Critical line in undirected Kauffman Boolean networks — the role of percolation

2008 J. Phys. A: Math. Theor. 41 224009

(http://iopscience.iop.org/1751-8121/41/22/224009)

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **41** (2008) 224009 (15pp) doi:10.1088/1751-8113/41/22/224009

Critical line in undirected Kauffman Boolean networks — the role of percolation

Piotr Fronczak and Agata Fronczak

Faculty of Physics and Center of Excellence for Complex Systems Research, Warsaw University of Technology, Koszykowa 75, PL-00-662 Warsaw, Poland

E-mail: fronczak@if.pw.edu.pl

Received 3 December 2007, in final form 31 January 2008
Published 21 May 2008
Online at stacks.iop.org/JPhysA/41/224009

Abstract

We show that to describe correctly the position of the critical line in Kauffman random Boolean networks one must take into account percolation phenomena underlying the process of damage spreading. For this reason, since the issue of percolation transition is much simpler in random undirected networks than in the directed ones, we study the Kauffman model in undirected networks. We derive the mean field formula for the critical line in the giant components of these networks, and show that the critical line characterizing the whole network results from the fact that the ordered behavior of small clusters shields the chaotic behavior of the giant component. We also show a possible attitude towards the analytical description of the shielding effect. The theoretical derivations given in this paper very much tally with the numerical simulations done for classical random graphs.

PACS numbers: 89.75.Hc, 89.75.−k, 64.60.Cn, 05.45.−a

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Almost 40 years ago Stuart Kauffman proposed random Boolean networks (RBNs) for modelling gene regulatory networks [1]. Since then, beside its original purpose, the model and its modifications have been applied to many different phenomena like cell differentiation [2], immune response [3], evolution [4], opinion formation [5], neural networks [6], and even quantum gravity problems [7].

The original RBNs were represented by a set of N elements, $\sum_t = \{\sigma_1(t), \sigma_2(t), \ldots, \sigma_N(t)\}$, each element σ_i having two possible states: active (1), or inactive (0). The value of σ_i was controlled by k other elements of the network, i.e.

$$\sigma_i(t+1) = f_i(\sigma_{i_1}(t), \sigma_{i_2}(t), \ldots, \sigma_{i_k}(t)), \tag{1}$$

where k was a fixed parameter. The functions f_i were selected so that they have returned values 1 and 0 with probabilities respectively equal to p and $1 - p$. The parameters k and p have determined the dynamics of the system (Kauffman network), and it has been shown that for a given probability p, there exists the critical number of inputs [13]

$$k_c = \frac{1}{2p(1-p)}, \tag{2}$$

below which all perturbations in the initial state of the system die out (*frozen phase*), and above which a small perturbation in the initial state of the system may propagate across the entire network (*chaotic phase*).

In fact, the behavior of the Kauffman model in the vicinity of the critical line $k_c(p)$ has become a major concern of scientists interested in gene regulatory networks. The main reason for this was the conjecture that living organisms operate in a region between order and complete randomness or chaos (the so-called *edge of chaos*) where both complexity and the rate of evolution are maximized [8–10]. The analogous behavior has been noticed in Kauffman networks, which in the interesting region described by equation (2) show stability, homeostatis, and the ability to cope with minor modifications when mutated. The networks are stable as well as flexible in this region.

Recently, when data from real networks have become available [11, 12], a quantitative comparison of the *edge of chaos* in these datasets and RBN models has brought an encouraging and promising message that even such simple models may quite well mimic characteristics of real systems.

Since, however, one has noticed that real genetic networks exhibit a wide range of connectivities, the recent modifications of the standard RBN take into consideration a distribution of nodes' degrees $P(k)$. It has been shown that if the random topology of the directed network is homogeneous (i.e. all elements of the network are statistically equivalent), then the network topology can be meaningfully characterized by the average in-degree $\langle k\rangle$, and the transition between frozen and chaotic phase occurs for [14]:

$$\langle k\rangle_c = \frac{1}{2p(1-p)}. \tag{3}$$

Several authors [17, 18] have provided a general formula for the edge of chaos in directed networks characterized by the joint degree distribution $P(k, q)$

$$\frac{\langle kq\rangle}{\langle q\rangle} = \frac{1}{2p(1-p)}, \tag{4}$$

where k and q correspond to in- and out-degrees of the same node, respectively. The formula (4) shows that the position of the critical line depends on the correlations between k and q in such networks. It is also easy to show that the previous results (2) and (3) immediately follow from (4) if one assumes the lack of correlations $P(k, q) = P_{\text{in}}(k)P_{\text{out}}(q)$.

