Continuous network models of gene expression in knock-out experiments: A preliminary study

A. ROLI and F. VERNOCCHI

DEIS, Campus of Cesena,
Alma Mater Studiorum *Università di Bologna*
Cesena, Italy
E-mail: andrea.roli@unibo.it, francesco.vernocchi@studio.unibo.it

R. SERRA

Dipartimento di scienze sociali, cognitive e quantitative,
Università di Modena e Reggio Emilia,
Reggio Emilia, Italy
E-mail: rserra@unimore.it

In this work we simulate gene knock-out experiments in networks in which variable domains are continuous and variables can vary continuously in time. This model is more realistic than other well-known switching networks such as Boolean Networks. We show that continuous networks can reproduce the results obtained by Random Boolean Networks (RBN). Nevertheless, they do not reproduce the whole range of activation values of actual experimental data. The reasons for this behavior very close to that of RBN could be found in the specific parameter setting chosen and lines for further investigation are discussed.

Keywords: Genetic networks, gene expression, knock-out gene, random boolean networks, Glass networks

1. Introduction

In previous studies,[1] it is shown that single gene knock-out experiments can be simulated in Random Boolean Networks (RBN), which are well-known simplified models of genetic networks.[2,3] The results of the simulations are compared with those of actual experiments in *S. cerevisiae*. The actual data are taken from the experiments described in a work by Hughes et al,[4] in which a genetic network of over 6000 genes is considered and a series of 227 experiments in which one gene is silenced are run on DNA microarrays. Genes are knocked-out (i.e. silenced) one at a time, and the variations in the

expression levels of the other genes, with respect to the unperturbed case (the *wild* type), are considered as the ratio of the activation in knock-out state (KO) and wild type (WT). Besides the ratios of KO/WT activation, *avalanches* can be defined that measure the size of the perturbation generated by knocking out a single gene.

In previous work on RBN,[1,5] it has been found that the distributions of avalanches are very robust, i.e. they are very similar in different random networks and the distribution of avalanches of the RBN models are close to those observed in actual experiments performed with *S. cerevisiae*. These results suggest that these distributions might be properties common to a wide range of genetic models and real genetic networks.

RBN are a very simplified model of genetic networks as they assume that a gene is either active or inactive, whilst in nature gene activation can range in a continuous domain. In this work we undertake an analogous study as in the case of RBN using a continuous network model, first proposed and discussed by Glass.[6,7] We show that Glass networks can reproduce the same results as RBN. Moreover, also the results of experiments with DNA microarrays are reproduced with a high level of accuracy. Nevertheless, with the parameter setting we used, this model is still not capable of capturing the whole range of KO/WT activation ratios.

Glass networks are described in Section 2 and their differences with Boolean Networks are outlined. Section 3 provides an overview of the main experimental settings and results are reported and discussed in Sections 4 and 5 in which we compare the results of Glass networks simulations with RBN and actual experiments, respectively. We conclude by discussing future work in Section 6.

2. Continuous networks

Glass networks[6,7] are continuous time networks in which node activation rate is regulated by a differential equation that includes a non-linear component that depends on a Boolean function of the node inputs. In these networks, time and gene activation are continuous, while the influence among genes is represented by switching functions. The activation of gene i is indicated with variable x_i ranging in a continuous domain. We associate to x_i a Boolean variable X_i defined as follows:

$$X_i(t) = \begin{cases} 0 \text{ , if } x_i(t) < \theta_i \\ 1 \text{ , otherwise} \end{cases}$$

In a network with N nodes, each with K inputs we define the activation rate of node i as:

$$\frac{dx_i}{dt} = -\tau_i x_i + f_i(X_{i_1}(t), X_{i_2}(t), \ldots, X_{i_K}(t)), \qquad i = 1, 2, \ldots, N$$

where f_i is a Boolean function of the inputs of node i.

Since the functions f_i change only when at least one variable crosses the threshold, the equations can be solved analytically in the intervals in which the Boolean functions are constant. Thus, if we denote with $T_s = \{t_1, t_2, \ldots, t_s\}$ the set of switching times, for each $x_i, i = 1, 2, \ldots, N$, and $t_j < t < t_{j+1}, t_j \in T_s$, we have:

$$x_i(t) = x_i(t_j)e^{-(t-t_j)\tau_i} + \frac{1}{\tau_i} f_i(X_{i_1}(t_j^*), X_{i_2}(t_j^*), \ldots, X_{i_K}(t_j^*))(1 - e^{-(t-t_j)\tau_i})$$

where $t_j^* \in (t_j, t_{j+1})$.

