Random Boolean Networks and Attractors of their Intersecting Circuits
Jacques Demongeot, Adrien Elena, Mathilde Noual, Sylvain Sené

To cite this version:

HAL Id: hal-00666589
https://hal.science/hal-00666589
Submitted on 11 Feb 2014

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.
Random Boolean Networks and Attractors of their Intersecting Circuits

J. Demongeot1*, A. Elena1, M. Noual2,4, S. Sene3,4

Abstract — The multi-scale strategy in studying biological regulatory networks analysis is based on two level of analysis. The first level is structural and consists in examining the architecture of the interaction graph underlying the network and the second level is functional and analyse the regulatory properties of the network. We apply this dual approach to the “immunetworks” involved in the control of the immune system. As a result, we show that the small number of attractors of these networks is due to the presence of intersecting circuits in their interaction graphs. We obtain an upper bound of the number of attractors of the whole network by multiplying the number of attractors of each of its strongly connected components. We detect first the strongly connected components in the architecture of the interaction digraph of the network. Secondly, we study the dynamical function of the attractors by looking further inside these components, notably when they form circuits (intersecting or not).

Keywords: genetic networks, immune networks, attractors counting, tangential circuits, intersected circuits

I. INTRODUCTION

The theory of biological regulatory networks is born parallelly to that of neural networks [1-3] and is now a research topic of intense activity for interpreting the "omic" data from the bio-arrays devices. The networks are made of elements (genes, proteins, neurons,...) in interaction and the central tool used for representing these interactions is a directed graph (digraph), called interaction graph, whose signed arrows (positive or negative) are related to the influence (activation or inhibition) exerted by an element (repressor or inducer) on another. Researches have focused on modelling real genetic networks of size \( n \), with about \( 2n \) interactions networks, for which many authors have noticed that the number of attractors observed experimentally (i.e., the number of their different possible temporal asymptotic behaviours such as fixed or cyclic configurations of expressed and/or silent genes) is relatively small, of the order of magnitude \( \sqrt{n} \). We will see in this paper that we can relate this small number of empirical attractors to the architecture of the network. The attractors number is crucial in applications because it is related to the number of differentiation fate possibilities of the cells controlled by the network [1].

In an arbitrary network, besides nodes belonging to circuits, all other nodes belong to trees that are either down-trees (i.e., that have no influence on the rest of the network), or up-trees (i.e., that undergo no influence from the rest of the network and only eventually act punctually on it by allowing or disallowing it to have its own dynamics), or trees that connect nontrivially strongly connected components and only serve as information pathways. Now the theory of the random networks [4, 5] enables us to predict exactly the number of isolated and intersecting circuits, as well as the number of up- and down-trees. Therefore, in order to estimate the number of attractors in an arbitrary network, one may focus on the dynamics that is induced by the underlying circuits of the network. In a Section 2, we will introduce basic concepts about random Boolean graphs and networks. After we will focus on the crucial problem of sampling these networks in order to count their attractors. In Section 4, we give new results about intersecting circuits and we apply these results to the “immunetworks”, i.e., the genetic regulatory networks used for creating and controlling the immune genome [6].

II. RANDOM GRAPHS AND RANDOM BOOLEAN NETWORKS

Let us focus now on how one may build the underlying digraph representing the architecture of a random network. Assuming we want to keep the connectivity \( c \) (mean number of oriented interactions coming to a node) of the digraph constant, it is possible to choose the interaction arcs according to various probability distributions that respect this connectivity [7, 8]. For instance, focusing on \( d \)-regular digraphs (the number of non-oriented interactions per node being constant equal to \( d \)), in order to obtain a digraph of connectivity \( c = d/3 \) [9], one may choose uniformly at random a \( d \)-regular digraph in the set of all \( d \)-regular digraphs. Using an \( d \)-regular non-oriented graph having \( m=nd/2 \) edges, one may also construct a digraph of connectivity \( c \) if \( d \) is chosen equal to \( 3c/2 \) [9]. Let \( X_n \) denote the random variable representing the number of \( d \)-regular cycles of length \( s \) in a random \( d \)-regular graph of size \( n \). In [10-12], it is shown that \( X_n \), for \( 3 \leq s \leq 2^d \), is asymptotically distributed as independent Poisson variable with mean equal to \( (d-1)^{2s-1}/2s!=\frac{3c-2}{2s} \) if \( c=2 \), with \( d=d(n) \) and \( g=g(n) \) allowed to increase with \( n \), provided that: \( (d-1)^{2s-1}=o(n) \). Hence we have: \( g=[Log_2(o(n))]+1/2 \), if \( c=2 \), which holds for \( g \leq 5 \), if \( n=22.10^3 \) and \( o(n)=n^{5/3} \). When \( d=3 \), the graph yields on average \( 2s-1/2 \) cycles of length \( s \). To construct a signed digraph from a \( d \)-regular non-oriented graph, each edge of the graph is replaced by
one or two signed arcs (loops are counted twice). Hence, each cycle of the $d$-regular graph may yield at most two circuits in the digraph.

