THE SIMULATION OF GENE KNOCK-OUT IN SCALE-FREE RANDOM BOOLEAN MODELS OF GENETIC NETWORKS

ROBERTO SERRA

Dipartimento di scienze sociali, cognitive e quantitative
Università di Modena e Reggio Emilia
Via Allegri 9, 42100 Reggio Emilia, Italia

MARCO VILLANI

Dipartimento di scienze sociali, cognitive e quantitative
Università di Modena e Reggio Emilia
Via Allegri 9, 42100 Reggio Emilia, Italia

ALEX GRAUDENZI

Dipartimento di scienze sociali, cognitive e quantitative
Università di Modena e Reggio Emilia
Via Allegri 9, 42100 Reggio Emilia, Italia

ANNAMARIA COLACCI

Excellence Environmental Carcinogenesis
Environmental Protection and Health Prevention Agency
Emilia-Romagna, viale Filopanti 22, Bologna, Italia

STUART A. KAUFFMAN

Institute for Biocomplexity and Informatics
University of Calgary, 2500 University Dr. NW
Calgary, AB T2N 1N4, Canada

ABSTRACT. This paper describes the effects of perturbations, which simulate the knock-out of single genes, one at a time, in random Boolean models of genetic networks (RBN). The analysis concentrates on the probability distribution of so-called avalanches (defined in the text) in gene expression. The topology of the random Boolean networks considered here is of the scale-free type, with a power-law distribution of outgoing connectivities. The results for these scale-free random Boolean networks (SFRBN) are compared with those of classical RBNs, which had been previously analyzed, and with experimental data on *S. cerevisiae*. It is shown that, while both models approximate the main features of the distribution of experimental data, SFRBNs tend to overestimate the number of large avalanches.

1. **Introduction.** While most of the studies in molecular biology are devoted to the analysis of detailed molecular mechanisms or cellular subsystems, it should be remarked that, in order to make sense of the wealth of available data, also system-level models are needed, which are able to capture selected key features of the observed phenomena, while adopting simplifications which make them tractable.

In previous works ([22],[21]) it has been shown that random Boolean networks (briefly, RBNs) can be useful to model the perturbations in gene expression levels which occur in cells of the yeast *S. cerevisiae* which are subject to the knock-out (i.e. permanent silencing) of a single gene.

While those works were based on "classical" RBNs, which have a Poissonian outdegree distribution, in this paper we investigate the influence of the topology of the genetic network on the distribution of avalanches. Present data concerning the whole genetic network of *S. cerevisiae* do not allow to conclusively determine its topology ([24],[9],[19],[8],[18]), so we decided to concentrate the analysis on scale-free networks, which have raised considerable interest as they provide a good approximation to several real networks, including metabolic and protein networks ([3],[15],[25],[14],[5]), at least within a range (one often finds an "exponential cutoff" at high degree values).

We have therefore introduced a model of scale-free random Boolean networks, which differs from most of the existing models because it allows the presence of nodes without outgoing connections. This is motivated by the fact that in "classical" RBNs there are indeed such nodes, and that it is likely that the same applies to the real biological gene network. We will briefly refer to such scale-free networks as SFRBN, while the shorthand RBN will be used exclusively for the classical ones from now on.

In section 2 the main properties of "classical" random Boolean networks will be recalled and in section 3 the main results of previous studies concerning gene expression avalanches in classical RBNs will be summarized. The way in which SFRBNs are built, and the properties which characterize them, are described in section 4, while in Section 5 we report the results concerning avalanches in gene expression in the case of SFRBNs. A comparison with experimental data, for different values of the threshold used to distinguish between affected and unaffected nodes, is also provided. Critical comments and indications for further works are given in the final section 6.

2. **Random Boolean networks.** Let us consider a network composed of N genes, or nodes, which can take either the value 0 (inactive) or 1 (active). Let $x_i(t) \in [0, 1]$ be the activation value of node i at time t, and let $X(t) = [x_1(t), x_2(t)...x_N(t)]$ be the vector of activation values of all the genes. In a classical RBN each node has the same number of incoming connections k_{in}, and its k_{in} input nodes are chosen at random with uniform probability among the remaining *N-1* nodes (self-coupling and multiple connections from the same node being prohibited). It then turns out that the distribution of outgoing connections per node follows a Poisson distribution ([16]):

$$p_{out}(k) = e^{-A}\frac{A^k}{k!} \tag{1}$$

where, since every connections must have both ends, $A =< k_{out} >= k_{in}$.

