

Research article

Open Access

## Specialized or flexible feed-forward loop motifs: a question of topology

Javier Macía\*<sup>†1</sup>, Stefanie Widder<sup>†1</sup> and Ricard Solé<sup>1,2</sup>

Address: <sup>1</sup>Complex Systems Lab (ICREA-UPF), Barcelona Biomedical Research Park (PRBB-GRIB), Dr. Aiguader 88, 08003 Barcelona, Spain and <sup>2</sup>Santa Fe Institute, 1399 Hyde Park Road, Santa Fe NM 87501, USA

Email: Javier Macía\* - javier.macia@upf.edu; Stefanie Widder - stefanie.widder@upf.edu; Ricard Solé - ricard.sole@upf.edu

\* Corresponding author †Equal contributors

Published: 31 August 2009

Received: 28 January 2009

BMC Systems Biology 2009, 3:84 doi:10.1186/1752-0509-3-84

Accepted: 31 August 2009

This article is available from: <http://www.biomedcentral.com/1752-0509/3/84>

© 2009 Macía et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### Abstract

**Background:** Network motifs are recurrent interaction patterns, which are significantly more often encountered in biological interaction graphs than expected from random nets. Their existence raises questions concerning their emergence and functional capacities. In this context, it has been shown that feed forward loops (FFL) composed of three genes are capable of processing external signals by responding in a very specific, robust manner, either accelerating or delaying responses. Early studies suggested a one-to-one mapping between topology and dynamics but such view has been repeatedly questioned. The FFL's function has been attributed to this specific response. A general response analysis is difficult, because one is dealing with the dynamical trajectory of a system towards a new regime in response to external signals.

**Results:** We have developed an analytical method that allows us to systematically explore the patterns and probabilities of the emergence for a specific dynamical response. The method is based on a rather simple, but powerful geometrical analysis of the system's nullclines complemented by an appropriate formalization of the response probability.

**Conclusion:** Our analysis allows to determine unambiguously the relationship between motif topology and the set of potentially implementable functions. The distribution probability distributions are linked to the degree of specialization or flexibility of the given network topology. The implications for the emergence of different motif topologies in complex networks are outlined.

### Background

Molecular networks in cells are highly complex and dynamic. The global behaviour of these webs and their behavioral patterns are far too complicated to intuitively understand their logic. One way to address this problem is to represent them in terms of simplified interaction graphs combining both biological data and mathematical methods [1-6].

Much effort has been devoted to extract some general features of such networks, dissect them into hierarchical levels, modules and motifs to understand their functionalities, dynamics and evolution [7-16]. Simple switches and oscillators have been shown to map to the core processes of biological decision-making, implemented by two- or three-gene network motifs and characterized by their behaviour around the systems' fixed points [17-22]. However, it is reasonable to think that not only the system's steady state is of interest, but also the

way such equilibrium is achieved. Such transient behavior might be characteristic, somehow representing the function performed by the genetic circuitry. In some circumstances, such as in stress responses, a quick change might be favorable [23], whereas in other occasions, e.g. cell-cell intercommunication, it might be more adequate to filter noisy signals and respond only under absolute certainty [24].

Transcriptional networks regulating cell responses exhibit several biochemical wiring patterns, termed network motifs, which appear at frequencies much higher than expected by chance, suggesting that they may have specific functions in the information processing performed by the network. Over the last years, powerful bioinformatic tools such as FANMOD [25] have been developed to detect motif distributions in complex transcriptional networks. One of these motifs is the feed-forward loop (FFL), defined by a transcription factor  $X$  that regulates a second transcription factor  $Y$ , such that both  $X$  and  $Y$  jointly regulate a target gene  $Z$  (figure 1a-b). Many examples of FFLs can be found in complex transcriptional networks. For example, in *E. Coli*, FFL is present in the L-arabinose system, where protein Crp is the general transcription factor ( $X$ ) and AraC is the specific transcription factor ( $Y$ ). This motif regulates 40 effector operons in 22 different systems in the network database [26]. A second example can be found in *Saccharomyces* network, where the protein *Mcm1* ( $X$ ) regulates the expression of *Swi4* ( $Y$ ). Both proteins *Mcm1* and *Swi4* regulate the final expression of *Clb2*. In the yeast network, 39 regulators have been found that are involved in 49 feedforward loops potentially controlling 240 genes [27]. In general, FFLs are known to be associated to multiple key regulations, exhibiting different functionalities, e.g. under conditions of glucose starvation (*CRP*), nitrogen limitation (*rpoN*), and noxious drugs (*rob*), these regulators act as  $X$  in a C1 type FFL. On the other hand, 11 type FFLs in yeast include anaerobic metabolism (*HAP1* as  $X$ ) and nitrogen starvation (*DAL80* or *GLN3* as  $X$ ) systems [28]. In this context, the question about the relation between the functional response implemented by FFLs and their topology arises. The study of the response of three-gene feed-forward loops upon external input shows that they are capable of either implementing transient pulsing (rapid) or, filtering (delayed grader) dynamics [28-35]. However, despite it seems clear that motif topology has an impact on its functionality, is the mapping between motif topology and the possible dynamics one-to-one? Some studies have demonstrated that topology does not necessarily determine function [13,36,37]. Most analysis focused on motif's function have been carried out considering single motif networks. However, recent studies [38,39] have provided evidence that for complex networks, the embedding of the motif

with the rest of the network needs to be taken into account.

Here we have developed a method to systematically study the different functions which can be implemented by each FFL motif and how the topology determines univocally the distribution of probabilities for these functions. As will be shown below, this distribution is correlated with the degree of specialization or flexibility of each motif, by taking into account the different likelihood to perform any function. In other words: topology determines the motif's level of functional specialization. Recently, a similar question on the context of genetic clocks has been addressed [40]. The conclusions of the study suggest that for these clocks topology does not determine dynamics univocally. Although our analysis focuses on single motifs, our results provide new insights to understand the different distribution of motifs in more complex networks, as we will discuss later.

In order to analyze the relation between FFL's functionalities and topology we will describe our biological systems in terms of a set of differential equations:

$$\begin{aligned}\dot{Y} &= g(X) \\ \dot{Z} &= h(X, Y)\end{aligned}\tag{1}$$

describing how concentration of different species  $Y$  and  $Z$  change during time. Here  $\dot{Y}$  and  $\dot{Z}$  represent the derivatives  $dY/dt$  and  $dZ/dt$ , respectively. The FFL topology is implicitly described by functions  $g(X)$  and  $h(X, Y)$ . Assuming that expression of  $X$  is unregulated, the dynamics of the system can be represented in a two-dimensional diagram displaying  $Y$  against  $Z$ , the so called phase space. In the absence of input, the system evolves towards a stable state, i.e. a specific set of values for the concentrations of  $Y$  and  $Z$  that remain constant over time. This stable state is determined by the crossing of the so-called nullclines of the system [41] described by the curves  $\dot{Y} = 0$  and  $\dot{Z} = 0$ , i.e.  $g(X) = 0$  and  $h(X, Y) = 0$  respectively. These curves define the points of the phase space where  $Y$  and  $Z$  do not change. The nullclines capture the essence of the dynamical potential of each component and the relevant chemical, physical or biological constraints. Their shapes reflect saturation effects, forbidden ranges of variables or how fast each component responds to perturbations. In other words, a careful analysis of nullclines allows us to understand the dynamics of the underlying systems and its biological implications. The crossings between both nullclines define stable fixed points, where concentrations of both  $Y$  and  $Z$  remain constant. Upon input the shape of the nullclines change, providing a new, different stable state. Hence, the system will evolve toward this new state

following a trajectory, i.e. a set of intermediate states (Y, Z) within the available phase space.

As will be shown below, we can analyse the geometrical requirements necessary in order to observe different responses to be generated and the probability for a certain FFL to implement a given function independent of the set of parameters. The main message of our work is that topological interactions encode the shape of the nullclines, which in turn determines the limits of possibilities. The functional response of the FFL depends on the parameter configuration within these limitations. In other words, motif structure does not determine its function, but encodes the probability of potential functions that can be implemented. This paper is organized as follows: We first introduce the general model of the FFL based on ODEs and their respective nullclines. Departing from the nullcline scenarios we determine the constraints imposed onto the dynamics of the response of the FFL. Finally, these constraints are formalized analytically in such a way that all feasible types of behaviour can be evaluated. This evaluation allows for the first time to draw generally valid conclusions on the relation between motif topology and function in FFLs.

## Results and Discussion

### General model of the FFL

The analysis is focused on the most general FFL formed by three genes. We assume that the gene circuit acts as a functional unit responding to an external input by producing output. In figure (1) we show a schematic representation of the FFL, depicting three genes  $G_X$ ,  $G_Y$  and  $G_Z$ , with regulatory interactions among each other via their corresponding proteins X, Y and Z.

Gene  $G_X$  (not shown) has a constant production of its protein which is independent of the regulation of the other proteins. We assume that X, however, is synthesized in its inactive form and needs an external effector (the input of the circuit) for its activation. The concentration of active X without input is negligible. Upon addition of effector, the activation of X proceeds very rapidly compared to synthesis and decay and hence, can be approximated by a step function

$$\begin{cases} X > 0 & \text{if effector is present} \\ X = 0 & \text{otherwise} \end{cases} \quad (2)$$

as in [28]. In the following, X will denote the active protein. Transcription of  $G_Y$  and  $G_Z$  is subjected to internal regulation. Production of Y depends on regulator X, whereas production of Z depends on both X and Y. The FFL can be described by the following ODEs

$$\dot{Y} = \gamma_Y \left( \frac{1 + \alpha^X \omega_Y^X X^n}{1 + \omega_Y^X X^n} \right) - d_Y Y \quad (3)$$

$$\dot{Z} = \gamma_Z \left( \frac{1 + \beta^X \omega_Z^X X^n + \beta^Y \omega_Z^Y Y^m + \beta^{XY} \omega_Z^{XY} X^n Y^m}{1 + \omega_Z^X X^n + \omega_Z^Y Y^m + \omega_Z^{XY} X^n Y^m} \right) - d_Z Z. \quad (4)$$