Very recently, it has been shown by finite-size scaling methods (FSS) that the critical connectivity $\langle k\rangle_c^{FSS}$ deviates significantly from the value established by equation (3), even for large system sizes [19]. More precisely, one observes that $\langle k\rangle_c^{FSS} < \langle k\rangle_c$. To support this observation the authors recall other studies, for example [20], which suggest that gene regulatory networks appear to be in the ordered regime and reside slightly below the phase transition between order and chaos in contrast to the theory which proposes the critical line to be an evolutionary attractor.

In the present paper, we suggest another explanation for the observed discrepancy. We show (both analytically and numerically) that the discrepancies are due to the percolation phenomena which become important in the region of small values of the parameter $\langle k\rangle$.

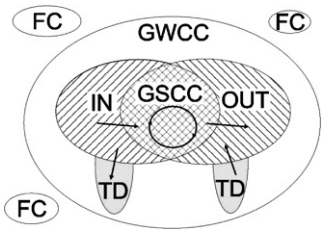

Figure 1. General structure of a directed network above the percolation threshold.

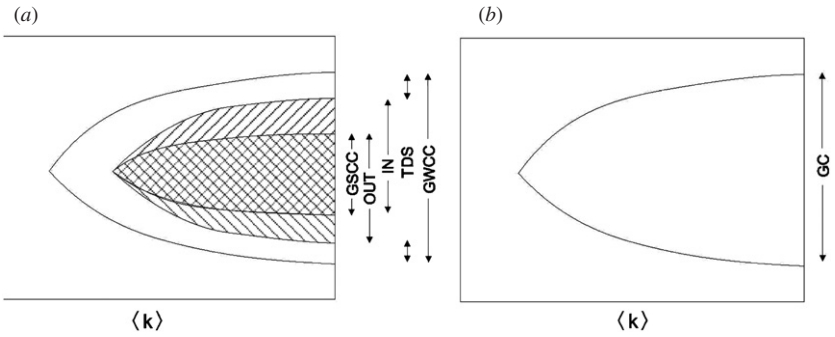

Figure 2. Schematic plot of sizes of network components as a function of average node degree in (a) directed ER graphs and (b) undirected ER graphs.

To understand the complexity of the percolation phenomena in directed graphs let us recall the structure of such a graph [23, 26]. In general, a directed graph consists of a giant weakly connected component (GWCC) and several finite components (FCs). In the GWCC every site is reachable from every other, provided that the links are treated as bidirectional. The GWCC is further divided into a giant strongly connected component (GSCC), consisting of all sites reachable from each other following directed links. All sites reachable from the GSCC are referred to as the giant OUT component, and the sites from which the GSCC is reachable are referred to as the giant IN component. The GSCC is the intersection of the IN and OUT components. All sites in the GWCC, but not in the IN and OUT components, are referred to as the tendrils (TDs) (see figure 1).

The size of all components listed above doubtlessly has an impact on propagation of perturbations in directed RBNs. Moreover, GSCC and GWCC start to form at different values of the parameter $\langle k\rangle$ (see figure 2(a)). Although it has been shown in [23, 26] how to find the relative sizes of the components (for example GWCC appears when $\langle kq\rangle \geqslant \langle q\rangle$), the problem of how to implement these results into the theory of perturbation spreading in RBNs is still far from being solved. As a first step in this direction, and to show the importance of the percolation phenomena on the dynamics of RBN we concentrate on an undirected case of the model. Although the original RBNs have been defined as directed ones, the study of undirected networks significantly reduces the complexity of the problem (see figure 2).

To this end, we organize our paper as follows. In the next section, we present numerical methodology and finite-size scaling of perturbation spreading in RBNs. In section 3 we derive general relation describing position of the critical line in undirected RBNs with arbitrary distribution of connections $P(k)$, in the analogy of the mean-field theory for

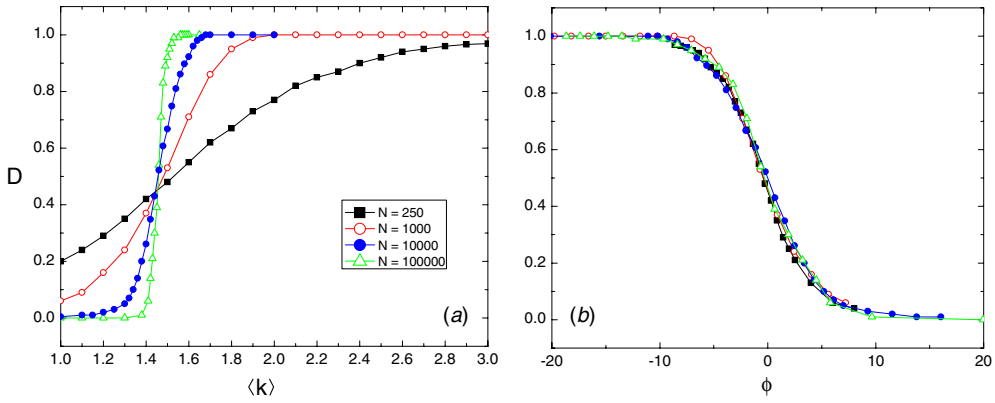

Figure 3. Probability D against (a) control parameter $\langle k\rangle$ and (b) rescaled parameter ϕ for $p = 0.5$.

