



**HAL**  
open science

## Dynamics of complex feedback architectures in metabolic pathways

Madalena Chaves, Diego A. Oyarzun

► **To cite this version:**

Madalena Chaves, Diego A. Oyarzun. Dynamics of complex feedback architectures in metabolic pathways. *Automatica*, 2019, 99, pp.323 - 332. 10.1016/j.automatica.2018.10.046 . hal-01936242

**HAL Id: hal-01936242**

**<https://hal.science/hal-01936242v1>**

Submitted on 27 Nov 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Dynamics of complex feedback architectures in metabolic pathways

Madalena Chaves<sup>a</sup>, Diego A. Oyarzún<sup>b,c</sup>

<sup>a</sup>*Université Côte d'Azur, Inria, INRA, CNRS, Sorbonne Université, Biocore team, Sophia Antipolis, France*

<sup>b</sup>*School of Informatics, University of Edinburgh, Edinburgh EH8 9AB, United Kingdom*

<sup>c</sup>*School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3JH, United Kingdom*

---

## Abstract

Cellular metabolism contains intricate arrays of feedback control loops. These are a key mechanism by which cells adapt their metabolism to survive environmental disturbances. Gene regulation, in particular, typically displays complex architectures that combine positive and negative feedback loops between metabolites and enzymatic genes. Yet because of strong nonlinearities and high-dimensionality, it is challenging to determine the closed-loop dynamics of a given feedback architecture. Here we present a novel technique for the analysis of metabolic pathways under gene regulation. Our theory blends ideas from timescale separation and piecewise affine dynamical systems, applied to a wide class of unbranched metabolic pathways under steep nonlinear feedback. We propose a systematic method to construct a state transition graph for a given regulatory architecture, from where candidate closed-loop dynamics can be singled out for further analysis. The method recasts a high-dimensional nonlinear system into a piecewise affine system defined on a polytopic partition of the state space. In its most general setup, our theory allows to characterize the dynamics of pathways of arbitrary length, with any number of regulators, and under any combination of positive and negative feedback loops. We illustrate our results on an exemplar system that displays a stable limit cycle and bistable dynamics. We also discuss the implications of our results for the design of gene circuits in synthetic biology and metabolic engineering.

Keywords: Metabolic pathways; synthetic biology; stability analysis; piecewise linear models; gene regulatory networks

---

## 1 Introduction

Cellular metabolism adapts to environmental fluctuations through networks of regulatory feedback loops [6]. In natural systems, such feedbacks control the expression of catalytic enzymes and match the activity of specific metabolic pathways to extracellular nutrients and the cellular energy budget. Last decades have witnessed striking examples on the construction of synthetic feedback control loops for metabolism, with the goal of increasing production of valuable chemicals [10,18,37] and the generation of entirely new metabolic dynamics such as oscillations [11] or cell-to-cell communication [13,31].

Regulatory motifs in metabolic systems include architectures with various combinations of feedback and feed-forward loops, whose signs and interactions determine

the overall pathway dynamics. A fundamental design problem is the characterization of control architectures that achieve a prescribed system response [33,23]. However, there are no general stability analysis methods that can deal with the type of nonlinearities and connectivity of metabolic systems. A number of recent reviews in metabolic engineering have highlighted the need for model-based approaches for design [3,14,21], while the use of control theoretical principles has gained substantial traction in the field [25,8].

In this paper we provide two contributions on the analysis of control architectures in metabolic pathways under gene regulation. First, we show that the interaction graph of a large family of feedback systems can be classified into few distinct classes of interaction graphs. Our findings suggest a catalogue of architectures in the form of interactions graphs with nested positive or negative loops, where the number of distinct classes depends only on pathway length (Section 3). This catalogue links each feedback architecture to a systematized feedback control network and its corresponding dynamical behavior.

---

\* This paper was not presented at any IFAC meeting. Corresponding author: M. Chaves.

*Email addresses:* [madalena.chaves@inria.fr](mailto:madalena.chaves@inria.fr) (Madalena Chaves), [d.oyarzun@ed.ac.uk](mailto:d.oyarzun@ed.ac.uk) (Diego A. Oyarzún).

ior. Second, we provide an algorithm to build a state transition graph for a given feedback architecture, from which candidate asymptotic dynamics can be identified and further analyzed (Section 4). The analysis relies on the combination of timescale separation and qualitative, piecewise-affine, models for gene regulation. Here we extend a technique originally described in [24] for pathways with one regulator, to a much larger class of control systems with any number of regulators and any combination of positive and negative feedback loops. We demonstrate our theory with an example motivated by [11] that illustrates the dynamical diversity within each class of feedback architectures (Section 5).

## 2 Metabolic pathways with multiple regulators

Metabolic pathways interconvert chemicals (metabolites) through sequences of biochemical reactions. Such reactions are catalyzed by specific proteins (enzymes) synthesized by the cell. By controlling the expression of genes that code for the enzymes, cells activate and deactivate the activity of pathways depending on the environmental conditions. Because enzyme synthesis requires energy, this control mechanism avoids the inefficient production of enzymes that are not needed in a specific environment. This is the case, for example, when key metabolites for growth are already present in the growth media and do not need to be produced by cells themselves.

Here we consider a general class of unbranched metabolic pathways with  $n$  metabolites and  $n + 1$  enzymes, as shown in Fig. 1. The conversion of each metabolite  $s_i$  into  $s_{i+1}$  is catalyzed by an enzyme  $e_{i+1}$ . The substrate  $s_0$  represents an extracellular nutrient and we assume it to remain constant in the timescale of the pathway. This assumption corresponds to cases in which, for example,  $s_0$  represents a pool of extracellular nutrient in continuous culture. Under a constant substrate, the pathway reaches a nonzero steady state flux and remains outside thermodynamic equilibrium, which is the typical operating regime of metabolism.

We further consider that enzyme synthesis is controlled by some of the metabolites in the pathway. We consider a large class of feedback architectures satisfying the following assumptions:

- A1.** Each enzyme is regulated by a single metabolite;
- A2.** There are  $1 \leq k \leq n$  regulatory metabolites.

Assumptions A1-A2 define a large class of feedback architectures that may include complex combinations of positive and negative feedback loops. This is a significant extension to previous work [24] that considered the case of only one regulatory metabolite ( $k = 1$ ). In particular, Assumption A2 allows for any number of regula-

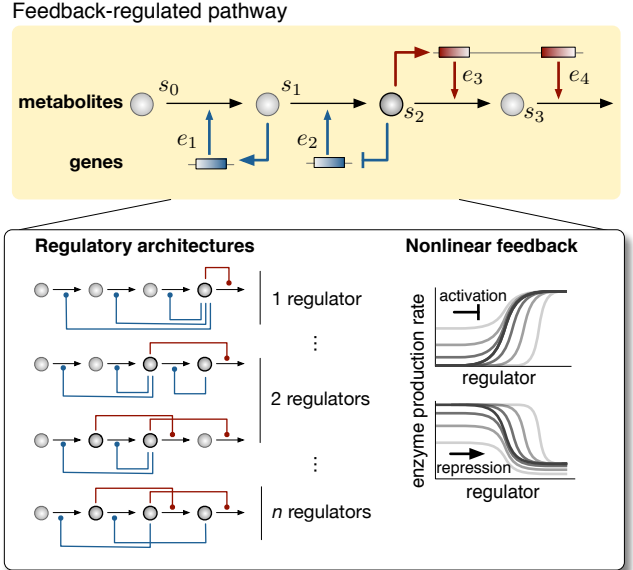


Fig. 1. Regulatory architectures in unbranched metabolic pathways. We consider a general class of nonlinear feedback systems with up to  $n$  regulatory metabolites ( $s_i$ ) with any combination of positive or negative regulation on the enzymatic genes ( $e_i$ ). The diagram illustrates assumptions A1, A2, and A2'. Regular arrows (resp., bar-headed arrows) represent activation (resp., repression). Dot-headed arrows represent either of these.

tory metabolites to control enzyme expression. In most natural instances, e. g. the lactose operon or amino acid metabolism [6], pathways have one regulatory metabolite ( $k = 1$ ), with the case of  $k > 1$  being particularly interesting for the design of regulatory architectures for metabolic engineering [21].

If  $s_i$  and  $e_i$  denote the concentration of metabolites and enzymes respectively, the system in Fig. 1 can be described by the nonlinear model:

$$\dot{s}_i = g_i(s_{i-1})e_i - g_{i+1}(s_i)e_{i+1}, \quad i = 1, \dots, n, \quad (1)$$

$$\dot{e}_i = \kappa_i^0 + \kappa_i^1 \sigma_i(s_{\ell_i}, \theta_i) - \gamma e_i, \quad i = 1, \dots, n + 1, \quad (2)$$

where  $s_{\ell_i}$  is the metabolite that regulates  $e_i$ ,  $g_i : \mathbb{R}_{\geq 0}^+ \rightarrow [0, M_i]$  are continuous and strictly increasing functions representing the kinetics of each enzyme [15], and  $\sigma_i \in \{\tilde{\sigma}_+, \tilde{\sigma}_-\}$  are increasing or decreasing sigmoidal functions describing the nonlinear regulation from metabolites to enzyme synthesis [22,38,7]. The positive parameters ( $\kappa_i^0, \kappa_i^1$ ) represent the minimal and maximal rates of enzyme synthesis, and  $\gamma > 0$  describes the dilution of protein concentrations caused by cellular growth.