This model still introduces strong simplifications, but it explicitly takes into account the continuous time dynamics and continuous values of actual genetic networks. A more simplified model, though able to capture relevant properties of genetic networks, is that of (Random) Boolean Networks.[2] Variables associated to nodes in RBN assume values in the binary domain $\{0, 1\}$ and the transition functions are Boolean functions of the inputs. Usually, a synchronous dynamics is imposed to RBN. RBN have been studied as a model of genetic networks in which a gene is either active or inactive and it has been shown that they can simulate important properties of genetic networks.[2,5]

In this work, we present a preliminary study in which Glass networks are used to simulate gene knock-out experiments, as previously done with RBN. Our goal is to check if these networks can reproduce the results of simulations by RBN and to what extent and under which hypotheses they can capture more realistic features of the actual genetic networks.

3. Experimental setting

We performed a series of experiments to study the influence of single knock-out genes in continuous networks. Node activation ranging in a continuous domain makes it possible to compare results both with RBN, by converting values x_i into Boolean ones, and directly with *DNA microarray* results.

We designed and implemented a software system to simulate Glass networks and alike. The software has been designed trying to find a trade-off

between performance, in terms of execution time, and extensibility. The tool can be configured so as to simulate models with different parameter settings and network characteristics, such as the topology. The simulator has has been implemented in C++.

The networks we generated have a unitary decay parameter, $\tau_i = \tau = 1$, for every node and threshold value $\theta_i = \theta = 0.5$ ($i = 1, 2, \ldots, N$), so as to have node values in the range $[0, 1]$. The other network parameters have been chosen according to previous work in which the results of simulations of RBN are compared with results in DNA microarrays. Thus, every node has two inputs, i.e., $K = 2$, randomly chosen among the other nodes. Boolean functions are assigned randomly to nodes, by picking them among the set of *canalizing* functions, i.e., functions in which at least one value of one of its inputs uniquely determines the output. For the case with $K = 2$, all the possible 16 functions except for *coimplication* and *exclusive or* are canalizing.

We generated 30 networks with 6000 nodes; each network is initialized with random values in the range $[0, 1]$ and its evolution in time is simulated until an attractor is reached.[a] The activation of a node is computed as the average value it assumes along the attractor. Then, 227 genes randomly chosen among the active ones, i.e., the ones with average activation greater than zero, were silenced in turn. The activation of the genes in the knock-out experiment were then evaluated in the new attractor and the ratio between the activation in the knock-out and wild type has been computed. Hence we obtain a 6000×227 matrix of real values that can be compared both with the corresponding matrix of experiments with RBN and real data. For each network we produce a matrix $E_{ij}, i = 1, \ldots, 6000\ ,\ j = 1, \ldots, 227$, in which E_{ij} is the ratio of the activations of gene i in experiment j.

4. Comparison with Random Boolean Networks

The first analysis we make is a comparison of the results of simulations of gene knock-out in Glass networks with those of RBN. The values of the matrix built from real data and from simulation of the continuous model have been processed as in previous work[1] by introducing a threshold θ_E to define the level of meaningful difference between the knock-out and wild type: the difference between KO and WT activations is considered significant if the ratio is greater than θ_E or less than $1/\theta_E$. Hence we obtain a Boolean matrix E'_{ij} defined as follows: $E'_{ij} = 1$ if $E_{ij} > \theta_E \vee E_{ij} < 1/\theta_E$,

[a] In the experiments we made we always found a cyclic attractor.

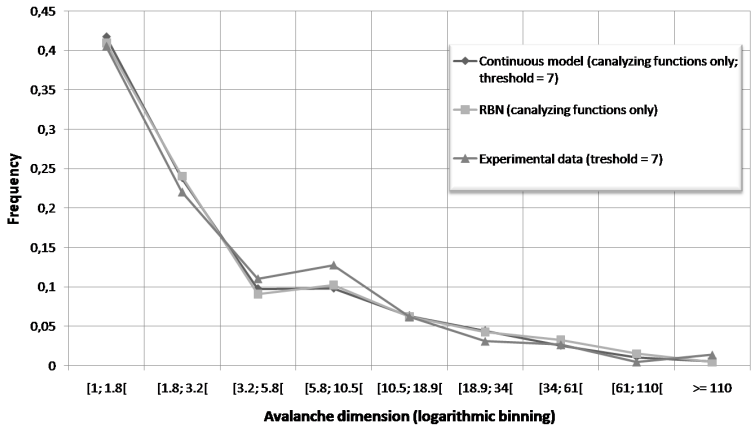

Fig. 1. Avalanche size frequency in simulations with continuous and Boolean network models and experiments in DNA microarray (linear binning)

$E'_{ij} = 0$, otherwise. As for results by RBN, any ratio not equal to 1 is considered as meaningful. The threshold θ_E has been set to 7, as from original work on RBN.[1]

Figures 1 and 2 plot the frequency of avalanche size of the two models and actual experimental data, in linear and logarithmic binning, respectively. This comparison shows that the results of the continuous model, when processed as the experimental ones, exhibit an avalanche frequency that closely approximates both the actual one and that of RBN.