directed RBNs [13]. Comparing the theory with numerical simulations we show significant deviations between both the approaches. Then an improved treatment including the percolation phenomena is presented in section 4. A summary of our findings is given in section 5.

2. Critical line in undirected random graphs — numerical simulations

In order to find the position of the critical line in RBN one has to examine the sensitivity of its dynamics with regard to initial conditions. In numerical studies such a sensitivity can be analyzed quite simply. One has to start with two initial states $\sum_0 = \{\sigma_1(0), \sigma_2(0), \ldots, \sigma_N(0)\}$ and $\widetilde{\sum}_0 = \{\widetilde{\sigma}_1(0), \widetilde{\sigma}_2(0), \ldots, \widetilde{\sigma}_N(0)\}$, which are identical except for a small number of elements, and observe how the differences between both configurations $\sum_t$ and $\widetilde{\sum}_t$ change in time. If a system is robust then the studied configurations lead to similar long-time behavior, otherwise the differences develop in time. A suitable measure to find the distance between the configurations is the overlap $x(t)$ defined as

$$x(t) = 1 - \frac{1}{N}\sum_{i=1}^{N} |\sigma_i(t) - \widetilde{\sigma}_i(t)|. \tag{5}$$

Note, that in the limit $N \to \infty$, the overlap becomes the probability for two arbitrary but corresponding elements, $\sigma_i(t)$ and $\widetilde{\sigma}_i(t)$, to be equal. Moreover, the stationary long-time limit of the overlap $x = \lim_{t\to\infty} x(t)$ can be treated as the order parameter of the system. If $x = 1$ then the system is insensitive to initial perturbations (frozen phase), while for $x < 1$, the initial perturbations propagate across the entire network (chaotic phase).

For numerical purposes we define probability D that the system is sensitive to perturbations

$$D = \frac{\sum_{x(t=T)<x(0)}^{R} 1}{R}, \tag{6}$$

where R is the number of generated networks, and T the number of system updates. In our simulations we take $RN = 10^6$ and $T = 200$. Figure 3(a) presents a typical example of D-dependence on our control parameter $\langle k\rangle$ for different network sizes. Then, we apply finite-size scaling method [24] to determine how probability D scales with the system size.

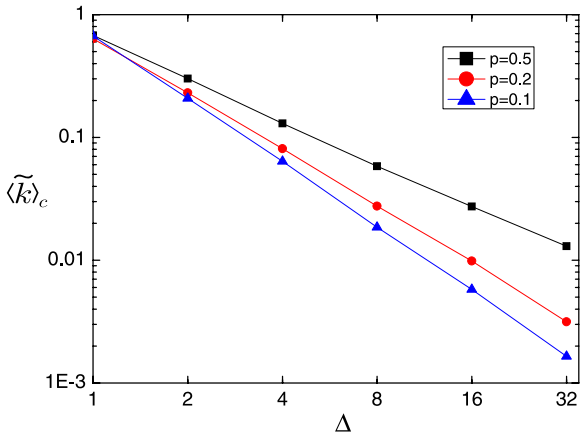

Figure 4. Normalized critical connectivity against the number of perturbed nodes in networks of $N = 1000$ elements. Lines are shown only for a better visibility of the presented dependence.

Around some critical point, we predict that systems of all sizes are indistinguishable except for a change of scale. This suggests

$$D(\langle k\rangle) = f(\phi), \tag{7}$$

where

$$\phi = \left(\frac{\langle k\rangle - \langle k\rangle_c}{\langle k\rangle_c}\right) N^{1/\nu}. \tag{8}$$

In equation (7), f is one of the functions shown in the figure 3(*a*), $\langle k\rangle_c$ is the critical point, and $N^{1/\nu}$ provides the change of scale. Figure 3(*b*) shows how probability D depends on the parameter ϕ with fitted parameters $\langle k\rangle_c = 1.45 \pm 0.04$ and $\nu = 2.2 \pm 0.1$.