Our objective in this paper is to characterize the asymptotic dynamics of nonlinear feedback systems described by equations (1)-(2). To this end, we will use two techniques for simplifying the models [24]: first, we approximate enzyme synthesis by piecewise linear functions; and

second, we exploit the inherent separation of timescales and the shape of the nonlinearities of the enzyme kinetics.

### 2.1 Piecewise linear models for gene regulation

Gene expression can often be approximated by two levels, corresponding to “high” and “low” levels of transcription. In such cases, the sigmoidal functions that describe enzyme synthesis can be simplified to step functions, following the approaches suggested by Thomas [34] and Glass and Kauffman [12]:

$$\sigma_+(r, \theta) = 1 - \sigma_-(r, \theta) = \begin{cases} 1, & r > \theta, \\ 0, & r < \theta, \end{cases}$$

where  $\sigma_+$  (resp.,  $\sigma_-$ ) corresponds to an activation (resp., inhibition), and the parameter  $\theta$  is a threshold of regulation. In this model, enzyme concentrations eventually converge to a bounded interval,  $E_i^{\text{off}} \leq e_i(t) \leq E_i^{\text{on}}$  for all  $t \geq t_0 > 0$ , with

$$E_i^{\text{off}} = \frac{\kappa_i^0}{\gamma}, \quad E_i^{\text{on}} = \frac{(\kappa_i^0 + \kappa_i^1)}{\gamma}. \quad (3)$$

The step function approximation leads to a qualitative description of the dynamics in terms of piecewise linear systems which, in turn, allows for more detailed analytical results (see examples for *Escherichia coli* metabolic networks [24,29] and genetic transcription and translation networks [9,16]). Piecewise linear systems induce a partition of the state space into well defined domains which can be naturally mapped onto the vertices of a state transition graph [2,4]; this graph provides a global view of the dynamics of the system and will be constructed in Section 4.

### 2.2 Timescale separation

Metabolic timescales are much faster than those of enzyme synthesis [15,29], so the system (1)-(2) can be simplified by assuming the metabolites to be in quasi-steady state. The variables  $s_i$  evolve fast towards a slow manifold of the form  $s = \tilde{G}(e)$ , where  $s$  and  $e$  are the vectors of metabolite and enzyme concentrations, respectively. By substitution of the metabolites  $s_i$  in the ODEs for the enzymes, the slow dynamics in (2) can be simplified to depend only on the enzyme concentrations  $e_i$ .

Such timescale separation ideas have been used for analysis of various enzymatic mechanisms [1,20], genetic transcription and translation networks [9], and justify the well known Michaelis-Menten kinetic mechanism [30]. Briefly, the goal is to rewrite (1)-(2) in the form:  $\varepsilon \dot{s} = G(s, e, \varepsilon)$ ,  $\dot{e} = F(s, e, \varepsilon)$ , where  $\varepsilon \ll 1$ . Under suitable conditions (see [19] for more details),

$G(s, e, 0) = 0$  defines the slow manifold  $s = \tilde{G}(e)$ , and Tikhonov’s Theorem guarantees that solutions of the slower dynamics  $\dot{e} = F(\tilde{G}(e), e, 0)$  remain close to those of the original system.

In the particular case where all enzymes follow Michaelis-Menten kinetics, that is,  $g_i(s_i) = k_{\text{cat } i} s_{i-1} / (K_{M,i} + s_{i-1})$  (see example in Section 5), the timescale separation can be applied to (1)-(2) provided that

$$\varepsilon_i = \max \left\{ \frac{\gamma K_{M,i}}{E_i^{\text{on}} k_{\text{cat } i}}, \frac{\gamma K_{M,i+1}}{E_i^{\text{on}} k_{\text{cat } i}} \right\} \ll 1, \quad (4)$$

for  $i = 2, \dots, n$ . These conditions can be obtained by applying an appropriate change of variables:  $\tilde{e}_i = e_i / E_i^{\text{on}}$ ,  $\tilde{s}_i = s_i / K_{M,i}$ , and  $\tilde{t} = \gamma t$ , similar to those used in [20,30]. The conditions typically hold for metabolic systems, since dilution by growth is much slower than enzyme turnover rates, and hence  $\gamma \ll k_{\text{cat } i}$ . The quasi-steady state assumption is:

**A3.** We assume that  $\dot{s}_i \approx 0$  for each  $t \geq 0$ , leading to

$$g_1(s_0) e_1 = g_i(s_{i-1}) e_i, \quad \text{for } i = 2, \dots, n+1. \quad (5)$$

Using assumption A3 we can substitute

$$s_{\ell_i} = g_{\ell_i+1}^{-1} \left( g_1(s_0) \frac{e_1}{e_{\ell_i+1}} \right) \quad (6)$$

into equation (2) for enzyme  $e_i$ . To guarantee the existence of solutions of (6), a sufficient condition is:

$$E_i^{\text{off}} \geq \frac{g_1(s_0)}{M_{\ell_i+1}} E_i^{\text{on}}, \quad \text{for } i = 2, \dots, n+1. \quad (7)$$

Now, let  $I_{\text{reg}} = \{1, \ell_1+1, \dots, \ell_k+1\}$  be the set of indices corresponding to the enzymes that catalyze the outgoing flux for each of the regulator metabolites together with index 1, the enzyme which catalyzes the constant input substrate; set  $I_{\text{nr}} = \{1, 2, \dots, n\} \setminus I_{\text{reg}}$ . From (6), it becomes clear that the subset of  $k+1$  equations corresponding to  $\{e_i : i \in I_{\text{reg}}\}$ , does not depend on the variables  $\{e_j, j \in I_{\text{nr}}\}$ , so it can be studied separately. Therefore, without loss of generality we will thus consider a modified version of Assumption **A2**:

**A2’.** Each metabolite in the pathway regulates at least one enzyme, or equivalently,  $n \equiv k$ .

Taken together points A1 and A2’ imply that there is only one metabolite (denoted as  $s_p$ ) that regulates two enzymes and all other metabolites regulate a single enzyme. It is then useful to establish an index correspondence:

$$\mathcal{I} : \{1, \dots, n+1\} \rightarrow \{1, \dots, n\} \quad (8)$$

with  $\mathcal{I}(i) = \ell_i$  and  $\mathcal{I}(p_1) = \mathcal{I}(p_2) = p$ . Its inverse is defined as:  $\mathcal{I}^{-1}(j) = q_j$  and  $\mathcal{I}^{-1}(p) = \{p_1, p_2\}$ .

### 2.3 Enzyme dynamics and regulatory motifs

The unbranched metabolic pathway dynamics can thus be reduced to the enzyme equations. In order to capture this genetic regulation, we will introduce, with a slight abuse of notation, the following definition of a *regulatory motif*:

$$\sigma = (\sigma_1, \dots, \sigma_n), \quad \sigma_i = \begin{cases} 1, & \sigma_i(r, \theta) = \sigma_+(r, \theta) \\ -1, & \sigma_i(r, \theta) = \sigma_-(r, \theta). \end{cases} \quad (9)$$

In addition, the quasi steady state assumption (6) and the fact that  $g_i$  are non-decreasing imply

$$s_{\ell_i} > \theta_i \Leftrightarrow g_1(s_0)e_1 > g_{\ell_i+1}(\theta_i)e_{\ell_i+1} \Leftrightarrow e_1 > \frac{1}{\beta_i}e_{\ell_i+1} \quad (10)$$

where we have defined the parameter  $\beta_i = g_1(s_0)/g_{\ell_i+1}(\theta_i)$ .

Under the relations in (10), the step functions satisfy

$$\sigma_i(s_{\ell_i}, \theta_i) = \sigma_i(s_{\ell_i} - \theta_i, 0) = \sigma_i(e_1 - \frac{1}{\beta_i}e_{\ell_i+1}, 0), \quad (11)$$

and therefore the enzyme dynamics in (2) can be rewritten as

$$\dot{e}_i = \kappa_i^0 + \kappa_i^1 \sigma_i \left( e_1 - \frac{1}{\beta_i} e_{\ell_i+1}, 0 \right) - \gamma e_i. \quad (12)$$

Equation (12) is a much simpler version of the original model in (1)–(2) and, as we show in the next sections, it is amenable to detailed analysis.

### 3 General structure of the interaction graph

Given a metabolic pathway such as the one in Fig. 1, one would like identify and catalogue the possible feedback networks induced by its specific regulatory motif  $\sigma$ . In this section we provide such classification in terms of the possible interactions graphs associated to a regulatory motif. The pattern of feedback loops and their signs determine the dynamical behavior of the system: for instance, it has been shown that a necessary condition for multistability is the existence of a positive loop, while a negative loop is required for stable oscillations [17,32,28,26].

An interaction graph describes the dependencies between model variables according to the sign pattern of the Jacobian matrix [7]. Although the Jacobian of (12)

has a non-constant sign pattern, next we describe a suitable change of variables under which the interaction graph can be readily built. We first note that the right hand side of (12) depends on sign of  $e_1 - \frac{1}{\beta_i}e_{\ell_i+1}$ , and thus we introduce the following definition:

**Definition 1** [*Cone membership function*] Let  $\beta > 0$ ,  $r \in \mathbb{R}_{>0}^{n+1}$ , and  $r_i$  the  $i^{\text{th}}$  component of  $r$ . Define the cone membership function by

$$\nu_i(\beta, r) : \mathbb{R}_{>0} \times \mathbb{R}_{>0}^{n+1} \longrightarrow \mathbb{R} \\ (\beta, r) \longrightarrow r_1 - \frac{1}{\beta} r_{i+1}.$$

In other words, sign of  $\nu_i$  determines whether or not the  $(n+1)$ -dimensional vector  $r$  belongs to the planar cone defined by a half-line with slope  $\beta$ .