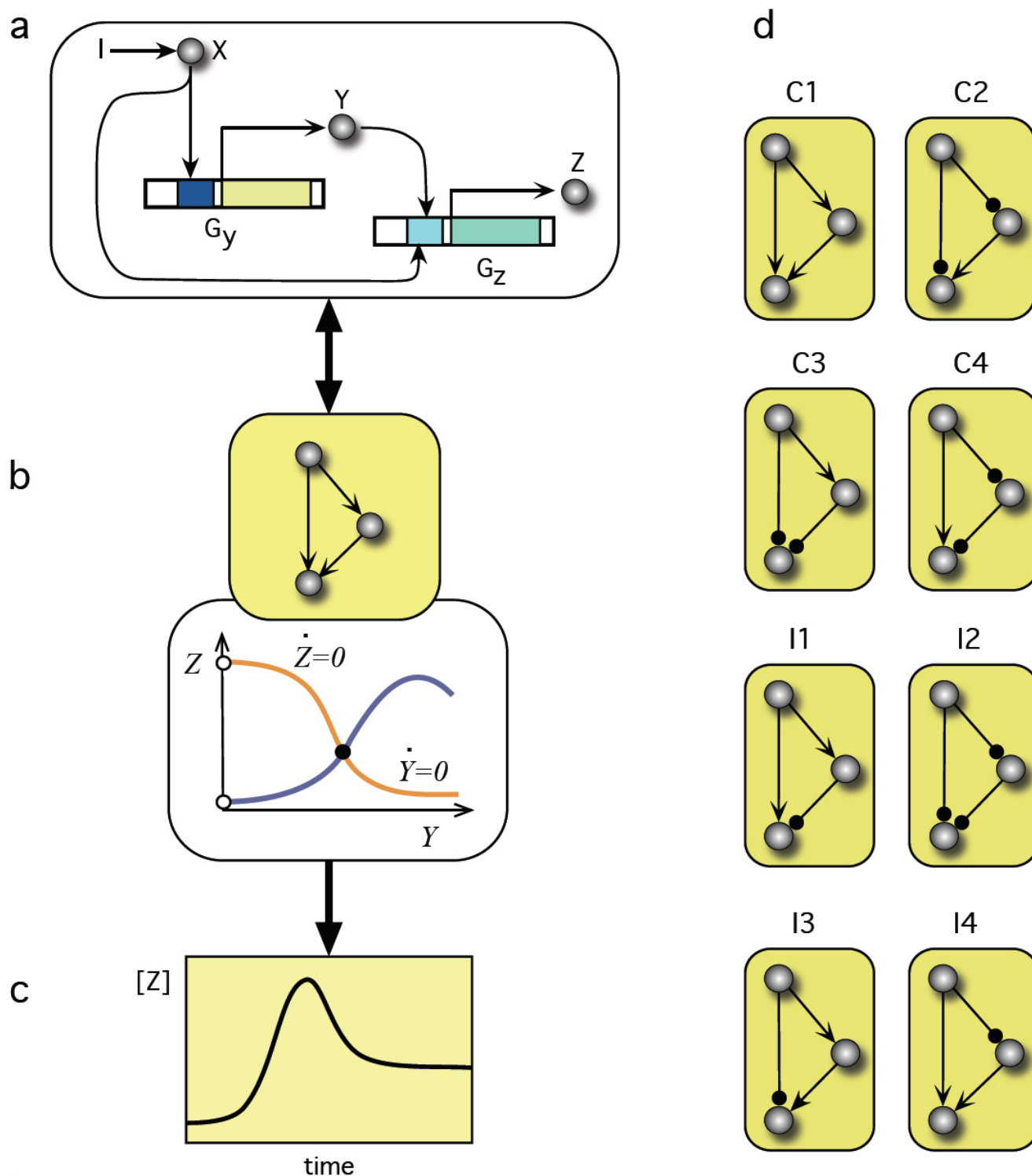
The parameter  $\gamma_i$  represents the basal production of protein  $i$ , where  $i = \{X, Y, Z\}$ . In this parameter we subsume the concentration of all biochemical elements which remain constant in time. The degradation rate of protein  $i$  is denoted as  $d_i$ . The binding equilibrium of the regulators  $j$  with the gene  $G_i$  are denoted by  $\omega_i^j$ , with  $j = \{X, Y, Z\}$ . The tunable positive parameters  $\alpha^x$  and  $\beta$  describe the type of regulatory interactions, i.e. activation or inhibition, for gene  $G_Y$  and  $G_Z$ , respectively, without any predefined specific assumptions. They provide the regulatory rates with respect to the basal transcription. Values  $< 1$  correspond to inhibitory regulation, whereas  $> 1$  accounts for activation. The parameter  $\beta^{XY}$  accounts for the simultaneous regulation of  $G_Z$ .

Traditionally, studies on FFL dynamics have been performed under the assumption of Boolean logic [28,42] for the control of the output regulation. The presented model includes all these cases as specific subsets of parameters. For example, assuming a Boolean AND logic for a circuit, where the output Z is positively regulated by Y and X, is described by the following parameter configuration:  $\beta^x = \beta^y = 1$ ,  $\beta^{XY} > 1$  and  $\omega_z^{XY} = \omega_z^X \omega_z^Y$ . The same circuit displaying OR logic requires  $\beta^x = \beta^y > 1$ ,  $\beta^{XY} = 0$  and  $\omega_z^{XY} = 0$ .

Finally,  $n$  and  $m$  denote the degree of multimerization of the regulators. The presented results, however, are considering the general case independent of the size of the regulatory factors.

### Nullclines' analysis: Changes of the nullclines upon input

The system's fixed points can be determined by inspection of the crossings of its nullclines. The studied input-output system will remain stable unless an external input pushes the FFL towards another stable equilibrium. The FFLs dynamical response upon this change depends on a specific configuration of circuit parameters. Notwithstanding, the possible response-dynamics are restricted by the shape of the nullclines. The general expressions for the



**Figure 1**  
**Schematic representation of the FFL motif and its functions.** Schematic representation of the FFL motif and its functions. In (a) the genetic organization is shown. Activation signal  $X$  regulates the expression of gene  $G_y$ , whereas gene  $G_z$  is regulated in joined mode by gene product  $Y$  and  $X$ . This small system can be formally represented in a graph (b) with a given topology and a set of dynamical equations describing the kinetics of its regulatory interactions. For a given set of parameters, a specific type of response (c) might be observed. In (d) all possible regulatory combinations are shown.

nullclines are obtained imposing  $\dot{Y} = 0$  and  $\dot{Z} = 0$  [41] and written as follows:

$$(Y)_{\dot{Y}=0} = \frac{\gamma Y}{d_Y} \left( \frac{1 + \alpha^X \omega_Y^X X^n}{1 + \omega_Y X^n} \right) \quad (5)$$

$$(Z)_{\dot{Z}=0} = \frac{\gamma Z}{d_Z} \left( \frac{1 + \beta^X \omega_Z^X X^n + \beta^Y \omega_Z^Y Y^m + \beta^{XY} \omega_Z^{XY} X^n Y^m}{1 + \omega_Z^X X^n + \omega_Z^Y Y^m + \omega_Z^{XY} X^n Y^m} \right) \quad (6)$$

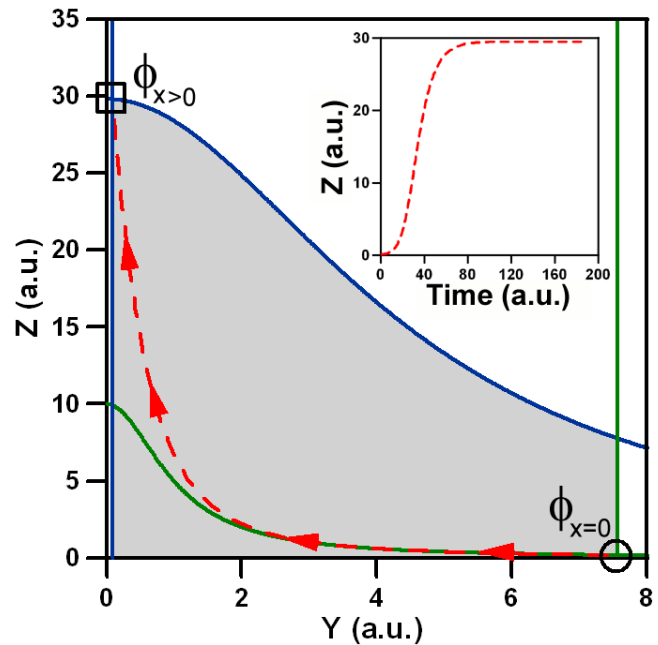
For a general analysis of the nullclines' shape independent of the specific parameters and the size of the multimeric regulators, we focus on a simple set of geometrical features. Within the phase space spanned by  $\{Y, Z\}$ , the nullcline (5) is a simple, vertical straight line. Nullcline (6) shows an horizontal asymptote located at

$$Z_{HA} = \frac{\gamma Z}{d_Z} \left( \frac{\beta^Y \omega_Z^Y + \beta^{XY} \omega_Z^{XY} X^n}{\omega_Z^Y + \omega_Z^{XY} X^n} \right) \quad (7)$$

and crosses the vertical axis at

$$Z_0 = \frac{\gamma Z}{d_Z} \left( \frac{1 + \beta^X \omega_Z^Y X^n}{1 + \omega_Z^Y X^n} \right) \quad (8)$$

Furthermore, expression (6) shows a single inflection point, i.e. a point where curvature changes from concave to convex or viceversa, in the biological domain defined by  $Y \geq 0$  and  $Z \geq 0$ , but no extrema (local maximum or minimum). In order to understand how the input triggers the dynamical response, we study the configuration of the two nullclines with  $X = 0$ , i.e. no input, versus  $X > 0$ , i.e. with input. In figure (2) we show a numerical example. In absence of input, the system is governed by a single stable fixed point, denoted by  $\phi_{X=0}$ , located at the crossing of the nullclines (5) and (6). On addition of input, the fixed point moves, because the nullcline configuration changes. Now the system shifts from  $\phi_{X=0}$  to the new stable fixed point  $\phi_{X>0}$ . The point  $\phi_{X>0}$  is determined by the crossing of the new location of the nullclines for  $X > 0$ . The dynamical response of the system, which is represented by the trajectory of the FFL in phase space, is generated by this change of the stable regime. A subset of four parameters, i.e.  $\{\alpha^X, \beta^X, \beta^Y, \beta^{XY}\}$ , describing the interactions of the proteins, classify the type of the circuits into coherent and incoherent according to [28]. Without input, only  $\beta^Y$  is relevant for the geometry of the nullcline  $\dot{Z} = 0$ . If  $\beta^Y > 1$  the nullcline rises, because  $Z_{HA} > Z_0$ , whereas for  $\beta^Y < 1$  the nullcline decreases as shown in figure (3a). Under the