The output (i.e. the new value of a node) corresponding to each set of values of the input nodes is determined by a Boolean function, which is associated to that

node, and which is also chosen at random, according to some probability distribution. The simplest choice is that of a uniform distribution among all the possible Boolean functions of k_{in} arguments. However, a careful analysis of some known real biological control circuits has shown that there is a strong bias in favour of the so-called canalyzing functions ([11]). A Boolean function is said to be canalyzing if at least one value of one of its input nodes uniquely determines its output, no matter what the other input values are (this notion has recently been generalized to that of nested Boolean functions ([17])).

Both the topology and the Boolean function associated to each gene do not change in time (using the so-called "quenched model" of RBN ([2])). The network dynamics is discrete and synchronous.

In order to analyze the properties of an ensemble of random Boolean networks, different networks are synthesized and their dynamical properties are examined. While individual realizations may differ markedly from the average properties of a given class of networks ([6],[23]), one of the major results is the discovery of the existence of two different dynamical regimes, an ordered and a disordered one, divided by a "critical zone" in parameter space. Attractors in the ordered regime are cycles whose length scales as a power of N, moreover in this regime the system is stable with respect to small perturbations of the initial conditions (e.g. if one modifies the value of one of the $x_k(0)$, the system usually relaxes to the same attractor). In the disordered regime the length of the cycles grows exponentially with N, and small changes in initial conditions often lead to different attractors. For fixed N, the most relevant parameter which determines the kind of regime is the connectivity per node, k: one typically observes ordered behaviour for small k, and a disordered one for larger k. The exact value of k at which the transition takes place depends upon the way in which the Boolean functions are chosen. For a more detailed discussion, the reader is referred to [16],[2],[6],[23].

3. **Avalanches in gene expression data.** We will briefly summarize here the main result of our previous studies. The experimental setup is described in [13], while the theoretical analyses are discussed in depth in [22] and [21]: the reader interested in a deeper understanding of these topics is referred to the original works.

[13] performed several experiments where a single gene of *S. cerevisiae* has been knocked-out, and compared the expression levels of all the genes, in cells with a knocked-out gene, with those in normal, wild type cells. In order to make precise statements about the number of genes perturbed in a given experiment, and to compare them with Boolean models, it is required that a threshold be defined, such that the difference is regarded as "meaningful" if the ratio of the expression of gene i in experiment j to the expression of gene i in the wild type cell is greater than the threshold. In order to describe the global features of these experiments, we introduced the notion of avalanche, which is the number of genes affected by the perturbation induced by a particular knock-out experiment.

The knock-out experiment can be simulated in silico by comparing the evolution of two RBNs which start from identical initial conditions, except for the fact that one gene is clamped permanently to the value 0 in the one which simulates knock-out. We consider a gene belonging to that avalanche if it differs in the final states of the two networks at least once in the attractor cycle. The initial simulations were performed using a classical RBN with 2 input connections per node, restricting the set of Boolean functions to the so-called canalyzing ones ([16]). The data set

concerns 6325 genes and 227 experiments (for details see [22]). The comparison with the experimental distribution of avalanches turns out to be good, if the threshold is ≥ 4, being optimal for $\Theta = 7$ (according the well-known χ^2 index ([21])).

The reason why such a simple model worked so well has been uncovered by analytical methods which have proven that the distribution of avalanches depends only upon the outdegree distribution, while the indegree distribution plays no role. Moreover, in the case of classical random Boolean networks, where the distribution of outgoing connections is Poissonian, it can be also proven that the distribution of small avalanches depends only upon a single parameter, the so-called Derrida exponent ([7],[21]) which is given by the equation:

$$\lambda \equiv (1 - q)A \tag{2}$$

where A is the average connectivity of the network and q is the probability that a chosen node does not change its value whether one (and only one) of its inputs has changed (note that q depends on the choice of the set of Boolean functions). λ had been introduced in the past in order to distinguish between ordered and disordered dynamical regimes (being 1 the critical value), and it turns out that it also rules the distribution of avalanches. It is interesting to observe that the best agreement with experimental data (given $\Theta = 7$) is provided by the case where $\lambda = 6/7$, slightly smaller than the critical value 1.