The other problem which should be noted here is the observation that $\langle k\rangle_c$ depends on the number of initially perturbed nodes. In the figure 4 we plot the dependence of normalized critical connectivity

$$\langle\widetilde{k}\rangle_c = \frac{\langle k\rangle_c(\Delta) - \langle k\rangle_c}{\langle k\rangle_c}, \tag{9}$$

against the number Δ of initially perturbed nodes in the network of $N = 1000$ elements. For further calculations we choose $\Delta = 0.032N$, since then the error in $\langle k\rangle_c$ is less than the error seen in finite-size scaling.

In figure 5, using the method described above, we show the numerically obtained values of $\langle k\rangle_c$ against the parameter p. For $p = 0.5$ critical connectivity is minimal, i.e. $\langle k\rangle_c = 1.45$. Please note that the size of the giant component for this connectivity is about one half of the whole network. One can expect that a large number of isolated nodes and clusters can significantly affect the perturbation spreading rate in this regime. Moreover, it has been demonstrated [25], that the giant component is correlated in sparse networks. In the following, we will show that a mean field theory which does not take into account these percolation and correlation issues, although correct for large values of $\langle k\rangle$, deviates from numerical results for $\langle k\rangle$ close to 1.

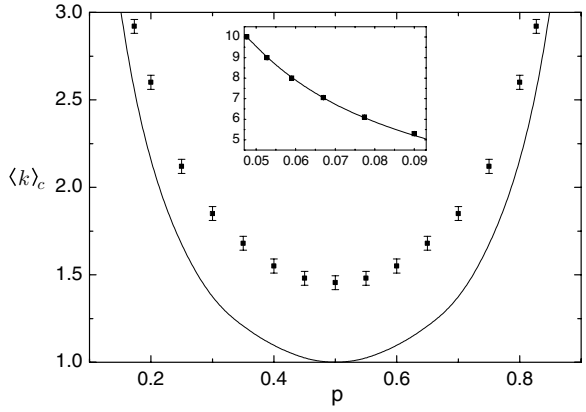

Figure 5. Phase diagram for the undirected RBN model. Points show results obtained by numerical simulations. The line is a solution of equation (19).

3. Damage spreading in undirected Kauffman RBN—a simple approach

In this section, we derive a mean field formula for the critical line characterizing Kauffman Boolean model in undirected and uncorrelated random graphs with arbitrary degree distributions $P(k)$. To this end, we partially reproduce and generalize a simple annealed calculations that have been for the first time carried out by Derrida and Pomeau [13]. The case of random directed networks has been studied by Aldana [15], and also by Lee and Rieger [17].

Thus, let $x_i(k,t)$ corresponds to the probability that a given element i of degree k possesses the same value in both configurations $\sum_t$ and $\widetilde{\sum}_t$ of the considered Boolean network, i.e. $\sigma_i(t) = \widetilde{\sigma}_i(t)$. It occurs either when all the k inputs of $\sigma_i(t)$ are equal to respective inputs of $\widetilde{\sigma}_i(t)$, or when the function f_i, cf (1), ascribed to the node i returns the same value for these two configurations. The first case happens with probability

$$X(q_1, q_2, \ldots, q_k, t-1) = x(q_1, t-1)x(q_2, t-1)\ldots x(q_k, t-1), \tag{10}$$

where $x(q_j, t-1)$ represents probability that in the previous time step $(t-1)$ the jth nearest neighbor of i having degree q_j was in the same state in the two considered configurations. It is also easy to see that the second case arises with probability $p^2 + (1-p)^2$, when at least one of the k inputs of σ_i differs from its counterpart in $\widetilde{\sigma}_i$ giving rise to the same values of σ_i and $\widetilde{\sigma}_i$. Such a situation, in turn, happens with probability equal to $1 - X(q_1, q_2, \ldots, q_k, t-1)$. Taking all above facts together we find that the probability $x_i(k,t)$ that $\sigma_i(t) = \widetilde{\sigma}_i(t)$ is given by

$$\begin{aligned} x_i(k, t+1) &= X(q_1, \ldots, t) + (p^2 + (1-p)^2)X(q_1, \ldots, t) \\ &= 1 - 2p(1-p)\,(1 - X(q_1, q_2, \ldots, q_k, t))\,, \end{aligned} \tag{11}$$

where $q_1, q_2, \ldots, q_k$ stand for degrees of nodes found in the nearest neighborhood of the node i.