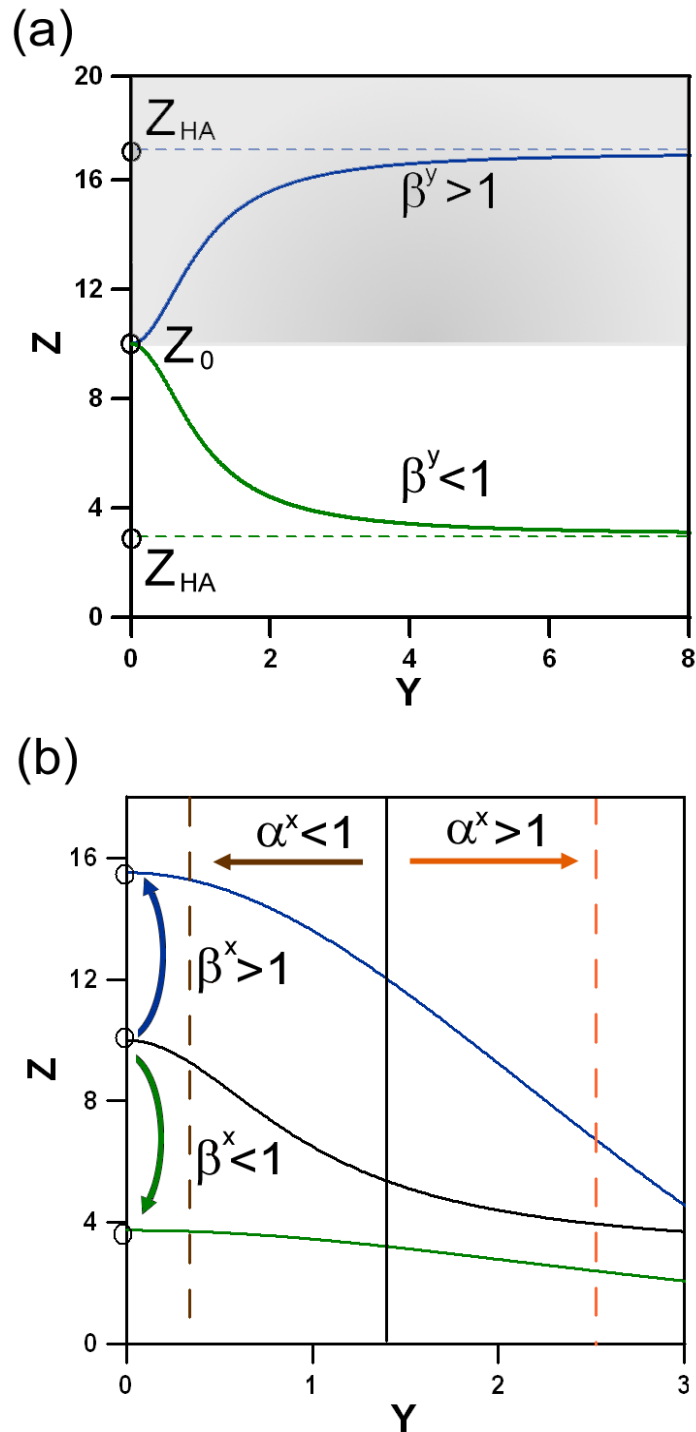


**Figure 2**

**Input effects on the nullclines.** The change of the nullclines upon input is illustrated in these examples. The green lines are the nullclines (5) and (6) without input crossing at point  $\phi_{X=0}$ . In absence of input, the system is stable in this point. Upon input the nullclines change and are depicted as blue lines, shifting the stable point to  $\phi_{X>0}$ . The change of the fixed point, forces the system to evolve towards the new regime. The red dashed line represents its trajectory in the phase space. The inset shows the corresponding time course of the system's output. The parameters of the simulations are:  $\gamma_Y = 1$ ,  $\alpha^X = 0$ ,  $\omega_Y^X = 100$ ,  $d_Y = 0.1$ ,  $\gamma_Z = 1$ ,  $\beta^X = 3$ ,  $\omega_Z^X = 100$ ,  $\beta^Y = 0$ ,  $\omega_Z^Y = 1$ ,  $\beta^{XY} = 0$ ,  $\omega_Z^{XY} = 100$  and  $d_Z = 0.1$ . In absence of input  $X = 0$  and in presence of input we consider  $X = 1$ .

presence of an external input the other parameters become relevant (figure (3b)):

1. The regulation of Y by X, described by  $\alpha^X$ , defines the direction of the shift of the straight nullcline (5) regarding its earlier position for  $X = 0$ .
2. The regulation of Z by X, described by  $\beta^X$ , displaces the crossing of the nullcline (6) with the vertical axis to a higher ( $\beta^X > 1$ ) or a lower ( $\beta^X < 1$ ) value.
3. Finally, the joint regulation  $\beta^{XY}$  of Z by X and Y, strongly influences the location of the horizontal asymptote  $Z_{HA}$ .



**Figure 3**

**The effects of the regulatory parameters on the geometry of the nullclines.** In (a) we show the impact of positive (grey area) or negative (white area) regulation of  $G_Z$  by  $Y$ . The geometrical effect associated to  $\beta^y$  is the only independent of presence of input. In (b) we depict input-dependent effects associated to other regulatory parameters. Nullcline (5) shifts are related to the regulatory interaction of  $X$  upon  $G_Y$ . The orange line represents the qualitative location for  $\alpha^x > 1$ , in brown the location for  $\alpha^x < 1$  is shown. Similarly, the crossing point of nullcline (6) with the vertical axis is shifted to higher (blue) or lower values (green) depending on the type of regulation by  $\beta^x$ .

As we will see in the following sections the geometrical features generated by the biological interactions will form the basis for the dynamical behaviour of the FFL.

**Response of the FFL**

To understand how the changes in the nullclines upon input confine the feasible dynamical response of the circuit, we will focus our attention on two representative cases, namely a coherent and an incoherent type ([28]). We choose circuits C4 and I1, respectively, which differ only in one regulatory interaction in  $G_Y$ , i.e.  $\alpha^X$ . Due to this interaction only nullcline (5) is affected and with it the direction of its shift as can be seen in figure (3b). The other interactions are the same for both circuit architectures ( $\beta^Y > 1$ ,  $\beta^X < 1$ ) and hence the possible nullcline's shapes associated to expression (6) are also the same. In figure (4) we show their subset of changes realized by the nullcline. Based on the positive interaction of  $\beta^X$  the crossing point  $Z_0|_{X>0}$  with the vertical axis always shifts toward higher values for cases with input compared to cases without input ( $Z_0|_{X=0}$ ).

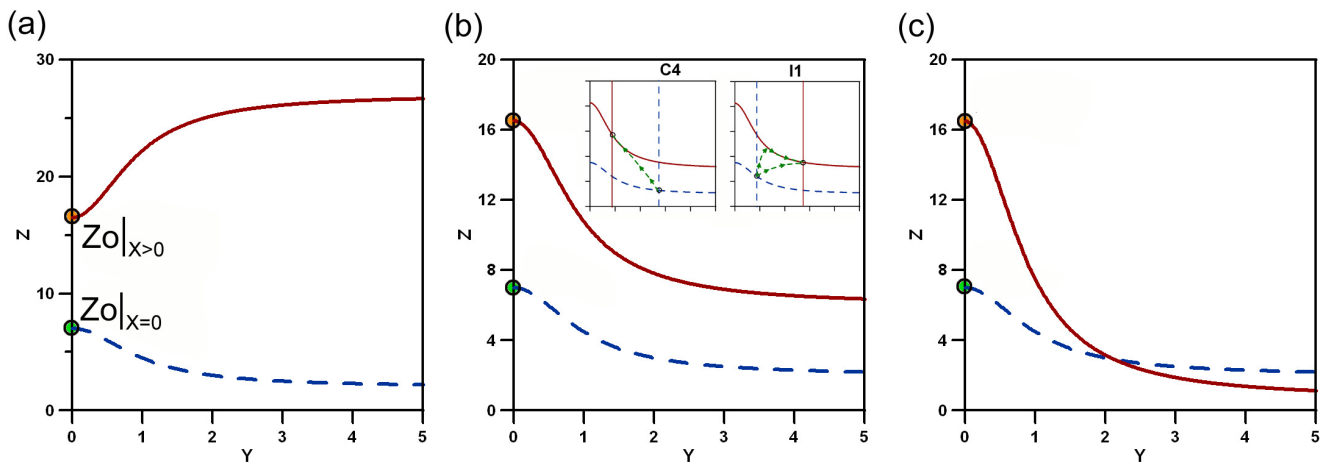
Two biological scenarios for the joint regulation of  $G_Z$  by  $Y$  and  $X$  are plausible: either the complex acts as an activator  $\beta^{XY} > 1$  as shown in figure (4a) or as inhibitor  $\beta^{XY} < 1$ . In (4b) we show the scenario associated with the conditions  $\beta^{XY} < 1$  and  $\beta^{XY} > \beta^Y$ . The nullcline moves down, but does not cross the original nullcline. In (c) conditions  $\beta^{XY} < 1$  and  $\beta^{XY} < \beta^Y$  lead to a single crossing. The same conditions can lead to a double crossing of the nullclines, how-

ever our numerical analysis (data not shown) indicates that the probability to find an adequate configuration of parameters has very low probability ( $< 0.3\%$ ). For simplicity we discard these cases (see Methods for a detailed example of our analysis applied to nullclines with double crossing). By using the relatively simple configuration of the nullclines shown in (b) we already find different possible behaviour of the two circuits C4 and I1 due to their opposing  $\alpha^X$ .

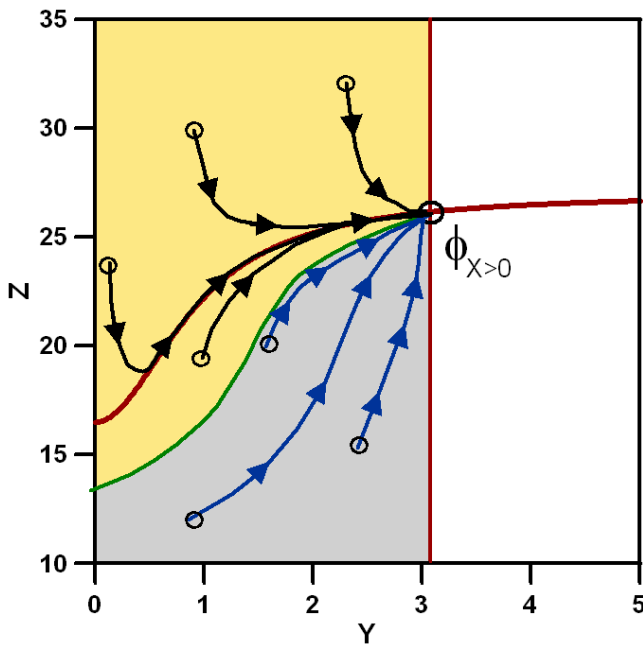
Whereas for C4 the nullcline's (5) shifts to higher values, allowing only for grader trajectories (inset 1), I1 may show instead two different functionalities. The shift direction is to lower values and hence either grader or pulser can be implemented depending on the set of parameters (see inset 2). The fact that a range of feasible functional scenarios can be intuitively deduced, demands for a method of unambiguous discrimination to resolve the problem.