4. **Scale-free random Boolean networks.** The above results hold for classical RBNs, with a constant indegree and Poissonian outdegree distribution, but other topologies of outgoing connections may however provide a closer approximation to the real one. Among them, scale-free networks, with a power-law distribution of node degrees, have attracted particular interest, as they have been found (within suitable approximations) in several different biological as well as artificial and social systems. The dynamical properties of these networks differ from the classical ones in many respects ([1],[10],[20]), so it would be very interesting to explore the distribution of avalanches in networks of this kind.

The well-known formula for a scale-free distribution of outgoing links is:

$$p_{out}(k) = \frac{1}{Z}k^{-\gamma} \tag{3}$$

$$Z(\gamma) = \sum_{k=1}^{k_{max}} k^{-\gamma} \tag{4}$$

where k can take values from 1 to a maximum possible value k_{max} (in our case, where self-coupling and multiple connections are prohibited, $k_{max} = N - 1$. Z (which coincides with Riemann zeta function in the limit $k_{max} \to \infty$) guarantees the proper normalization. It is well known that such scale-free networks also show two regimes, an ordered and a disordered one, separated by a curve which determines the critical slope of the exponent γ ([1]). The average value of k is:

$$< k > \equiv \sum_{k=1}^{k_{max}} kp(k) \cong \sum_{k=1}^{\infty} kp(k) = \frac{1}{Z(\gamma)} \sum_{k=1}^{\infty} k^{-\gamma+1} = \frac{Z(\gamma - 1)}{Z(\gamma)} \tag{5}$$

The condition for a critical network with $\rho = 0.5$ is that the last term on the r.h.s. of Eq. 5 be equal to 2 ([1]). Thus, the average value of the connectivity for critical SFRBNs coincides with that for classical RBNs; consequently, we concentrate our

analysis on this kind of networks. As we did with RBNs, we exclude the so-called non canalyzing functions ([11]); in such a way the network should be slightly subcritical, exactly as in the RBN case. We also performed some experiments using all the Boolean functions, whose results are briefly summarized in section 5.

Since, as long as the avalanches are much smaller than the total number of genes, the distribution of avalanches does not depend on the indegree distribution ([21]), the SFRBNs were generated keeping the number of ingoing connections equal to two for each node.

Note that the requirement of Eq. 5, together with the fact that we are considering a finite network of 6325 nodes and the requirement that the sum of all the probabilities equals 1, would uniquely determine γ, which turns out to be 2.469, if there were no node without outgoing links.

However, it should be stressed that the synthetic RBNs discussed in the previous sections indeed have some nodes without outgoing links (a feature which might hold also for real genetic networks). So, in order to analyze the effects of changes of the form of the distribution of outdegrees, it is necessary to extend Eq. 3 to the case $k = 0$. Of course, a direct extension would lead to a meaningless divergence. The simplest generalization of Eq. 3 capable to include the value to $k = 0$ is then:

$$\begin{cases} p_{out}(k) = \frac{1}{Z'} k^{-\gamma} \ if \ k \neq 0 \\ p_{out}(0) = p_0 \end{cases} \tag{6}$$

We therefore need to introduce another parameter $p_{out}(0)$. In the following we will refer also to the distribution given by Eq. 6 as a "scale-free" distribution. The normalization coefficient is now different from Z in Eq. 4. Indeed the normalization condition which defines Z' is:

$$Z' = \frac{\sum_{k=1}^{k_{max}} k^{-\gamma}}{1 - p_0} \tag{7}$$

Note that as $p_{out}(0)$ grows the slope γ decreases. Hence, we have simulated different networks, with different $p_{out}(0)$ ranging from 0.1 to 0.9. The corresponding γ's, for a network with 6325 nodes[1], are listed below in table 1.

$p_{out}(0)$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
γ	2.469	2.39	2.317	2.247	2.179	2.11	2.037	1.956	1.857	1.711

TABLE 1. Values of γ for different values of $p_{out}(0)$ in SFRBNs.

The SFRBNs were generated as follows. To every node of the network a specific number of outgoing connections is assigned, which is determined according to the distribution of outgoing links given by Eq. 6. The destination of the connections is randomly chosen among all the other N-1 nodes, with uniform probability, but the maximum number of ingoing connections per node is fixed to two, self and multiple connections being forbidden. When a given node has received its two input links, it cannot receive any more links. The final total number of outgoing links must be equal to the number of incoming links which, given the way in which the network is built, is 2N. If, due to rounding, one gets a slightly different total number of

---

[1]Here the criticality condition of [1] doesn't holds, because of the presence of the parameter $p_{out}(0)$; in order to allow an easier comparison with the other results of this work, during all the simulations we maintain $\rho = 0.5$.

connections, some links are adjusted in order to respect the constraint on the total number. No preferential attachment procedure is used.