Equation (11) describes dynamics of a single node i of degree k. In order to study Boolean dynamics of the whole network one has to average the equation, first over the nearest neighborhood of i, then over the whole network. The first step simply means averaging over

the distribution $P(q_1, q_2, \ldots, q_k/k)$, which describes probability that nearest neighbors of i have degrees respectively equal to $q_1, q_2, \ldots, q_k$

$$x(k, t+1) = 1 - 2p(1-p)\left(1 - \sum_{q_1,\ldots,q_k} X(q_1, \ldots, t) P(q_1, \ldots/k)\right), \quad (12)$$

whereas the second step corresponds to averaging of the last equation over the node degree distribution $P(k)$ characterizing the whole network. Note, that we have omitted the subscript i at $x(k, t+1)$ in equation (12). After averaging, $x(k, t+1)$ refers to the set of nodes having the same degree k.

At the moment, before we proceed with our calculations let us outline structural properties of the studied networks. At the beginning let us remind that the assumed lack of higher-order correlations (e.g. three-point or four-point correlations) means that a given link $\{i, j\}$ does not influence other links of the considered nodes i and j. It translates to the fact that the conditional probability $P(q_1, q_2, \ldots, q_k/k)$ factorizes

$$P(q_1, q_2, \ldots, q_k/k) = P(q_1/k)P(q_2/k)\ldots P(q_k/k), \quad (13)$$

where $P(q_j/k)$ describes probability that a node of degree q_j is the nearest neighbor of a node having degree k. Given the formulas (10), (13) and (15), the equation (12) further simplifies as follows

$$x(k) = 1 - 2p(1-p)\left(1 - \left(\sum_q x(q)P(q/k)\right)^k\right), \quad (14)$$

where, since we are interested in the stationary (i.e. for $t \to \infty$) solutions of this equation, we have omitted dependence on time t.

Now, assuming the lack of two point correlations, i.e.

$$P(q_j/k) = \frac{q_j}{\langle k\rangle} P(q_j), \quad (15)$$

which causes that the nearest neighborhood of each node is the same (in statistical terms), and then multiplying both sides of equation (14) by k, and finally averaging the resulting equation over the node degree distribution $P(k)$, we get the desired mean-field equation which describes stationary states of the Kauffman model defined on undirected and uncorrelated random networks with arbitrary degree distributions

$$\frac{\langle kx\rangle}{\langle k\rangle} = M\left(\frac{\langle kx\rangle}{\langle k\rangle}\right) = 1 - 2p(1-p)\left(1 - \sum_k \left(\frac{\langle kx\rangle}{\langle k\rangle}\right)^k \frac{k}{\langle k\rangle} P(k)\right), \quad (16)$$

where $\langle kx\rangle = \sum_k kx(k)P(k)$.

At the moment, note that the state $\langle kx\rangle = \langle k\rangle$, which in fact corresponds to the set of conditions $x(k) = 1$ for all nodes' degrees k, is always a solution of the last equation, see figure 6. Note also, that this solution may be stable or unstable depending on properties of the considered map $y = M(y)$, where $y = \langle kx\rangle/\langle k\rangle$ (16). In fact, one can show that the solution loses its stability, when another solution $\langle kx\rangle < \langle k\rangle$ of this equation appears. For the first time we see this when

$$\lim_{y\to 1^-} \frac{\mathrm{d}M(y)}{\mathrm{d}y} = 1, \quad (17)$$

where the limit $y \to 1^-$ is equivalent to $\langle kx\rangle \to \langle k\rangle^-$. Substituting (16) into (17) we get the condition for the phase transition between ordered and chaotic behavior of the Kauffman

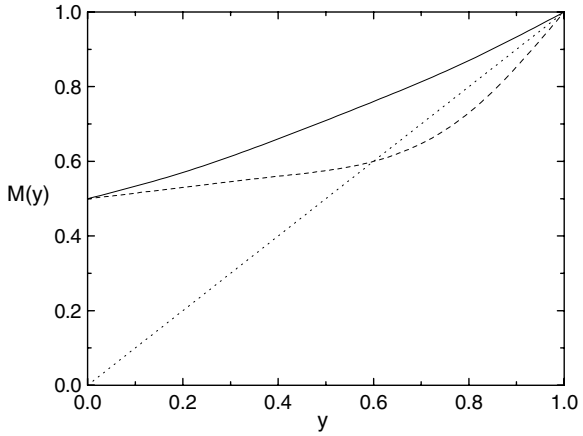

Figure 6. The map $y = M(y)$ considered in the text. The solid line corresponds to the situation when the only stable solution is $\langle kx\rangle = \langle k\rangle$, i.e. $x(k) = 1$ for all values of k. The dashed line shows the case when the second solution $\langle kx\rangle < \langle k\rangle$ appears.

model defined on undirected and uncorrelated random network

$$\frac{\langle k^2\rangle}{\langle k\rangle} = \frac{1}{2p(1-p)}, \tag{18}$$

where $\langle k\rangle$ and $\langle k^2\rangle$ stand for the first and the second moment of the degree distribution $P(k)$, respectively. In the following we analyze briefly the formula for the critical line (18) in classical random graphs. The case of scale-free networks $P(k) \sim k^{-\gamma}$, for which the second moment $\langle k^2\rangle$ of the degree distribution becomes important, has been analyzed in [31].

