Supplementary information: The combination of the functionalities of feedback circuits is determinant for the number and size of attractors of molecular networks

Eugenio Azpeitia[1,*], Stalin Muñoz[2,+], Daniel González-Tokman[3,+], Mariana Esther Martínez-Sánchez[4,5,+], Nathan Weinstein[6,+], Aurélien Naldi[7], Elena R. Álvarez-Buylla[5,8], David A. Rosenblueth[2], and Luis Mendoza[9]

[1]INRIA project-team Virtual Plants, joint with CIRAD and INRA, Montpellier Cedex 5, France.
[2]Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, Universidad Nacional Autónoma de México, Apdo. 20-126, 01000 México, D.F., México.
[3]Investigador Cdras CONACYT, Instituto de Ecología, A. C., Antiguo camino a Coatepec 351, El Haya, 91070 Xalapa, Veracruz, México.
[4]Programa de Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México, México.
[5]Instituto de Ecología, Universidad Nacional Autónoma de México, México.
[6]ABACUs-Centro de Matemáticas Aplicadas y Cómputo de Alto Rendimiento, Departamento de Matemáticas, Centro de Investigación y de Estudios Avanzados CINVESTAV-IPN, Carretera México-Toluca Km 38.5, La Marquesa, Ocoyoacac, Estado de México, 52740 México.
[7]DIMNP UMR CNRS 5235, Université de Montpellier. France.
[8]Centro de Ciencias de la Complejidad, Universidad Nacional Autónoma de México, México.
[9]Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México, México.
[+]These authors contributed equally to this work.
[*]Corresponding author: eugenio.azpeitia-espinosa@inria.fr

Supplementary Methods

Boolean networks basic definitions

Let $\mathbb{B} = \{0,1\}$ and $\mathbb{N}_{\leq n} = \{1,2,\ldots,n\}$, an initial segment of natural numbers. A *synchronous Boolean network* is a triple (V,σ,f), where $V = \{v_1, v_2, \ldots, v_n\}$ is a set of variables ranging over $\mathbb{B}$, V represents the set of modeled molecules, σ an ordering of the variables and f defines the *dynamics* of the network.

A *state* of the network is a tuple $x = (x_1, x_2, \ldots, x_n)$ such that $x \in \mathbb{B}^n$. The dependency of the state on the discrete time parameter t is denoted as $x(t)$ and obeys the evolution rule given by f. That is for all $t \in \mathbb{Z}$

$$x(t+1) = f(x(t)). \tag{1}$$

The state x represents the activation state of every variable in the network , while the activation value of a particular variable is denoted by the indexing function σ, if $x_{\sigma(v_i)} = 1$ the gene represented by v_i is active, whilst if $x_{\sigma(v_i)} = 0$ the gene represented by v_i is inactive.

For synchronous Boolean networks, an *attractor* is a set of states $A \subseteq \mathbb{B}^n$ where for all $x(0) \in A$ there exist $l > 0$ such that $l = \min_k x(k) = x(0)$ and for all $s \in \mathbb{N}_{\leq l}$, $x(s) \in A$; l is the *size* of the attractor.

For simplicity, we refer to the variable v_j by its position in the vector x, that is variable i means $i = \sigma(v_j)$. For a state $x \in \mathbb{B}^n$ and a variable i we denote as $x \sim i$ the vector resulting from replacing the value of the variable i in x with its opposite value.

Given two variables j and i and the update function of variable i, namely f_i, j *activates* i if there exists a pair of network states x, y that differ only in the state of activation of variable j, that is $y = x \sim j$, $x_j = 0$ and $y_j = 1$, such that $f_i(y) - f_i(x) > 0$. Conversely, j *inhibits* i if there exists a pair of network states x, y that differ only in the state of activation of variable j, that is $y = x \sim j$, $x_j = 0$ and $y_j = 1$, such that $f_i(y) - f_i(x) < 0$. An *interaction*, denoted as the pair (i,j), $i,j \in \mathbb{N}_{\leq n}$ is *functional* if variable j activates or inhibits node i. Note that according to this definition, it is possible for variable j to both activate

and inhibit variable i. A non-functional interaction does not provide useful information about the biological system and it is an accepted convention that interactions, where a molecule activates and inhibits the same gene are scarce in molecular regulations[1,2]. Thus, we excluded both non-functional regulations and regulations where variable j both activate and inhibit variable i.