In the simulations of knock-out experiments the knocked-out gene was chosen at random, with uniform probability, among all the genes which were not always in the 0 ("inactive") state in their own attractor cycle.

5. **Avalanches in SFRBNs: comparison with avalanches distribution in classical RBNs and with experimental data.** As it might be expected, the value of $p_{out}(0)$ affects the avalanche distribution $p(n)$ (the probability that an avalanche size is n), in such a way that the number of the smallest avalanches increases as $p_{out}(0)$ increases: this is due to the increasing number of nodes without outgoing connections, which prevent the avalanche to grow. As it could also be expected, large avalanches can be the result of knocking-out a hub (fig. 1) or a poorly connected gene which however affects a hub located downstream (fig. 2).
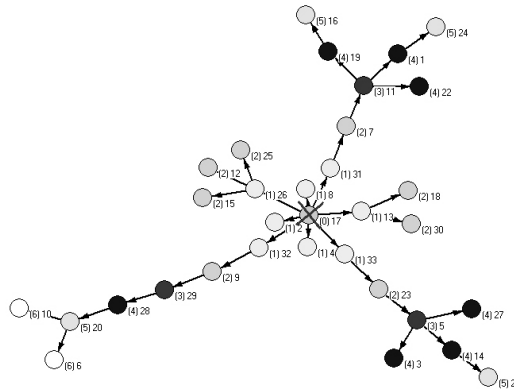

FIGURE 1. Avalanche dimension = 33, depth level = 6. The root node is a hub (marked with a cross) of the network, radial diffusion of the perturbation. The number close to the nodes is its identifier; the number among brackets its depth level.

In order to compare the avalanches in SFRBNs with those in classical RBNs, we have computed the average fraction of nodes which have no outgoing links in the RBN case (i.e. 0.1353) and have set $p_{out}(0)$ in SFRBNs equal to that value (for this kind of networks γ=2.36). Note that, for both RBNs and SFRBNs, the set of Boolean functions is restricted to canalyzing Boolean functions only.

This value of $p_{out}(0)$ might be considered as small on biological grounds, but two remarks are in order.

First, this value has been chosen to ensure that the same fraction of nodes without outgoing links is used in comparing classical RBNs and SFRBNs (and in the former case this number is an outcome of the usual way of generating the network).

Second, in order to interpret this value from a biological viewpoint, it is necessary to observe that the number of functional dependencies which can be observed is different from the number of links in the graph. This is due to the fact that some Boolean functions do not actually depend upon the value of some or all of their inputs. Consider for example a node A which has a link to a node B (which has two
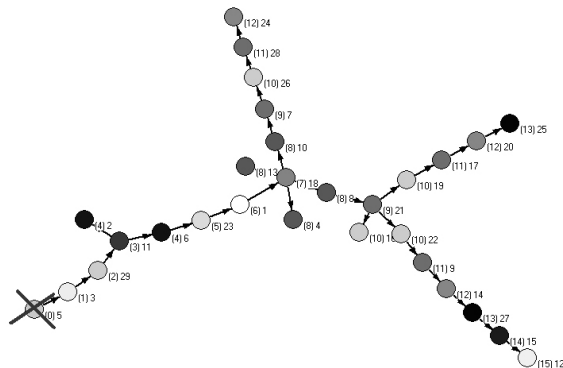
FIGURE 2. Avalanche dimension = 29, depth level = 15. The root node (marked with a cross) in the down-left corner is not a hub and the perturbation spreads along the branch of the tree; only two hubs are involved in the perturbation.

inputs): in this case there are four allowed Boolean functions for B which make it actually independent of the state of A: the function which is always 1, always 0, the one which always takes the same value of the other input or the opposite one. So, if non canalyzing functions are excluded, in $4/14 \approx 28\%$ of the cases there would be no real influence - and every experimental test would show no influence of A on B - even if the formalism of RBNs would show a formal link between the two.

