Eur. Phys. J. B (2016) 89: 33
DOI: 10.1140/epjb/e2015-60530-6

THE EUROPEAN PHYSICAL JOURNAL B

Regular Article

Coevolution of information processing and topology in hierarchical adaptive random Boolean networks

Piotr J. Górski[1], Agnieszka Czaplicka[2], and Janusz A. Hołyst[1,3,a]

[1] Faculty of Physics, Center of Excellence for Complex Systems Research, Warsaw University of Technology, Koszykowa 75, 00-662 Warsaw, Poland
[2] Instituto de Física Interdisciplinar y Sistemas Complejos IFISC (CSIC-UIB), Palma de Mallorca, Spain
[3] ITMO University, 19, Kronverkskiy av., 197101 Saint Petersburg, Russia

Received 30 June 2015 / Received in final form 30 October 2015
Published online 8 February 2016

Abstract. Random Boolean Networks (RBNs) are frequently used for modeling complex systems driven by information processing, e.g. for gene regulatory networks (GRNs). Here we propose a hierarchical adaptive random Boolean Network (HARBN) as a system consisting of distinct adaptive RBNs (ARBNs) – subnetworks – connected by a set of permanent interlinks. We investigate mean node information, mean edge information as well as mean node degree. Information measures and internal subnetworks topology of HARBN coevolve and reach steady-states that are specific for a given network structure. The main natural feature of ARBNs, i.e. their adaptability, is preserved in HARBNs and they evolve towards critical configurations which is documented by power law distributions of network attractor lengths. The mean information processed by a single node or a single link increases with the number of interlinks added to the system. The mean length of network attractors and the mean steady-state connectivity possess minima for certain specific values of the quotient between the density of interlinks and the density of all links in networks. It means that the modular network displays extremal values of its observables when subnetworks are connected with a density a few times lower than a mean density of all links.

1 Introduction

Many complex systems consist of evolving, constantly changing elements and complicated functional relations between them. Simplified models have been created to study and compare certain properties of these structures. One example of such approaches is random Boolean networks (RBNs). RBNs are generic and that is why they have been applied in many different fields [1–4] including gene regulatory networks (GRNs).

GRNs are models describing the structure and behavior of the transcriptional network responsible for regulating the gene expression in a cell [5]. Although gene expression levels corresponding to GRN can be described as a continuous process, they are also studied using discrete variables, like in the ensemble approach [6,7]. Nodes of RBNs represent genes as on-off devices and directed connections in RBNs represent physical or regulatory interactions. GRNs are both robust and stochastic [8]. Robustness signifies the ability to reproduce identical traits despite changing environment and stochasticity refers to fluctuations in a population of similar cells. GRNs must respond to diverse stimuli and they are highly modular, e.g. hierarchical regulatory interactions in the transcriptional networks of yeast [9]. The terms – hierarchy and modularity in GRNs – can be understood in many ways. One of them is assigning particular biological processes to some modules [5]. Specific groups of genes are then responsible for responding to specific stimuli. In another approach GRN elements can be divided into three layers which differ in the level of noise in their output [8]. Generally there exist genes, whose expression needs to be precise in order to avoid some lethal mutations and there exist genes whose expression can be more variable and modulated (i.e. noisier) in response to environmental cues. Essential genes display lower variability and a lower level of noise. Variable genes are more sensitive to mutations. The analysis of GRNs has demonstrated that diverse stimuli may cause modifications of gene interactions and network topology [10,11]. During GRN evolution edges are created and destroyed, and in this way the gene expression is altered.

Evolution is a process noticed and studied in different branches of science [11–23], including biology, biochemistry, social sciences and economics. In the constantly changing environment individuals must adapt and evolve to survive. In other words, evolution means searching for the most optimal configuration [4]. However, in order not to lose diversity and to explore the whole space of states individuals stay only near the possible best state [24]. Therefore, dynamics of evolutionary systems

[a] e-mail: jholyst@if.pw.edu.pl

can be considered to be critical [24] since there are broad spectra of system observables (e.g. power law distributions of attractor lengths). Models of macroevolution form a specific class of models [13] and they exhibit certain unique features and phenomena. Examples are punctuated equilibria combined with self-organized criticality, as well as cascades of changes producing power law distribution [12,13,18–20,24]. It is often noticed that dynamics of evolutionary systems consists of two parts: the local and global rules [12,13,16,18,21–23]. In the former states of system elements interact and change, in the latter – the whole system evolves.

RBNs have already been used as a model of evolution, e.g. [4,11,22,23]. In reference [22] Liu and Bassler propose an RBN modification, where network topology is time dependent and the observed coevolution encompasses some effects noticed in the GRN. Here, we extend the original Liu and Bassler's model, most of all by incorporating the concept of hierarchy [25–28] understood in terms of connection density that is in general higher for links connecting nodes belonging to the same subnetwork (module) and lower for links connecting different subnetworks. In our approach there are no robust genes, but a part of network edges are assumed to be robust, i.e. they stay unchanged during the system evolution. In this sense constant interlinks are similar to the described above essential genes and in this way we are extending the meaning of the standard RBN models when genes correspond to nodes only.

Selected features of possible modular structures of RBNs have already been analysed in references [29,30]. In reference [29] the authors have investigated the modular structure of RBNs consisting of nodes that have been identified as *relevant nodes*, i.e. the only nodes that influence the network's dynamics. A different idea of RBN modularity has been used in reference [30]. Poblanno-Balp and Gershenson have connected separate RBNs with additional links and analysed statistical properties of such systems (average number of attractors, average attractor lengths and average percentage of states in attractors) and their sensitivity to initial conditions. The network topology proposed in our paper is closer to [30], however we assume different kind of system dynamics and other observables are investigated.