III. NETWORK SAMPLING FROM MINIMAL REPRESENTATION

In order to avoid the over-representation of certain graphs in the random sampling of Boolean Hopfield-like threshold networks, a representation using minimal vectors made of the parameters (weights and thresholds) has been proposed [13-18]. This sampling methodology allows a non biased representation of the attractors in a given category of networks. For example, Table 1 gives the number of attractors for networks of size $n=3,\ldots,6$ for the different categories: $C_y$ (resp. $F_i$; $M_i$) networks having only limit cycles (resp. only points remaining fixed; both fixed points and cycles) as attractors for any updating modes going from the synchronous to the sequential ones, and $E_v$ networks whose nature of attractors (cycles or fixed points) depends on the updating mode and therefore passing through the previous categories when this mode evolves.

<table>
<thead>
<tr>
<th>$n$</th>
<th>Pts Fixed</th>
<th>$C_y$</th>
<th>$F_i$</th>
<th>$M_i$</th>
<th>$E_v$</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Cycles</td>
<td>1,15</td>
<td>0</td>
<td>1,19</td>
<td>1,36</td>
<td>1,00</td>
</tr>
<tr>
<td></td>
<td>Attract.</td>
<td>1,15</td>
<td>1,93</td>
<td>2,26</td>
<td>1,69</td>
<td>1,63</td>
</tr>
<tr>
<td>4</td>
<td>Pts Fixed</td>
<td>0</td>
<td>2,30</td>
<td>1,35</td>
<td>1,56</td>
<td>1,00</td>
</tr>
<tr>
<td></td>
<td>Cycles</td>
<td>1,25</td>
<td>0</td>
<td>1,17</td>
<td>0,37</td>
<td>0,80</td>
</tr>
<tr>
<td></td>
<td>Attract.</td>
<td>1,25</td>
<td>2,20</td>
<td>2,52</td>
<td>1,93</td>
<td>1,70</td>
</tr>
<tr>
<td>5</td>
<td>Pts Fixed</td>
<td>0</td>
<td>2,80</td>
<td>1,49</td>
<td>1,73</td>
<td>1,00</td>
</tr>
<tr>
<td></td>
<td>Cycles</td>
<td>1,36</td>
<td>0</td>
<td>1,30</td>
<td>0,45</td>
<td>0,94</td>
</tr>
<tr>
<td></td>
<td>Attract.</td>
<td>1,36</td>
<td>2,80</td>
<td>2,79</td>
<td>2,18</td>
<td>1,94</td>
</tr>
<tr>
<td>6</td>
<td>Pts Fixed</td>
<td>0</td>
<td>3,49</td>
<td>1,63</td>
<td>1,86</td>
<td>1,00</td>
</tr>
<tr>
<td></td>
<td>Cycles</td>
<td>1,47</td>
<td>0</td>
<td>1,45</td>
<td>0,57</td>
<td>1,09</td>
</tr>
<tr>
<td></td>
<td>Attract.</td>
<td>1,47</td>
<td>3,49</td>
<td>3,08</td>
<td>2,43</td>
<td>2,00</td>
</tr>
</tbody>
</table>

Figure 1: number of attractors for networks of size 3 to 6 depending on their attractor class

Let us notice that the order of magnitude of $\sqrt{n}$ for the attractor number is respected, despite of the fact that the networks have been sampled in the whole set of networks of size $n$, without respecting the connectivity $2$.

IV. ATTRACTORS OF INTERSECTING CIRCUITS

4.1. Tangential circuits

The attractor number of isolated circuit has been extensively studied in [4] and applied in [9].