Circuits functionality analyses

A feedback circuit is defined as a set of directed interactions forming a closed path. Feedback circuits can be positive or negative. The sign of a circuit is given by the signs of its interactions. A circuit is positive if it has an even number of negative interactions, it is negative otherwise. It is important to note that the sole presence of a circuit in a network does not guarantee the appearance of the corresponding dynamical behavior (i.e., oscillations or multistability). Thus, a circuit is considered functional if at least one combination of the states of external regulators of its members allows all interactions of the circuit to be functional together[3].

In more formal terms, the *functionality context of the interaction*(i,j) of a Boolean network $F=(V,\sigma,f)$ is the set $\Phi(f,i,j)$ defined by:

$$\Phi(f,i,j)=\{x \mid f_j(x)\neq f_j(x\sim i) \text{ and } x\in\mathbb{B}^n\} \tag{2}$$

For a Boolean network $F=(V,\sigma,f)$ we say that $\mathscr{G}_F$ is its *structure* or *interaction graph* $\mathscr{G}_F=\langle V,\mathscr{I}_f^+,\mathscr{I}_f^-\rangle$, where: $\mathscr{I}_f^+$ is its set of *positive interactions* defined by

$$\mathscr{I}_f^+=\{(\sigma^{-1}(i),\sigma^{-1}(j)) \mid \forall x\in\Phi(f,i,j)\ x_i=f_j(x)\} \tag{3}$$

and $\mathscr{I}_f^-$ is its set of *negative interactions* defined by

$$\mathscr{I}_f^-=\{(\sigma^{-1}(i),\sigma^{-1}(j)) \mid \forall x\in\Phi(f,i,j)\ x_i\neq f_j(x)\} \tag{4}$$

For a circuit $C=(c_1,c_2,\ldots,c_k)$ (simple cycle with no shortcuts) of an interaction graph $\mathscr{G}_F$, where $c_i\in\mathbb{N}_{\leq n}$, $c_i\neq c_j$ if $i\neq j$, we define the *functionality context of the circuit* C, denoted $\Phi(f,C)$ as follows:

$$\Phi(f,C)=\bigcap_{i=1}^{k}\Phi(f,c_i,c_{(i \bmod k)+1})\,. \tag{5}$$

The circuit C is functional if $\Phi(f,C)$ is not empty.

The *functionality context for a circuit C with shortcuts* $S=\{(c_i,c_j) \mid |(j \bmod k)-i|\neq 1\}$ is defined by:

$$\Phi(f,C,S)=\Phi(f,C)-\bigcup_{(i,j)\in S}\{x \mid x\in\Phi(f,C)\wedge x\sim i\notin\Phi(f,C)\}\,. \tag{6}$$

As with the previous case the circuit C is functional if $\Phi(f,C,S)$ is not empty.

A *functional circuit description* is given by the triple (C,n,ς) where C is a circuit, $n=|\Phi(f,C,S)|$ the cardinality of its functionality context, an its sign $\varsigma\in\mathbb{S}$, where $\mathbb{S}=\{+,-\}$ is the set of signs.

The *combination of the functionality of (feedback) circuits* of a Boolean network F is defined as the set of functional circuit descriptions for all circuits in its interaction graph $\mathscr{G}_F$.

Networks structural and dynamical distances

We define the *adjacency matrix* of a graph $G=\langle V,E\rangle$ with $V=\{v_1,\ldots,v_n\}$ and $E\in V\times V$ as $\mathscr{A}(G)=(a_{ij})$, $(a_{ij})\in\mathbb{B}^{n\times n}$ with entries satisfying

$$a_{ij}=\begin{cases}1 & \text{if } (v_i,v_j)\in E\\ 0 & \text{otherwise}\end{cases} \tag{7}$$