The paper is organised as follows: in Section 2 we describe our model, among others presenting the algorithm of network evolution (Sect. 2.2). Section 3 discusses the applied information measures. Section 4 contains simulations description, results and discussion. Section 5 gives a summary and conclusions. Appendices contain additional analysis (Appendix A) and error estimations for simulation results (Appendix B).

2 Model of hierarchical adaptive random Boolean network (HARBN)

Our study extends the idea of adaptive rewiring algorithm that was for the first time introduced in random threshold networks [21] and then it was applied to RBNs in reference [22]. However, the terminology used in this paper is slightly different, following the paper [11]. Let us consider a system of coupled nodes in a form of a directed network that corresponds to a GRN. For simplicity we will assume that internal nodes' variables $\sigma_n(t)$ are 0 or 1 and they can evolve in time according to randomly chosen Boolean functions f_n. Arguments of the function f_n are internal variables $\sigma_{n_i}(t)$ of all nodes n_i such that there is a connection from a node n_i to the node n and $i = 1, 2, \ldots, m$, if the node n possesses m inputs.

$$\sigma_n(t+1) = f_n(\sigma_{n_1}(t), \sigma_{n_2}(t), \ldots, \sigma_{n_m}(t)). \quad (1)$$

Let us note that none of the nodes n_i can be the node n itself.

During evolution of adaptive RBN (ARBN) changes of the network state and topology are measured. Simulations consist of an a priori-defined number of epochs. In each epoch the network's attractor is found (i.e. a periodic orbit of all variables $\sigma_n(t)$ that is reached by the system after a transient time), one node is randomly chosen and according to the activity-dependent rewiring rule (ADRR) (described in Sect. 2.1) the number of incoming connections to this node is changed.

We have extended Liu and Bassler's model as follows. We consider a small number of ARBNs that we shall call *subnetworks*. Different subnetworks are connected by directed edges called *interlinks*. Our model is built in such a way that a density of interlinks is different than a corresponding density of edges within the subnetworks, i.e. *intralinks*. Therefore, when the density of interlinks is much smaller than the density of intralinks, different subnetworks are sparsely connected [31,32] and the created network is hierarchical in terms of connection density. Here, there are two hierarchy levels. But the introduced concept can be easily extended to create subsequent levels. During a single epoch, a network's attractor is found and for each subnetwork a node is chosen randomly. Connections of such a node are changed according to the ADRR. New edges may arise only between nodes in the same subnetwork. On the other hand, interlinks are permanent edges: they are created at the beginning of each simulation and they cannot be changed. There are two reasons for this. Firstly, interlinks can be viewed as weak links – connections, that are believed to stabilize the system [30,33,34]. Moreover, interlinks are the only robust network elements, thus in the sense of GRNs one can identify them with essential genes. Secondly simple allowing interlinks to rewire destroys the modular network structure. To solve this problem, different probabilities of wiring to the node from the same subnetwork and to the node from the others should be implemented. This would, however, complicate the model. Therefore, we decided the interlinks should stay fixed in a single run of simulation algorithm (Sect. 2.2). Different realizations lead to different interlink structures.

2.1 Activity-dependent rewiring rule

Equation (2) defines node's mean state $\langle \sigma_n \rangle$ during the attractor. T is the attractor length and t_0 is the number

of iterations sufficient to reach the attractor.

$$\langle \sigma_n \rangle = \frac{1}{T} \sum_{t=t_0+1}^{t_0+T} \sigma_n(t). \quad (2)$$

If the node's mean state $\langle \sigma_n \rangle$ is 0 or 1, then this node is considered to be frozen and one new incoming edge is added to this node. The new edge starts in a random node chosen from the group of nodes from the same subnetwork that does not already have a link to the considered node. Otherwise, this node is considered to be active and one of its incoming edges is randomly chosen and deleted.

Although the presence of the ADRR in biological systems has not been scientifically proven, it includes a very natural idea that invokes both the system stability and flexibility [11]. There exist also models assuming other rules of evolution that can be based, for instance, on centrality [19] or local information transfer [23]. However, dependence on node's mean state seems to be simple and natural. A frozen node in GRN is useless because it does not serve any purpose. This can be changed by adding to this node a new incoming link. On the other hand, also an overactive node that exhibits permanent switching inefficiently responds to received stimuli. Removing one of its connections may stabilize its output.

2.2 Adaptive algorithm for hierarchical RBNs

We make use of the algorithm proposed in reference [22]. Let N denote the total number of nodes. Then let $\mathbf{S}(t)$ define a state of the whole network: $\mathbf{S}(t) = \{\sigma_1(t), \sigma_2(t), \ldots, \sigma_N(t)\}$. Our algorithm for the coevolution of hierarchical RBNs is as follows:

1. Generate M uniform Boolean networks (subnetworks) containing N_M nodes each with K_{ini} directed incoming edges starting in a randomly chosen group of nodes belonging to the same subnetwork. Then generate K_M edges between subnetworks (interlinks) and for each node n randomly select a Boolean function f_n giving 0 and 1. The function f_n generates the output signal of the node n that reaches all other nodes m that are dependent on the node n (i.e. there is a directed link from n to m). We assume that the distribution of these functions is symmetrical, i.e. on average there is the same number of 0 and 1 as output values.
2. Generate a random initial state $\mathbf{S}(0)$ of all internal node variables $\sigma_n(0)$ and find the network's attractor length using the algorithm introduced in reference [22]. Its idea is to memorize a network state $\mathbf{S}(T_k)$, then to update the states of all the nodes of the network and to compare new network states $\mathbf{S}(t)$ with $\mathbf{S}(T_k)$. Equality means a system attractor has been reached. In order to skip transient parts of trajectories a new network state is memorized after a certain number of initial iterations.
3. Choose a node from each subnetwork and calculate its mean state during one attractor cycle.
4. According to the ADRR, change the topology of each subnetwork. Do not delete interlinks. In case there are no connections to delete, randomly choose another node and repeat this step.
5. Generate new Boolean functions for each node.
6. If the predefined maximum number of epochs is not reached, go back to step 2 [22].

