: 1) Each RNA participates in at most two subcomponents: we identify this case as a *neighbor* interconnection, as shown in Figure 2 (b); 2) Each RNA participates in the creation of $n-1$ subcomponents: we identify this case as a *handshake* interconnection, as shown in Figure 2 (c).

We note that in these three cases, all RNA outputs are used in the same number of complexes. Thus, we would like them to be produced and degraded at comparable rates, given that their downstream utilization is the same. We introduce feedback in these circuits, to compensate imbalances in the

concentration of templates and match the transcription rate of the RNA outputs. We will present architectures based on negative and positive feedback, which scale up previously proposed two–gene networks [11], [12].

In the following, using mass action kinetics we derive ordinary differential equation (ODE) models for the different interconnection cases above. We use the MATLAB `ode23` routine to solve the nonlinear ODEs. We focus on $n = 3$ genes: in this case, the neighbor and handshake connection coincide; we consider $n = 4$ for the negative feedback architecture.

C. Parameters

Here we report the parameter values used in the numerical analysis that follows. Parameters were chosen consistently with the literature to represent synthetic gene network dynamics. For the handshake/neighbor connection we used $k_{ij} = 2 \cdot 10^3$ /M/s and for single product connection $k = 6 \cdot 10^3$/M/s. Negative feedback rates were $\delta_i = 5 \cdot 10^3$ /M/s, positive feedback rates $\delta_{ij} = 50$ /M/s. In all cases $\alpha_i = 3 \cdot 10^{-4}$ /s, $\beta_i = 1 \cdot 10^{-2}$ /s. The total genelets amounts are: $[\mathrm{T}_1^{\mathrm{tot}}] = 100$ nM, $[\mathrm{T}_2^{\mathrm{tot}}] = 200$ nM, $[\mathrm{T}_3^{\mathrm{tot}}] = 300$ nM; $[\mathrm{T}_4^{\mathrm{tot}}] = 150$ nM for the 4 genes simulations.

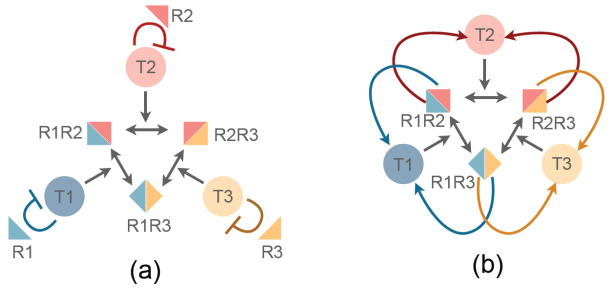


Fig. 3: Handshake/neighbor interconnection for $n = 3$, (a) Negative feedback, (b) Positive feedback

D. Negative feedback architectures

Negative feedback is implemented as follows: when an RNA output species is in excess relative to the effectively used amount (i.e. RNA output bound to other RNA species to form a product), it down–regulates its own production rate. Inactive genes T_i^* spontaneously revert to the active state at rate α_i, $\mathrm{T}_\mathrm{i}^* \overset{\alpha_i}{\rightharpoonup} \mathrm{T}_\mathrm{i}$ and $[\mathrm{T}_\mathrm{i}^{\mathrm{tot}}] = [\mathrm{T}_\mathrm{i}] + [\mathrm{T}_\mathrm{i}^*]$ (see Section II-A). Negative feedback occurs by self-repression: $\mathrm{R_i} + \mathrm{T_i} \overset{\delta_i}{\rightharpoonup} \mathrm{T}_\mathrm{i}^*$, where δ_i is the strength of the negative feedback. Regardless of the output interconnection the template dynamics are:

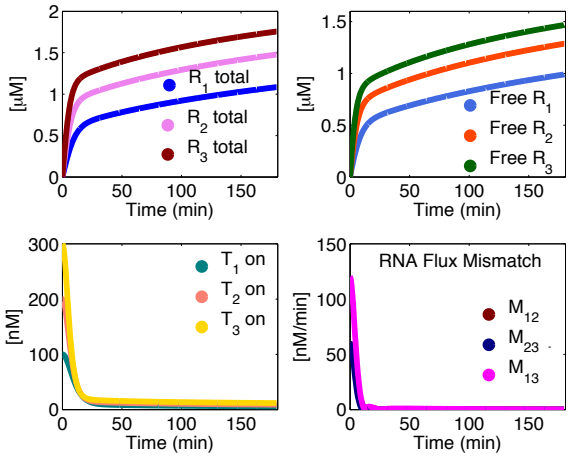
$$\frac{\mathrm{d}[\mathrm{T_i}]}{\mathrm{dt}} = \alpha_\mathrm{i}\left([\mathrm{T}_\mathrm{i}^{\mathrm{tot}}] - [\mathrm{T_i}]\right) - \delta_\mathrm{i}\,[\mathrm{R_i}][\mathrm{T_i}], \quad i = 1, ..., n \quad (1)$$

1) Single product: a single product P is produced by the simultaneous interaction of n RNA outputs: $\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{R_i} \overset{k}{\rightharpoonup} \mathrm{P}$:

$$\frac{\mathrm{d}[\mathrm{R_i}]}{\mathrm{dt}} = \beta_\mathrm{i}\,[\mathrm{T_i}] - \delta_\mathrm{i}\,[\mathrm{R_i}][\mathrm{T_i}] - \mathrm{k}\prod_{\mathrm{i}=1}^{\mathrm{n}}[\mathrm{R_i}], \qquad \frac{\mathrm{d}[\mathrm{P}]}{\mathrm{dt}} = \mathrm{k}\prod_{\mathrm{i}=1}^{\mathrm{n}}[\mathrm{R_i}] \quad (2)$$

with $[\mathrm{R}_\mathrm{i}^{\mathrm{tot}}] = [\mathrm{R_i}] + [\mathrm{T}_\mathrm{i}^*] + [\mathrm{P}]$.

(a) single product, $n = 3$



(b) single product, $n = 4$

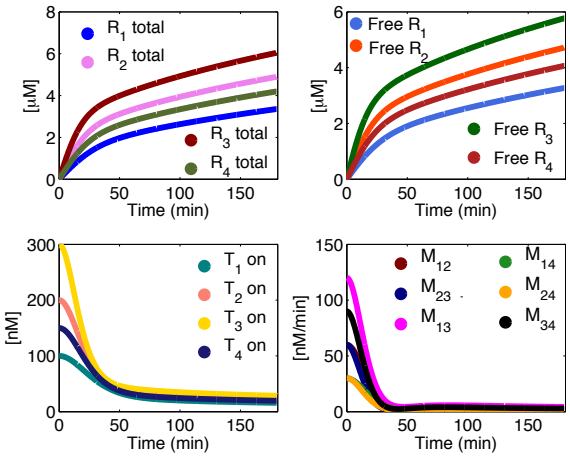


Fig. 4: Simulations: single product, negative autoregulation.

Figure 4 shows the numerical solution to the ODEs for $n = 3$ and $n = 4$. Even though we have different total amounts of genes, the flux mismatches (bottom right panel), namely the differences in absolute value between any two production rates, considerably reduce with a fast time response. The single product interconnection is slower in the four genes case; this interconnection also leads to a significant amount of free RNA, which can be considered waste because it is not used in the product formation.

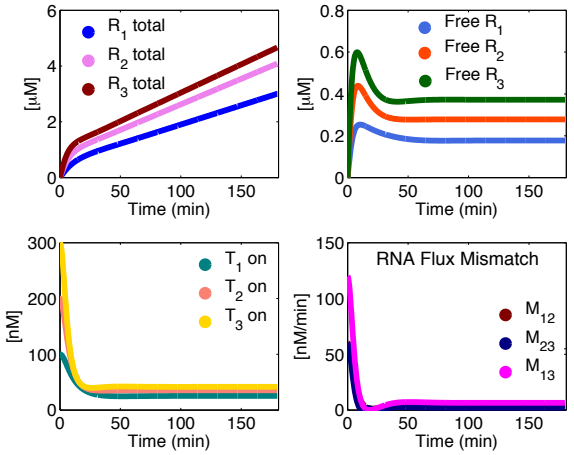
2) Handshake and neighbor connection: each pair–wise product $\mathrm{P_{ij}}$ is generated at rate k_{ij}: $\mathrm{R_i} + \mathrm{R_j} \overset{k_{ij}}{\rightharpoonup} \mathrm{P_{ij}}$ and thus