Let us call the links which are built according to the formal procedure for RBNs or SFRBNs formal links. By excluding the ineffective links discussed above (on average, the 28% of the total number of formal links), one is left with the subset of active functional outgoing links. Taking into account the actual distribution of *formal links* in our networks, one finds that about 32% of the nodes have no *active functional outgoing links* - a number which differs significantly from the number of nodes without any formal outgoing link (13%).
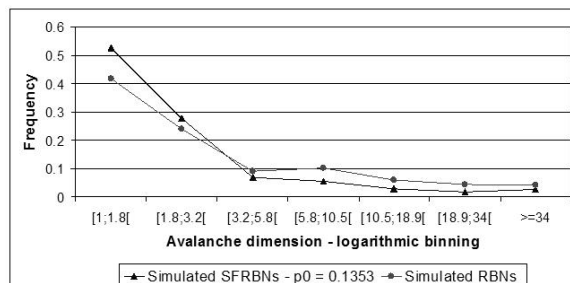

FIGURE 3. Comparison between the distribution of the avalanches in classical RBNs (mean on 50 simulated networks) and SFRBNs with $p_{out}(0)$=0.1353 (mean on 10 simulated networks) - logarithmic binning.

In the comparison with classical RBNs (see fig. 3), the avalanche distribution differs significantly, since smaller avalanches are largely more frequent in SFRBNs (with $p_{out}(0) = 0.1353$). Moreover, it is interesting to observe that, while the value of the average avalanche is similar in the two cases (8.17 vs. 7.14), the mean maximum avalanche is much larger in SFRBN case (403.7) than in RBN (132.8). SFRBNs also show a higher variance on the dimension of the avalanches. All these phenomena are likely due to the same reasons highlighted above: the larger number of poorly connected nodes compared to the RBN case (where the outdegree distribution is Poissonian) increases the frequency of very small avalanches (including those of size 1) in SFRBNs; on the other hand, the presence of hubs makes large avalanches more frequent than in the case of Poisson distribution, where highly connected nodes are very unlikely.

The distribution of avalanches in experimental data plainly depends on the choice of the threshold (see section 3): small avalanches become relatively more frequent as the chosen threshold increases. Consequently, the choice of a too large threshold would lead to a trivial avalanches distribution, since the great majority of the genes would turn out to be unchanged. Hence, we decided to compare SFRBNs with experimental data with thresholds within the reasonable range 4 - 15.

As explained above, in SFRBNs smaller avalanches are relatively more frequent than in the case of analogous classical RBNs, and increase as $p_{out}(0)$ becomes larger. As a consequence, the best fit among SFRBNs and experimental data with low thresholds is obtained using SFRBNs with low values of $p_{out}(0)$. For all of the thresholds in the range 4 - 15 the best agreements are with SFRBNs whose $p_{out}(0)$ is smaller than or equal to 0.1. In particular, the best statistical match is between experimental data with a threshold equal to 11 and SFRBNs with $p_{out}(0) = 0.1$
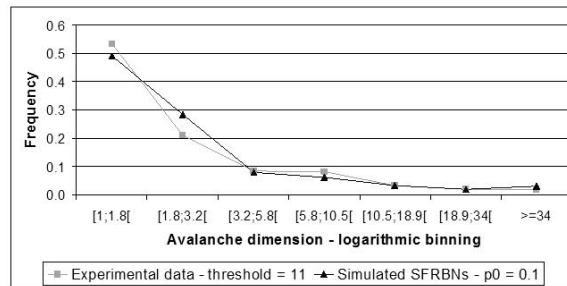

FIGURE 4. Comparison between experimental data with Θ=11 and the results of simulations on SFRBNs with $p_{out}(0) = 0.1$ (mean on 10 simulated networks), logarithmic binning. Θ = 11 provides optimal fit of SFRBNs to experimental data, according the well-known index χ^2.

Figure 4 shows the comparison between the simulated and experimental distributions, using the optimal (for SFRBNs with $p_{out}(0) = 0.1$) threshold 11.