**Separatrix**

Sometimes, two qualitatively different dynamics are possible for a given set of nullclines (figure (4b)). In that case, if we consider any arbitrary initial point  $(Y, Z)$  in phase space (i.e. any arbitrary concentration values of  $Y$  and  $Z$ ) phase space can be divided into two different areas. As shown in figure (5) this partition of phase space is characterized by a different functional behaviour of the trajectories upon input: starting in one part (yellow area) the trajectory will first join/intersect the nullcline  $\dot{Z} = 0$  and



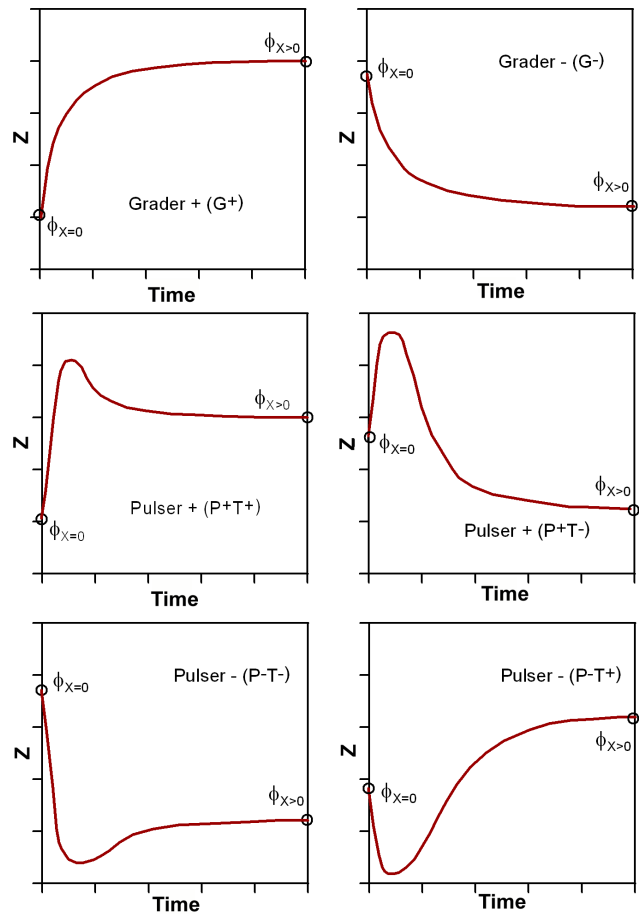
**Figure 4**  
**Possible shapes of nullcline.** The set of possible shapes for the nullcline  $\dot{Z} = 0$  with (brown) and without (blue) input. Figures (a), (b) y (c) correspond to different biological regulations. In (a) the joined regulation by proteins  $X$  and  $Y$  enhances the production of the output  $Z$  ( $\beta^{XY} > 1$ ). In (b) and (c) the production of  $Z$  is reduced ( $\beta^{XY} < 1$ ), but in (b) the reduction is smaller due to  $\beta^{XY} > \beta^Y$ , whereas in (c) we apply strong inhibition  $\beta^{XY} < \beta^Y$ . Insets in (b) show the feasible trajectories for C4 and I1, respectively. For C4 the functionality of the trajectory is uniquely defined as grader, whereas I1 can implement pulser or grader dynamics, depending on the specific parameters.



**Figure 5**  
**Schematic representation of two areas characterizing the behaviour of the trajectory.** The nullclines of the system with input are depicted in brown. Any trajectory starting in the pink area will reach the nullcline  $\dot{Z} = 0$  first and then follow its path to the fixed point  $\varphi_{X>0}$ . Any trajectory starting in the blue area will reach the fixed point directly without intersecting the nullcline first. The frontier between these two regions, the separatrix, is drawn in green. As example some trajectories have been depicted to illustrate their qualitatively different course.

then following the nullcline until it reaches the fixed point. Starting from an arbitrary point in the other region (grey area) the trajectory will only reach the nullcline in the fixed point  $\varphi_{X>0}$ . This qualitative difference in the trajectories gives rise to a different functional behaviours of the circuit (see figure (6)).

The system without input is determined by its fixed point  $\varphi_{X=0}$ . The circuit will exhibit a given dynamic behavior depending on which part of phase space  $\varphi_{X=0}$  is located. The location of the boundary between the two dynamical areas, termed separatrix, allows classifying the functionality of the trajectory. The key elements for the discrimination of the different possible functionalities, for a given geometry of the nullclines, is the relative position of three points: the crossing of the vertical nullcline  $= 0$  without input, with: i) the nullcline  $\dot{Z} = 0$  without input ( $\varphi_{X=0}$ ), ii) the separatrix (S) and with iii) the nullcline  $\dot{Z} = 0$  with input ( $\xi$ ). We illustrate this with an example (circuit I1) in figure (7a-c) where the relative positions of the crossing points and their implication for the time course is out-



**Figure 6**  
**Qualitative representation of all possible timecourse responses realized by FFLs.** Here, 'G' indicates grader and 'P' pulser response of the circuit. The initial slope is represented by '+' or '-' in our nomenclature. The difference of the initial versus the target concentration is denoted as 'T+' or 'T-'.

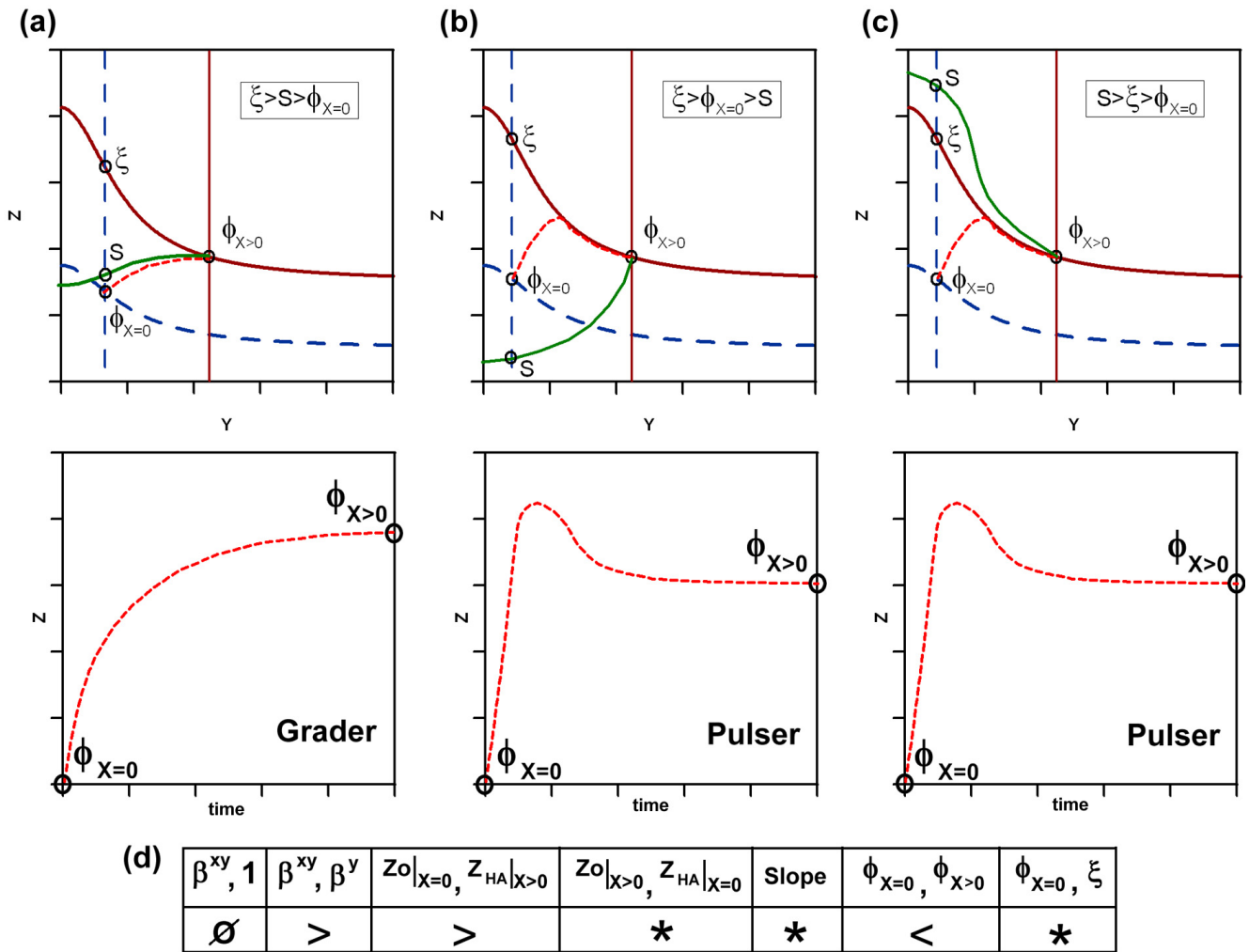
lined. An analytical estimation of the separatrix and all the possible combinations of these elements with its resulting dynamics are presented in Supplementary Materials.

Our analysis can be generally applied to all three-gene FFLs. We have shown that both examples show enough plasticity in the dynamics to implement different type of response. In the next section we will focus on the probability of emergence of the different feasible types of dynamics.

**Probability of emergence of different FFL's dynamics**

In the previous section we have shown that in both examples, C4 and I1, more than one possible dynamic behavior can be obtained depending on the specific set of parameters chosen. Generally we can list six types of





**Figure 7**

**Positional effect of the separatrix.** Here we illustrate the positional effect of the separatrix  $S$ ,  $\xi$  and  $\phi_{X=0}$  on the functional behaviour of the FFL (a-c). The relative position between the three crossing points with the vertical nullcline  $= 0$  without input determines the functionality of the time course. Positioning shown in (a) results in grader, whereas (b) and (c) display pulser dynamics. The respective relative location is outlined in the boxes. In (d) the unique backbone of parametric requirements defining univocally the shape of the nullclines is shown. Each column describes the relation ( $>$  or  $<$ ) between two characteristic elements (head column). The single values are described in the text. Column 'Slope' refers to the slope of the nullcline  $\dot{Z} = 0$  with input ( $X > 0$ ) and can be positive or negative. The columns containing ' $\emptyset$ ' or ' $*$ ' do not need to be explicitly addressed, because of parametric inter-dependencies (see text).

dynamics, namely positive or negative graders ( $G^+$ ,  $G^-$ ) and positive or negative pulsers ( $P^+$ ,  $P^-$ ). Grader response corresponds to an increase (+) or decrease (-) of the concentration of  $Z$  characterized by a transient from the initial to the final state where the concentration of  $Z$  never is higher ( $G^+$ ) or lower ( $G^-$ ) than the final concentration. In general, grader responses are related with responses able to filter noise and respond only on absolute certainty [24]. On the other hand, a pulser response is characterized by a rapid increase ( $P^+$ ) or decrease ( $P^-$ ) of the concentration of  $Z$  reaching higher (+) or lower (-) values of  $Z$  before they

reach the final state. Note that for the pulsers independently of the pulse direction ( $P^+$  or  $P^-$ ) the final concentration of output protein  $Z$  can be higher ( $T^+$ ) or lower ( $T^-$ ) than the initial concentration. Hence we separately analyze four different pulser dynamics, namely  $P^+(T^+)$ ,  $P^+(T^-)$ ,  $P^-(T^+)$ ,  $P^-(T^-)$ . The time courses of the different functionalities are outlined in figure (6). A specific subset of dynamics can be determined for each FFL frequently containing functions, which cannot be implemented. Based on this results the main question is: Are all the feasible dynamics equally probable? To address this question we have per-

formed an analytical study of the parametric requirements necessary to implement a given type of dynamical response.