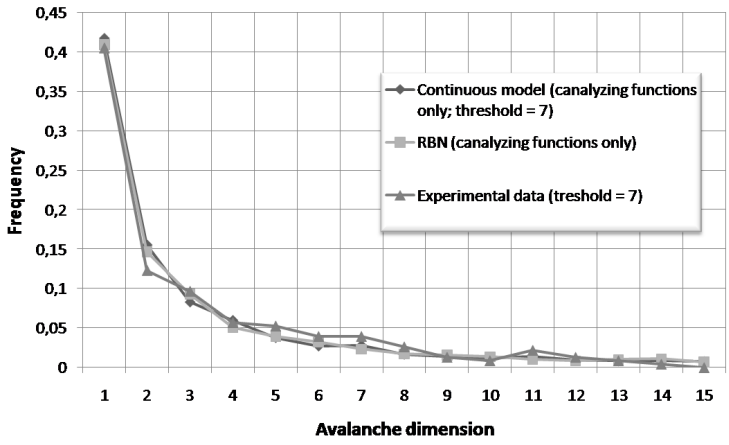

Fig. 2. Avalanche size frequency in simulations with continuous and Boolean network models and experiments in DNA microarray (logarithmic binning)

5. Comparison with microarrays experiments

Activation values of the continuous model can be directly compared against the experimental data from DNA microarrays. In a first analysis, we simply ordered the ratios and compared the curves plotted from actual experimental data and simulations. In Figure 3 we plot the data of knock-out experiments.

As for the continuous model, a typical case is plotted in Figure 4. The ratios produced by simulation of Glass networks approximately range in the same interval as the experimental data, nevertheless they have not the same distribution. Indeed, one can note that the values of the simulations by Glass networks are clustered around the extremes, while the values from experiments in DNA microarrays are more sparse.

We also considered a measure of the avalanche produced by a gene knock-out that does not depend upon a threshold. For each experiment we summed up the deviation from 1 of each gene, obtaining an array A defined as follows:

$$A_j = \sum_{i=1}^{6000} |1 - E_{ij}| \ , \ j = 1, \ldots, 227$$

The array A obtained from simulations by continuous networks is the

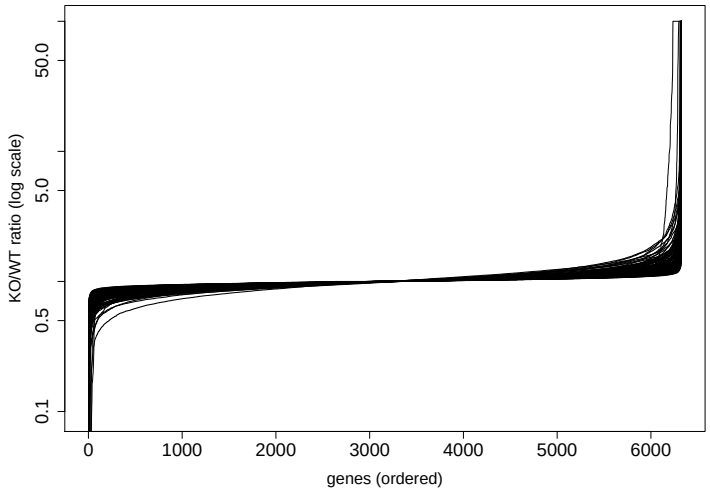

Fig. 3. KO/WT activation ratio in real data. In the x-axis, genes are ordered in non decreasing values of KO/WT activation ratio.

average over the 30 experiments.

In Figures 5 and 6 the cumulative distribution of the deviations is plotted in the case of experiments and simulations via continuous networks, respectively. The difference between the two distributions is apparent, as in the previous analysis. The discrepancy we observe could be ascribed to the network parameters chosen, that might keep the network in a 'quasi-Boolean' regime, thus preventing the nodes from assuming the whole range of values in the attractors.