Accordingly, the *structural distance* $D_{\text{str}}(f,g)$ between two Boolean networks $F=(V,\sigma,f)$ y $G=(V,\sigma,g)$ is defined by

$$D_{\text{str}}(f,g)=\left\|\mathscr{A}(\langle V\mathscr{I}_f^+\cup\mathscr{I}_f^-\rangle)-\mathscr{A}(\langle V,\mathscr{I}_g^+\cup\mathscr{I}_g^-\rangle)\right\|_1 \tag{8}$$

where $\|\cdot\|_p$ is the matrix entrywise norm defined by

$$\left\|(a_{ij})\right\|_p = \left(\sum_i \sum_j |a_{ij}|^p\right)^{1/p} \tag{9}$$

The *signed structural distance* $D^{\varsigma}_{\text{str}}(f,g)$ between two Boolean networks $F=(V,\sigma,f)$ y $G=(V,\sigma,g)$ is defined by

$$D^{\varsigma}_{\text{str}}(f,g) = \sum_{\varsigma_1\in\mathbb{S}} \sum_{\varsigma_2\in\mathbb{S}} \#_{\varsigma_1}\cdot\#_{\varsigma_2}\cdot \left\| \mathscr{A}(\langle V,\mathscr{I}_f^{\varsigma_1}\rangle) - \mathscr{A}(\langle V,\mathscr{I}_g^{\varsigma_2}\rangle)\right\|_1 \tag{10}$$

where $\#:\mathbb{S}\to\{-1,1\}$, is defined by

$$\#_{\varsigma} = \begin{cases} -1 & \text{if } \varsigma = - \\ 1 & \text{if } \varsigma = + \end{cases} \tag{11}$$

The *dynamical distance* between two Boolean networks $F=(V,\sigma,f)$ y $G=(V,\sigma,g)$ is defined by

$$D_{\text{dyn}}(f,g) = \sum_{x\in\mathbb{B}^n} \sum_{i\in\mathbb{N}_{\leq n}} |f_i(x)-g_i(x)| \tag{12}$$

We compared each pair of networks A and B of size N using the three distances described above, implementing the necessary algorithms in Python. The dynamical distance clustering analysis was done using the scipy/linkage function (ward)[4] for the dynamical distance. We considered that only distances below a certain threshold where valid edges (0 and 8 for the structural and dynamical distance, respectively). In the resulting networks each node represents a PLN and the edges' weight corresponds to the dynamical or structural distance. We analyzed the network properties using. python/networkx[5].

PLNs simulations and queries

For the first two sections of the Results, PLNs were simulated using R/BoolNet 2.1.1[6]. There were 9×10^3 biologically meaningful 1-PLNs, all of which were analyzed. As for biologically meaningful 2-PLNs, due to their astronomical number, we used samples of sizes between 10^4 and 32.8×10^6 of them to analyze their properties.

For the third and fourth Results sections, the search of PLNs with the epistasis expected set of attractors (see section 3.3) was done with Griffin[7]. Griffin is a software tool that uses symbolic algorithms for the inference of Boolean networks. Griffin transforms the set of constraints into a Boolean sentence, which in turn using a Tseitin transformation[8,9] is converted into an equisatisfiable conjunctive normal form sentence. This sentence is then provided as an input to a SAT4j[10] solver instance. When the solver finds an assignment of the Boolean variables that make the sentence true, this assignment is returned to Griffin. Griffin then decodes the assignment into a set of Boolean functions corresponding to the network dynamics.

Certain biological constraints were added to Griffin to formulate the epistasis queries. First, we use a set of *generalized interactions* which are a set of gene interaction constraints that corresponded to the MUS,OUS and MP interactions. The expected fixed point attractors of the 2-PLNs required partially defined fixed points and the double mutant experiment required multiple genes mutations with partially defined states. All of these biological constraints were transformed by Griffin into the Boolean sentence representing the query. Finally, we prohibited networks that exhibit cyclic trajectories in the state space. As it is computationally intractable to add this constraint a priori, Griffin performs a posteriori refinement of cycles using Dubrova and Teslenko's SAT based algorithm[11]. Any satisfying assignment will be decoded to a biologically meaningful Boolean network. The extended version of Griffin containing all the biological constraints is available under request.