Remarks:

- The number of interlinks is settled. The interlinks are generated as follows: first the highest possible number of interlinks is equally distributed between ordered pairs of subnetworks (due to directed edges each pair of subnetworks appears twice); secondly, the remaining interlinks are assigned between random pairs. For example, if the network structure consists of 3 subnetworks and 20 interlinks, then there are 6 ordered pairs of subnetworks and each pair is given 3 interlinks. The remaining 2 interlinks are assigned between randomly chosen pairs.
- The described method allows to find the attractor length equal to 90 000 at most. If the attractor is not found then it is assumed to be equal to 90 000.
- A single run of the adaptive algorithm is called a realization. Advancing through the steps 2–6 is called an epoch (after [22]). Updating the states of all nodes in step 2 is called an iteration.

3 Information measures of RBNs

Measuring information amount transmitted by a network is an important tool that facilitates exploring the features of a system. There are many approaches to do it, see for instance [35–40]. In reference [35] self-overlap past-future mutual information and in reference [38] average pairwise mutual information are investigated. In the former the states of the same node in consecutive iterations are compared and in the latter the states of all pairs of nodes are compared. It has been proven that both of these measures are maximal for critical RBNs. The authors of [36] propose entropy of the avalanche size distribution as a measure for information propagation and show that is is maximized for critical RBNs. Krawitz and Shmulevich [37] view the network as a classifier and introduce basin entropy to measure the uncertainty of network's dynamical behavior. In reference [39] information storage and information transfer have been used to study RBNs. Information storage is similar to previously mentioned self-overlap mutual information as it is the amount of information of a node from its past that is needed to predict its future. The authors have found that information storage is dominating in RBNs in ordered phase, information transfer is dominating in chaotic phase and both of these measures are balanced for small-world RBNs. Snyder et al. [40] have focused on a particular computational task and used mutual information between the network input and output to estimate necessary number of connections coming from input.

Here, we define network *activity information* I as a sum of transformed activities of each node. Activity A_n of a node n is the number of changes of the node's state during the network attractor divided by the attractor length. There is no information stored in frozen nodes (nodes with $A_n = 0$) and in nodes whose state changes in every iteration ($A_n = 1$). Such nodes should not contribute to the network information. On the other hand, nodes which change their state the same number of times as they keep the old state ($A_n = 0.5$) contribute mostly because of their unexpected behavior. Node activity can be identified with probability of changing the node's state. Then the activity A_n links with the parameter a_n described in reference [35] as the system's "self-overlap": $A_n = 1-a_n$. Here, we introduce a natural definition of node *activity information* as:

$$I_n = -A_n \log(A_n) - (1 - A_n) \log(1 - A_n). \qquad (3)$$

This definition is similar to the definition of mutual information used in reference [35] with equal probabilities of generating zeros and ones in Boolean functions. Equation (3) means that $I_n = 0$ for periods 1 or 2, when $A_n = 0$ or $A_n = 1$, respectively. It also means that we can completely predict the evolution of the node state in next time steps when $I_n = 0$ and we know the first two steps of the evolution, that specify whether the orbit is of period 1 or 2. Let us emphasize that $A_n = 0$ only for orbits of period 1 and $A_n = 1$ only for orbits of period 2. For orbits of longer periods there is no such explicitness. For instance for period 3, $A_n = \frac{2}{3}$. However, this value can be received for plenty of other orbits, e.g. of period 6. Therefore, for longer orbits the value of A_n is not specific for the orbit type. Moreover, the largest number of possible different orbits is for $A_n = 0.5$.

The total network activity information is given as $I = \sum_n I_n$. For brevity, further in this paper network activity information will be replaced by *network information*.

4 Simulations

Simulations of different structures of HARBN model consisted of 1000 epochs (maximum number) repeated 100 times (100 realizations having different structures of interlinks). Let (M, N_M, K_M) denote a network consisting of M subnetworks each of N_M nodes and linked by K_M interlinks. Here we show results for networks that contain 60 or 80 nodes. Our simulations were performed in two parts. In the beginning we reproduced results for ARBNs achieved in reference [22] ($M = 1$). Then, we divided networks into two, three, four subnetworks and linked them by $0 - 80$ interlinks, that formed HARBN. We compared results for ARBNs and HARBNs. An example structure of a simulated network after 1000 epochs is shown in Figure 1.