$$\frac{\mathrm{d}[\mathrm{R_i}]}{\mathrm{dt}} = \beta_\mathrm{i}\,[\mathrm{T_i}] - \delta_\mathrm{i}\,[\mathrm{R_i}][\mathrm{T_i}] - \sum_\mathrm{j} \mathrm{k_{ij}}\,[\mathrm{R_i}][\mathrm{R_j}],$$
$$\frac{\mathrm{d}[\mathrm{P_{ij}}]}{\mathrm{dt}} = \mathrm{k_{ij}}\,[\mathrm{R_i}][\mathrm{R_j}], \quad [\mathrm{R}_\mathrm{i}^{\mathrm{tot}}] = [\mathrm{R_i}] + [\mathrm{T}_\mathrm{i}^*] + \sum_\mathrm{j}[\mathrm{P_{ij}}]. \quad (3)$$

In the handshake case $i, j = 1, ..., n$, $j \neq i$; in the neighbor case $i = 1, ..., n$, $j = i-1, i+1$ and when $i = 1$, $i-1 = n$,

when $i = n$, $i+1 = 1$, to close the loop. Figure 3 (a) shows a schematic representation for the case $n = 3$, when the two connections coincide. Figure 5 shows the numerical solution to the ODEs for $n = 3$ and for $n = 4$ in the handshake connection case. As for the single product formation, even though we have different total amounts of genes, the flux mismatches (bottom right panel) are considerably reduced with a fast time response; moreover, the response for $n = 4$ is faster with this kind of connection and there is much less waste.

(a) handshake/neighbor, $n = 3$



(b) handshake, $n = 4$

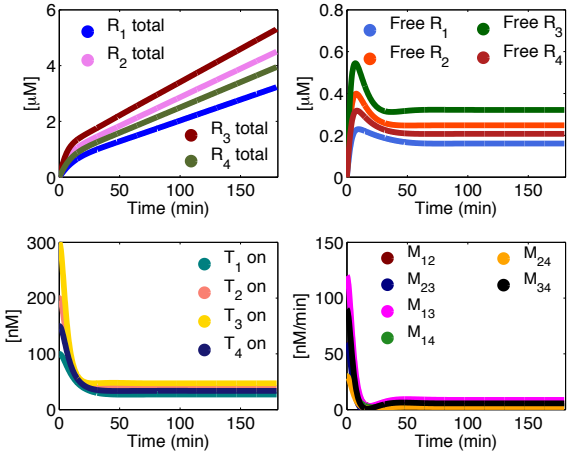


Fig. 5: Simulations: negative feedback schemes.

For all the connections with negative feedback, the mismatch decreases if δ increases and increases if α increases, as shown in Figures 7 (a) and (b).

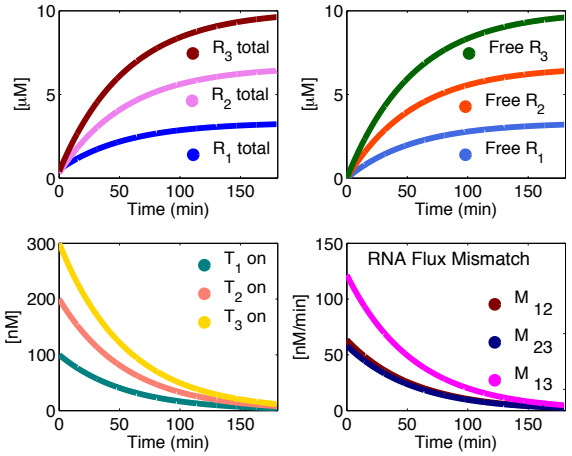
E. Positive feedback architectures

Positive feedback is implemented with a cross–activation scheme. When a reagent is in excess (i.e. it is not used in the product formation), it increases the generation rate of all the other reagents it is reacting with. The active $\mathrm{T_i}$ is assumed to naturally inactivate with rate α_i: $\mathrm{T_i} \xrightarrow{\alpha_i} \mathrm{T_i^*}$ and $[\mathrm{T_i^{tot}}] = [\mathrm{T_i}] + [\mathrm{T_i^*}]$.