Thus, since in classical random graphs $\langle k^2\rangle = \langle k\rangle^2 + \langle k\rangle$, the formula (18) simplifies as

$$\langle k\rangle_c = \frac{1}{2p(1-p)} - 1. \tag{19}$$

In figure 5 one can see numerical simulations of the Kauffman Boolean model defined on these graphs as compared with the expression (19). In our previous paper [31] we have suggested that the visible discrepancy between numerical calculations and their theoretical prediction for $\langle k\rangle \to 1$ (i.e. for $p \to 0.5$) may result from the fact that $\langle k\rangle = 1$ corresponds to the percolation threshold in these networks. A simple heuristic argument behind this statement was the following: because the size of the largest component near $\langle k\rangle = 1$ is significantly smaller than the network size (the network is divided into several disconnected components), any perturbation cannot propagate across the entire system, and the frozen phase is easily achieved. It means that the closer percolation threshold $\langle k\rangle = 1$ we are, the more crumbled network (separated pieces of the whole system) we analyze, and the theoretical prediction given by equation (19) works worse and worse. In fact, comparing the general formula $\langle k^2\rangle/\langle k\rangle = 2$ [32] for the percolation threshold in arbitrary undirected and uncorrelated random network with the general expression for the critical line (18), one can show that the arguments exposed in relation to classical random graphs should apply also for the whole class of the considered networks.

In the next section we show how to adjust the approach presented in this section in order to correctly describe properties of the analyzed systems in the whole range of parameters, also in the vicinity of the percolation transition.

4. The effect of percolation phenomena on damage spreading

In the following, in order to address correctly the problem of damage spreading in the vicinity of percolation transition, that has been outlined at the end of the previous section, we use a few important results on percolation phenomena in the considered class of networks. To begin with, we recall these results. As we move to directly (i.e. in the course of numerical simulations) check our derivations in classical random graphs, together with general formulas describing behavior of arbitrary undirected and uncorrelated random networks we also provide the respective formulas for these graphs.

Thus, as we have already mentioned, random graph with a given node degree distribution $P(k)$ does not need to be connected. However, if

$$\frac{\langle k^2\rangle}{\langle k\rangle} > 2, \tag{20}$$

that in classical random graphs translates into

$$\langle k\rangle > 1, \tag{21}$$

the giant component GC emerges gathering a finite fraction of all nodes and links. The size of the giant component S, i.e. the probability that an arbitrary node belongs to GC, is given by the formula

$$S = 1 - G_0(u), \tag{22}$$

where u is the solution of the self-consistency equation

$$u = G_1(u), \tag{23}$$

and $1 - u^2$ is the probability that a link belongs to the giant component. The functions $G_0(u)$ and $G_1(u)$ correspond to generating functions of the node degree distribution $P(k)$, and the conditional distribution $P(q_j/k)$ (15), respectively. Since in classical random graphs $G_0(x) = G_1(x) = \mathrm{e}^{\langle k\rangle(x-1)}$, the formula (22) for these networks significantly simplifies

$$S = 1 - \mathrm{e}^{-\langle k\rangle S}, \tag{24}$$

and the expression for u becomes

$$u = 1 - S. \tag{25}$$

The general results for percolation transition in random undirected and uncorrelated networks outlined in the previous paragraph are already well known. They have been derived by several authors using different theoretical approaches, see e.g. [30, 32]. Recently, however, new interesting results, adding to our knowledge in this subject comprehensively, have been obtained by Białas and Oleś [25]. The authors have shown that the neighboring nodes in the giant connected components are disassortatively correlated. They have also derived analytic formulas for the node degree distribution

$$P^*(k) = P(k)\frac{1-u^k}{S}, \tag{26}$$

and the joint nearest-neighbor degree distribution

$$P^*(k,q) = P(k,q)\left(\frac{1-u^{k+q-2}}{1-u^2}\right) = \frac{kP(k)qP(q)}{\langle k\rangle^2}\left(\frac{1-u^{k+q-2}}{1-u^2}\right), \tag{27}$$

characterizing the giant component. Let us note, that in the limit $u \to 0$, when the giant component covers the whole network $S \to 1$, the both distributions $P^*(k)$ and $P^*(k,q)$ respectively converge to distributions $P(k)$ and $P(k,q)$, which characterize random

uncorrelated networks. The formulas (26) and (27) are crucial for the further developments of this paper, as they show that although in average the considered networks are uncorrelated, in the vicinity of percolation transition their giant components are disassortative (note that we still do not know anything about higher-order correlations in GC s). Now, since we know that this type of correlation makes different spreading-like phenomena more difficult [21], we expect that disassortativity of the giant component is partially responsible for the discrepancy observed in figure 5, with the crumbling of the system as a whole being the second reason. Below, we show that taking these effects into consideration significantly improves theoretical prediction for the critical line in the Kauffman model defined on random uncorrelated networks.