Each realization of ARBN and HARBN demonstrates different structure and information parameters. Let the mean steady-state (m.s.s.) of the parameter denote a mean value of asymptotic states reached by a system. The average is performed over initial conditions of network topology and internal nodes variables. Identical networks tend to fluctuate (after the initial transient period) near the same mean steady-state (m.s.s.) levels regardless the initial incoming connectivity (see Fig. A.1 in Appendix). On the other hand, networks of a different number of nodes, subnetworks or interlinks tend to reach different m.s.s. values (see Fig. A.2 in Appendix). Therefore, for each network type we define: a m.s.s. incoming connectivity K_{ss}, a m.s.s. network information I, a m.s.s. node information IPN (calculated as information I per node), a m.s.s. edge information IPE (calculated as information I per edge) and a geometric m.s.s. attractor length T. If not stated otherwise, the arithmetic mean is assumed. In order to determine all above parameters 200 beginning epochs of each realization have been discarded as transient periods. K_{ss} includes both intra – and interlinks and is calculated as follows:

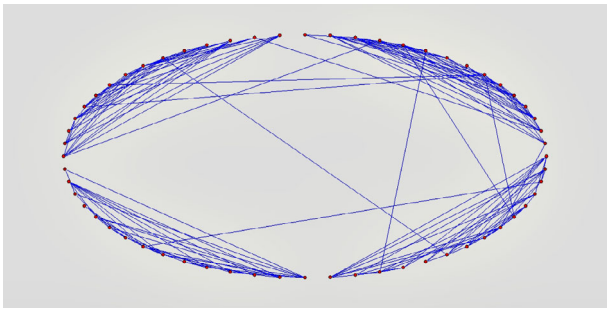

Fig. 1. An example structure after 1000 epochs of a HARBN system $(4, 15, 10)$, i.e. consisting of 4 subnetworks of 15 nodes each connected altogether by 10 interlinks.

- Calculate mean incoming connectivity in an epoch → K_{inc}.
- Calculate mean K_{inc} starting from the 201th epoch in a realization → $\langle K_{inc} \rangle$.
- Calculate mean $\langle K_{inc} \rangle$ over all realizations → K_{ss}.

The information parameters I and IPN are calculated likewise. IPE is computed by dividing IPN by K_{ss}.

The last parameter is a geometric mean attractor length T. The geometric mean has been used because of a broad distribution of attractor lengths.

Additional analyses can be found in Appendix A. Let us stress that observed networks adapt their configuration in such a way that they evolve towards a critical point [11] what can be seen by observing power law distributions of attractor lengths (Fig. A.4).

Along with the computation of m.s.s. data, deviations have been calculated. See Appendix B for details.

Now let us discuss the influence of network topology on system properties. We are aware of the fact that because of a small number of investigated networks additional large-scale simulations will be necessary to confirm our findings.

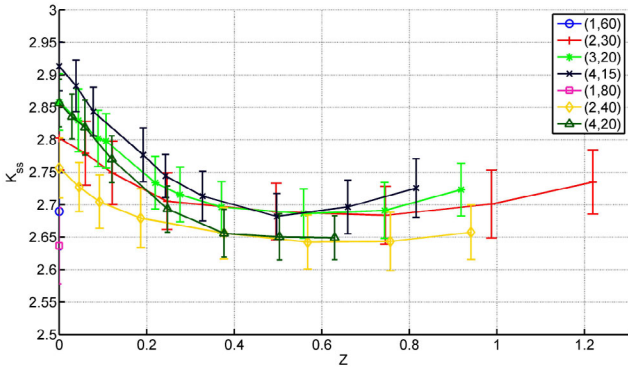

Fig. 2. The mean steady-state connectivity K_{ss} as a function of the quotient between the density of interlinks and the density of all links in the network for various structures.

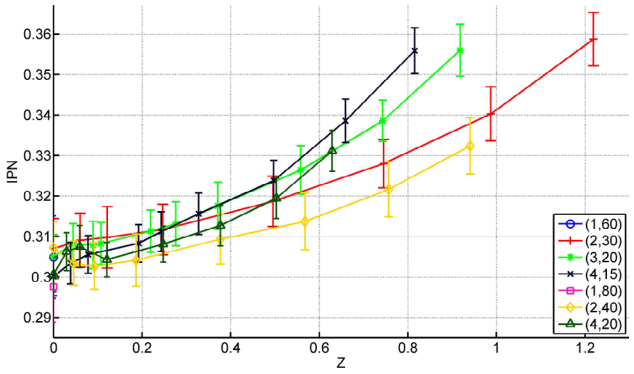

Fig. 3. The mean node information IPN as a function of the quotient between the density of interlinks and the density of all links in the network for various structures.

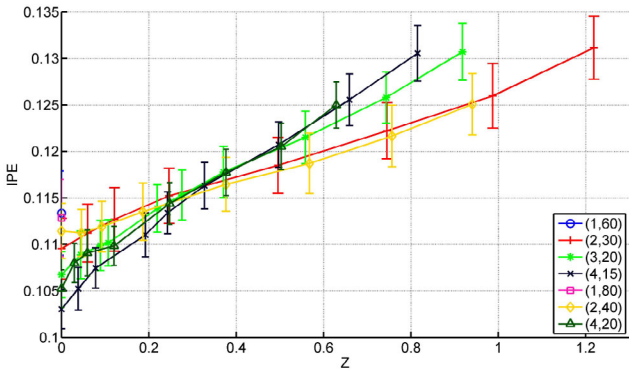

Fig. 4. The mean edge information IPE as a function of the quotient between the density of interlinks and the density of all links in the network for various structures. IPE has been calculated by dividing the mean node information IPN by the mean steady-state in-degree connectivity K_{ss}.

4.1 Mean steady-state connectivity

Let Z denote the quotient between the density of interlinks (defined as a quotient of the number of existing interlinks by the number of all possible interlinks) and the density of all links in the network.

$$\begin{cases} Z = \dfrac{K_M}{M(M-1)N_M^2} \Big/ \dfrac{K_{ss}}{N-1} \approx \\ \qquad \approx \dfrac{K_M}{N_M(M-1)K_{ss}} & \text{if } M \neq 1 \\ Z = 0, & \text{if } M = 1. \end{cases} \quad (4)$$

This parameter is considered here as an indicator of the network modularity. $Z = 0.25$ signifies that the density of interlinks is four times smaller compared to the density of all links and when $Z = 1$ then different subnetworks are connected as if there were no modular structure. In all of the following figures the standard deviations of Z are not shown because error bars are smaller than the marker size. The quotient $\dfrac{\Delta Z}{Z}$ is approximately 2%, at most.