**Backbone of requirements for the FFL response**

We can deduce geometrically all possible sets of trajectories. We discriminate the necessary parametric conditions for the emergence of one specific dynamic among all the possible, due to the relative position of the initial point  $\varphi_{x=0}$  of the trajectory in relation to the separatrix  $S$ ,  $\xi$  and the specific shape of the nullclines. The parametric conditions, which determine the shape of the nullcline, can be systematically formalized. The key geometrical elements of the nullclines can be described by a set of exclusive parametric combinations defined in a string, which we call the *backbone of requirements*. Each position in this sequence contains the solution for two possible parametric states.

The example shown in figure (7) is meant to illustrate the procedure. The cases in (a-c) display the same geometrical shape and can be described by a single backbone of requirements shown in (d). Note that not all the elements of the parametric sequence need to be defined explicitly. In certain occasions some of them are uniquely defined by previous conditions, denoted by \*, or otherwise do not have an impact on the geometrical scenario, represented by '∅'. For example position one is denoted '∅'. If condition two ( $\beta^{XY} > \beta^Y$ ) is satisfied, position one is not relevant, because both solutions ( $\beta^{XY} > 1$  or  $\beta^{XY} < 1$ ), provide the same geometry. On the other hand position four, denoted as '\*', is always solved as '>', because it can be deduced from the second condition given that  $\beta^X > 1$ .

**Quantification of the dynamical probabilities**

Within the framework of our model, we assume that the feasibility of given dynamical behaviour (transient response) is heavily dependent on the number of exclusive, parametric configurations necessary for the realization of this behaviour. These are subsumed in the backbone of requirements. We have generated all possible backbone sequences for the circuits C4 and I1. For each of these sequences the separatrix discriminates among different dynamics as we have shown in figure (7a-c). In circuit C4 we have found 18 different sequences implementing five different types of dynamics, whereas for circuit I1, 21 sequences implement four different behaviors. The detailed list of backbone sequences are shown in Supplementary Materials. The total number  $N_C$  of possible requirements captured within one sequence  $i$  within a circuit  $k = \{C4, I1\}$  is constant and follows the sum:

$$N_C = N_{ik}^\emptyset + N_{ik}^* + N_{ik}^{suf} \tag{9}$$

Here,  $N_{ik}^\emptyset$  represents the number of requirements having no impact on the nullclines' geometry. For this kind of elements of the sequence, both possible solutions ('>' or '<') are valid and hence discrimination is unnecessary.

Therefore,  $2^{N_{ik}^\emptyset}$  different combinations of parameters are described by the same backbone sequence providing the same geometry for the nullclines. The second term  $N_{ik}^*$ , is the number of elements predefined by other conditions of the backbone. Finally,  $N_{ik}^{suf}$  is the number of requirement actually necessary to determine univocally the shape of the nullcline. In other words, in order to implement a determined backbone sequence it is sufficient to properly establish the conditions  $N_{ik}^{suf}$ .

Once these conditions are set, the rest of the sequence is determined. This allows to calculate the probability for a given set of parameters to implement a certain backbone sequence  $i$  as

$$p_{ik} = \frac{1}{2^{N_{ik}^{suf}}} = \frac{2^{N_{ik}^\emptyset}}{2^{N_C - N_{ik}^*}} \tag{10}$$

For a given FFL topology it is possible to find different sets of parameters compatible with such topology but implementing different dynamics among the six possibilities. Each set of parameters fits in a specific backbone sequence. For each backbone sequence there exist three different relative positions of the separatrix discriminating among the different possible dynamics. This idea is illustrated in figure (7). The same backbone sequence shown in 7d can implement two different dynamics, namely  $G^+$  and  $P^+ T^+$ . The location of the separatrix shown in figures (7a-c) depends on the numerical values of the parameters. If these specific values locate the separatrix as depicted in figure (7a) the FFL motif will implement a grader dynamics, whereas other locations (see figures (7b, c)) will provide a pulser dynamics. The probability of circuit  $k$  to implement a given dynamic  $j = \{G^+, G^-, P^+(T^+), P^+(T^-), P^-(T^+), P^-(T^-)\}$  described by the backbone sequence  $i$  can, thus, be calculated as:

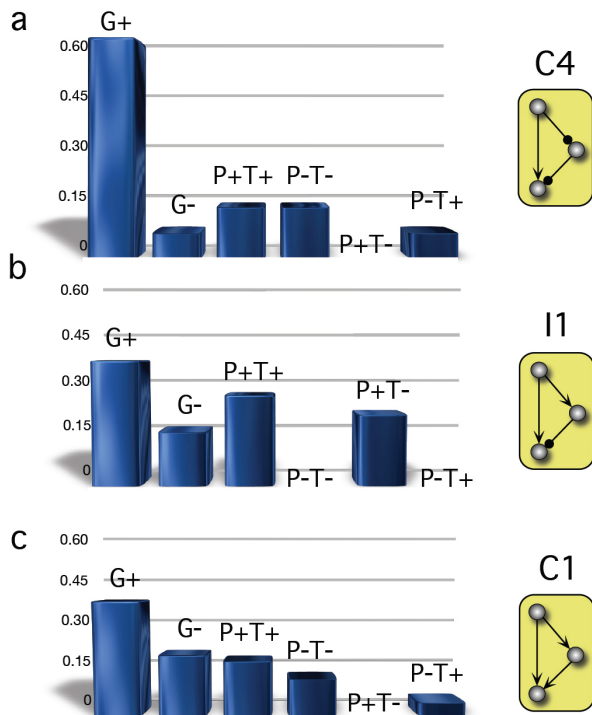
$$P_{ij} = \Omega_k \sum_{i=1}^{T_{jk}} p_{ik} \Psi_{ijk} \tag{11}$$

where  $\Psi_{ijk}$  is the number of equal dynamical outcomes  $j$  for a given backbone  $i$  in the  $k$ -th FFL and  $T_{jk}$  the total number of backbone sequences implementing dynamic  $j$ .

The normalization constant  $\Omega_k$  is defined by imposing the condition  $\sum_{j=1}^6 P_{jk} = 1$  and gives:

$$\Omega_k = \frac{1}{\sum_{j=1}^6 \sum_{i=1}^T \Psi_{ijk} \left( \frac{2^{N_{ik}^\emptyset}}{2^{N_C - N_{ik}^*}} \right)} \quad (12)$$

By calculating the probability of implementation of the different types of dynamics for the different FFLs we obtain an interesting picture regarding our initial question, whether the topology of the FFL implements its function. As shown in figure (8a) the coherent FFL C4 implements five different types of dynamics. Their probability (given an arbitrary set of parameters) is however very different. We find that C4 is significantly specialized for G+ dynamics. The degree of specialization is equally reflected in its robustness versus parametric change. On the other hand, the incoherent FFL I1 potentially implements four different types of dynamics, as shown in figure



**Figure 8**  
**The probability for emergence of specific functionalities.** In (a) the distribution for the coherent FFL C4 is shown. The probability to implement G+ dynamics is significantly higher than to implement any other possible function. In (b) we plot the distribution of the incoherent FFL I1. All possible functions range at similar probability values. The two FFLs differ in a single regulatory interaction.

(8b). The probabilities for I1 to implement these functionalities are within the same range, unlike in C4. In figure (8c) the distribution of probabilities of C1 (the most abundant FFL) obtained performing the same analysis is shown. The specific topology allows for flexibility at the cost of less specialization. We see that the topology of the FFL does not implement its function, but instead the probability of a certain function to arise. If, for a given FFL, the distribution accentuates a certain function, the FFL is said to be specialized for this function. If the probabilities are located within the same range, the topology implements flexibility. Both aspects are equally relevant in terms of adaptation and evolvability.

The highest level of specialization of a given motif would correspond to a distribution of probabilities displaying a single peak, whereas the maximum level of flexibility would correspond to a flat distribution of probabilities, where all the potential dynamics would be equally probable. In order to compare the level of specialization of different FFL topologies we propose to measure the *peakedness* of the probability distribution, i.e. the kurtosis [43] (see Methods for details about kurtosis). The values of kurtosis are  $K_{C1} = 4.6$ ,  $K_{C4} = 7.8$  and  $K_{I1} = 1.28$ , respectively indicating that C1 has an intermediate level of specialization with respect to C4 (more specialized) and I1 (more flexible). The same qualitative distributions are obtained by numerical simulation choosing random sets of parameters and counting the frequency of each dynamics (data not shown).

**Conclusion**

FFL motifs appear frequently in cellular regulatory networks. Despite the efforts devoted to understand how FFLs encode their functionalities, the question about the relation exact between function and topology remained open. In this work we have presented a new analytical formalism based on the geometric analysis of the system's nullclines to elucidate this question. We found that the dynamical response triggered by the external input can be analysed in terms of: i) the nullcline's geometry as described by a backbone sequence of parametric conditions and ii) the specific location of the initial stable state in the phase space with respect to the nullclines and the separatrix. This puts us into the position to unambiguously enumerate the probability of a given FFL to implement a certain function. Our results support this view topology does not define a unique functionality, as have been previously discussed, ([36]). Circuits with the same topology can implement different functions, yet not all of these possible dynamics are equally probable. However, topology defines univocally the distribution of probabilities for the emergence of the different feasible responses.

For illustrative purposes we have analyzed two interesting examples, namely circuits C4 and I1, exhibiting the same regulatory interactions except a single regulation of the gene  $G_Y$  by the protein  $X$ . In these cases we found two paradigmatic scenarios: circuit C4 can implement  $G^+$  response with significantly higher probability than the other feasible dynamics, whereas, I1 exhibit more uniform distribution of probabilities. These results demonstrate that C4 has a specialized topology, optimal for the implementation of grader response, whereas I1 has a high degree of flexibility among different dynamics. Under an evolutionary perspective, a trade-off between these different features, flexibility and specialization, is likely to play an important role. This problem will be investigated elsewhere.