Statistical Analyses

All statistical analyses were carried out in R version 3.2.3[12].

To test the relationship between number of attractors and attractor average size we carried out a non-parametric Spearman rank correlation, given that the assumptions of parametric correlation were violated (Fligner-Killeen test for homogeneity of variances $X^2=526.784$, $d.f.=9$, $P=2.2x10^{16}$).

The differences between the circuits and structural properties in the singles and 2-PLNs were analyzed with generalized linear models (GLM) with Poisson error structure and log link function. In GLMs with overdispersion (overdispersion test $P<0.05$;[13]) we used models with quasipoisson errors[14]. For analyzing attractor sizes and ratio of positive/negative circuits we used generalized least squares (GLS) to account for heterogeneous variances found by type (Fligner-Killeen test $P<0.001$)[15,16].

Finally, in order to test the frequency distribution of networks ($N=6.3X10^7$ networks) of the attractors size and number we used Kolmogorov-Smirnov tests for log-normal, exponential, normal and Poisson distributions implemented in R package nortest. To test for power law distributions we used a bootstrapping procedure with 30 simulations in R package poweRlaw[17].

Statistical analyses results

Attractors properties

The attractors size was significantly larger for 2-PLNs than for 1-PLNs (GLS $F1,18998 = 217.63$,$P < 0.001$; $C.I.95\%$: 2-PLNs= 1.62 ± 0.012; 1-PLNs = 1.49 ± 0.012).

The number of attractors was significantly larger for 2-PLNs than for 1-PLNs (Poisson GLM $z = 103.2$, $d.f. = 1,18998$, $P < 0.001$; $C.I.95\%$: 2PLN = 6.22 ± 0.049; 1-PLNs = 2.91 ± 0.035).

Feedback circuits and PLNs structure

In 1-PLNs, the number of negative feedback circuits was significantly larger for the complete set of PLNs than for n+s– (Poisson GLM $z = 31.92$, $P < 0.001$) and for n–s– ($z = 3.27$, $P = 0.001$) and lower than for n–s+ ($z = 19.70$, $P < 0.001$; $d.f. = 3,12596$). The number of positive feedback circuits was significantly larger for the complete set of PLNs than for n–s+ (Poisson GLM $z = 23.36$, $P < 0.001$) and lower than for n+s– ($z = 28.43$, $P < 0.001$) and n–s– ($z = 3.22$, $P = 0.001$; $d.f. = 3,13795$). The total number of feedback circuits did not differ significantly between the complete set of PLNs and n+s– (GLM $z = 0.13$, $P = 0.897$), n–s+ ($z = 1.74$, $P = 0.083$) or n–s– ($z = 0.02$, $P = 0.988$; $d.f. = 3,13795$). The positive/negative feedback circuits ratio was significantly larger for the complete set of PLNs than for n–s+ (GLS $t = 19.56$, $P < 0.001$), significantly lower than n+s– ($t = 24.83$, $P < 0.001$) and did not differ significantly from n–s– ($t = 0.64$, $P = 0.520$; $d.f. = 3,8215$).

In 2-PLNs, the number of negative feedback circuits was significantly larger for the complete set of PLNs than for n+s– (Quasipoisson GLM $t = 38.31$, $P < 0.001$) and for n–s– ($t = 2.62$, $P = 0.009$) and lower than for n+s+ ($t = 5.17$,$P < 0.001$) and n–s+ ($t = 27.92$, $P < 0.001$; $d.f. = 4,39480$). The number of positive feedback circuits was significantly larger for the complete set of PLNs than for n+s– (Quasipoisson GLM $t = 27.66$, $P < 0.001$) and for n–s– ($t = 2.27$, $P = 0.023$) and lower than for n+s+ ($t = 14.79$, $P < 0.001$) and n–s+ ($t = 15.78$, $P < 0.001$; $d.f. = 4,39480$). The total number of feedback circuits was significantly larger for the complete set of PLNs than for n+s– (Quasipoisson GLM $t = 39.65$, $P < 0.001$) and for n–s– ($t = 2.81$, $P = 0.005$) and lower than for n+s+ ($t = 11.37$, $P < 0.001$) and n–s+ ($t = 25.13$, $P < 0.001$; $d.f. = 4,39480$). The positive/negative feedback circuits ratio was significantly larger for the complete set of PLNs than for n–s+ ($GLSt = 9.29$, $P < 0.001$), lower than for n+s+ ($t = 3.47$, $P < 0.001$) and n+s– ($t = 19.83$, $P < 0.001$) and did not differ from n–s– ($t = 1.22$,$P = 0.223$; $d.f. = 4,35536$).