For small Z values different network types differentiate themselves in terms of K_{ss} by the size of a subnetwork (Fig. 2) and K_{ss} decreases with the growth of the subnetwork size. With the increase of Z, the subnetworks become more and more linked. From $Z \approx 0.4$ different network types start to differentiate themselves by the size of the whole network, i.e. subnetworks are strongly mutually dependent. The resulting K_{ss} values are independent from the subnetwork size. K_{ss} reaches a minimal value approximately in the range $0.5 < Z < 0.7$. Afterwards K_{ss} increases slightly.

4.2 Information per node and information per edge

Calculated information per node IPN exhibits ordering by number of the subnetworks for large Z values (Fig. 3). When $Z > 0.2$ then the information IPN grows steeper for the networks with higher number of subnetworks. Moreover, IPN is larger in smaller networks. In the interval $0 < Z < 0.2$, especially for 80-node networks, ambiguous behaviour is observed. However, the statistical fluctuations do not allow detailed analysis. Afterwards, for $Z > 0.2$ a stable increase is visible.

The behavior of the third network parameter – IPE (Fig. 4) includes some features of the previous two observables. For $Z \approx 0$, networks differentiate themselves by the number of distinct subnetworks (the less subnetworks, the higher IPE). Neglecting the first two points for $(2, 40)$ the parameter IPE grows with the increase of Z. Around $Z \approx 0.3$ information per edge is equal for all network structures. It does not depend on the total network size or the number of subnetworks or the size of the single subnetwork. For higher Z values, strong ordering by the number of subnetworks is observed. That is there is no difference in IPE values for the network structures having the same number of subnetworks.

4.3 Geometric mean attractor length

Figure 5 shows the relationship between the geometric mean attractor length T and the quotient Z. The lengths of the attractors grow with the increase of the total number of nodes. Moreover, structures of the same size with more subnetworks tend to possess longer attractors.

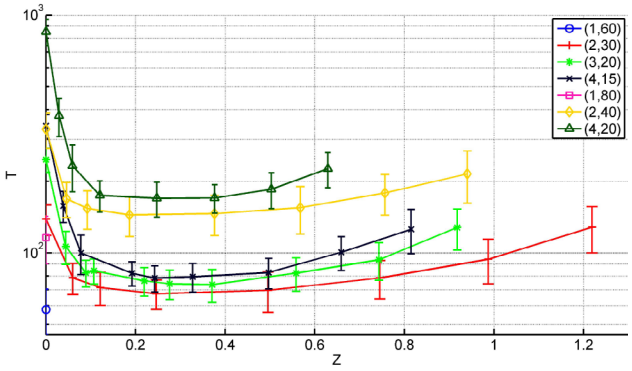

Fig. 5. The geometric mean attractor length T as a function of the quotient between the density of interlinks and the density of all links in the network for various structures.

Dividing a network into subnetworks creates separate unconnected parts ($Z \approx 0$). Such structures reach the highest T values. It is related to the attractor search subalgorithm: the attractor is found for the whole network and its separate parts multiply and elongate the attractor length. Interlinks connect the subnetworks, but interlinks are not flexible. Interlinks significantly decrease the attractor length leading to higher flexibility, because the system can quickly explore one short orbit and then switch to another one due to stochastic perturbations [41]. The shortest attractors are achieved for Z ranging from around 0.2 to around 0.4. Further increase of the quotient leads to longer lengths T, as interlinks stiffen the network.

4.4 Discussion

Edge rewiring is the basis of the considered model of network evolution. At first there are separate subnetworks which evolve independently. Adding constant interlinks interferes with this process. When the quotient Z of the density of interlinks to the density of all links is approximately 0.4 the connection between modules is so strong that the mean steady-state connectivity K_{ss} is similar as in nonmodular systems. However, before reaching this point, for $Z \in (0; 0.4)$, the impact of linkage between distinct subnetworks is noticeable but it is not so effective (see Fig. 2).

In the above range of the parameter Z two interesting effects have been observed. More and more interlinks lead to smaller geometric mean attractor lengths and for $Z \in (0.2; 0.4)$ the minimum of the observable T is noticed. Smaller values of geometric mean attractor length T mean that the network more often falls into shorter attractors and less frequently exhibits chaotic behavior. This makes the whole system more stable [4]. The second important feature is the behavior of the information per edge IPE. The number of links evolves due to the process of edge rewiring. The mean information processed by a single link may depend on the network structure, but is almost not influenced by the number of nodes in the network. In other words, adding new nodes to the network will change the total number of links, and this number will evolve in such a way that the mean edge information will stay constant. Moreover for $Z \approx 0.3$, the observable IPE does not depend on the subnetwork size and the number of subnetworks. It means that the information processed by a single link does not depend on the network structure for $Z \approx 0.3$. Such an effect could be a sign of the critical point, as in reference [42], however we have not observed any phase transition in our system.