1) Single product: δ_i is the strength of the positive feedback on gene i due to all the others, k is the generation rate of the unique product P, $\sum_{\mathrm{i=1}}^{\mathrm{n}} \mathrm{R_i} \xrightarrow{k} \mathrm{P}$. Thus

$$\begin{aligned}
\frac{\mathrm{d}[\mathrm{T_i}]}{\mathrm{dt}} &= -\alpha_{\mathrm{i}}\,[\mathrm{T_i}] + \delta_{\mathrm{i}}\,([\mathrm{T_i^{tot}}] - [\mathrm{T_i}]) \prod_{\mathrm{j\neq i}} [\mathrm{R_j}] \\
\frac{\mathrm{d}[\mathrm{R_i}]}{\mathrm{dt}} &= \beta_{\mathrm{i}}\,[\mathrm{T_i}] - \mathrm{k} \prod_{\mathrm{i=1}}^{\mathrm{n}} [\mathrm{R_i}] - \delta_{\mathrm{i}}\,[\mathrm{R_i}] \prod_{\mathrm{j\neq i}} ([\mathrm{T_j^{tot}}] - [\mathrm{T_j}]) \qquad (4) \\
\frac{\mathrm{d}[\mathrm{P}]}{\mathrm{dt}} &= \mathrm{k} \prod_{\mathrm{i=1}}^{\mathrm{n}} [\mathrm{R_i}], \qquad [\mathrm{R_i^{tot}}] = [\mathrm{R_i}] + \sum_{\mathrm{j\neq i}} [\mathrm{T_j}] + [\mathrm{P}].
\end{aligned}$$

(a) single product, $n = 3$



(b) handshake/neighbor, $n = 3$

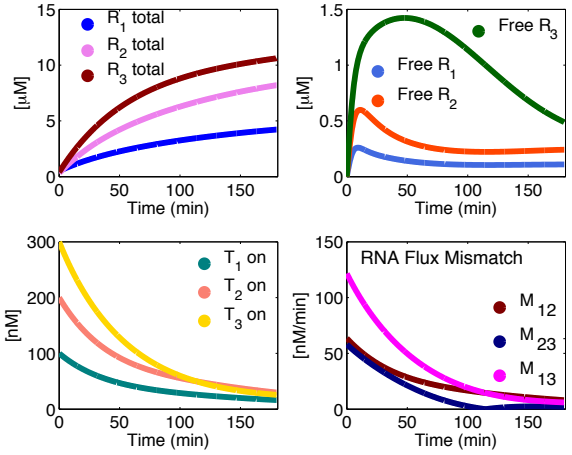


Fig. 6: Simulations: positive–feedback architectures.

The cross–activation scheme asymptotically eliminates the flow mismatches, but the response time is much longer than in the single product negative feedback case.

2) Handshake and neighbor connection: now δ_{ij} is the strength of the positive feedback on gene i due to gene j, $\mathrm{R_i} + \mathrm{T_j^*} \xrightarrow{\delta_{ij}} \mathrm{T_j}$. The generation of products $\mathrm{P_{ij}}$ occurs at rates

k_{ij}, $\mathrm{R_i} + \mathrm{R_j} \overset{k_{ij}}{\rightharpoonup} \mathrm{P_{ij}}$. A scheme is shown in Figure 3 (b). Thus

$$\begin{aligned}
\frac{\mathrm{d[T_i]}}{\mathrm{dt}} &= -\alpha_\mathrm{i}\,[\mathrm{T_i}] + \sum_\mathrm{j} \delta_\mathrm{ij}\,[\mathrm{R_j}]([\mathrm{T_i^{tot}}] - [\mathrm{T_i}]) \\
\frac{\mathrm{d[R_i]}}{\mathrm{dt}} &= \beta_\mathrm{i}\,[\mathrm{T_i}] - \sum_\mathrm{j} \mathrm{k_{ij}}\,[\mathrm{R_i}][\mathrm{R_j}] - \sum_\mathrm{j} \delta_\mathrm{ji}\,[\mathrm{R_i}]([\mathrm{T_j^{tot}}] - [\mathrm{T_j}]) \\
\frac{\mathrm{d[P_{ij}]}}{\mathrm{dt}} &= \mathrm{k_{ij}}\,[\mathrm{R_i}][\mathrm{R_j}], \qquad [\mathrm{R_i^{tot}}] = [\mathrm{R_i}] + \sum_\mathrm{j}[\mathrm{T_j}] + \sum_\mathrm{j}[\mathrm{P_{ij}}].
\end{aligned} \tag{5}$$