The number of combinations of the functionalities of feedback circuits was significantly larger for 2-PLN than for n–s– 2-PLN (Quasipoisson GLM $t = 53.96$, $d.f. = 30768$, $P < 0.001$; $C.I.95\%$: 2PLN = 14.24 ± 0.199; n–s– 2-PLNs = 2.36 ± 0.146).

The number of combinations of the functionalities of feedback circuits was not significantly different for 1-PLNs and 1-n–s– PLNs (Quasipoisson GLM $t = 1.423$, $d.f. = 19$, $P = 0.172$).

The combinations of the functionalities of feedback circuitrs were contained in a significantly different number of PLNs structures by Type (Quasipoisson GLM $Res.Dev. = 280865$, $d.f. = 254468$, $P < 0.001$). A priori contrasts show that 2-PLNs have lower values (1.46 ± 0.010) than 2-PLNs with the n–s– functionalities (2.60 ± 0.060) ($t = 46.896$,$P < 0.001$) and similar values to n–s– 2-PLNs (1.43 ± 0.044) ($t = 0.973$, $P = 0.331$); finally 2-PLNs with the n–s– functionalities has higher values than n–s– 2-PLNs ($t = 29.95$, $P < 0.001$).

Relations

In 1-PLNs, the relationship between the number of attractors and the attractors average size was significant and negative (Poisson GLM: $Res.Dev. = 1139.8$, $d.f. = 8998$, $z = -22.59$, $P < 0.001$). In 1-PLNs, the relationship between the number of attractors and the number of negative loops was significant and negative (Poisson GLM: $Res.Dev. = 5683.8$, $d.f. = 8998$, $z = -30.45$, $P < 0.001$). The relationship between the number of attractors and the number of positive loops was significant and positive (Poisson GLM: $Res.Dev. = 5633.5$, $d.f. = 8998$, $z = 34.29$, $P < 0.001$). The relationship between the attractors average size and the number of negative loops was significant and positive (Poisson GLM: $Res.Dev. = 5428.5$, $d.f. = 8998$, $z = 38.15$, P ¡ 0.001). The relationship between the attractors average size and the number of positive loops was significant and negative (Poisson GLM: $Res.Dev. = 4975.5$, $d.f. = 8998$, $z = -37.45$, $P < 0.001$).

In 2-PLNs to test the relationship between the number of attractors and the attractors average size we took a random sample of 10,000 networks and carried out a generalized linear model (GLM) with Poisson errors and log link function. We found a significant, negative relation between the number of attractors and the attractors average size (Poisson GLM: $Res.Dev. = 1968.9$, $d.f. = 9998$, $z = -14.00$,$P < 0.001$). The relationship between the number of attractors and the number of negative loops was significant and negative (Poisson GLM: $Res.Dev. = 146.46$, $d.f. = 9998$, $z = -11.79$, $P < 0.001$). The relationship between the number of attractors and the number of positive loops was significant and positive (Poisson GLM: $Res.Dev. = 9305.5$, $d.f. = 9998$, $z = 39.60$, $P < 0.001$). The relationship between the attractors average size and the number of negative loops was significant and positive (Poisson GLM: $Res.Dev. = 9607.10$, $d.f. = 9998$, $z = 39.45$, $P < 0.001$). The relationship between the attractors average size and the number of positive loops was significant and negative (Poisson GLM: $Res.Dev. = 10621.00$, $d.f. = 9998$, $z = -11.53$, $P < 0.001$).