The model introduced in this paper is generic and conclusions can also be applied in other fields where RBNs are used. We have shown and described the main application – GRNs. Modular structures are often noticed in real networks. If one is modeling such systems, the decision is necessary which values of the density ratio (that is parameter Z) should be applied. We do not strongly claim that certain Z values (e.g. $Z = 0.8$) should not be used. But we point out why certain ranges ($Z \in (0.2; 0.4)$) may be a better choice. The above range leads to the minimal mean attractor length for HARBNs and we suppose this property can be also valid for other modular systems, not just GRN models. Additional studies are needed to confirm or reject this hypothesis.

5 Conclusions

The model of a hierarchical adaptive random Boolean network (HARBN) has been introduced and numerically explored. The system consists of subnetworks connected by fixed interlinks, where the internal topology of the subnetworks can evolve depending on individual node activity. We have observed that the main natural feature of ARBNs, i.e. their adaptability, is preserved in HARBNs and they can evolve towards critical configurations. This feature has been documented by power law distributions of network attractor lengths. Adding just a few interlinks to separated subnetworks changes the values of all observables and the effect is especially striking for geometric mean attractor lengths. When the quotient Z of the density of interlinks to the density of all links in the network is of the order of $Z \approx 1.0$, then the interlinks efficiently re-connect the whole network (as if separate modules did not exist). However such a large system is less flexible and displays longer attractors. The shortest attractors are observed for intermediate values of Z ranging from 0.2 to 0.4 depending on the network structure. On the other hand, minimum values of mean steady-state connectivity K_{ss} are for Z ranging from 0.5 to 0.7. It means that the modular network displays extremal values of its observables when subnetworks are connected with a density a few times smaller than a mean density of all links.

The mean node information (IPN) and the mean edge information (IPE) grow in HARBN with the increase of the quotient Z and IPE tends to achieve the same value for all network types for $Z \approx 0.3$. Adding a new node to the network decreases IPN and, moreover, leads to fewer incoming connections per node. Adding a new interlink increases IPE values. We suggest that the introduced HARBNs can be a step forward towards modeling systems, such as GRNs, where a modular structure is present.

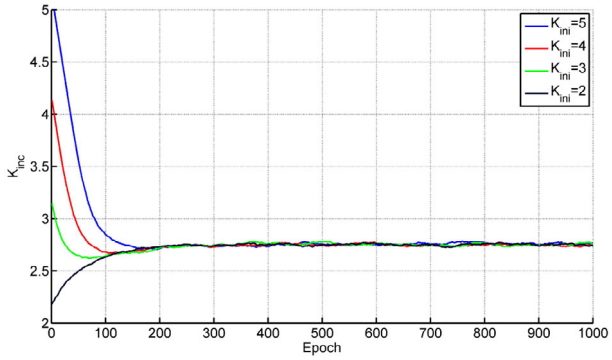

Fig. A.1. Evolution of incoming connectivity K_{inc} of a network $(2, 30, 10)$ for different initial incoming connectivities inside each subnetwork K_{ini}. Simulations have been averaged over 100 realizations. Let us note that $K_{inc}(0)$ is not equal to $K_{ini}(0)$ because K_{inc} includes interlinks, whereas K_{ini} does not.

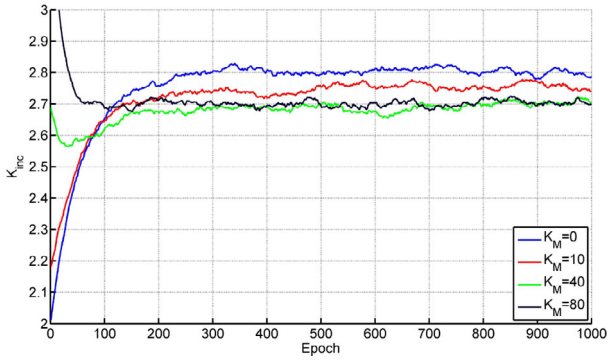

Fig. A.2. Evolution of incoming connectivity K_{inc} of a network $(2, 30)$ for different numbers of interlinks. The initial incoming intralink connectivity in each subnetwork is 2.

Author contribution statement

All authors have taken part in the conceiving of the system dynamics. Piotr J. Górski has designed and performed the simulations and drafted the paper. All authors have read and approved the final manuscript.

We are thankful to Dr Krzysztof Suchecki for a useful discussion. The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under Grant Agreement No. 317534 (the Sophocles project) and from the Polish Ministry of Science and Higher Education Grant No. 2746/7.PR/2013/2. J.A.H. has been also partially supported by Russian Scientific Foundation, proposal #14–21–00137 and by European Union COST TD1210 KNOWeSCAPE action.

Appendix A: HARBN dynamics

The goal of this section is to show details related to time evolution and criticality of HARBN. A similar network analysis for ARBN can be found in reference [22].

In the beginning of the network construction each node is given an equal number of incoming connections coming from other nodes from the same subnetwork (*internal connectivity*, K_{ini}). Then during network evolution due to the rewiring process, m.s.s. values of connectivity are reached (K_{ss}). Figure A.1 shows that the initial value of internal connectivity does not influence the K_{ss} values. After approximately 200 epochs all starting points converge to the same value of K_{inc}.

On the other hand, different structure types (i.e. those that are characterised by different symbol (M, N_M, K_M)) achieve, in general, different m.s.s. levels. Figure A.2 demonstrates that the m.s.s. level of K_{inc} is dependent on the number of interlinks K_M.