In the handshake case $i, j = 1, ..., n$, $j \neq i$; in the neighbor case $i = 1, ..., n$, $j = i-1, i+1$ and when $i = 1$, $i - 1 = n$, when $i = n$, $i + 1 = 1$, to close the loop. The flux mismatches decrease, but the response time is still longer than in the negative feedback architecture. The handshake/neighbor connection generates less waste (unused $\mathrm{R_i}$ species) than the single product interconnection.

For all the connections with positive feedback, the mismatch increases if δ increases (apart from the single product case, when the mismatch is independent of δ) and decreases if α increases, as shown in Figures 7 (c) and (d).

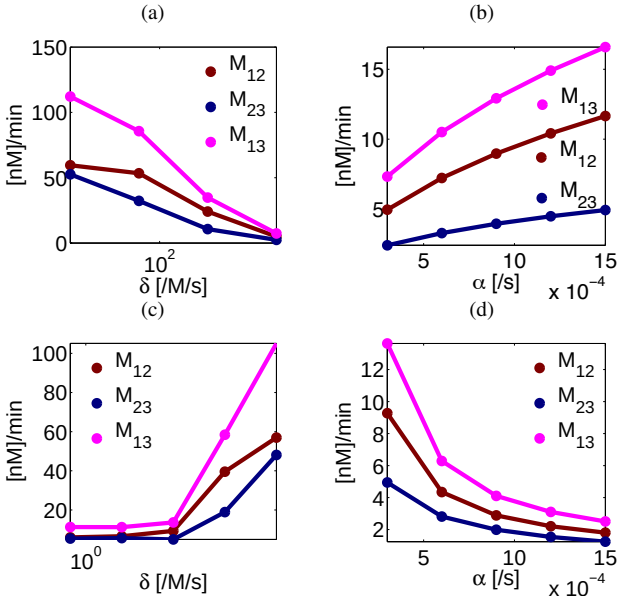


Fig. 7: RNA mean flow mismatch after 10 hours. Negative feedback (a) for different orders of magnitude of δ, (b) as a function of α. Positive feedback (c) δ, (d) α.

F. Negative and positive feedback: a comparison

Negative feedback schemes have a fast time response, are easy to scale up (each gene controls its own production rate) and, loosely speaking, stabilize the system. However, negative feedback does not enhance the output production rate: the overall output flux maintains low levels. Positive feedback schemes have a slower response time, are more difficult to scale up (each gene is controlled by others, thus increasing the number of genes involved leads to a growing number of interactions) and the feedback constant must be kept very small to avoid unbounded increase of

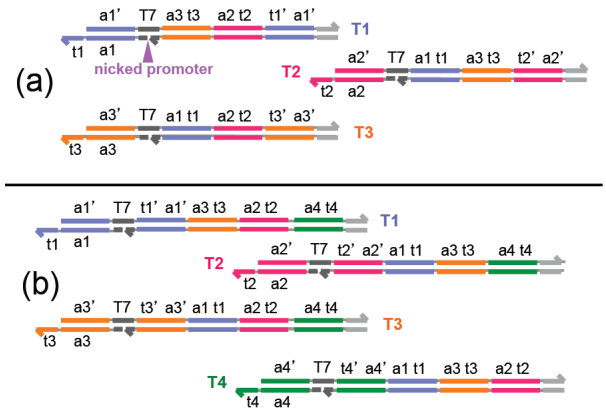


Fig. 8: DNA-domain implementation of the (a) three and (b) four genes negative feedback interconnection. Complementary domains have the same color. Nicked T7 promoters are in dark gray, terminator domains in light gray. The RNA output of each genelet is designed to be complementary to its activator strand. RNA species are pairwise complementary.

product. However, positive feedback maximizes the output flux. Thus, negative feedback is the best control strategy if genes are not highly required, to avoid the accumulation of potentially harmful excess of unused reagents; while positive feedback is better for genes in high demand, to maximize the production rate [14]. The output production fluxes can be matched also when we use both positive and negative feedback. With simulations, we can see that the system shows an intermediate behavior, which is more similar to pure negative or positive feedback, depending on which feedback constant δ is the strongest.