Nevertheless, although the overall agreement is undoubtedly significant, there is an important aspect which is hidden by this representation: given the presence of hubs, the average maximum avalanche increases with $p_{out}(0)$ (see fig. 5) and even when $p_{out}(0)=0$ it is already much higher (583.2) than the observed experimental value in *S. cerevisiae* (which decreases from 219 to 150 as the threshold is increased from 4 to 15).
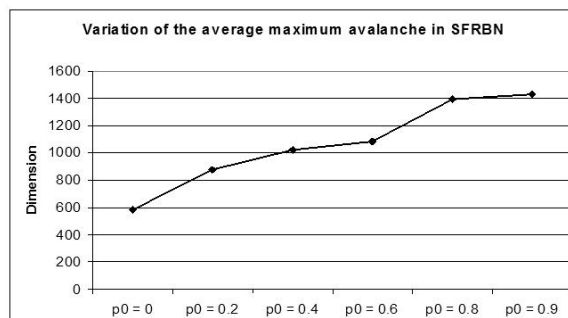
FIGURE 5. Variation of the maximum observed avalanche, averaged on 10 different networks for each distinct value of $p_{out}(0)$; set of the Boolean functions restricted to the canalyzing functions only. The considered values of $p_{out}(0)$ are: 0, 0.2, 0.4, 0.6, 0.8 and 0.9. The simulations are realized with a particularly high number of knockout perturbations (2000), much larger than the experimental one (227).

Further analyses have been made on SFRBNs with critical slope (with, in particular, $p_{out}(0) = 0$ and $\gamma = 2.469$), but without the restriction on canalyzing Boolean functions only. Networks of this kind are exactly critical ([1]). The results of the simulations showed a much higher variance on the distribution of avalanches compared to the sub-critical case, recalling an analogous behaviour which had been observed analysing different RBNs ([22]). Also for SFRBNs, the restriction on canalyzing functions clearly reduces the dispersion of the distributions.

In addition, even though SFRBNs with $p_{out}(0) = 0$ and no constraint on the set of Boolean functions provide indeed a good approximation to the distribution of experimental small avalanches, the mismatch regarding the maximum avalanche is even larger than the one observed with SFRBNs with canalyzing functions only (the value of the average maximum avalanche being 1296).

6. **Conclusions.** The previous studies on classical RBNs demonstrated that the distribution of avalanches, provided that they are small with respect to the total number of nodes, depends only upon the Derrida parameter. The only topological feature which enters this parameter is the average degree, while other topological parameters are irrelevant (as long as the outdegree distribution is Poissonian). On the other hand, the analyses presented here on scale-free random Boolean networks show that the form of the outdegree distribution function does indeed affect the way in which the perturbations induced by knock-outs spread in the gene network, i.e. the distribution of avalanches.

In particular, due to the presence of hubs with a large number of connections and of a high number of poorly connected nodes, scale-free networks display larger maximum avalanches than RBNs. In contrast, if the number of knock-outs is much smaller than the number of nodes, it may happen that a hub is never knocked-out or reached by a propagating avalanche. Therefore, when examining the maximum avalanche in SFRBNs, one observes a high variance in the case of a small number of knock-outs.

Under the hypotheses of the presented study, discussed above, there are two parameters whose values affect the comparison of simulations versus experimental data: the value of $p_{out}(0)$, which determines the distribution of simulated avalanches, and the threshold, which determines that of the real ones. When the parameters are chosen in a way which guarantees that the two distributions are close, one observes that the maximum simulated avalanche is typically much larger than the experimental one. Since a similar effect has not been found in classical RBNs, the present results would seem to favour them with respect to SFRBNs for the description of the response of the network to knock-out perturbations.

However, the high variance of the maximum avalanche in SFRBNs suggests a more cautious approach. Note that the data concerning simulations refer to a number of knock-outs (2000) which is much larger than the experimental one. Indeed, when simulating a much smaller number of knock-outs, of the order of the experimental one (227), one occasionally observes values of the maximum avalanche close to the one actually observed in *S. cerevisiae*. Therefore the mismatch on the value of the maximum avalanche might be related to the limited size of the experimental sample: in 227 knock-outs one might have never have hit a large hub. While this might be regarded as an unlikely event, it cannot be ruled out until further experimental data become available.

Note that in this study some topological features have been supposed *a priori*, and the consequences of different hypotheses (RBNs, SFRBNs) have been compared. It should be remarked that it is unlikely that real gene networks are purely random or purely scale-free without cut-off. Moreover, there are different ways to obtain scale-free networks, with different properties ([4]) and in this work we explored the behaviour of a particular kind of network of this type. It has been however shown that data on gene knock-out can provide useful indications, and accumulating further data might reduce the uncertainties which are still there. A different, very promising approach which might provide suggestions or constraints on the network topologies is the one which takes into account the evolutionary process which led to present-day gene networks (see e.g. [12] for a recent review).

REFERENCES

[1] M. Aldana, *Boolean dynamics of networks with scale-free topology*, Physica D, **185** (2003), 45–66.