In reference [22] it has been noticed that for ARBN in the limit $N \to \infty$, the steady state value of mean number of incoming connections K_{ss} decreases to 2 as a power law

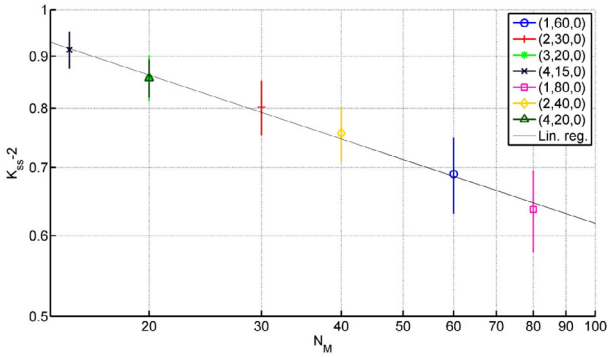

Fig. A.3. Finite size scaling for mean steady-state connectivity K_{ss} for networks without interlinks ($K_M = 0$) as a function of the subnetwork size N_M. The straight line possesses a slope -0.21 and it has been fitted from the linear regression analysis.

function of the system size. For HARBN without interlinks this feature can be also observed (see Fig. A.3). As it has been indicated before, separate subnetworks behave similarly as multiple realizations of a single smaller network. Using the linear regression with the least squares method the following scaling has been observed ($R^2 = 0.993$)

$$(K_{ss} - 2) = AN_M^{-\gamma} \tag{A.1}$$

where with 95% confidence level: $A = 1.61 \pm 0.11$ and $\gamma = 0.208 \pm 0.020$. The scaling exponent γ is close to the value $\gamma = 0.26$ reported in reference [22].

Figure A.4 shows distributions of attractor lengths of the network $(2, 40)$ in a steady-state with 0, 10, 40 and 80 interlinks. The distributions are very similar to each other and the only difference is seen for very short attractors. The probability of such attractors increases with the number of interlinks growing. In all cases the power law distribution is observed for attractor lengths $T > 5$ which indicates that analysed networks are critical in corresponding steady-states.

$$P(T) = BT^{-\beta}. \tag{A.2}$$

The absence of interlinks structure prolongs the attractor lengths (see Sect. 4.3) leading to smaller probabilities

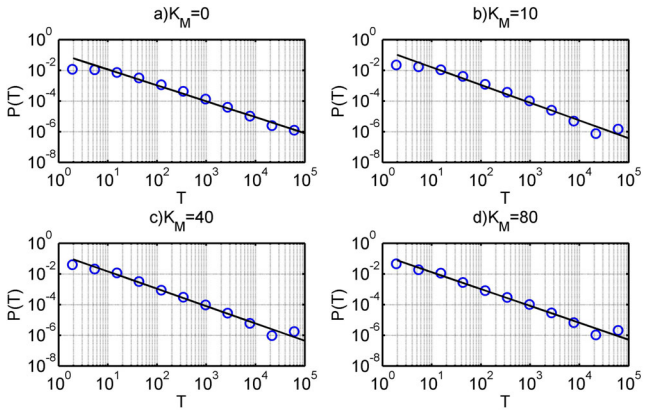

Fig. A.4. The distribution of attractor lengths of a network $(2, 40)$ in a steady-state for different numbers of interlinks. The power law distribution is observed for attractor lengths $T > 5$. The increase at the last data point is caused by the time limit of 90 000 in the attractor search subalgorithm. The straight lines are calculated from the linear regression analysis with the exclusion of the first data point at each subfigure.

of very short attractors. With more and more interlinks added this probability slightly rises. Using the linear regression with the least squares method the scaling has been found. For 0, 10, 40 and 80 interlinks, the coefficient of determination R^2 is 0.99, 0.98, 0.98 and 0.98, respectively. With 95% confidence level the scaling parameters are: $\{B_0 = \exp(-2.09 \pm 0.62), \beta_0 = 1.038 \pm 0.089\}$, $\{B_{10} = \exp(-1.5 \pm 1.0), \beta_{10} = 1.16 \pm 0.14\}$, $\{B_{40} = \exp(-1.64 \pm 0.84), \beta_{40} = 1.13 \pm 0.12\}$ and $\{B_{80} = \exp(-1.82 \pm 0.85), \beta_{80} = 1.10 \pm 0.12\}$. Let us note that the observed scaling exponents β are close to the slope 1 observed in reference [22]. We have also used the test introduced by Clauset et al. [43]. Unfortunately the test is not confirming the power law fitting.

Appendix B: Data error analysis

Let all errors (standard deviations) be denoted by a proper observable name preceded by Δ. Since an average over $\langle K_{inc} \rangle$ from different realizations will result in K_{ss}, thus ΔK_{ss} is estimated as a standard deviation of K_{ss} for 100 realizations of network structures having different structures of interlinks.

A similar method has been applied to m.s.s. node information IPN and geometric m.s.s. attractor length T. For some observables (e.g. K_{ss}) it is possible to calculate standard deviations of observable in a single realization when deviations are observed over time trajectories and over individual nodes in a network. In the case of observables K_{ss} and T [44] this approach would decrease the deviations, approximately by half and 10 times, respectively. Such a method cannot be applied however to the variable IPN, because node information values averaged over all nodes in consecutive epochs are not Gaussian and they can drastically change from 0 (e.g. for epoch with attractor length $T = 1$) to some larger values. Therefore, to avoid the above ambiguity an identical procedure has been applied to all directly measured observables.