III. Negative feedback: a new model for a viable DNA strand implementation

Here we describe a possible DNA implementation for our three or four genes handshake negative feedback architecture, with artificial gene networks (transcriptional circuits) [3]. The working principles of these circuits were described in Section II-A. Referring to Figure 8, domains of the genes are represented as sequences of bases with different colors. RNA outputs of each gene are not shown, but their domains are identical to the transcribed regions of the genes (downstream of the promoter, dark gray). Domains with the same color are complementary and are expected to bind. The domain annotated as $\mathrm{t_1 a_1}$ on $\mathrm{T_1}$, for instance, is an activator strand which can be displaced by the RNA output $\mathrm{R_1}$, which has the domain $\mathrm{t_1' a_1'}$. Output RNAs $\mathrm{R_i}$ are designed to be complementary to their own activator strands and pairwise complementary to one another, so that they can bind to form products. Once $\mathrm{R_i}$ and $\mathrm{R_j}$ form $\mathrm{P_{ij}}$, the complex is inert and all the regulatory domains for negative auto–regulation are covered. This design is an extension of that proposed in [11] for a two-gene interconnection. This choice of the domains introduces, in addition to the desired self–inhibition loops, an undesired binding between $\mathrm{T_i}$ and $\mathrm{R_j}$. The resulting complex can be considered as another off state of the gene: the complex obtained is a substrate for RNaseH and the RNA

strand is degraded by the enzyme, releasing the activating strand.

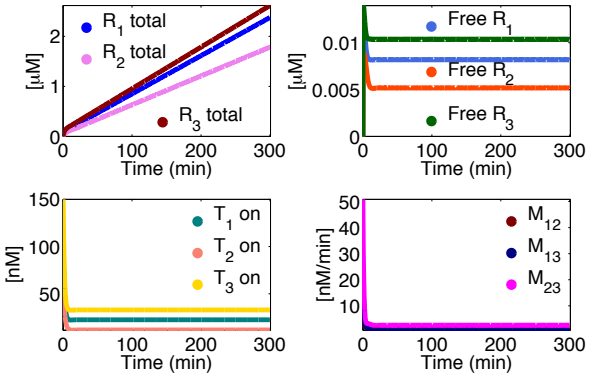
We built a detailed model of this system, based on the expected domain interactions. Each gene $\mathrm{T_i}$ can have three possible states: the on state, in which activator and template are bound and form the complex $\mathrm{T_iA_i}$; the off state given by free $\mathrm{T_i}$; the off state represented by $\mathrm{R_j}$ bound to $\mathrm{T_i}$ (thus forming $\mathrm{T_iR_j}$). To be sure that the inhibition rate is the same for all the genes, it is better to place the self–inhibition domains $\mathrm{t_i'\ a_i'}$ in the same position inside the strand; for example, first (near the $3'$ end) or last (near the $5'$ end) domain (in Figure 8, self-inhibition domains are at the $3'$ end). In the case of more than two genes, when the complex $\mathrm{R_iR_j}$ forms, we cannot avoid the formation of loops or torsions for some values of i, j. For example, referring to Figure 8, the $\mathrm{R_2}$ and $\mathrm{R_1}$ complex (binding of the domains indexed 1 and 2) occurs by formation of a loop in the domain $\mathrm{a_3t_3}$ on $\mathrm{R_2}$. It is evident that it is possible to form a complex between $\mathrm{R_n}$ and $\mathrm{R_m}$ iff, going through one strand in the arrow direction (from 5' to 3') you find at first $\mathrm{t_m'\ a_m'}$ and then $\mathrm{a_n\ t_n}$, while going through the other strand you find at first $\mathrm{t_n'\ a_n'}$ and then $\mathrm{a_m\ t_m}$. Placing all the self–inhibition domains $\mathrm{t_i'\ a_i'}$ either at the beginning or at the end of the RNA strand assures that this condition is satisfied and thus the binding can successfully occur. So, in the n genes case, there can be $2 \cdot ((n-1)!)^n$ different domain level designs for a DNA strand implementation of the negative feedback architecture for rate–regulation: the self–inhibition domains can be in the first position inside the strand or in the last; in each of these 2 cases, the possible permutations of the other segments are $((n-1)!)^n$ because there are $(n-1)!$ possible different configurations in each of the n strands.