Thus, let us study damage spreading within the giant component of the considered networks. Knowing properties of this cluster, we can start our analysis from equation (14), which is valid for the general class of networks with two-point correlations. The conditional probability $P^*(q/k)$ for the giant component can be calculated from the standard expression [33]

$$P^*(q/k) = \frac{\langle k\rangle^* P^*(k,q)}{kP^*(k)}, \tag{28}$$

where

$$\langle k\rangle^* = \sum_k kP^*(k) = \langle k\rangle \frac{1-u^2}{S}, \tag{29}$$

is the average degree characterizing this component. Inserting (26) and (27) into (28) we get

$$P^*(q/k) = P(q/k)\left(\frac{1-u^{k+q-2}}{1-u^k}\right), \tag{30}$$

where $P(q/k)$ is given by (15). The last formula (30) can be also written in the equivalent form

$$P^*(q/k) = P^*(q)\left(\frac{q}{\langle k\rangle}\frac{S}{(1-u^q)}\right)\left(\frac{1-u^{k+q-2}}{1-u^k}\right), \tag{31}$$

which turns out to be useful for further deductions.

Now, let us apply the equation (14) to the giant component

$$x^*(k) = 1-2p(1-p)\left(1-\left(\sum_q x^*(q)P^*(q/k)\right)^k\right). \tag{32}$$

Due to the complicated form of the conditional distribution $P^*(q/k)$ (30), it is impossible to deduce on possible solutions of the equation (32) in the same way as we have done it for the case of uncorrelated networks. However, substituting (31) into (32) we obtain

$$x^*(k) = 1-2p(1-p)\left(1-\left(\sum_q \kappa(q)w(q,k)P^*(q)\right)^k\right), \tag{33}$$

where

$$\kappa(q) = x^*(q)\frac{S}{1-u^q}\frac{q}{\langle k\rangle}, \tag{34}$$

and

$$w(q,k) = \frac{1-u^{q+k-2}}{1-u^k}. \tag{35}$$

Next, applying a mean field approximation to equation (33)

$$\langle \kappa(q) w(q,k) \rangle^* = \sum_q \kappa(q) w(q,k) P^*(q)$$
$$\simeq \left(\sum_q \kappa(q) P^*(q) \right) \left(\sum_q w(q,k) P^*(q) \right) = \kappa^* w^*(k), \tag{36}$$

we get the simplified equation

$$x^*(k) = 1 - 2p(1-p)(1 - (\kappa^* w^*(k))^k), \tag{37}$$

which after some algebra, consisting of multiplying both sides of this equation by $(k/\langle k \rangle)(S/(1-u^k))$ and then averaging it over $P^*(k)$, further simplifies and becomes equivalent to equation (16)

$$\kappa^* = M^*(\kappa^*) = 1 - 2p(1-p) \left(1 - \sum_k (\kappa^* w^*(k))^k \frac{k}{\langle k \rangle} P(k) \right). \tag{38}$$

The equivalence of the two equations (16) and (38) is visible when $u \to 0$ (i.e. $S \to 1$). Then, the parameter κ^*, see equations (34) and (36), simplifies as follows

$$\kappa^* = \sum_k x^*(k) \frac{k}{\langle k \rangle^*} \frac{S}{(1-u^q)} P^*(k) \tag{39}$$

$$\simeq \sum_k x^*(k) \frac{k}{\langle k \rangle^*} P^*(k) \simeq \frac{\langle xk \rangle}{\langle k \rangle}, \tag{40}$$

where the averages $\langle \ldots \rangle^*$ and $\langle \ldots \rangle$ have their standard meaning (in our calculations $'*'$ always refers to the giant component). This equivalence, also makes possible a similar analytical treatment of equation (38), as the one performed in the reference case of uncorrelated networks, compare equations (14)–(18).