In single motif networks a given function will be implemented with high probability by the most specialized topology. However, in evolved, complex networks other aspects need to be considered. In order to obtain reliable networks, i.e. robust and with high fault tolerance, complex topologies can emerge as a result of the evolutionary process. An evolved and fit network is not necessarily the sum of its optimal sub-modules. In order to provide redundancy and degeneracy, flexible sub-modules, able to change their functionalities with minimal cost are often a good solution to reliability [44]. Future work will be devoted to analyze the implications of these two characteristics in the natural emergence of current biological networks.

**Methods**

**Analytical estimation of the separatrix**

The unique discrimination of the circuit's dynamical behaviour is determined by the location of the initial point  $\varphi_{X=0}$  with respect to the nullclines and the separatrix. The separatrix defines the boundary between two different areas defining the dynamical outcome in phase space. If the initial point of the trajectory lies within one part (in the following  $A_1$ ) it will reach the nullcline  $\dot{Z}=0$  directly at the fixed point ( $\varphi_{X>0}$ ). If it lies in the other part ( $A_2$ ) the trajectory will reach the nullcline before the fixed point and follow its path towards the final state. In order to obtain an analytical expression for the separatrix of phase space, we examine any arbitrary point and its corresponding vector field described by the system's ODEs (see article). Each arbitrary point is defined by its two components ( $Y, Z$ ) in phase space, as is the final point  $\varphi_{X>0} = (Y_f, Z_f)$ . The relation between the horizontal distance  $Y_f - Y$  and the horizontal component of the field and the corresponding vertical elements determine the area. We assume that each point ( $Y, Z$ ) in one area satisfies the local condition

$$\frac{Y_f - Y}{Z_f - Z} > \frac{\dot{Y}}{\dot{Z}} \tag{13}$$

whereas the opposite is true for each point in the other area. The separatrix is defined by the set of points satisfying

$$\frac{Y_f - Y}{Z_f - Z} = \frac{\dot{Y}}{\dot{Z}} \tag{14}$$

Using the system's ODEs we obtain the following expression for the separatrix:

$$Z(Y) = \left( \frac{1}{1 - \frac{dY}{dZ}} \right) \left[ \frac{YZ}{dz} \left( \frac{1 + \beta^X \omega_z^X X^n + \beta^Y \omega_z^Y Y^m + \beta^{XY} \omega_z^{XY} X^n Y^m}{1 + \omega_z^X X^n + \omega_z^Y Y^m + \omega_z^{XY} X^n Y^m} \right) - \frac{dY}{dZ} Z_f \right] \tag{15}$$

In order to test this expression we have performed numerical simulations with random sets of parameters and multiple random initial points. This separatrix expression defines properly the frontier between  $A_1$  and  $A_2$ .

**Relation between dynamics and geometrical elements**

The relation between the geometrical elements  $S$ ,  $\xi$  and  $\varphi_{X=0}$  and the slope of the nullcline  $\dot{Z}=0$  upon input defines the function of the FFL. The combinations are summarized in tables (1) and (2).

**Analysis of the relaxation dynamics after induction**

Our method based on the analysis of nullcline's geometry allows to determine the dynamics of the system upon input as well as to determine the relaxation dynamics when the external input disappears. Initially, in absence of input the system remains stable in a fixed point determined by the crossing between nullclines in the phase space ( $\varphi_{X=0}$ ). After input addition nullclines change providing a new stable fixed point ( $\varphi_{X>0}$ ). The system evolves from the initial point to this new final state following a trajectory in the phase space constrained to the specific geometry of the nullclines upon input. If the external input is removed the nullclines recover the initial geometry, and hence the system evolves towards the initial stable fixed point. Here, the dynamics is constrained to the specific geometry of the nullclines under no input. However, in both cases the same mathematical analysis can be applied. For illustrative purposes a particular example is considered in figure (9). Figure (9a) shows the evolution from the initial state without input, determined by the crossing between nullclines (blue nullclines), to the new stable state upon input (brown nullclines). The dynamics is determined by the location of the separatrix (green line) calculated upon input. In this example, the dynamics is a pulser ( $P^+ T^+$ ). If the input is removed, the system will

**Table 1: Relation between geometry of the nullclines and functionality**

Sl.	Relation	Function	Relation	Function
+	$S > \varphi_{X=0} > \xi$	$P^-$	$\xi > S > \varphi_{X=0}$	$G^+$
-	$S > \varphi_{X=0} > \xi$	$G^-$	$\xi > S > \varphi_{X=0}$	$G^+$ if $Z_f > Z_i$ , else impossible
+	$\varphi_{X=0} > \xi > S$	$P^-$	$\xi > \varphi_{X=0} > S$	$G^+$
-	$\varphi_{X=0} > \xi > S$	$G^-$	$\xi > \varphi_{X=0} > S$	$P^+$
+	$\varphi_{X=0} > S > \xi$	$G^-$ if $Z_f < Z_i$ , else impossible	$S > \xi > \varphi_{X=0}$	$G^+$
-	$\varphi_{X=0} > S > \xi$	$G^-$	$S > \xi > \varphi_{X=0}$	$P^+$

Possible combinations of nullcline and geometrical elements for the evaluation of the FFL functionality considering  $\alpha^Y > 1$ . 'Sl.' denotes the slope of the nullcline  $Z=0$  upon input,  $Z_f$  is the concentration of protein Z in its equilibrium  $\varphi_{X>0}$  and  $Z_i$  its concentration at the initial point  $\varphi_{X=0}$ .

return to the initial state. Figure (9b) shows the location of the separatrix (grey line) calculated for the system without input, where the relaxation dynamics (green dashed line) is a grader ( $G^-$ ). The conditions determining the dynamics are different with ( $\xi > \varphi_{X=0} > S$ ) and without input ( $\varphi_{X>0} > S > \xi$ ) providing different dynamics. The same mathematical analysis has been applied in both cases, only considering  $X > 0$  or  $X = 0$  to analyze induction and relaxation dynamics respectively.

**Analysis of the dynamics in nullclines with two crossings**

In order to simplify or study, geometrical scenarios with two crossings between nullclines have been not considered due to their low probability of emergence (see main text). However, the mathematical analysis applied can be easily extended to these scenarios with double crossings. For illustrative proposes the possible dynamics of the circuit C1 with two crossing between its nullclines is analyzed. Figure (10) shows the four possible combinations of nullclines with and without input. With the same geometrical conditions determined by the value of the parameters that determine the biological interactions of

the circuit, different dynamics can emerge depending on the relative location of the initial and final points. The phase space can be split into three different regions, namely I, II and III. Regions I and III show the same relative position between initial and final point, i.e.  $\varphi_{X=0} < \varphi_{X>0}$  (see figures (10a) and (10b)). In these scenarios, according with table 1 only  $G^+$  dynamics can exist (independently of the location of the separatrix). On the other hand, in the region II there are two different possibilities, namely  $\varphi_{X=0} < \varphi_{X>0}$  (figure (10c)) or  $\varphi_{X=0} > \varphi_{X>0}$  (figure (10d)). In the first case, according with table 1 only  $P^-(T^+)$  dynamics can emerge, whereas in the second case, the unique possible dynamics is  $P^-(T^-)$ . In summary, the example analyzed with double crossing geometry shows how the phase space can be divided in three independent regions and the possible dynamics in each region can be determined using the same analysis performed in phase space with a single crossing between nullclines.

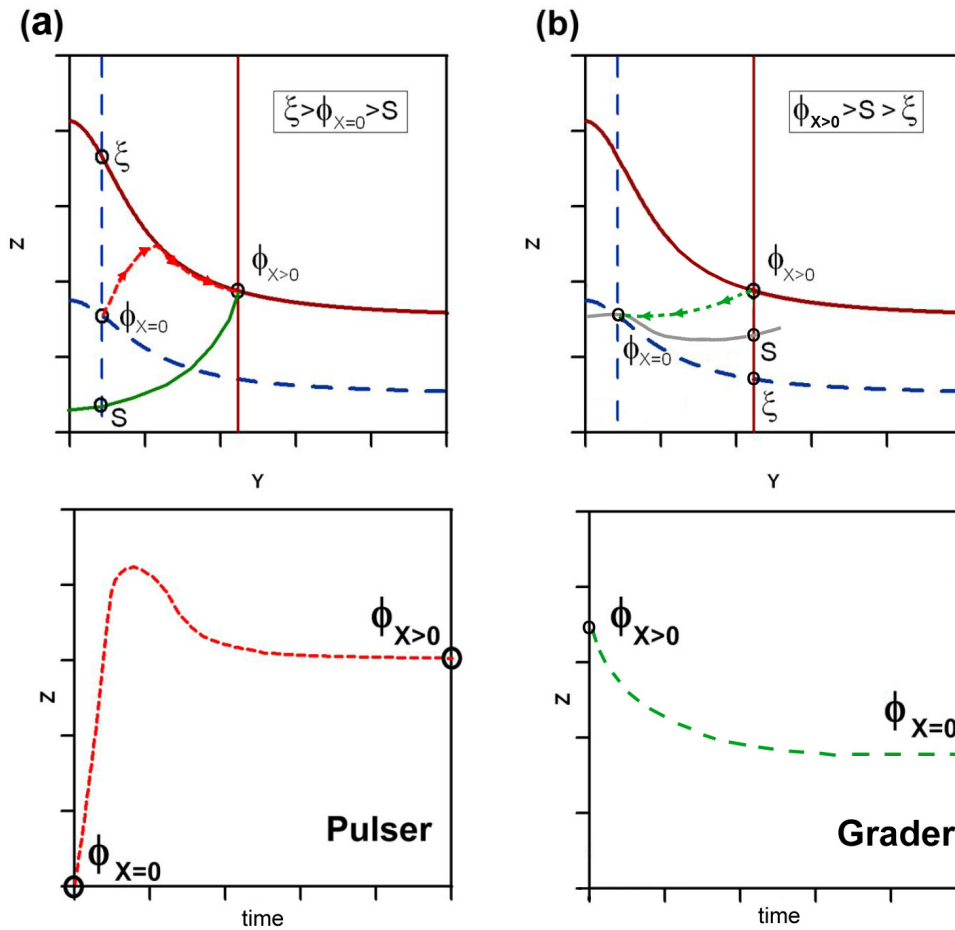
**Backbone sequences and the associated dynamics**

Within the backbone sequence, all geometrical requirements for the full description of the nullcline scenario and

**Table 2: Relation between geometry of the nullclines and functionality**

Sl.	Relation	Function	Relation	Function
+	$S > \varphi_{X=0} > \xi$	$G^-$	$\xi > S > \varphi_{X=0}$	$G^+$ if $Z_f > Z_i$ , else impossible
-	$S > \varphi_{X=0} > \xi$	$P^-$	$\xi > S > \varphi_{X=0}$	$G^+$
+	$\varphi_{X=0} > \xi > S$	$G^-$	$\xi > \varphi_{X=0} > S$	$P^+$
-	$\varphi_{X=0} > \xi > S$	$P^-$	$\xi > \varphi_{X=0} > S$	$G^+$
+	$\varphi_{X=0} > S > \xi$	$G^-$	$S > \xi > \varphi_{X=0}$	$P^+$
-	$\varphi_{X=0} > S > \xi$	$G^-$ if $Z_f < Z_i$ , else impossible	$S > \xi > \varphi_{X=0}$	$G^+$

Possible combinations of nullcline and geometrical elements for the evaluation of the FFL functionality considering  $\alpha^Y < 1$ .