Including all reactions occurring in the system [11], we can build the following ODEs, where $i = 1, ..., n$,

$$\begin{aligned}
\frac{\mathrm{d[T_i]}}{\mathrm{dt}} &= -\mathrm{k_{T_iA_i}[T_i][A_i]} + \mathrm{k_{R_iT_iA_i}[R_i][T_i \cdot A_i]} \\
&\quad - \mathrm{k_{R_jT_i}[R_j][T_i]} + \sum_{\mathrm{j}} \mathrm{k_{catH_{ji}}[RNaseH \cdot R_j \cdot T_i]} \\
\frac{\mathrm{d[A_i]}}{\mathrm{dt}} &= -\mathrm{k_{T_iA_i}[T_i][A_i]} - \mathrm{k_{R_iA_i}[R_i][A_i]} + \mathrm{k_{catH_i}[RNaseH \cdot R_i \cdot A_i]} \\
\frac{\mathrm{d[R_i]}}{\mathrm{dt}} &= -\sum_{\mathrm{j}} \mathrm{k_{R_iR_j}[R_i][R_j]} + \mathrm{k_{R_iT_iA_i}[R_i][T_i \cdot A_i]} \\
&\quad - \mathrm{k_{R_iT_j}[R_i][T_j]} - \mathrm{k_{R_iA_i}[R_i][A_i]} + \mathrm{k_{catONi}[RNAP \cdot T_i \cdot A_i]} \\
&\quad + \mathrm{k_{catOFF_i}[RNAP \cdot T_i]} + \sum_{\mathrm{j}} \mathrm{k_{catOFF_{ji}}[RNAP \cdot R_j \cdot T_i]} \\
\frac{\mathrm{d[R_i \cdot R_j]}}{\mathrm{dt}} &= +\mathrm{k_{R_iR_j}[R_i][R_j]} \\
\frac{\mathrm{d[R_j \cdot T_i]}}{\mathrm{dt}} &= +\mathrm{k_{R_jT_i}[R_j][T_i]} - \mathrm{k_{catH_{ji}}[RNaseH \cdot R_j \cdot T_i]}.
\end{aligned}$$

Enzyme species are not modeled as separate states in our analysis and we assume that the Michaelis–Menten quasi–steady–state approximation holds. For example, for the RNAP reactions, we define $\mathrm{P} = 1 + \sum_{\mathrm{i=1}}^{\mathrm{n}} \frac{\mathrm{[T_i \cdot A_i]}}{\mathrm{k_{MONi}}} + \sum_{\mathrm{i=1}}^{\mathrm{n}} \frac{\mathrm{[T_i]}}{\mathrm{k_{MOFFi}}} + \sum_{\mathrm{j \neq i}} \frac{\mathrm{[R_i \cdot T_j]}}{\mathrm{k_{MOFFij}}}$ and we find that $\mathrm{[RNAP \cdot T_i \cdot A_i] = [RNAP^{tot}]\ [T_i \cdot A_i]/P \cdot k_{MONi}}$. We can derive similar expressions for other terms involving enzyme species. The system is solved numerically for the cases of three and four genes, with parameters taken from [2]. Numerical solutions are shown in Figure 9; the negative feedback rate–regulation appears effective as well. This more complete and accurate model can be experimentally implemented and tested by means of transcriptional circuits.

$n = 3$, complete model



$n = 4$, complete model

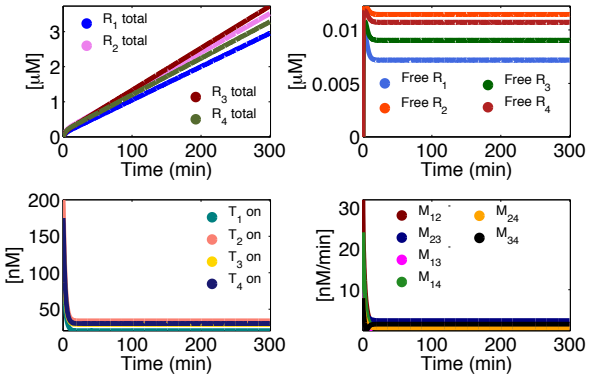


Fig. 9: Transcriptional circuit implementation, negative feedback scheme; ODE solutions for 3 and 4 genes.