The new variables can now be defined as:

$$x_j = \nu_{j-1}(\beta_{q_{j-1}}, e), \quad j \in \{2, \dots, n+1\} \setminus \{p+1\}, \\ z_1 = \nu_p(\beta_{p_1}, e), \quad z_2 = \nu_p(\beta_{p_2}, e),$$

From the model in (12), it follows that:

$$\dot{x}_j = \tilde{\kappa}_1^0 + \kappa_1^1 \sigma_1(x_{\ell_1+1}, 0) - \frac{\kappa_j^1}{\beta_{q_{j-1}}} \sigma_j(x_{\ell_j+1}, 0) - \gamma x_j, \\ j \in \{2, \dots, n+1\} \setminus \{p+1\} \quad (13) \\ \dot{z}_j = \tilde{\kappa}_{p_j}^0 + \kappa_1^1 \sigma_1(x_{\ell_1+1}, 0) - \frac{\kappa_{p+1}^1}{\beta_{p_j}} \sigma_{p+1}(x_{\ell_{p+1}+1}, 0) - \gamma z_j, \\ j = 1, 2.$$

where  $\tilde{\kappa}_j^0 = \kappa_1^0 - \kappa_j^0/\beta_{q_{j-1}}$  and  $\tilde{\kappa}_{p_j}^0 = \kappa_1^0 - \kappa_{p+1}^0/\beta_{p_j}$ .

The Jacobian of (13) is sign-constant and therefore we can build the interaction graph for the new variables by adding a directed arrow from  $x_j$  to  $x_i$  if  $\partial(\dot{x}_i)/\partial x_j \neq 0$ , and (for simplicity) omitting all arrows related to the natural degradation terms  $-\gamma x_j$  or  $-\gamma z_j$ . This interaction graph has some immediate properties:

- R1 all nodes have exactly two incoming arrows;
- R2 node  $x_{\ell_1+1}$  has  $n+1$  outgoing arrows, signed  $\sigma_1$ ;
- R3 node  $x_{\ell_{p+1}+1}$  has exactly two outgoing arrows, both signed  $-\sigma_{p+1}$ , towards  $z_1$  and  $z_2$ ;
- R4 all other nodes  $x_{\ell_j+1}$  have exactly one outgoing arrow, signed  $-\sigma_j$ , towards  $x_j$ . In particular, node  $z_j$  has an outgoing arrow, signed  $-\sigma_{p_j}$ , towards  $x_{p_j}$ .

A self arrow counts as both an incoming and an outgoing arrow for the node. If a node has an outgoing arrow to  $x_{p_1}$ , then it has necessarily an outgoing arrow to  $x_{p_2}$ , hence that node is either  $x_{\ell_1+1}$  or  $x_{\ell_{p+1}+1}$ .

The above properties suggest that the large family of feedback systems described by equations (1)–(2) can be classified into relatively few distinct interaction graphs, based on the identification of four special nodes:

- i the node  $w = x_{\ell_1+1}$  with  $n + 1$  outgoing arrows;
- ii the node  $y = x_{\ell_{p+1}+1}$  with two outgoing arrows;
- iii the two nodes  $z_1$  and  $z_2$  (which can coincide with either  $w$  or  $y$ , when  $p_j = \ell_{p+1} + 1$ , for some  $j \in \{1, 2\}$ );
- iv finally, list the remaining nodes,  $x_{\ell_j+1}$  for  $j \notin \{p+1, 1, 1\}$ , which can be re-labeled  $x_a, x_b$ , etc.

To obtain all the distinct interaction graphs, consider all combinations of  $z_1$  and  $z_2$  according to (iii), and add an edge between two of these nodes according to rules (R1)–(R4). The possible interaction graphs are listed in Fig. 2, for pathways with three or four enzymes. Next we consider a specific example to illustrate the construction of the interaction graphs.

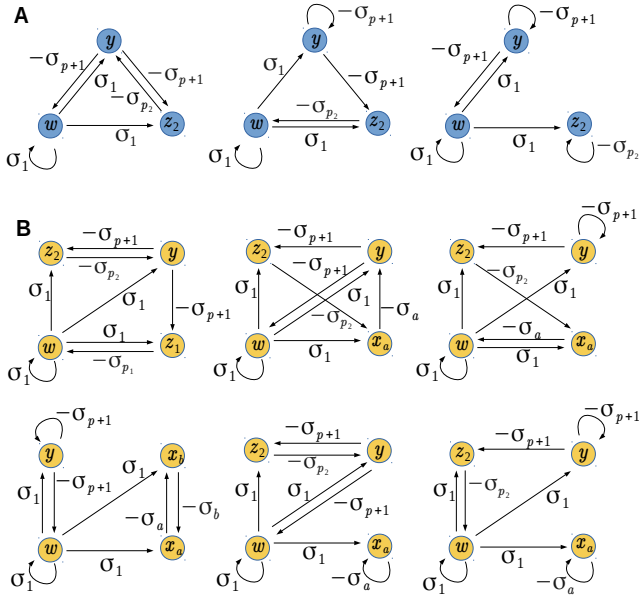


Fig. 2. Interaction graphs of the transformed system in (13). (A) All possible interaction graphs (up to re-ordering) for pathways of length  $n = 3$ . Left to right:  $w = z_1$ ;  $y = z_1$ ;  $w = z_1$  and  $y = z_2$ . (B) All possible interaction graphs (up to re-ordering) for pathways of length  $n = 4$ . Left to right and top to bottom:  $w, y, z_1, z_2$  all distinct;  $w = z_1$ ;  $y = z_1$ ;  $w = z_1$  and  $y = z_2$ ; alternative to  $w = z_1$ ; alternative to  $y = z_1$ .

**Example 2** Consider the pathway in Fig. 1 (top), which illustrates assumption A2'. There are only two regulators,  $s_1$  and  $s_2$ , so it follows from (6) that  $e_4$  will not influence the dynamics of  $e_i$ ,  $i = 1, 2, 3$ . Since  $s_2$  regulates two enzymes, it follows that  $p = 2$  and, from the definition in (8), the index correspondence is

$$(\ell_1, \ell_2, \ell_3) = (1, 2, 2), \quad (q_1, p) = (1, \{2, 3\}).$$

The change of variables is as follows:

$$x_2 = e_1 - \frac{1}{\beta_1}e_2, z_1 = e_1 - \frac{1}{\beta_3}e_3, z_2 = e_1 - \frac{1}{\beta_2}e_3,$$

with  $w = x_2$  and  $y = z_1$ , yielding an interaction graph as in Fig. 2A(middle).

To classify the interaction graphs, a first observation is that all (strongly connected) graphs with  $n$  nodes exhibit  $n$  nested loops of growing lengths, starting with the self loop on  $w$ , each loop adding only one new variable. For instance, for the first case in Fig. 2A:

$$w \circlearrowleft; w \rightleftharpoons y; w \rightarrow z_2 \rightarrow y \rightarrow w,$$

and for the first case in Fig. 2B

$$w \circlearrowleft; w \rightleftharpoons z_1; w \rightarrow y \rightarrow z_1 \rightarrow w; w \rightarrow z_2 \rightarrow y \rightarrow z_1 \rightarrow w.$$

In fact, this result can be generalized as follows:

**Proposition 3** [Nested loops] Consider the  $n + 1$  dimensional system (13). If the interaction graph is formed of a single strongly connected component, and if either  $w \neq z_j$  or  $y \neq z_j$ , for all  $j \in \{1, 2\}$ , then it contains  $n + 1$  nested loops,  $\{L_\ell\}_{\ell=1}^{n+1}$ , defined by:

$$\begin{aligned} L_1 &= \{w\}, \\ L_2 &= \{w, v_1 : w \rightarrow v_1, v_1 \rightarrow w\}, \\ L_{\ell+1} &= \{w, v_1, \dots, v_\ell : w \rightarrow v_\ell, v_{j+1} \rightarrow v_j, v_1 \rightarrow w\}, \end{aligned}$$

where  $j = 1, \dots, \ell - 1$  in  $L_{\ell+1}$ . If the interaction graph contains more than one strongly connected component, then one of these components has  $j + 1$  nodes and contains nested loops  $L_1, \dots, L_{j+1}$ .

**Proof.** This structure follows directly from the properties (R1)–(R4) and (i)–(iv). First, note that  $w$  has an arrow toward every node, including itself, so  $w$  forms a trivial first loop,  $L_1$ . Second, choose the node  $v_1 \neq w$  which has an edge towards  $w$ : by point (R1), each node has exactly two incoming arrows (with one of them  $w$ ), hence there is a unique choice for  $v_1$ . This gives a 2-node circuit  $L_2$ .