**Figure 9**  
**Schematic example of induction and relaxation dynamics.** In (a) the system evolves from the initial stable state  $\phi_{X=0}$  to the final state  $\phi_{X>0}$  upon addition of an external input. In this case the geometrical constraints of the nullclines upon input (brown) allow for a pulser dynamics ( $P^+T^+$ ). In (b) after input deletion the system evolves from  $\phi_{X>0}$  to  $\phi_{X=0}$ . The relaxation dynamics is determined by the shape of the nullclines without input (blue). Here the dynamic is grader (G). The same mathematical analysis has been performed in both cases, only considering  $X > 0$  or  $X = 0$  to analyze induction and relaxation dynamics respectively.

the resulting dynamics are listed. In the presented tables (3) and (4) all possible sequences for C4 and I1 are shown.

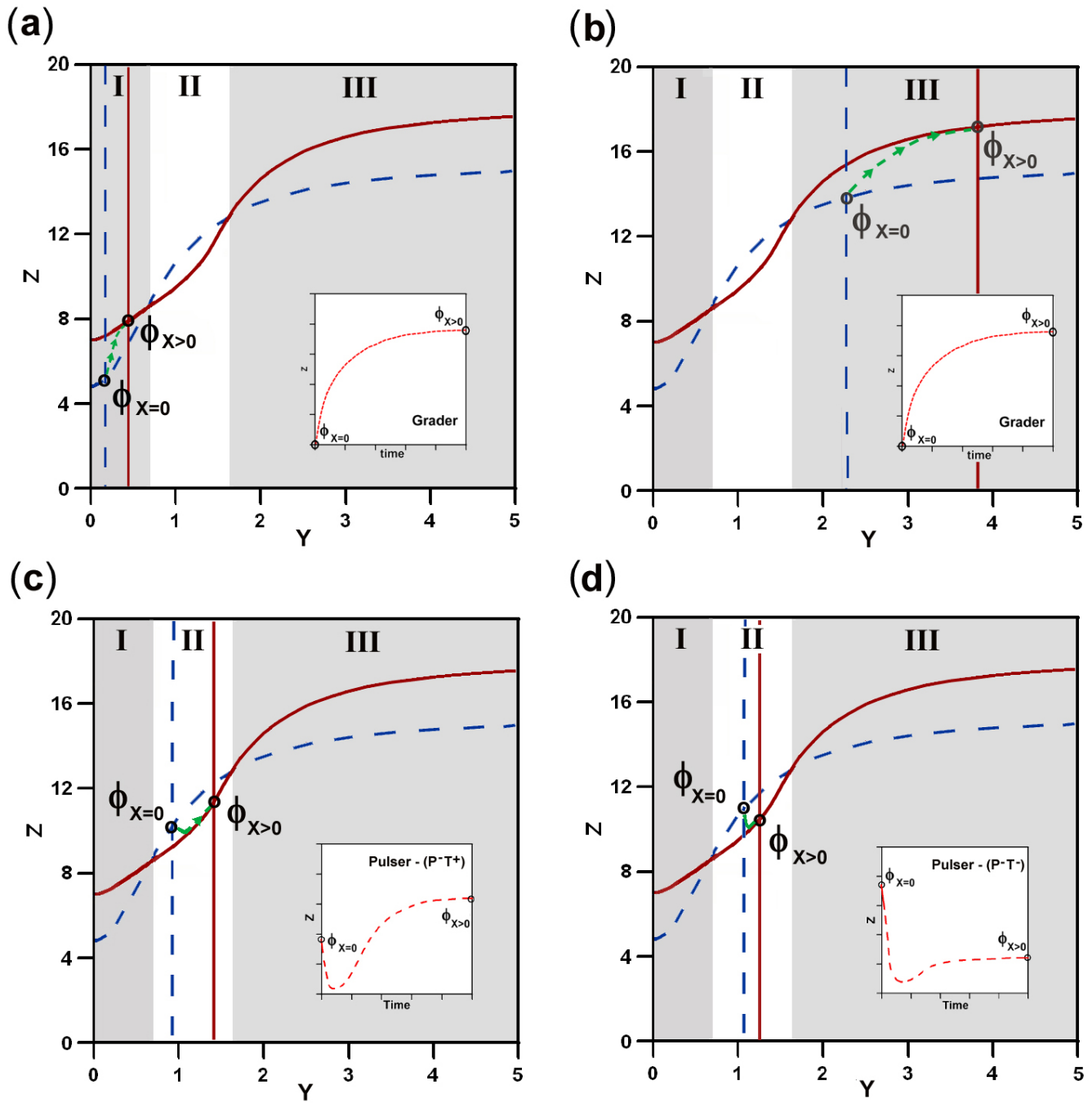
**Kurtosis as a measure of FFL specialization**

Our results indicated that, despite the topology of FFL does not determine univocally its dynamics, it is responsible for the probability of generating a given dynamics. Some topologies exhibit higher degrees of specialization, i.e. certain dynamics are more likely than others. On the other hand, there are topologies where the probabilities of emergence of the different dynamics are more similar. In order to quantify the degree of specialization we propose to measure the kurtosis of the probability distributions for each topology. Theoretically, the maximum degree of spe-

cialization would correspond to a motif able to implement only a single dynamic, i.e. its distribution of probabilities would be single peaked, and hence display maximum kurtosis. The maximum flexibility would correspond to a system where all possible dynamics have the same probability, i.e. a flat distribution of probabilities and hence minimal kurtosis. Kurtosis is defined as the fourth standardized moment [43],

$$K = \frac{\mu_4}{\sigma^4} - K_0 \tag{16}$$

where  $\mu_4$  if the fourth moment about the mean, and  $\sigma$  is the standard deviation. Here  $K_0$  is a reference value known as excess kurtosis. In general  $K_0 = 3$  in order to make the



**Figure 10**  
**Possible dynamics of CI in a geometrical scenario with two crossing between nullclines.** The phase space can be divided in three different regions I, II and III delimited by the crossings between nullclines. Figures (a) and (b) have the same qualitative relations between parameters, i.e..  $\varphi_{X=0} < \varphi_{X>0}$ , and hence a single dynamics  $G^+$  can emerge. Figures (c) and (d) show the two possible scenarios  $\varphi_{X=0} < \varphi_{X>0}$  and  $\varphi_{X=0} > \varphi_{X>0}$ , providing  $P(T^+)$  and  $P(T^-)$  dynamics respectively. Note that in (a) and (b) the location of  $\varphi_{X=0}$  determines the dynamics independently of the location of  $\varphi_{X>0}$ . However, scenario (c) constrains the locations of  $\varphi_{X>0}$  to the region II. Finally, scenario (d) can be found with  $\varphi_{X>0}$  located in region II or III. The insets represent the corresponding time courses.

**Table 3: Backbone sequences for C4**

$\beta^{XY}, I$	$\beta^{XY}, \beta^Y$	$Z_{0 X=0}, Z_{HA X>0}$	$Z_{0 X>0}, Z_{HA X=0}$	Slope	$\varphi_{X=0}, \varphi_{X>0}$	$\varphi_{X=0}, \xi$	Relation	Function
>	*	*	*	+	*	*	$\xi > S > \varphi_{X>0}$	$G^+$
>	*	<	*	-	*	*	$\xi > S > \varphi_{X>0}$	$G^+$
>	*	<	*	-	*	*	$\xi > \varphi_{X>0} > S$	$G^+$
>	*	<	*	-	*	*	$S > \xi > \varphi_{X>0}$	$G^+$
$\emptyset$	>	>	*	*	*	*	$\xi > S > \varphi_{X>0}$	$G^+$
$\emptyset$	>	>	*	*	*	*	$\xi > \varphi_{X>0} > S$	$G^+$
$\emptyset$	>	>	*	*	*	*	$S > \xi > \varphi_{X>0}$	$G^+$
*	<	*	*	*	*	<	$\xi > S > \varphi_{X>0}$	$G^+$
*	<	*	*	*	*	<	$\xi > \varphi_{X>0} > S$	$G^+$
*	<	*	*	*	*	<	$S > \xi > \varphi_{X>0}$	$G^+$
*	<	*	*	*	>	*	$\varphi_{X>0} > S > \xi$	$G^-$
>	*	*	*	+	*	*	$\xi > \varphi_{X>0} > S$	$P^+T^+$
>	*	*	*	+	*	*	$S > \xi > \varphi_{X>0}$	$P^+T^+$
*	<	*	*	*	>	*	$S > \varphi_{X>0} > \xi$	$P^-T^-$
*	<	*	*	*	>	*	$\varphi_{X>0} > \xi > S$	$P^-T^-$
*	<	*	*	*	<	>	$\varphi_{X>0} > S > \xi$	Impossible
*	<	*	*	$\emptyset$	<	>	$\varphi_{X>0} > \xi > S$	$P^-T^+$
*	<	*	*	$\emptyset$	<	>	$S > \varphi_{X>0} > \xi$	$P^-T^+$

All possible backbone sequences for C4 and its relation to the circuit's function.