[2] M. Aldana, S. Coppersmith and L. P. Kadanoff, *Boolean dynamics with random couplings*, in "Perspectives and Problems in Nonlinear Science" (eds. E. Kaplan, J.E. Marsden and K.R. Sreenivasan), Springer Applied Mathematical Sciences Series, (2003).

[3] L. A. N. Amaral, A. Scala, M. Barthelemy and H. E. Stanley, *Classes of small-world networks*, PNAS, **97** (2000), 11149–11152.

[4] A. Banerjee and J. Jost, *Spectral plots and the representation and interpretation of biological data*, Theory. Biosci., **126** (2007), 15–21.

[5] A. L. Barabasi, "Linked," Perseus, New York, 2002.

[6] U. Bastolla and G. Parisi, *The modular structure of Kauffman networks*, Physica D, **115** (1998), 219–233.

[7] B. Derrida and Y. Pomeau, *Random networks of automata: a simple annealed approximation*, Europhys. Lett. 1, **2** (1986), 45–49.

[8] I. Farkas, H. Jeong, T. Vicsek, A.-L. Barabasi and Z. N. Oltvai, *The topology of the transcription regulatory network in the yeast, Saccharomyces cerevisiae*, Physica A, **318** (2003), 601–612.

[9] D. E. Featherstone and K. Brodie, *Wrestling with pleiotropy: genomic and topological analysis of the yeast gene expression network*, BioEssays, **24** (2002), 267–274.

[10] J. J. Fox and C. C. Hill, *From topology to dynamics in biochemical networks*, Chaos, **11** (2001), 809–815.

[11] S. E. Harris, B. K. Sawhill, A. Wuensche and S. A. Kauffman, *A model of transcriptional regulatory networks based on biases in the observed regulation rules*, Complexity, **7** (2002), 23–40.

[12] E. L. Haseltine and F. H. Arnold, *Synthetic gene circuits: design with directed evolution*, Annual Review of Biophysics and Biomolecular Structurel, **36** (2007), 1–19.

[13] T. R. Hughes et al., *Functional discovery via a compendium of expression profiles*, Cell, **102** (2000), 109–126.

[14] H. Jeong, S. P. Mason, A.-L. Barabasi and Z. N. Oltvai, *Lethality and centrality in protein networks*, Nature, **411** (2001).

[15] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai and A.-L. Barabasi, *The large- scale organization of metabolic networks*, Nature, **407** (2000), 651–654.

[16] S. A. Kauffman, "The Origins of Order," Oxford University Press, Oxford, 1993.

[17] S. A. Kauffman, C. Peterson, B. Samuelsson and C. Troein, *Genetic networks with canalyzing Boolean rules are always stable*, PNAS, **101** (2004), 17102–17107.

[18] S. A. Kauffman, C. Peterson, B. Samuelsson and C. Troein, *Random Boolean networks models and the yeast transcriptional networks*, PNAS, **100** (2003), 14796–14799.

[19] T. I. Lee et al., *Transcriptional regulatory networks in Saccharomyces cerevisiae*, Science, **298** (2002), 799–804.

[20] R. Serra, M. Villani and L. Agostini, *On the dynamics of random Boolean networks with scale-free outgoing connections*, Physica A, **339** (2004), 665–673.

[21] R. Serra, M. Villani, A. Graudenzi and S. A. Kauffman, *Why a simple model of genetic regulatory networks describes the distribution of avalanches in gene expression data*, Journal of Theoretical Biology, **246** (2007), 449–460.

[22] R. Serra, M. Villani and A. Semeria, *Genetic network models and statistical properties of gene expression data in knock-out experiments*, Journal of Theoretical Biology, **227** (2004), 149–157.

[23] J. E. S. Socolar and S. A. Kauffman, *Scaling in ordered and critical random Boolean networks*, Phys. Rev. Let., **90** (2003).

[24] A. Wagner, "Estimating Coarse Gene Network Structure from Large-Scale Gene Perturbation Data," Santa Fe Institute Working Paper, 01-09-051, 2001.

[25] A. Wagner and S. Fell, "The Small World Inside Large Metabolic Networks," Santa Fe Institute Working Paper, 00-07-041, 2000.

E-mail address: `rserra@unimore.it`

E-mail address: `mvilani@unimore.it`

E-mail address: `alex.graudenzi@unimore.it`

E-mail address: `annamaria.colacci@unibo.it`

E-mail address: `skauffman@ucalgary.ca`