Now, to use an induction argument, suppose we have chosen  $j \geq 1$  distinct nodes,  $v_1, \dots, v_j$ , defining loops  $L_1$  up to  $L_{j+1}$ . Note that, by construction, each of the nodes  $v_i$ ,  $i < j$  has exactly two incoming arrows,  $w$  and  $v_{i+1}$ . If  $j = n$ , all the loops have been found and the interaction graph is formed of a single strongly connected component. If  $j < n$ , choose next the node  $v_{j+1} \neq w$  which has an edge towards  $v_j$ : again, by point (R1) there can be only one such node. If  $v_{j+1} \neq v_j$ , then we have found loop  $L_{j+2}$  and can continue. Otherwise,  $v_{j+1} = v_j$ , which means that  $v_j = y$  since it has two

outgoing arrows. But it also has two incoming arrows,  $w$  and  $v_j$ . This means that the set of nodes  $w, v_1, \dots, v_j$  (with  $j < n$ ) cannot have any further incoming arrows and thus forms a strongly connected component, which contains the nested loops  $L_1, \dots, L_{j+1}$ . Finally, the remaining nodes ( $v_{j+1}, \dots, v_n$ ) can be organized into at least one more strongly connected component. This finishes the proof of the Proposition.

**Remark 4** *In the case both  $w = z_1$  and  $y = z_2$  it follows that  $w \rightleftharpoons y$ , as well as self-arrows for both  $w$  and  $y$ , thus preventing the generation of  $n + 1$  nested loops. This is illustrated in Fig. 2B bottom left.*

As a corollary, it follows that an  $n + 1$  dimensional system (13) whose interaction graph has more than one strongly connected component can in fact be studied as a (lower)  $(j + 1)$ -dimensional system with nested loops  $L_1, \dots, L_{j+1}$ . Indeed, the dynamics of this sub-system  $w, v_1, \dots, v_j$  can be studied as an autonomous system, whose behaviour is an input to the system formed by the remaining variables  $v_{j+1}, \dots, v_n$ . Examples are shown in Fig. 2B (bottom middle and right) which can both be decomposed into a 3-dimensional system  $(w, y, z_2)$  which controls  $x_a$ . Interestingly, these 3-dim systems share the same graphs of Fig. 2A (left and middle). In other words, the systems whose graphs are represented in Fig. 2B(bottom middle and right) can be further simplified and studied as a two-enzyme metabolic pathway.

In addition, *the number of distinct strongly connected interactions graphs* can be exactly computed, as shown in the next Proposition.

**Proposition 5** *[Number of distinct interaction graphs] System (13) has at most four (for  $n = 2$ ) and nine (for  $n \geq 3$ ) distinct strongly connected interaction graphs with  $n + 1$  nested loops, up to re-ordering the  $x_i$  nodes:*

- (a) *two distinct graphs, if  $w = z_1$  or  $w = z_2$ ;*
- (b) *two distinct graphs, if  $y = z_1$  or  $y = z_2$ ;*
- (c) *five distinct graphs if  $n \geq 3$  and all  $w, y, z_1, z_2$  are different.*

**Proof.** (a) In this case, the first two loops are  $L_1 = \{w\}$  and  $L_2 = \{w, y\}$  (i.e.,  $v_1 = y$ , in the notation of Prop. 3). Choosing the sequence of nodes  $v_i$  as in Prop. 3, recall that  $v_j = z_2$  implies  $v_{j+1} = y$ , by definition of  $y$ . To guarantee that the graph has a single strongly connected component, it must be that  $v_n = z_2$ . Otherwise, if  $v_j = z_2$  for  $j < n$ , then  $v_{j+1} = y$ , which implies that  $L_{j+1} \equiv L_{j+1} = \{w, v_1, \dots, v_j\}$  is a strongly connected component and so the full graph cannot be strongly connected.

(b) In this case, assume  $y = z_1$  without loss of generality. Node  $y$  has a self arrow (which counts as both an outgoing and an incoming arrow), an incoming arrow from

node  $w$  and an outgoing arrow towards  $z_2$ . To guarantee that the graph has a single strongly connected component, it must be that  $v_n = y$  and  $v_{n-1} = z_2$ . Otherwise, by the same argument as in case (b), if  $v_j = y$  for some  $j < n$ , the set of nodes  $\{w, v_1, \dots, v_j\}$  forms a strongly connected component and so the full graph cannot be strongly connected.

(c) In this case, the second incoming arrow for node  $w$  can originate either from one of the variables  $z_i$  or from one of the  $n - 3$  remaining variables  $x_i$ , up to re-ordering. Suppose first that  $z_1 \rightarrow w$  (or  $z_2 \rightarrow w$ ). Then, as in case (b), we need  $v_n = y$  and  $v_{n-1} = z_2$  to guarantee a strongly connected interaction graph (count 2 distinct graphs). Suppose next that  $x_a \rightarrow w$  and  $z_1 \rightarrow y$  (or  $z_2 \rightarrow y$ ). Then, we again need  $v_n = y$  and  $v_{n-1} = z_2$  (count 2 distinct graphs). Suppose finally that  $x_a \rightarrow w$  and  $x_b \rightarrow y$ . The remaining nodes can be assigned in any order, according to rules (R1)-(R4) (count only 1 distinct graph).

Proposition 5 is illustrated in Fig. 2. The interaction graphs associated with (13) give an indication of the possible dynamics of the feedback system (for instance, multistability is expected whenever all loops are positive [35]).

## 4 Global dynamics

As a counterpart to the interaction graph, next we develop a *state transition graph* that allows for analysis of the global dynamics. To this end, the system in equation (12) will be represented by a *discrete system* formed by a set of vertices  $V$  and a set of directed edges  $\mathcal{E}$ : each vertex represents one region of the state space and there exists an edge between two vertices if there is at least one trajectory connecting the corresponding regions. The resulting state transition graph provides a complete qualitative view of the global dynamics of (12), for a given range of parameters.

The system in equation (12) can be written as the piecewise linear system (see also [4])

$$\dot{e}_i = \begin{cases} \gamma(E_i^{\sigma_i} - e_i), & e_1 > \frac{1}{\beta_i} e_{\ell_i+1} \\ \gamma(E_i^{-\sigma_i} - e_i), & e_1 < \frac{1}{\beta_i} e_{\ell_i+1} \end{cases}, \quad (14)$$

where  $\sigma = (\sigma_1, \dots, \sigma_n)$  represents the regulatory motif, and  $(E_i^{-1}, E_i^1) = (E_i^{\text{off}}, E_i^{\text{on}})$  are the minimal and maximal enzyme concentrations in equation (3). In this new description, the state space is partitioned into several conic regions defined by the two planes  $z_j = 0$  and the  $n - 1$  planes  $x_j = 0$  (see (13)). Each region has a *focal point*  $\phi = (\phi_1, \dots, \phi_n)'$  with  $\phi_i = E_i^{\sigma_i}$  if  $\nu_{\ell_i}(\beta_i, e) > 0$  and  $\phi_i = E_i^{-\sigma_i}$  otherwise. If a trajectory crosses the

boundary between two regions, a switch in the vector field – and hence in the focal point – occurs, leading to a new direction of the trajectory. The conic regions and its focal points are the main elements in the algorithm to construct the state transition graph.

#### 4.1 Vertices of the transition graph

As shown in the next definition, each vertex of the graph represents one region  $R^a$ , characterized in terms of the cone membership functions, and the nonnegative sign function  $\sigma_+(r, 0)$ .

**Definition 6** [Polytopic regions] Let  $p \in \{1, \dots, n\}$  be the index defined in (8). A region  $R^a$ , with  $a = (a_1, \dots, a_n)'$  with  $a_p \in \{0, 1, 2\}$  and  $a_i \in \{0, 1\}$  for  $i \in \{1, \dots, n\} \setminus \{p\}$ , is defined by:

$$R^a = \{ e \in \mathbb{R}_{\geq 0}^{n+1} : \\ a_i = \sigma_+(\nu_i(\beta_{q_i}, e), 0), \\ a_p = \sigma_+(\nu_p(\beta_{p_1}, e), 0) + \sigma_+(\nu_p(\beta_{p_2}, e), 0) \}.$$

The set of vertices of the graph is thus given by

$$V = \{(a_1, \dots, a_n) : a_p \in \{0, 1, 2\}, a_i \in \{0, 1\} \text{ for } i \neq p\}.$$

#### 4.2 Edges of the transition graph

Each edge represents a crossing between two adjacent conic regions: if there is a trajectory starting in some region  $R^a$  and going through a neighbouring region  $R^b$ , then there is a directed edge from vertex  $a$  to vertex  $b$ . The path of a trajectory is determined by the location of focal points. In general, the focal point  $\phi^a$  may be contained in some other region, say  $R^b$  with  $a \neq b$ , so trajectories starting in  $R^a$  will eventually cross to a neighbouring region.

Here, an *asynchronous strategy* will be considered in the sense that if an edge joins two vertices, say  $a \rightarrow b$ , then these vertices differ on exactly one coordinate  $j$  ( $a_j \neq b_j$  and  $a_i = b_i$  for all  $i \neq j$ ). To determine the edges, the next steps are the calculation of the focal points,  $\phi^a = (\phi_1^a, \dots, \phi_n^a)$ , for each region  $R^a$  and their location  $R^b$ , i.e.  $\phi^a \in R^b$ , with  $a = (a_1, \dots, a_n)$  and  $b = (b_1, \dots, b_n)$ .