The observable IPE is not measured directly. Its standard deviation ΔIPE is given by the following equation:

$$\Delta IPE = \sqrt{\left(\frac{\Delta IPN}{K_{ss}}\right)^2 + \left(\frac{IPN}{K_{ss}^2}\Delta K_{ss}\right)^2}. \quad \text{(B.1)}$$

The parameter Z defined in Section 4.1 (Eq. (4)) is a function of K_{ss}. Thus the standard deviation ΔZ can be calculated as:

$$\Delta Z = Z\frac{\Delta K_{ss}}{K_{ss}}. \quad \text{(B.2)}$$

References

1. C. Gershenson, Introduction to Random Boolean Networks, in *Workshop and Tutorial Proceedings, Ninth International Conference on the Simulation and Synthesis of Living Systems* (ALife IX, 2004), pp. 160–173
2. B. Drossel, Rev. Nonlinear Dyn. Complex. **1**, 69 (2008)
3. D. Cheng, H. Qi, Z. Li, in *Analysis and control of Boolean networks: a semi-tensor product approach* (Springer, London, 2011), pp. 431–450
4. M. Aldana, S. Coppersmith, L.P. Kadanoff, in *Perspectives and Problems in Nonlinear Science* (Springer, 2003), pp. 23−89
5. L.D.F. Costa, F.A. Rodrigues, A.S. Cristino, Genet. Mol. Biol. **31**, 591 (2008)
6. S.A. Kauffman, J. Theor. Biol. **22**, 437 (1969)
7. S. Kauffman, Physica A **340**, 733 (2004)
8. L.T. MacNeil, A.J. Walhout, Genome Res. **21**, 645 (2011)
9. R. Jothi, S. Balaji, A. Wuster, J.A. Grochow, J. Gsponer, T.M. Przytycka, L. Aravind, M.M. Babu, Mol. Syst. Biol. **5**, 294 (2009)
10. N.M. Luscombe, M.M. Babu, H. Yu, M. Snyder, S.A. Teichmann, M. Gerstein, Nature **431**, 308 (2004)
11. T. Rohlf, S. Bornholdt, in *Adaptive Networks* (Springer, 2009), pp. 73–106
12. P. Bak, K. Sneppen, Phys. Rev. Lett. **71**, 4083 (1993)
13. R.V. Solé, S.C. Manrubia, Phys. Rev. E **54**, R42 (1996)
14. S. Jain, S. Krishna, Phys. Rev. Lett. **81**, 5684 (1998)
15. S. Jain, S. Krishna, Proc. Natl. Acad. Sci. **99**, 2055 (2002)
16. M.G. Zimmermann, V.M. Eguíluz, M. San Miguel, Phys. Rev. E **69**, 065102 (2004)
17. P. Klimek, S. Thurner, R. Hanel, J. Theor. Biol. **256**, 142 (2009)
18. S. Thurner, P. Klimek, R. Hanel, New J. Phys. **12**, 075029 (2010)
19. M.D. König, C.J. Tessone, Phys. Rev. E **84**, 056108 (2011)
20. Y.P. Gunji, T. Sakiyama, H. Murakami, Biosystems **123**, 99 (2014)
21. S. Bornholdt, T. Rohlf, Phys. Rev. Lett. **84**, 6114 (2000)
22. M. Liu, K.E. Bassler, Phys. Rev. E **74**, 041910 (2006)
23. T. Haruna, S. Tanaka, On the Relationship between Local Rewiring Rules and Stationary Out-degree Distributions in Adaptive Random Boolean Network Models, in *ALIFE 14: The Fourteenth Conference on the Synthesis and Simulation of Living Systems* (2014), Vol. 14, pp. 420–426

24. P.M.C. de Oliveira, Theory in Biosciences **120**, 1 (2001)
25. *Hierarchy in natural and social sciences*, edited by D. Pumain, Vol. 3 of *Methodos series* (Springer, Dordrecht, 2006), http://opac.inria.fr/record=b1123148
26. A. Czaplicka, J.A. Holyst, P.M. Sloot, Sci. Rep. **3**, 1223 (2013)
27. A. Czaplicka, J.A. Hołyst, P.M. Sloot, Eur. Phys. J.: Special Topics **222**, 1335 (2013)
28. A. Czaplicka, K. Suchecki, B. Minano, M. Trias, J.A. Hołyst, Phys. Rev. E **89**, 062810 (2014)
29. U. Bastolla, G. Parisi, Physica D **115**, 219 (1998)
30. R. Poblanno-Balp, C. Gershenson, Artif. Life **17**, 331 (2011)
31. K. Suchecki, J.A. Hołyst, Phys. Rev. E **74**, 011122 (2006)
32. K. Suchecki, J.A. Hołyst, Phys. Rev. E **80**, 031110 (2009)
33. P. Csermely, Trends Biochem. Sci. **29**, 331 (2004)
34. P. Csermely, in *Weak Links* (Springer, 2009), pp. 101–116
35. B. Luque, A. Ferrera, Complex Systems **12**, 241 (2000)
36. P. Rämö, S. Kauffman, J. Kesseli, O. Yli-Harja, Physica D **227**, 100 (2007)
37. P. Krawitz, I. Shmulevich, Phys. Rev. Lett. **98**, 158701 (2007)
38. A.S. Ribeiro, S.A. Kauffman, J. Lloyd-Price, B. Samuelsson, J.E. Socolar, Phys. Rev. E **77**, 011901 (2008)
39. J.T. Lizier, S. Pritam, M. Prokopenko, Artif. Life **17**, 293 (2011)
40. D. Snyder, A. Goudarzi, C. Teuscher, Phys. Rev. E **87**, 042808 (2013)
41. M.E. Gáspár, P. Csermely, Briefings Funct. Genomics **11**, 443 (2012)
42. P. Holme, M.E. Newman, Phys. Rev. E **74**, 056108 (2006)
43. A. Clauset, C.R. Shalizi, M.E. Newman, SIAM Rev. **51**, 661 (2009)
44. T.B. Kirkwood, Biometrics **35**, 908 (1979)