**Table 4: Backbone sequences for II**

$\beta^{XY}, I$	$\beta^{XY}, \beta^Y$	$Z_{0 X=0}, Z_{HA X>0}$	$Z_{0 X>0}, Z_{HA X=0}$	Slope	$\varphi_{X=0}, \varphi_{X>0}$	$\varphi_{X=0}, \xi$	Relation	Function
>	*	*	*	+	*	*	$\xi > S > \varphi_{X>0}$	$G^+$
>	*	*	*	+	*	*	$\xi > \varphi_{X>0} > S$	$G^+$
>	*	*	*	+	*	*	$S > \xi > \varphi_{X>0}$	$G^+$
>	*	<	*	-	*	*	$\xi > S > \varphi_{X>0}$	$G^+$
$\emptyset$	>	>	*	*	<	*	$\xi > S > \varphi_{X>0}$	$G^+$
*	<	*	*	*	<	*	$\xi > S > \varphi_{X>0}$	$G^+$
*	<	*	*	*	*	>	$S > \varphi_{X>0} > \xi$	$G^-$
*	<	*	*	*	*	>	$\varphi_{X>0} > \xi > S$	$G^-$
*	<	*	*	*	*	>	$\varphi_{X>0} > S > \xi$	$G^-$
>	*	<	*	-	*	*	$\xi > \varphi_{X>0} > S$	$P^+T^+$
>	*	<	*	-	*	*	$S > \xi > \varphi_{X>0}$	$P^+T^+$
$\emptyset$	>	>	*	*	<	*	$\xi > \varphi_{X>0} > S$	$P^+T^+$
$\emptyset$	>	>	*	*	<	*	$S > \xi > \varphi_{X>0}$	$P^+T^+$
*	<	*	*	*	<	*	$\xi > \varphi_{X>0} > S$	$P^+T^+$
*	<	*	*	*	<	*	$S > \xi > \varphi_{X>0}$	$P^+T^+$
$\emptyset$	>	>	*	*	>	*	$\xi > S > \varphi_{X>0}$	Impossible
$\emptyset$	>	>	*	*	>	*	$\xi > \varphi_{X>0} > S$	$P^+T^-$
$\emptyset$	>	>	*	*	>	*	$S > \xi > \varphi_{X>0}$	$P^+T^-$
*	<	*	*	*	>	<	$\xi > S > \varphi_{X>0}$	Impossible
*	<	*	*	*	>	<	$\xi > \varphi_{X>0} > S$	$P^+T^-$
*	<	*	*	*	>	<	$S > \xi > \varphi_{X>0}$	$P^+T^-$

All possible backbone sequences for II and its relation to the circuit's function.

kurtosis of the normal distribution equal to zero. This allows obtaining positive kurtosis, i.e. distributions with higher *peakdeness* than the normal distribution, or negative kurtosis flatter than the normal distribution. Without loss of generality we can consider  $K_0 = 0$ . In this context, kurtosis is defined positive and the kurtosis values can be directly compared: systems with higher kurtosis will have higher degree of specialization

### Authors' contributions

JM: study conception, research design, mathematical analysis, manuscript writing. SW: study conception, research design, mathematical analysis, manuscript writing. RS: study conception, manuscript writing.

### Acknowledgements

We thank the members of the Complex Systems Lab for fruitful discussion. This work was supported by the EU grant CELLCOMPUT, the EU 6th Framework project SYNLET (NEST 043312), the James McDonnell Foundation and by the Santa Fe Institute.

### References

- Barabási AL, Oltvai ZN: **Network biology: understanding the cell's functional organization.** *Nat Rev Genet* 2004, **5**:101-13.
- Endy D, Brent R: **Modelling cellular behaviour.** *Nature* 2001, **409**:391-5.
- Huang S: **Back to the biology in systems biology: what can we learn from biomolecular networks?** *Brief Funct Genomic Proteomic* 2004, **2**:279-97.
- Bornholdt S, Schuster HG, Eds: *Handbook of Graphs and Networks - From the Genome to the Internet* Berlin: Wiley-VCH; 2002.
- Koonin EV, Wolf YI, Karev GP, Eds: *Power Laws, Scale-Free Networks and Genome Biology (Molecular Biology Intelligence Unit)* Springer; 2006.
- Képes F, Ed: *Biological networks* Complex Systems and Interdisciplinary Science, World Scientific Publishing Company, Incorporated; 2008.
- Hartwell LH, Hopfield JJ, Leibler S, Murray AW: **From molecular to modular cell biology.** *Nature* 1999, **402**:C47-52.
- Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, Alon U: **Network motifs: simple building blocks of complex networks.** *Science* 2002, **298**:824-7.
- Conant GC, Wagner A: **Convergent evolution of gene circuits.** *Nat Genet* 2003, **34**:264-6.
- Wolf DM, Arkin AP: **Motifs, modules and games in bacteria.** *Curr Opin Microbiol* 2003, **6**:125-34.
- Kashtan N, Itzkovitz S, Milo R, Alon U: **Topological generalizations of network motifs.** *Phys Rev E Stat Nonlin Soft Matter Phys* 2004, **70**:031909.
- Solé RV, Valverde S: **Are network motifs the spandrels of cellular complexity?** *Trends Ecol Evol* 2006, **21**:419-22.
- Cordero OX, Hogeweg P: **Feed-forward loop circuits as a side effect of genome evolution.** *Mol Biol Evol* 2006, **23**:1931-6.
- Kim D, Kwon YK, Cho KH: **Coupled positive and negative feedback circuits form an essential building block of cellular signaling pathways.** *Bioessays* 2007, **29**:85-90.
- Kim D, Kwon YK, Cho KH: **The biphasic behavior of incoherent feed-forward loops in biomolecular regulatory networks.** *Bioessays* 2008, **30**:1204-11.
- Ishihara S, Fujimoto K, Shibata T: **Cross talking of network motifs in gene regulation that generates temporal pulses and spatial stripes.** *Genes Cells* 2005, **10**:1025-38.
- Goldbeter A, Dupont G, Berridge MJ: **Minimal model for signal-induced  $Ca^{2+}$  oscillations and for their frequency encoding through protein phosphorylation.** *Proc Natl Acad Sci USA* 1990, **87**:1461-5.
- Ferrell JE, Machleder EM: **The biochemical basis of an all-or-none cell fate switch in *Xenopus* oocytes.** *Science* 1998, **280**:895-898.
- Cherry JL, Adler FR: **How to make a biological switch.** *J Theor Biol* 2000, **203**:117-133.
- Smolen P, Baxter DA, Byrne JH: **A reduced model clarifies the role of feedback loops and time delays in the *Drosophila* circadian oscillator.** *Biophys J* 2002, **83**:2349-59.
- Pomerening JR, Sontag ED, Ferrell JE: **Building a cell cycle oscillator: hysteresis and bistability in the activation of *Cdc2*.** *Nat Cell Biol* 2003, **5**:346-351.
- Chickarmane V, Troein C, Nuber UA, Sauro HM, Peterson C: **Transcriptional dynamics of the embryonic stem cell switch.** *PLOS Comput Biol* 2006, **2**:1080-1092.
- Macía J, Regot S, Peeters T, Conde N, Posas F, Solé R: **Dynamic signaling in the *Hog1* MAPK pathway relies on high basal signal transduction.** *Science Signal* 2009, **2**(63):ra13.
- Yu RC, Pesce CG, Colman-Lerner A, Lok L, Pincus D, Serra E, Holl M, Benjamin K, Gordon A, Brent R: **Negative feedback that improves information transmission in yeast signalling.** *Nature* 2008, **456**:755-61.
- Wernicke S, Rasche F: **FANMOD: a tool for fast network motif detection.** *Bioinformatics* 2006, **22**(9):1152-1153.
- Shen-Orr SS, Milo R, Mangan S, Alon U: **Network motifs in the transcriptional regulation network of *Escherichia coli*.** *Nat Genet* 2002, **31**:64-68.
- Lee TI, Rinaldi NJ, Robert F, Odom DT, Bar-Joseph Z, Gerber GK, N MH, Harbison CT, Thompson CM, Simon I, Zeitlinger J, Jennings EG, Murray HL, Gordon DB, Ren B, Wyrick JJ, Tagne J, Volkert TL, Fraenkel E, Gifford DK, Young RA: **Transcriptional Regulatory Networks in *Saccharomyces cerevisiae*.** *Science* 2002, **298**:799-804.
- Mangan S, Alon U: **Structure and function of the feed-forward loop network motif.** *Proc Natl Acad Sci USA* 2003, **100**:11980-5.
- Mangan S, Zaslaver A, Alon U: **The Coherent Feedforward Loop Serves as a Sign-sensitive Delay Element in Transcription Networks.** *J Mol Biol* 2003, **334**:197-204.
- Ghosh B, Karmakar R, Bose I: **Noise characteristics of feed forward loops.** *Phys Biol* 2005, **2**:36-45.
- Klemm K, Bornholdt S: **Topology of biological networks and reliability of information processing.** *Proc Natl Acad Sci USA* 2005, **102**:1841-9.
- Hayot F, Jayaprakash C: **A feedforward loop motif in transcriptional regulation: induction and repression.** *J Theor Biol* 2005, **234**:133-43.
- Mangan S, Itzkovitz S, Zaslaver A, Alon U: **The Incoherent Feed-forward Loop Accelerates the Response-time of the gal System of *Escherichia coli*.** *J Mol Biol* 2006, **356**:1073-81.
- Kremling A, Bettenbrock K, Gilles ED: **A feed-forward loop guarantees robust behavior in *Escherichia coli* carbohydrate uptake.** *Bioinformatics* 2008, **24**:704-10.
- Entus R, Aufderheide B, Sauro HM: **Design and implementation of three incoherent feed-forward motif based biological concentration sensors.** *Syst Synth Biol* 2007, **1**:119-28.
- Ingram PJ, Stumpf MP, Stark J: **Network motifs: structure does not determine function.** *BMC Genomics* 2006, **7**:108.
- Prill RJ, Iglesias PA, Levchenko A: **Dynamic properties of network motifs contribute to biological network organization.** *PLoS Biol* 2005, **3**:e343.
- Knabe JF, Nehaniv CL, Schilstra MJ: **Do motifs reflect evolved function? No convergent evolution of genetic regulatory network subgraph topologies.** *Biosystems* 2008, **94**:68-74.
- Voigt CA, Wolf DM, Arkin AP: **The *Bacillus subtilis* *sin* operon: an evolvable network motif.** *Genetics* 2005, **169**:1187-202.
- Conrad E, Mayo AE, Ninfa AJ, Forger DB: **Rate constants rather than biochemical mechanism determine behaviour of genetic clocks.** *J Roy Soc Interface* 2008, **6**(5 Suppl 1):S9-15.
- Abramowitz M, Stegun IA, Eds: *Handbook of mathematical functions with formulas, graphs and mathematical tables.* NBS Applied Mathematics Series 55 Washington, DC: National Bureau of Standards; 1964.
- Silva-Rocha R, de Lorenzo V: **Mining logic gates in prokaryotic transcriptional regulation networks.** *FEBS Lett* 2008, **582**:1237-44.
- Joanes D, Gill C: **Comparing measures of sample skewness and kurtosis.** *J Roy Stat Soc (Series D): The Statistician* 1998, **47**(1):183-189.
- Macía J, Solé RV: **Distributed robustness in cellular networks: insights from synthetic evolved circuits.** *J R Soc Interface* 2008, **442**:259-264.