##### 4.2.1 Focal points

The coordinates of the focal point  $\phi^a$  for region  $R^a$  can be explicitly calculated recalling (14) and using Def. 6:

$$\phi_{q_i}^a = a_i E_{q_i}^{\sigma_{q_i}} + (1 - a_i) E_{q_i}^{-\sigma_{q_i}}, \quad \text{if } i \neq p, \quad (15)$$

and in the case  $i = p$ :

$$\phi_{p_1}^a = \max(a_p - 1, 0) E_{p_1}^{\sigma_{p_1}} + (1 - \max(a_p - 1, 0)) E_{p_1}^{-\sigma_{p_1}}, \\ \phi_{p_2}^a = \min(a_p, 1) E_{p_2}^{\sigma_{p_2}} + (1 - \min(a_p, 1)) E_{p_2}^{-\sigma_{p_2}}. \quad (16)$$

##### 4.2.2 Location of the focal points

Computation of the region  $R^b$  is straightforward, as it suffices to check the sign of the functions  $\nu_i$  when computed at the (already known) focal points:

$$b_i = \sigma_+(\nu_i(\beta_{q_i}, \phi^a), 0) \\ b_p = \sigma_+(\nu_p(\beta_{p_1}, \phi^a), 0) + \sigma_+(\nu_p(\beta_{p_2}, \phi^a), 0). \quad (17)$$

Observe that the focal points and their location in the state space depend on the parameters of the system  $(\kappa_i^{0,1}, \gamma, \theta_i)$ . Thus, one single regulatory motif (9) may lead to different dynamical behavior depending on its parameters. However, as seen below in Section 4.4, once the state transition graph is constructed, it is easy to obtain a large range of parameters for which the graph holds. For simplicity, we assume that no focal point belongs to the boundaries of a cone, that is

$$\mathbf{A4.} \quad \nu_i(\beta_{q_i}, \phi^a) \neq 0 \text{ for all } i \text{ and all } a \in V.$$

##### 4.2.3 Construction of the graph edges

There are two cases to consider, depending on whether the given region contains its own focal point. First, assume that  $\phi^a \notin R^a$ . Then it can be shown that there is at least one boundary hyper-plane  $e_{i+1} = \beta_{q_i} e_1$  through which trajectories can escape  $R^a$ .

**Proposition 7** Consider a region  $R^a$  of system (14) with focal point  $\phi^a \in R^b$ , with  $a, b \in V$  and  $a \neq b$ . For each  $i \in \{1, \dots, n\}$  such that  $a_i \neq b_i$ , there exist trajectories crossing from  $R^a$  to  $R^c$  where the coordinates of  $c$  are given by  $c_i = a_i + \text{sign}(b_i - a_i)$  and  $c_k = a_k$  for  $k \neq i$ .

**Proof.** Assume that  $i \neq p$  (the case  $i = p$  is similar). The assumption  $a \neq b$  implies that, for at least one coordinate  $i$ :

$$\nu_i(\beta_{q_i}, e) \cdot \nu_i(\beta_{q_i}, \phi^a) < 0 \quad \text{for all } e \in R^a. \quad (18)$$

Next, equation (14) implies  $w_1(t) = \phi_1 - e_1(t) \rightarrow 0$  and  $w_{i+1}(t) = \phi_{i+1} - e_{i+1}(t) \rightarrow 0$ , as  $t \rightarrow \infty$ , which leads to

$$(\forall \varepsilon > 0)(\exists t_\varepsilon > 0) : |\nu_i(\beta_{q_i}, e(t)) - \nu_i(\beta_{q_i}, \phi^a)| < \varepsilon, \quad (19)$$

for all  $t > t_\varepsilon$ . Applying (18) yields

$$|\nu_i(\beta_{q_i}, e(t))| + |\nu_i(\beta_{q_i}, \phi^a)| < \varepsilon.$$



Since  $\nu_i(\beta_{q_i}, \phi^a) = \hat{\nu}_i$  is a constant, for  $\varepsilon = \hat{\nu}_i$  condition (19) implies  $|\nu_i(\beta_{q_i}, e(t_b))| = 0$  for some finite time instant  $t_b$ , for which the trajectory exactly reaches the boundary. In conclusion, trajectories may indeed cross the boundary from  $R^a$  to a neighboring region  $R^c$ , characterized by  $\nu_i(\beta_{q_i}, e) \cdot \nu_i(\beta_{q_i}, \phi^a) > 0$  (the opposite of (18)).

Next, assume that region  $R^a$  contains its focal point,  $\phi^a \in R^a$ . In the more “classical” piecewise linear systems with hyper-rectangular regions [4], it is well known that regions that contain their own focal points are forward invariant. However, this is not necessarily true for conic regions (as already discussed in [24]), unless all degradation rates are equal. In our case, we have the following result.

**Proposition 8** *Consider a region  $R^a$  of system (14) containing its own focal point  $\phi^a$ ,  $a \in V$ . Then,  $R^a$  is forward-invariant.*

**Proof.** Assumption  $\phi^a \in R^a$  implies:

$$\nu_i(\beta_{q_i}, e) \cdot \nu_i(\beta_{q_i}, \phi^a) > 0 \quad \text{for all } i \text{ and all } e \in R^a. \quad (20)$$

To check whether trajectories may leave region  $R^a$  through some coordinate  $i$ , consider the vector field along that boundary. Note that, for all  $e \in R^a$ , the enzyme dynamics may be written in terms of the focal point (14):  $\dot{e}_i = \gamma(\phi_i^a - e_i)$ . Now define a new variable

$$r = \frac{\nu_i(\beta_{q_i}, e)}{\nu_i(\beta_{q_i}, \phi^a)} \Rightarrow \dot{r} = \gamma(1 - r),$$

by reordering the terms  $d/dt(\nu_i(\beta_{q_i}, e)) = \gamma(\phi_1^a - e_1) - \frac{1}{\beta_{q_i}}\gamma(\phi_{i+1}^a - e_{i+1})$ . By (20),  $r > 0$  for all  $e \in R^a$ . At the boundary  $\nu_i(\beta_{q_i}, e) = 0$ , it holds that  $r = 0$  and  $\dot{r} = \gamma > 0$ . Since  $\nu_i(\beta_{q_i}, \phi^a)$  is constant, we can conclude that  $r$  remains positive and never crosses the boundary, that is, solutions  $e(t)$  of (14) remain in  $R^a$ .

The set of edges can now be constructed:

$$\mathcal{E} = \{a \rightarrow c : c_j = a_j + G_j, \text{ some } j \text{ and } c_k = a_k, k \neq j\}$$

with  $G_j = \text{sign}(b_j - a_j)$ , where  $R^b$  is the region that contains  $\phi^a$ . The procedure to compute the set  $\mathcal{E}$  is summarized in Algorithm 1.

### 4.3 The state transition graph

Propositions 7 and 8 determine all possible transitions from a given region  $R^a$  and establish the basis for an exact algorithmic representation of the global qualitative dynamics of system (12). The coefficients  $c_j \neq 0$  identify

the boundary crossing corresponding to a threshold in variable  $a_j$ . We note that step 4 in Algorithm 1 reflects an *asynchronous updating rule* for the evolution of the discrete trajectories, since *only one coordinate* is allowed to change at each transition.

**Algorithm 1.** State transition graph for system (12).

- 
0. **Input:** set of vertices  $V$ .
  1. Initialize  $\mathcal{E}$ .
  2. Pick  $a \in V$  and compute  $\phi^a$  from (15) and (16).
  3. Compute  $b$  such that  $\phi^a \in R^b$  from (17).  
If  $b \equiv a$  label node  $a$  by adding a symbol,  $\hat{a}$ .
  4. For  $j = 1$  until  $n$  do:
    - 4.1. Compute  $c_j = a_j + \text{sign}(b_j - a_j)$ ,
    - 4.2. If  $c_j \neq a_j$  add an edge  $a \rightarrow c$  to  $\mathcal{E}$ , with  $c_k = a_k$  for  $k \neq j$ .
  5. Repeat steps 2 to 4 for each vertex in  $V$ .
  6. **Output:** set of edges  $\mathcal{E}$ .
- 

The state transition graph can be analyzed with graph theoretical tools to obtain properties such as the existence of one or more steady states, or oscillatory behavior. In Algorithm 1, a node labeled  $\hat{a}$  represents a region  $R^a$  which contains its own focal point (step 3). In this case, by Prop. 8,  $\phi^a$  is a regular equilibrium point and a (locally stable) steady state of the system.

### 4.4 Parameter range for the state transition graph

Steps 2 and 3 of Algorithm 1 compute the focal points and their location using (15) and (17) which, in principle, require knowledge of the set of parameters. However, the computation of the edges in Step 4 uses only the signs:

$$\delta_{i,a} = \delta_{i,a}(\kappa_i^0, \kappa_i^1, \beta_{q_i}, \gamma) = \text{sign}(\nu_i(\beta_{q_i}, \phi^a)).$$

So, in fact, the state transition graph computed for a particular set of parameters remains valid *for all sets of parameters which yield exactly the same family of signs*. We can formally write

$$\mathcal{Q} = \left\{ \pi \in \mathbb{R}_{>0}^{4(n+1)} : \delta_{i,a}(\kappa_i^0, \kappa_i^1, \beta_{q_i}, \gamma) \equiv \bar{\delta}_{i,a}, \right. \\ \left. \forall i = 1, \dots, n+1 \text{ and } a \in V \right\} \quad (21)$$

where  $\pi = (\kappa_1^0, \kappa_1^1, \beta_1, \dots, \kappa_{n+1}^0, \kappa_{n+1}^1, \beta_{n+1}, \gamma)$ . Therefore, given a first set of realistic parameters, the algorithm leads to a full characterization of  $\mathcal{Q}$ .