Distributions

For the number of attractors, the frequency distribution differed significantly from a normal distribution ($D = 0.205$, $P < 0.001$) or from a Poisson distribution ($D = 0.256$, $P < 0.001$). For the attractors size, the frequency distribution differed significantly from a normal distribution ($D = 0.265$, $P < 0.001$), from a Poisson distribution ($D = 0.420$, $P < 0.001$) or from a power law distribution ($KS = 0.013$, $Xmin = 6$, $Scaling = 10.278$, $P < 0.001$). For the number of attractors, the frequency distribution differed significantly from a log-normal distribution ($D = 0.853$, $P < 0.001$) and from an exponential distribution ($D = 0.474$, $P < 0.001$). The attractors size frequency distribution also differed significantly from a log-normal ($D = 0.307$, $P < 0.001$) and an exponential distribution ($D = 0.401$, $P < 0.001$).

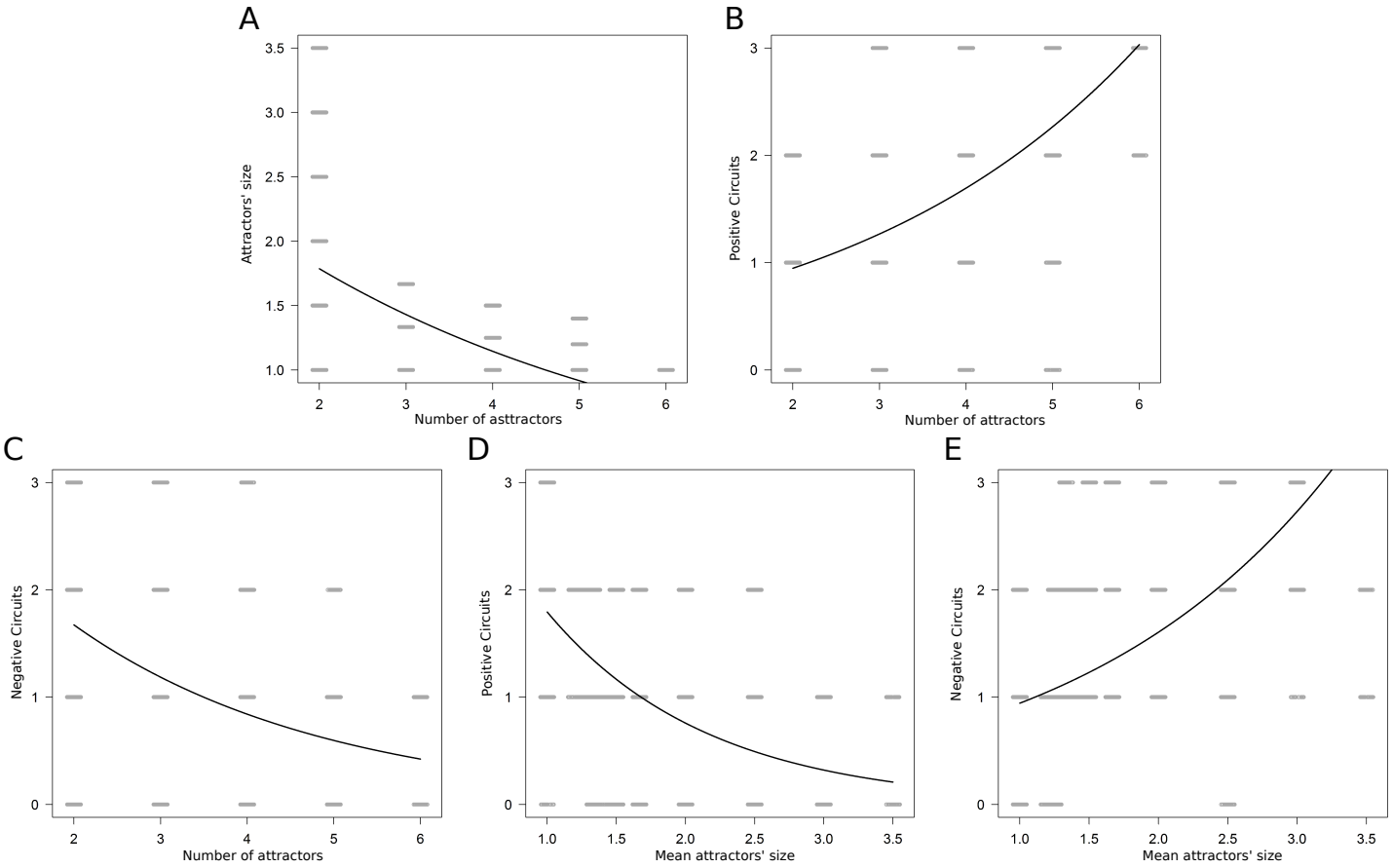
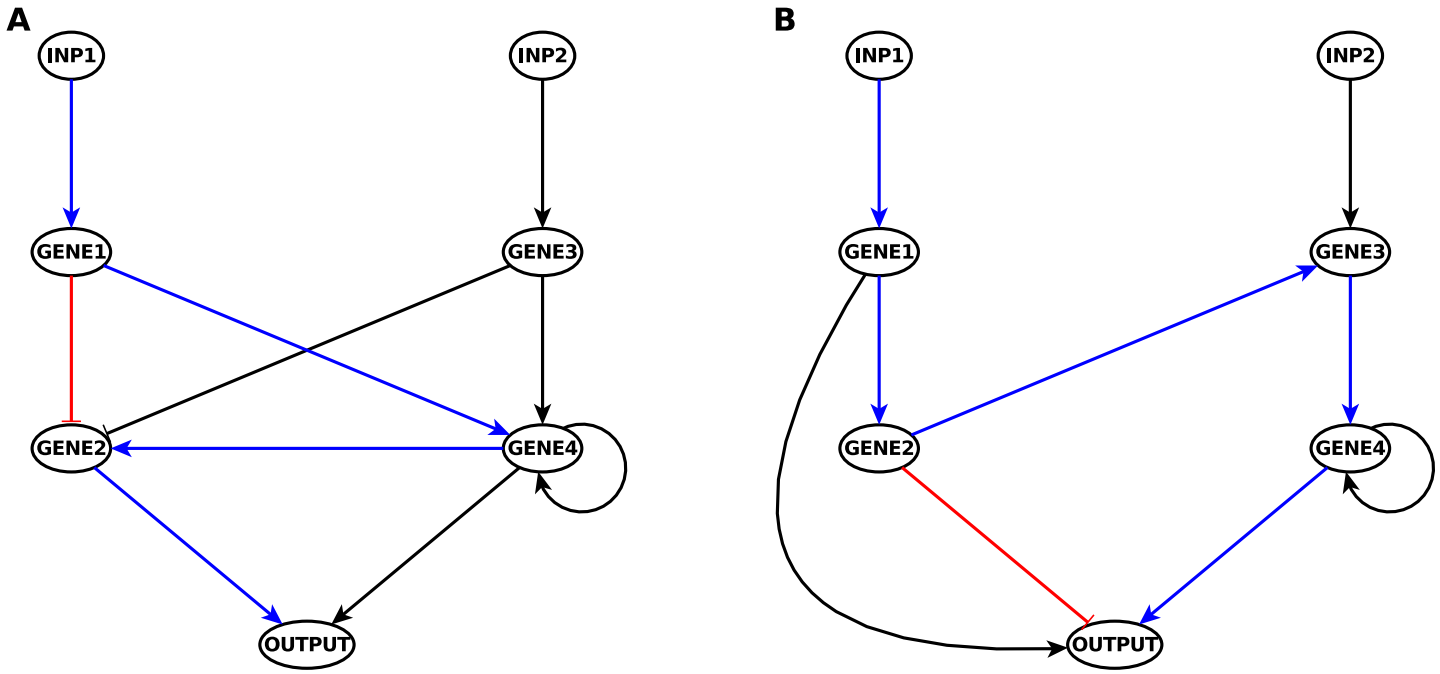