4.2. Intersecting circuits

The attractor number of intersecting circuit has been extensively studied in [5] and applied in [9].

By looking on the networks of the Figure 1, we see that each is made of 4 main paths of opposite sense: two are up, A and C, and two down, B and D, of respective lengths $\ell_A$, $\ell_B$, $\ell_C$, $\ell_D$, and parity $s_A$, $s_B$, $s_C$, $s_D$, equal to 1 (resp. -1), if they have an even (resp. odd) number of negative arcs: $s_a = \Pi_{a \in \partial A} s_a$, where the sign $s_a$ of the arc $a$ of $A$ is equal to -1 if $a$ is negative (inhibition) and 1 if $a$ is positive (activation). For example, in Figure 1, the path A of $N_1$ is such as $\ell_A=3$ and $s_A=-1$. The main paths have in $N_1$ two common nodes, 10 and 13, and in $N_2$ only one common node 23. Finally, $N_1$ having only 4 paths having two common nodes, there are 6 combinations two by two of the 4 possible circuits (A,B), (B,C), (C,D) and (A,D), each circuit circuit like (A,C) having the parity $s_{a,C} = s_A s_C$. As a conjecture, we can say that two intersecting circuits both managing to cycle together may not induce more attractors than would any of the two circuits if they were isolated. In addition, the period of any attractor induced by the intersection is no greater than the longest period of an attractor induced by a circuit of size that of the largest circuit encircling the two intersecting circuits (if it were isolated). The Propositions 4.1. to 4.3. below make more precise the context of the Conjecture, which could become: the number of attractors of $N_1$ is less than the number of attractors of $N_2$, whose two circuits (tangential in 23) have respectively as signs $sup(s_A,s_B)s_C$ and $sup(s_A,s_B)s_D$, with $x_{300}(t) = x_{20}(t)$, and the proof is then given in the Proposition 2.

Let us now consider the Hamiltonian $H$ of a circuit $(A,C)$ as defined by:

$$H(A,C)(t) = \sum_{t(A,C)}(x_A(t) - x_A(t-1))^2 = \sum_{t(A,C)}(x_A(t) - x_A(t-1))^2 = \sum_{t(A,C)}(x_A(t) - x_A(t-1))^2$$

The Hamiltonian $H(A,C)$ is equal to twice the sum of the kinetic energies over all nodes of the circuit $(A,C)$, and is equal to the frustration of $(A,C)$, i.e., to the number of couples of successive nodes whose states are such as the second one is not equal to the predicted by the first one. The conservation of kinetic energy is for example ensured in continuous Hamiltonian systems like the simple pendulum, for which the dynamical differential equations are:

$$dx/dt = y \text{ and } dy/dt = -x$$
and its Hamiltonian equals $(x^2+y^2)/2$, i.e., the kinetic energy. For example, for a positive circuit of size $n=8$, we have the following even values of frustration (they are odd for a negative circuit) and the number of frustration configurations (Table 2), with the corresponding number and length of attractors (fixed state configurations or limit cycles of state configurations):

<table>
<thead>
<tr>
<th>$H$ (frustration)</th>
<th>Configuration Nb</th>
<th>Attractor Nb</th>
<th>Attractor Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2 fixed points</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>$4 = C_{n-H+2/2}^{H-1}$</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>$10 = C_{n-H+2/2}^{H-1}$</td>
<td>3</td>
<td>4 (symmetrized)</td>
</tr>
<tr>
<td>6</td>
<td>$4 = C_{n-H+2/2}^{H-1}$</td>
<td>16</td>
<td>8</td>
</tr>
<tr>
<td>8</td>
<td>$1$</td>
<td>2 fixed points</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 2: Values of Hamiltonian energy, attractor number and length ($n=8$)

For a positive circuit of size $n=6$, we have the following even values of frustration and the number of frustration configurations (Table 3), with the corresponding number and length of attractors (fixed state configurations or limit cycles of state configurations):

<table>
<thead>
<tr>
<th>$H$ (frustration)</th>
<th>Configuration Nb</th>
<th>Attractor Nb</th>
<th>Attractor Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2 fixed points</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>$3 = C_{n-H+2/2}^{H-1}$</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>$3 = C_{n-H+2/2}^{H-1}$</td>
<td>2</td>
<td>3 (symmetrized)</td>
</tr>
<tr>
<td>6</td>
<td>$1$</td>
<td>1 fixed points</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 3: Values of Hamiltonian energy, attractor number and length ($n=6$)