## 5 Case study: a two-regulator pathway

### 5.1 Model definition

To illustrate the application of our results to a metabolic pathway with a complex regulatory architecture, we consider the system shown in Fig. 3 with  $n = 2$  metabo-

lites and three enzymes. In this system, both metabolites control enzyme synthesis, and thus we have  $k = 2$  regulators. The regulatory architecture resembles the “metabolator”, a metabolic oscillator implemented in the *Escherichia coli* bacterium by Fung et al [11]. The metabolator is based on the metabolic reactions that link glycolysis and the respiration cycle, which include two metabolites and three enzymes, and contains positive and negative feedback regulation of two pathway enzymes.

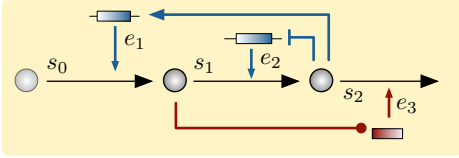


Fig. 3. Example pathway with two regulatory metabolites. We consider the cases in which metabolite  $s_1$  activates or represses the production of enzyme  $e_3$ , denoted by a dot-headed arrow.

The full 5-dimensional model for the system in Fig. 3 reads:

$$\begin{aligned} \dot{s}_1 &= g_1(s_0)e_1 - g_2(s_1)e_2, \\ \dot{s}_2 &= g_2(s_1)e_2 - g_3(s_2)e_3, \\ \dot{e}_1 &= \kappa_1^0 + \kappa_1^1 \tilde{\sigma}_+(s_2, \theta_1) - \gamma e_1, \\ \dot{e}_2 &= \kappa_2^0 + \kappa_2^1 \tilde{\sigma}_-(s_2, \theta_2) - \gamma e_2, \\ \dot{e}_3 &= \kappa_3^0 + \kappa_3^1 \tilde{\sigma}_3(s_1, \theta_3) - \gamma e_3. \end{aligned} \quad (22)$$

In the notation of Section 2, the model in (22) has the subindices  $p = 2$ ,  $(\ell_1, \ell_2, \ell_3) = (2, 2, 1)$ , and  $(q_1, p) = (3, \{1, 2\})$ . We assume that the metabolic reactions follow standard Michaelis-Menten kinetics of the form

$$\begin{aligned} g_1(s_0) &= \frac{k_{\text{cat } 1} s_0}{K_{M1} + s_0}, \\ g_2(s_1) &= \frac{k_{\text{cat } 2} s_1}{K_{M2} + s_1}, \\ g_3(s_2) &= \frac{k_{\text{cat } 3} s_2}{K_{M3} + s_2}. \end{aligned} \quad (23)$$

We fix the kinetic parameters to realistic values  $k_{\text{cat } 1} = 60 \text{ min}^{-1}$ ,  $k_{\text{cat } 2} = 180 \text{ min}^{-1}$ ,  $k_{\text{cat } 3} = 480 \text{ min}^{-1}$ , with a Michaelis-Menten constant  $K_{M_i} = 1 \mu\text{M}$  for all three enzymatic reactions ( $i = 1, 2, 3$ ). The degradation rate  $\gamma = 0.0116 \text{ min}^{-1}$  corresponds to a doubling time of  $\sim 60 \text{ min}$ , typical of the bacterium *Escherichia coli*. We set the expression rates  $\kappa_i^0$  and  $\kappa_i^1$  so that the minimal and maximal enzyme concentrations are  $E^{\text{off}} = [1, 4, 1.5]$  and  $E^{\text{on}} = [15, 20, 30]$ , both in units of nM concentrations. The regulatory functions in (22) are assumed to

be Hill functions of the form

$$\tilde{\sigma}_+(s, \theta) = \frac{s^h}{\theta^h + s^h}, \quad \tilde{\sigma}_-(s, \theta) = \frac{\theta^h}{\theta^h + s^h}, \quad (24)$$

for a Hill coefficient  $h \gg 1$  and regulatory thresholds  $\theta_1 = 0.047 \mu\text{M}$ ,  $\theta_2 = 0.055 \mu\text{M}$ , and  $\theta_3 = 0.22 \mu\text{M}$ . Note that in the model (22) we have not specified the type of regulation for  $\sigma_3$  because we will use our approach for both cases  $\sigma_3 = \sigma_+$  and  $\sigma_3 = \sigma_-$ .

After the separation of timescales,  $\dot{s}_i \approx 0$  for all  $t \geq 0$  and using step regulatory functions, the model in (22) reduces to

$$\begin{aligned} \dot{e}_1 &= \kappa_1^0 + \kappa_1^1 \sigma_+(e_1/e_3, 1/\beta_1) - \gamma e_1, \\ \dot{e}_2 &= \kappa_2^0 + \kappa_2^1 \sigma_-(e_1/e_3, 1/\beta_2) - \gamma e_2, \\ \dot{e}_3 &= \kappa_3^0 + \kappa_3^1 \sigma_3(e_1/e_2, 1/\beta_3) - \gamma e_3, \end{aligned} \quad (25)$$

where the  $\beta_i$  parameters are

$$\beta_1 = \frac{g_1(s_0)}{g_3(\theta_1)}, \quad \beta_2 = \frac{g_1(s_0)}{g_3(\theta_2)}, \quad \beta_3 = \frac{g_1(s_0)}{g_2(\theta_3)}. \quad (26)$$

## 5.2 State transition graphs

Assuming that the regulatory functions in (25) can be well approximated by step functions, the reduced model is equivalent to a 3-dimensional piecewise affine system defined on  $3 \times 2^1 = 6$  disjoint polytopes. Following the notation of Section 4, we associate each polytope to a pair  $a = (a_1, a_2)$  with  $a_1 \in \{0, 1\}$  and  $a_2 \in \{0, 1, 2\}$ . Since  $s_2$  regulates the production of two enzymes, we have that  $p = 2$ . Moreover, under the chosen parameter sets, we have that  $\beta_2 < \beta_1$  and thus the subindices are  $p_1 = 2$  and  $p_2 = 1$ . The regulatory motif vector is  $\sigma = [1, -1, 1]$  or  $\sigma = [1, -1, -1]$ , depending on whether  $\sigma_3 = \sigma_+$  or  $\sigma_3 = \sigma_-$ , respectively. Using equations (15)-(16) we get the focal points for each of the six polytopes in both cases, shown in Fig. 4A.

Fig. 4B shows the six polytopes for a fixed nutrient concentration ( $s_0$ ). To illustrate the impact of model parameters on the graph structure, we used Algorithm 1 to compute the state transition graphs for increasing values of  $s_0$ . The results, shown in Fig. 4C–D, reveal salient differences between the resulting graphs. Some graphs have one or two sink nodes, while others display various cycles. We observe changes in the graphs for a fixed regulatory architecture and increasing various values of  $s_0$ , as well for a fixed  $s_0$  and different regulatory architectures.

One particular case, highlighted in yellow in Fig. 4C–D, reveals that the regulatory sign of  $\sigma_3$  switches the graph from having two sink nodes to having two cycles.

We investigated these two cases further to examine the bistable behaviour of the model and the possibility of a periodic orbit along the graph cycles. As described in Section 4.4, we can determine the set  $\mathcal{Q}$  in the parameter space that leads to both these graphs

$$E_1^{\text{off}} < \frac{1}{\beta_3} E_2^{\text{on}}, \quad E_1^{\text{on}} < \frac{1}{\beta_3} E_2^{\text{on}}, \quad E_1^{\text{on}} > \frac{1}{\beta_3} E_2^{\text{off}}, \quad (27)$$

$$E_1^{\text{off}} > \frac{1}{\beta_1} E_3^{\text{off}}, \quad E_1^{\text{on}} > \frac{1}{\beta_2} E_3^{\text{off}}, \quad E_1^{\text{on}} < \frac{1}{\beta_1} E_3^{\text{on}}, \quad (28)$$

$$E_1^{\text{off}} < E_1^{\text{on}} \quad (29)$$

Note that the above conditions link together the enzyme expression parameters ( $E_i^{\text{off}}$  and  $E_i^{\text{on}}$ ) with the enzyme kinetic parameters ( $k_{\text{cat } i}$  and  $K_{M i}$  appearing in the  $\beta_i$  parameters).

In the case of repression, i. e.  $\sigma_3 = \sigma_-$ , the interaction graph in Fig. 4C contains only positive loops and therefore the system cannot exhibit oscillatory behavior [35]. Accordingly, the state transition graph (yellow) has two sinks at nodes 1 and 6, and thus we expect two distinct locally stable steady states in polytopes  $R^{00}$  and  $R^{12}$ . As shown in Fig. 5A, simulations of the full model in (22) with a Hill coefficient  $h = 4$  verify the existence of two stable state states located at the predicted polytopes. Note that because of our piecewise affine approximation, the exact location of the steady states of the continuous model differs slightly from the focal points  $\phi^{00}$  and  $\phi^{12}$  in Fig. 4A.