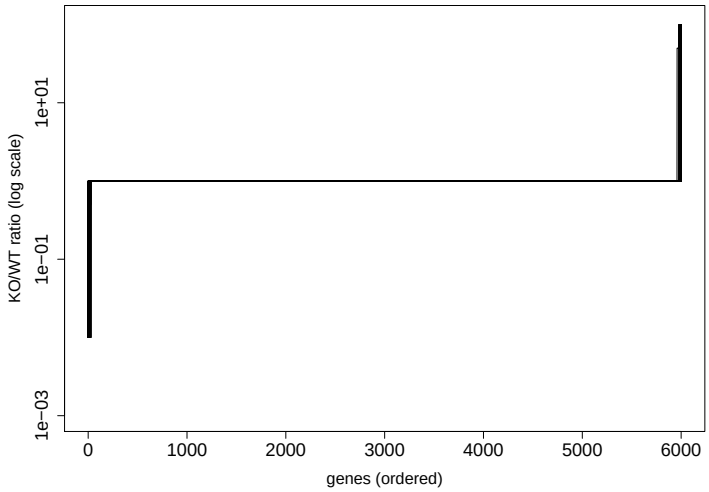

Fig. 4. KO/WT activation ratio in Glass networks. In the x-axis, genes are ordered in non decreasing values of KO/WT activation ratio.

6. Discussion and future work

In this work we have presented a preliminary investigation of the simulation of gene knock-out experiments via continuous networks. We have studied the frequency of avalanches, defined as the number of genes significantly affected by a single gene knock-out, and we have shown that this model can not only reproduce with accuracy the results of simulations by RBN, but also the results of experiments in DNA microarrays.

We have also observed that the distribution of continuous KO/WT activation values produced in our simulations via Glass networks seems still not very close to that of actual experiments. However, this is a preliminary study and further analyses are planned in which crucial parameters of Glass networks will be varied in order to try a more accurate tuning of the model. First of all, delays τ_i and thresholds θ_i can be varied and be different across the nodes. In addition, different equations regulating the expression rate can be studied. Finally, the topology of the network can be changed and networks with more realistic topologies can be studied.

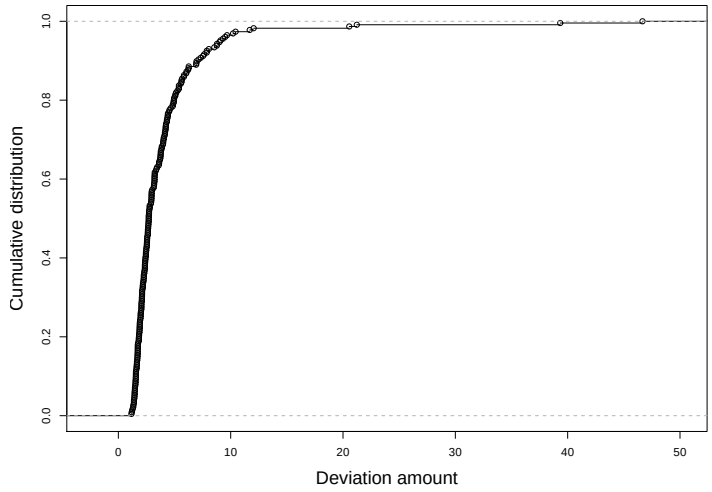

Fig. 5. Cumulative distribution of the deviations A_j in data from experiments.

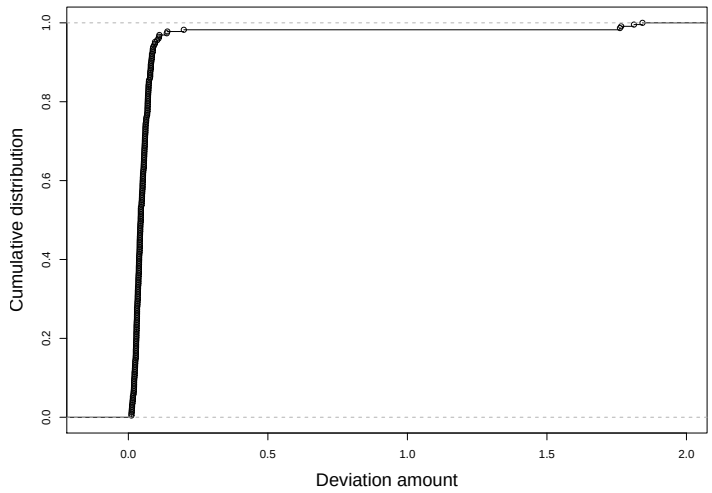

Fig. 6. Cumulative distribution of the deviations A_j in data from simulations by continuous networks.

References

1. R. Serra, M. Villani and A. Semeria, *Journal of Theoretical Biology* , 149 (2004).
2. S. A. Kauffman, *The Origins of Order: Self-Organization and Selection in Evolution* (Oxford University Press, 1993).
3. S. Kauffman, *Current topics in dev. biol.* **6** (1971).
4. T. R. Hughes et al, *Cell* **102**, 109 (2000).
5. R. Serra, M. Villani, A. Graudenzi and S. A. Kauffman, *Journal of Theoretical Biology* , 449 (2007).
6. K. Kappler, R. Edwards and L. Glass, *Signal Processing* (2003).
7. R. Edwards, *Physica D* , 165 (2000).