For a positive circuit of size $n=4$, we have the following even values of frustration and the number of frustration configurations (Table 4), with the corresponding number and length of attractors (fixed state configurations or limit cycles of state configurations):

<table>
<thead>
<tr>
<th>$H$ (frustration)</th>
<th>Configuration Nb</th>
<th>Attractor Nb</th>
<th>Attractor Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2 fixed points</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>$2 = C_{n-H+2/2}^{H-1}$</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>$1$</td>
<td>2 fixed points</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 4: Values of Hamiltonian energy, attractor number and length ($n=4$)

We will try now to solve in the Propositions 4.1. to 4.3. the two questions: Are the attractors of the network $N_2$, with the coupling $x_{200}(t) = x_{20}(t)$, identical to the attractors of $N_1$? What are the constraints about the period of these attractors?

**Proposition 4.1.** The attractors of the network $N_2$, with the coupling $x_{200}(t) = x_{20}(t)$, are attractors of $N_1$.

**Proof.** From an initial condition identical for $N_1$ and $N_2$, where $x_{200}(0) = x_{20}(0) = x_{10}(0)$, the trajectories are the same for all nodes, if $x_{200}(t) = x_{20}(t)$, for any $t \geq 1$. For the node 10 of the network $N_1$, we have, if the mixing rule is monotonic, e.g., $v$ in $N_1$ and $N_2$ (the reasoning would be the same, if the rule is $\land$ or any composition of $v$ and $\land$):

\[ x_{10}(t) = [s_y \times x_{12}(t-\ell_b)] \lor [s_D \times x_{13}(t-\ell_D)] \]

\[ x_{20}(t) = s_y \times x_{23}(t-\ell_b) \text{ and } x_{20}(t) = s_y \times x_{23}(t-\ell_D) \]

By imposing $x_{200}(t) = x_{20}(t)$, for any $t \geq 1$, then $s_y \times x_{23}(t-\ell_b) = s_D \times x_{23}(t-\ell_D)$, and we have in the network $N_1$:

\[ x_{10}(t) = [s_y \times x_{13}(t-\ell_b)] \lor [s_D \times x_{13}(t-\ell_D)] = x_{20}(t) = x_{20}(t) \]

By recurrence on $t$, this common value for $x_{10}(t)$, $x_{200}(t)$ and $x_{20}(t)$ is equal to:

\[ s_y \times x_{13}(0) = s_y \times x_{23}(0), \text{ for any } t=k\ell_b, \text{ with } k \geq 0. \]

The same reasoning can be made for $t = 1, ..., \ell_b - 1 \pmod{\ell_b}$. Then, the trajectories being the same, the attractors of $N_2$, with the coupling $x_{200}(t) = x_{20}(t)$, are attractors of $N_1$.

Remark: let us note that, if $\ell_b = \ell_b$, that imposes, for observing the equality $x_{200}(t) = x_{20}(t)$, the constraint $s_D = s_y$; if not, the attractors of $N_1$ are not attractors of $N_2$, with $x_{200}(t) = x_{20}(t)$, and the converse of the Proposition 1. is then not available. But we have the following result:

**Proposition 4.2.** The attractors of $N_1$ are the attractors of $M_2$, whose two circuits (tangential in 23) have respectively as signs $\text{sup}(s_4, s_y, s_C)$ and $\text{sup}(s_4, s_y, s_D)$, with $x_{200}(t) = x_{20}(t)$.

**Proof.** Let us consider an attractor of $N_1$, for which $x_{10}(t) = [s_y \times x_{13}(t-\ell_b)] \lor [s_D \times x_{13}(t-\ell_D)]$. If we identify $x_{10}(t)$ and $x_{20}(t)$, then, if $x_{10}(0) = x_{20}(0)$, this attractor is an attractor of $M_2$, where the 2 circuits (tangential in 23) have respectively as signs $\text{sup}(s_4, s_y, s_C)$ and $\text{sup}(s_4, s_y, s_D)$ and $x_{200}(t) = x_{20}(t)$.