Figure 1. 1-PLNs properties.(A) Number of attractor vs. attractors mean size. (B) and (C) number of attractors vs. quantity of negative and positive feedback circuits, respectively. (D) and (E) size of attractors vs. quantity of negative and positive feedback circuits, respectively. As observed, negative and positive feedback circuits have opposite effects, just as in 2-PLNs. Each point represents a single 2-PLN data. Points are displaced in the X axis only for visual purpose. The lines are predicted by Poisson GLMs.




```
INP1 = true
GEN1 = INP1
GEN2 = GEN4 | ! GEN1 & ! GEN3
INP2 = true
GEN3 = INP2
GEN4 = GEN4 & GEN3 | GEN1 & GEN4 | GEN1 & GEN3
OUTPUT = GEN2 & GEN4
```

```
INP1 = true
GEN1 = INP1
GEN2 = GEN1
INP2 = true
GEN3 = GEN2 & INP2
GEN4 = GEN4 | GEN3
OUTPUT = GEN4 & ( GEN1 | ! GEN2)
```

Figure 2. Two examples of wrongly inferred interactions using epistasis analysis. Examples of the cases where the interactions from GEN1 to GEN2 (A) and the interaction from GEN2 to OUTPUT (B) are wrongly inferred. In both cases the expected pathway variant is ++. The orange edge is the incorrect inferred interaction and the blue edges are an alternative pathway that contains the expected signs of interactions between INP1, GEN1, GEN2 and OUTPUT with some extra intermediary interactions. Below the Boolean functions for each of these PLNs.

References

1. Raeymaekers, L. Quantitative epistasis analysis and pathway inference from genetic interaction data. *Journal of Theoretical Biology* **218**, 331–341 (2002).
2. Azpeitia, E., Benítez, M., Padilla-Longoria, P., Espinosa-Soto, C. & Alvarez-Buylla, E. R. Dynamic network-based epistasis analysis: boolean examples. *Frontiers in Plant Sciences* **15**, 2:92 (2010).
3. Naldi, A., Thieffry, D. & Chaouiya, C. *Decision Diagrams for the Representation and Analysis of Logical Models of Genetic Networks*, vol. 4695 of *Computational Methods in Systems Biology. Lecture Notes in Computer Science*, 233–247 (Springer Berlin Heidelberg, 2007).
4. Müllner, D. fastcluster: Fast hierarchical, agglomerative clustering routines for r and python. *Journal of Statistical Software* **53**, 1–18 (2013).
5. Hagberg, A., Schult, D. & Swart, P. Networkx reference (2012).
6. Müssel, C., Hopfensitz, M. & Kestler, H. A. Boolnet – an R package for generation, reconstruction and analysis of Boolean networks. *Bioinformatics* **26**, 1378–1380 (2010).
7. Rosenblueth, D. A., Muñoz, S., Carrillo, M. & Azpeitia, E. Inference of Boolean Networks from Gene Interaction Graphs Using a SAT Solver. In *Algorithms for Computational Biology*, 235–246 (Springer International Publishing, 2014).
8. Prestwich, S. D. CNF encodings. *Handbook of Satisfiability* **185**, 75–97 (2009).
9. Tseitin, G. S. On the complexity of derivation in propositional calculus. In *Automation of reasoning*, 466–483 (Springer, 1983).
10. Le Berre, D. & Parrain, A. The sat4j library, release 2.2, system description. *Journal on Satisfiability, Boolean Modeling and Computation* **7**, 59–64 (2010).
11. Dubrova, E. & Teslenko, M. A SAT-based algorithm for finding attractors in synchronous Boolean networks. *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)* **8**, 1393–1399 (2011).
12. R Development Core Team. *R: a Language and Environment for Statistical Computing* (Austria, 2015).
13. Kleiber, C. & Zeileis, A. *Applied Econometrics with R* (New York, U.S.A, 2008).
14. Crawley, M. J. *The R Book* (London, U.K., 2007), 1st edn.
15. Venables, W. N. & Ripley, B. D. *Modern Applied Statistics with S* (New York, U.S.A, 2002), fourth edn.
16. Zuur, A. F., Ieno, E. N., Walker, A. A., N. J.and Saveliev & Smith, G. M. *Mixed Effects Models and Extensions in Ecology with R* (New York, U.S.A, 2009).
17. Gillespie, C. S. Fitting heavy tailed distributions: The poweRlaw package. *Journal of Statistical Software* **64**, 1–16 (2015). URL http://www.jstatsoft.org/v64/i02/.