In the case of positive regulation of enzyme  $e_3$ , i. e.  $\sigma_3 = \sigma_+$ , the interaction graph in Fig. 4D contains three negative loops and thus the system may exhibit periodic oscillations [35]. Indeed, the state transition graph (Fig. 4D, yellow) has a short cycle through nodes 2-3-6-5, and long cycle through nodes 1-2-3-6-5-4. Upon a close examination of the piecewise affine model along these cycles, we can prove that [5]: a) the short cycle corresponds to a Filippov-type equilibrium at the intersection of the four polytopes  $R^{01} \cap R^{02} \cap R^{12} \cap R^{11}$ ; b) the long cycle corresponds to a stable periodic orbit along the polytopes in the cycle. We verified the latter prediction through simulations of the full continuous model in (22) for increasing values of the Hill coefficient. As observed in Fig. 5B, we found a periodic orbit for  $h \geq 10$ , which qualitatively matches the predicted orbit from the piecewise model for increasing values of the Hill coefficient. From the time course simulations for the metabolites Fig. 4B (bottom), we also observe that the shape and period of the oscillations corresponds well with the metabolic oscillator in [11] (see, for instance, Figs. 1,2 therein), thus suggesting the periodic orbit is within physiological ranges and potentially observable in experiments.

## 6 Conclusions

Cellular metabolism is controlled by networks of transcriptional regulatory loops, often masked by complex feedback architectures between metabolic intermediates and enzymatic genes. Determining the closed-loop dynamics of a specific architecture is challenging because of the strong nonlinearities and the large number of chemical species involved. In this paper we have proposed a formalism for the qualitative analysis of closed-loop dynamics in a large class of regulation systems for unbranched pathways. Our analysis relies on time scale separation and piecewise affine models to accomplish: (i) a complete classification of all possible feedback interaction networks associated to a metabolic pathway; and (ii) an algorithmic construction of the state transition graph that describes the global system dynamics. In a synthetic biology and metabolic engineering context, this feedback network catalogue provides an indication of the components (metabolites, enzymes, promoters) required to assemble a metabolic pathway exhibiting a desired circuitry and dynamics [22].

Our analytical results are obtained for a new class of piecewise linear systems in conic regions, derived by approximating Hill functions by step functions. Among other benefits, this framework allows establishing existence of an asymptotically stable periodic orbit for a system whose regulatory motif is reminiscent of the *metabolator* [11]. In fact, theoretical proofs of oscillatory behavior are difficult to obtain in general, and the result for the metabolator remains unclear [27].

Despite their usefulness, piecewise linear systems provide a possibly less realistic description than their continuous counterparts. Large Hill coefficients are typically associated with strong cooperativity effects [15], while gene expression profiles can often be described by steep sigmoidal curves [36]. To verify the suitability of the piecewise linear approximation in our case, we performed simulations for Hill coefficients in a wide interval. The differences observed in the dynamical behavior reside essentially in the form and amplitude of the time solutions, both in the case of multiple equilibria and in the presence of periodic behavior (see Fig. 5).

One of our goals in this work was to devise an analysis technique applicable to a wide range of regulatory architectures. As a result, for tractability we made a number of simplifying assumptions. In particular, we have not considered regulatory mechanisms such as product inhibition, allosteric and other post-translational effects that are common in metabolic pathways. A key step in our method is that the solution of quasi steady state equation in (5) depends only on enzyme concentrations. In such case, the complete system (1)–(2) can be projected onto the state space of enzymes, partitioned into suitably defined polytopes. We note that this does require the

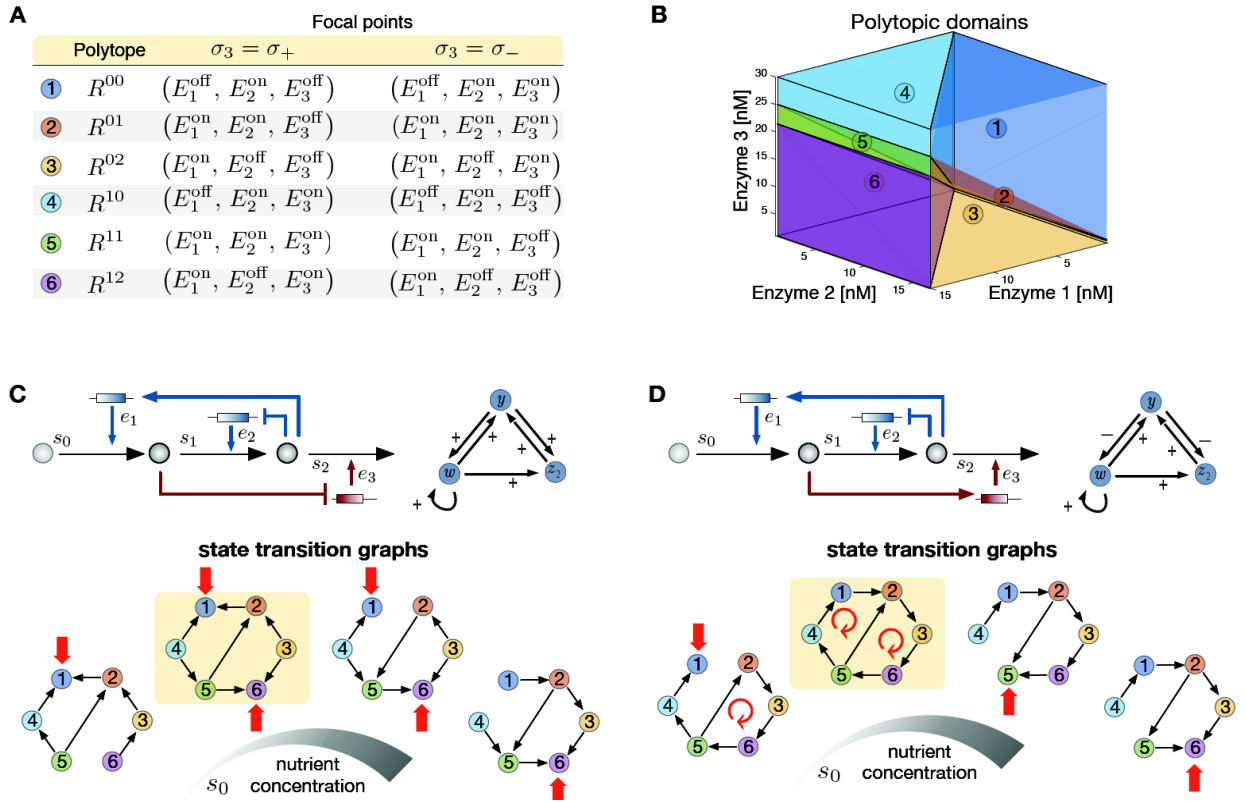


Fig. 4. State transition graphs for the example in Fig. 3 and equation (22). **(A)** Polytopic domains and their focal points as defined in equations (15) and (16). **(B)** Positive orthant partitioned into the six polytopic domains in Definition 6. **(C-D)** Pathway diagram and state transition graphs for the system with negative ( $\sigma = [1, -1, -1]$ ) (C) and positive ( $\sigma = [1, -1, 1]$ ) (D) regulation of enzyme  $e_3$ . The interaction graphs are given for the transformed variables defined in Fig. 2a(left), i.e.  $w = z_1 = e_1 - e_3/\beta_1$ ,  $y = e_1 - e_2/\beta_3$ , and  $z_2 = e_1 - e_3/\beta_2$ . The state transition graphs were computed with Algorithm 1 for increasing nutrient concentrations  $s_0 = \{0.5, 1.5, 4, 10\} \mu\text{M}$ ; arrows denote sink nodes or graph cycles, respectively. The two highlighted graphs are for  $s_0 = 1.5, \mu\text{M}$  and suggest the bistable and oscillatory dynamics.

explicit solution of the quasi steady state equation: it is sufficient to guarantee the existence of a unique solution for the regulatory metabolites in terms of enzyme concentrations. Specific types of post-translational mechanisms may lead to quasi steady state equations that are solvable and respect these conditions. In such cases our approach could be extended with minor modifications, taking into account the geometry of the resulting polytopic partition. A similar reasoning applies to other potential extensions of our method, including the analysis of branched pathways.

In conclusion, our results draw a link between various architectures of transcriptional regulation and classes of prototypical networks, with all the positive and negative feedback loops explicitly identified. This allowed us to single out candidate global dynamics using graph analyses, and thus uncover the enormous dynamical diver-

sity that can emerge in metabolic pathways under gene regulation.

## Acknowledgements

M.C. was partly supported by the French agency for research through project ICycle ANR-16-CE33-0016-01. D.O. was partly supported by the Human Frontier Science Program through a Young Investigator Grant (RGY0076-2015).

## References

- [1] V. Baldazzi, D. Ropers, J. Geiselmann, D. Kahn, and H. de Jong. Importance of metabolic coupling for the dynamics of gene expression following a diauxic shift in *Escherichia coli*. *J. Theor. Biol.*, 295:100–115, 2012.