**Proposition 4.3.** The attractors of $N_2$, with the coupling $x_{200}(t) = x_{20}(t)$, are characterized by the property (P) on their length (or period) $p$:

- if $s_y \times s_D = 1$ (resp. $s_x \times s_C = 1$), we have:
  - $p$ divides $(\text{sup}(\ell_b, \ell_D) - \text{inf}(\ell_b, \ell_D))$ (resp. $p$ divides $(\text{sup}(\ell_a, \ell_D) - \text{inf}(\ell_a, \ell_D))$)\n  - $-p$ divides $(\text{sup}(\ell_b, \ell_D) - \text{inf}(\ell_b, \ell_D))$ (resp. $p$ divides $(\text{sup}(\ell_a, \ell_D) - \text{inf}(\ell_a, \ell_D))$)\n
- if $s_y \times s_D = -1$ (resp. $s_x \times s_C = -1$), $p$ does not divide $(\text{sup}(\ell_b, \ell_D) - \text{inf}(\ell_b, \ell_D))$ and $p$ divides $(\text{sup}(\ell_b, \ell_D) - \text{inf}(\ell_b, \ell_D))$\n
**Proof.** If $p$ denotes the length (or period) of an attractor of $N_2$, with the coupling $x_{200}(t) = x_{20}(t)$, then we have:

$\forall t \geq 1$, $x_{200}(t) = x_{20}(t)$ $\iff$ $\forall t \geq 1$, $s_y \times x_{23}(t-\ell_b) = s_D \times x_{23}(t-\ell_D)$

$\iff$ $\forall t \geq 1, x_{23}(t) = s_y \times s_D \times x_{23}(t+\ell_b - \ell_D)$, and

$\forall t \geq 1, x_{200}(t) = x_{20}(t)$ $\iff$ $\forall t \geq 1$, $s_y \times x_{23}(t-\ell_a) = s_C \times x_{23}(t-\ell_C)$

$\iff$ $\forall t \geq 1, x_{23}(t) = s_y \times s_C \times x_{23}(t+\ell_a - \ell_C)$

Hence, we have, if $s_y \times s_D = 1$, $x_{200}(t) = x_{20}(t)$ $\iff$ $\forall t \geq 1$, $x_{23}(t) = x_{23}(t+\text{sup}(\ell_b, \ell_D) - \text{inf}(\ell_b, \ell_D))$\n
$\iff$ $\forall t \geq 1, p \mid (\text{sup}(\ell_b, \ell_D) - \text{inf}(\ell_b, \ell_D))$.

The proof in the case $s_y \times s_D = -1$ is similar.
Figure 2: Immunetworks (from [19]) ruling RAG-1, a Recombination Activating protein G responsible in human and mouse for the regulation of the rearrangements V(D)J creating the genes of the lymphocyte T receptors TCRα [20-25]

Table 5: Top: Number of attractors for intersecting negative circuits of respective size \( r \) and \( l \). Number of attractors of length \( p \) for isolated circuits of size \( n \)
V. APPLICATION TO IMMUNETWORKS

The previous results can be applied to real biological genetic networks like the “immunetworks” controlling the building of the immune genome in mammals (made of the genes of the TCR-α and -β and of the chains CD3-β, -ε and -γ). The creation of the immune genes come from rearrangements of the chromosome 14 ruled by an enzyme, the RAG-1 (Recombination Activating protein G), controlled by a specific network having two negative circuits (of sign -1, i.e. having an odd number of inhibitions), one of length 6 (in blue on Figure 2) and the other of length 2 (in red on Figure 2). By calculating the number of attractors brought by these circuits intersecting on gene PU.1, we see that this number is reduced from 6 (blue on Table 5 bottom) if the circuits were isolated, or 16 (green on Table 5 bottom) if the circuits are added to form an unique circuit, to only 1 (green on Table 5 top). The prediction by the model is then in this case that the only attractor brought by the network controlling the RAG-1 expression has been possibly selected during the evolution for having only two behaviours: i) the RAG-1 is expressed and ii) the RAG-1 is inhibited by micro-RNAs, notably has-miR 30 in human.

VI. CONCLUSION

We given in this paper mathematical results necessary to understand the asymptotic (attractor) behaviour of genetic networks. Some behaviours induce difficult mathematical problems, e.g., because the expression of RAG-1 authorizes the updating or not of the up-tree controlling the genes TCR-α (in red in Figure 2), its implies a state-dependent updating (Figure 3) constituting a real mathematical challenge as perspective of the present work.

REFERENCES