IV. Conclusions

We considered different feedback architectures to regulate the production of RNA species in a synthetic n–gene system, where these species interact with one another to produce one or more complexes. This problem is relevant in the context of *in vitro* synthetic biology and nanotechnology, where synthetic gene networks are useful as circuits producing components that assemble in nanostructures or that orchestrate dynamic behaviors for molecular computations. Our numerical analysis for $n = 3$ and $n = 4$ revealed that negative autoregulation guarantees better scalability and faster response than positive feedback based architectures. Finally, we analyzed the performance of negative–autoregulated three and four gene systems, proposing a viable DNA strand implementation. Our results provide useful predictions for the future experimental construction of these *in vitro* genetic networks.

ACKNOWLEDGEMENTS

The authors acknowledge financial support by the National Science Foundation (NSF) grant CCF-0832824 (The Molecular Programming Project) and the Bourns College of Engineering at UC Riverside.

REFERENCES

[1] F. Blanchini and E. Franco, "Structurally robust biological networks," *Bio Med Central Systems Biology*, vol. 5, no. 1, p. 74, 2011.

[2] E. Franco, "Analysis, design, and *in vitro* implementation of robust biochemical networks," Ph.D. dissertation, California Institute of Technology, 2011.

[3] J. Kim, K. S. White, and E. Winfree, "Construction of an *in vitro* bistable circuit from synthetic transcriptional switches," *Molecular Systems Biology*, vol. 2, p. 68, 2006.

[4] J. Kim and E. Winfree, "Synthetic *in vitro* transcriptional oscillators," *Molecular Systems Biology*, vol. 7, p. 465, 2011.

[5] P. Subsoontorn, J. Kim, and E. Winfree, "Ensemble bayesian analysis of bistability in a synthetic transcriptional switch," *ACS Synthetic Biology*, 2012.

[6] E. Franco, E. Friedrichs, J. Kim, R. Jungmann, R. Murray, E. Winfree, and F. C. Simmel, "Timing molecular motion and production with a synthetic transcriptional clock," *Proceedings of the National Academy of Sciences*, vol. 108, no. 40, pp. E784–E793, 2011.

[7] K. A. Afonin, E. Bindewald, A. J. Yaghoubian, N. Voss, E. Jacovetty, B. A. Shapiro, and L. Jaeger, "*In vitro* assembly of cubic RNA–based scaffolds designed in silico," *Nature Nanotechnology*, vol. 5, September 2010.

[8] S. M. Douglas, I. Bachelet, and G. M. Church, "A logic–gated nanorobot for targeted transport of molecular payloads," *Science*, vol. 335, pp. 831–834, 2012.

[9] H. Gu, J. Chao, S.-J. Xiao, and N. C. Seeman, "Dynamic patterning programmed by DNA tiles captured on a DNA origami substrate," *Nature Nanotechnology*, vol. 4, no. 4, pp. 245–248, 2009.

[10] A. Phillips and L. Cardelli, "A programming language for composable DNA circuits," *Journal of The Royal Society Interface*, 2009.

[11] E. Franco, P.-O. Forsberg, and R. M. Murray, "Design, modeling and synthesis of an *in vitro* transcription rate regulatory circuit," in *Proceedings of the American Control Conference*, 2008.

[12] E. Franco and R. M. Murray, "Design and performance of *in vitro* transcription rate regulatory circuits," in *Proceedings of the IEEE Conference on Decision and Control*, 2008.

[13] B. Yurke and A. P. Mills, "Using DNA to power nanostructures," *Genetic Programming and Evolvable Machines*, vol. 4, pp. 111–122, 2003.

[14] M. A. Savageau, "Design of molecular control mechanisms and the demand for gene expression," *Proceedings of the National Academy of Sciences of the USA*, vol. 74, pp. 5647–5651, 1977.