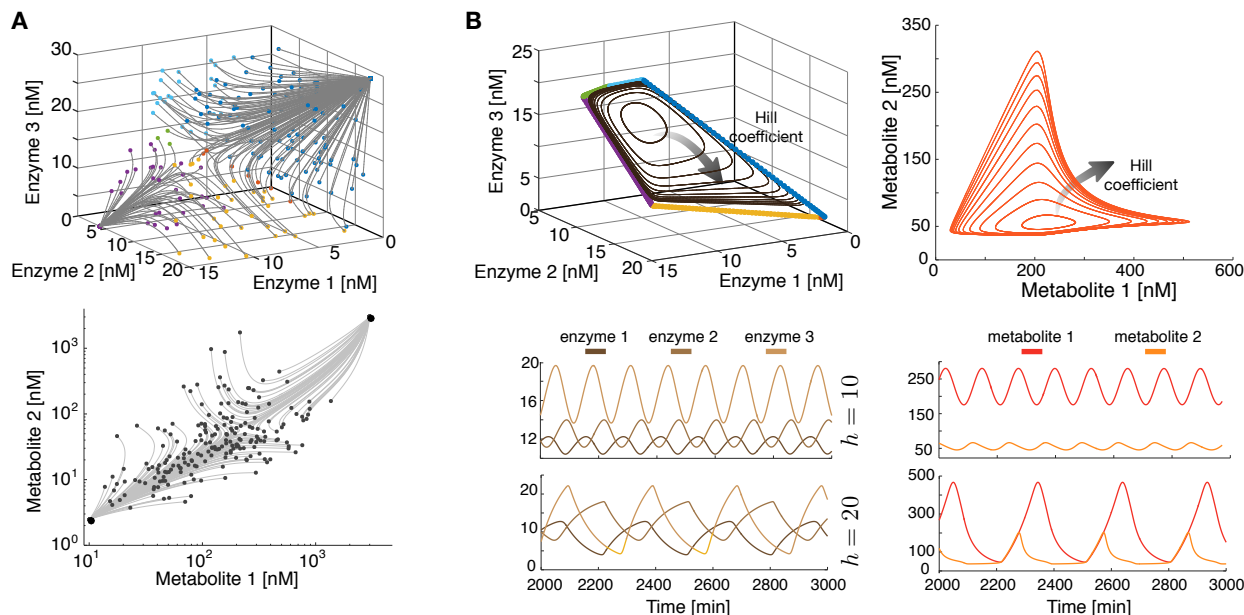


Fig. 5. Simulations of the example system in Fig. 3 and equation (22). **(A)** For the case of negative regulation of  $e_3$  (Fig. 4C), we observe bistable dynamics as predicted by the state transition graph highlighted in Fig. 4C. Plots show the state trajectories for 500 randomized initial conditions and Hill coefficient  $h = 4$ . The enzyme initial conditions are coloured according to the polytopic domain to which they belong, defined in Fig. 4B. **(B)** For the case of positive regulation of  $e_3$  (Fig. 4D), we found a stable periodic orbit as predicted by the state transition graph highlighted in Fig. 4D and our theoretical analysis. Top plots show the limit cycle for high and increasing values of the Hill coefficient, starting from  $h = 10$ . For large values of  $h$ , the orbit approaches the one predicted by the piecewise affine model under timescale separation; we have coloured the limiting orbit according to the polytopic domains defined in Fig. 4B). Bottom plots shows the time courses for  $h = 10$  and  $h = 20$ .

- [2] G. Batt, C. Belta, and R. Weiss. Temporal logic analysis of gene networks under parameter uncertainty. *IEEE Transactions on Automatic Control*, 53:215–229, 2008.
- [3] Irene M Brockman and Kristala L J Prather. Dynamic metabolic engineering: New strategies for developing responsive cell factories. *Biotechnology Journal*, 10(9):1360–1369, sep 2015.
- [4] R. Casey, H. de Jong, and J.-L. Gouzé. Piecewise-linear models of genetic regulatory networks: equilibria and their stability. *J. Math. Biol.*, 52:27–56, 2006.
- [5] M. Chaves, D. A. Oyarzún, and Jean-Luc Gouzé. Analysis of a genetic-metabolic oscillator with piecewise linear models. Submitted, 2018.
- [6] Victor Chubukov, Luca Gerosa, Karl Kochanowski, and Uwe Sauer. Coordination of microbial metabolism. *Nature reviews. Microbiology*, 12(5):327–40, may 2014.
- [7] H. De Jong. Modeling and simulation of genetic regulatory systems: A literature review. *Journal of Computational Biology*, 9:67–103, 2002.
- [8] D. Del Vecchio, A.J. Dy, and Y. Qian. Control theory meets synthetic biology. *Journal of The Royal Society Interface*, 13(120), 2016.
- [9] R. Edwards, A. Machina, G. McGregor, and P. van den Driessche. A modelling framework for gene regulatory networks involving transcription and translation. *Bull. Math. Biol.*, 77:953–983, 2015.
- [10] W.R. Farmer and J.C. Liao. Improving lycopene production in *Escherichia coli* by engineering metabolic control. *Nature Biotechnology*, 18:533–537, 2000.
- [11] Eileen Fung, Wilson W. Wong, Jason K. Suen, Thomas Bulter, Sun gu Lee, and James C. Liao. A synthetic gene-metabolic oscillator. *Nature*, 435(7038):118–122, May 2005.
- [12] Leon Glass and Stuart A. Kauffman. The logical analysis of continuous, nonlinear biochemical control networks. *J. Theor. Biol.*, 39:103–129, 1973.
- [13] Apoorv Gupta, Irene M Brockman Reizman, Christopher R Reisch, and Kristala L J Prather. Dynamic regulation of metabolic flux in engineered bacteria using a pathway-independent quorum-sensing circuit. *Nature biotechnology*, 35(3):273–279, mar 2017.
- [14] F. He, E. Murabito, and H.V. Westerhoff. Synthetic biology and regulatory networks: where metabolic systems biology meets control engineering. *Journal of The Royal Society Interface*, 13(117), 2016.
- [15] R. Heinrich and S. Schuster. *The regulation of cellular systems*. Chapman & Hall, 1996.
- [16] D. Hudson and R. Edwards. Dynamics of transcription-translation networks. *Physica D*, 331:102–11, 2016.
- [17] J.-L. Gouzé. Positive and negative circuits in dynamical systems. *J. Biological Systems*, 6(1):11–15, 1998.
- [18] J.D. Keasling. Synthetic biology and the development of tools for metabolic engineering. *Metabolic Engineering*, 14:189–195, 2012.
- [19] H.K. Khalil. *Nonlinear systems*. Prentice Hall, New Jersey, 2002.
- [20] J. Kuntz, D.A. Oyarzún, and G.-B. Stan. Model reduction of genetic-metabolic systems using timescale separation. In *System Theoretic Approaches to Systems and Synthetic Biology*. Springer-Verlag, 2014.

- [21] Di Liu, Ahmad A. Mannan, Yichao Han, Diego A. Oyarzún, and Fuzhong Zhang. Dynamic metabolic control: towards precision engineering of metabolism. *Journal of Industrial Microbiology & Biotechnology*, 45(7):1–9, jan 2018.
- [22] A. A. Mannan, D. Liu, F. Zhang, and D. A. Oyarzún. Fundamental Design Principles for Transcription-Factor-Based Metabolite Biosensors. *ACS Synthetic Biology*, 6(10):1851–1859, 2017.
- [23] D. A. Oyarzún and M. Chaves. Design of a bistable switch to control cellular uptake. *Journal of The Royal Society Interface*, 12(20150618), 2015.
- [24] D. A. Oyarzún, M. Chaves, and M. Hoffmeyer-Zlotnik. Multistability and oscillations in genetic control of metabolism. *J. Theor. Biol.*, 295:139–153, 2012.
- [25] D. A. Oyarzún and G.-B. V. Stan. Synthetic gene circuits for metabolic control: design trade-offs and constraints. *Journal of the Royal Society Interface*, 10(78), 2013.
- [26] E. Remy, P. Ruet, and D. Thieffry. Graphics requirement for multistability and attractive cycles in a Boolean dynamical framework. *Advances in Applied Mathematics*, 41:335–350, 2008.
- [27] E. Reznik, T.J. Kaper, and D. Segrè. The dynamics of hybrid metabolic-genetic oscillators. *Chaos*, 23:013132, 2013.
- [28] A. Richard and J.-P. Comet. Necessary conditions for multistationarity in discrete dynamical systems. *Discrete Applied Mathematics*, 155:2403–2413, 2007.
- [29] D. Ropers, V. Baldazzi, and H. de Jong. Model reduction using piecewise-linear approximations preserves dynamic properties of the carbon starvation response in *Escherichia coli*. *IEEE/ACM Trans. Comp. Biology and Bioinformatics*, 8:166–181, 2011.
- [30] L.A. Segel. On the validity of the steady state assumption of enzyme kinetics. *Bull. Math. Biol.*, 50:579–593, 1988.
- [31] Y. Soma and T. Hanai. Self-induced metabolic state switching by a tunable cell density sensor for microbial isopropanol production. *Metabolic engineering*, 30:7–15, 2015.
- [32] C. Soulé. Graphical requirements for multistationarity. *ComplexUs*, 1:123–133, 2003.
- [33] Jason T. Stevens and James M. Carothers. Designing RNA-based genetic control systems for efficient production from engineered metabolic pathways. *ACS Synthetic Biology*, 4(2):107–115, 2015.
- [34] R. Thomas. Boolean formalization of genetic control circuits. *J. Theor. Biol.*, 42:563–585, 1973.
- [35] R. Thomas and M. Kaufman. Multistationarity, the basis of cell differentiation and memory. I. structural conditions of multistationarity and other nontrivial behavior. *Chaos*, 11:165–179, 2001.
- [36] B. Wang, R.I. Kitney, N. Joly, and M. Buck. Engineering modular and orthogonal genetic logic gates for robust digital-like synthetic biology. *Nature Communications*, 2:580, 2011.
- [37] P. Xu, L. Li, F. Zhang, G. Stephanopoulos, and M. Koffas. Improving fatty acids production by engineering dynamic pathway regulation and metabolic control. *Proceedings of the National Academy of Sciences*, 111(31):11299–11304, aug 2014.
- [38] G. Yagil and E. Yagil. On the relation between effector concentration and the rate of induced enzyme synthesis. *Biophysical Journal*, 11:11–27, 1971.