Thus, in order to find condition for the transition between ordered and chaotic phase of the Kauffman model defined in giant components of random uncorrelated networks we have to check when the solution $\kappa^* = 1$ (39), corresponding to $x^*(k) = 1$ for all nodes' degrees, becomes unstable. In fact, it happens when

$$\lim_{\kappa^* \to 1^-} \frac{\mathrm{d}M^*(\kappa^*)}{\mathrm{d}\kappa^*} = 1. \tag{41}$$

From the equation (38) it follows that the condition has a very simple form

$$\frac{\langle k^2 w^*(k)^k \rangle}{\langle k \rangle} = \frac{1}{2p(1-p)}, \tag{42}$$

where

$$w^*(k) = \sum_q w(q,k) P^*(q) = \sum_q \frac{1-u^{q+k-2}}{1-u^k} P^*(q) \tag{43}$$

is defined in equation (36). At the moment, let us note that in the limiting case of $u \to 0$, the parameter $w^*(k) \to 1$, and the formula (42) simplifies to the previous condition (18).

It is easy to check, that in the simplest case of classical random graphs the parameter $w^*(k)$ (43) is given by

$$w^*(k) = \frac{1 - u - u^{k-1} + u^{k+u-1}}{(1-u)(1-u^k)}. \tag{44}$$

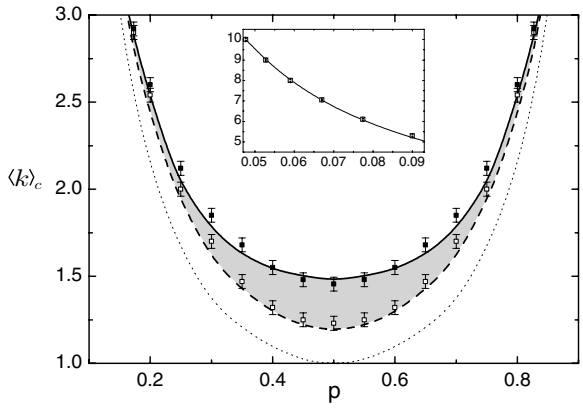

Figure 7. Phase diagram for undirected RBN model in classical random graphs. Dotted line is a solution of basic equation (19). Filled points represent numerical simulations made for the whole network (the same data are shown in figure 5). Open points and dashed line correspond respectively to numerical simulations and analytic prediction of equation (42) for the Kauffman model defined in giant components only. Solid line is the solution of final equation (47). Gray area emphasizes the set of parameters where the chaotic behavior, although present in the giant component, is not yet visible in the whole network.

Inserting (44) into (42), and then solving numerically the resulting equation for $\langle k\rangle$ we obtain a theoretical prediction for the critical line of the Kauffman model in giant components of these graphs. In figure 7 one can see that numerical simulations very much tally with the theoretical prediction of equation (42). Given figure 7, we would also like to take note of two other interesting effects related to the analyzed problem. First, the critical line characterizing the giant component significantly differs from the curve described by the formula (18). It is shifted towards the numerically obtained critical line characterizing the whole network. The observation is in some sense promising, as it partially confirms the main proposition of this paper, which states that the percolation transition is responsible for discrepancies observed in figure 5. The second effect concerns mutual relationship between the behavior of the giant component and the behavior of the whole network. Since one knows that the giant component makes up a macroscopic part of the network (it grows linearly with the network size N, and becomes infinite in the thermodynamic limit $N \to \infty$) one could expect that dynamics of the whole network should reflect behavior of the giant component. Thus, the question is, why the numerically obtained critical line characterizing the whole network differs from the theoretical prediction for the giant component. In other words, why, for the set of parameters marked by the light gray area in figure 7, the chaotic behavior of the giant component is not visible through out the whole network.

To solve the problem stated at the end of the last paragraph, let us briefly recall what the numerical simulations of the Kauffman model consist in. Thus, in numerical studies we check how the initial perturbation of the system $x(0) \equiv 1-\Delta$ (5), where $\Delta \ll 1$, develops over time. In general, when the parameter $x(t = T) < x(0)$ we identify the system as the chaotic one. On the other hand, when $x(0) \leqslant x(t = T) \leqslant 1$ we treat it as being in the ordered phase. In reality, however, due to the fact that in the vicinity of the percolation transition the considered Kauffman networks are strongly heterogenous, they consist of the giant component which is escorted by a number of small tree-like clusters and isolated nodes, the systems should be treated more carefully.

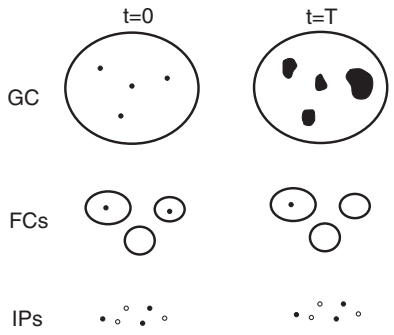

Figure 8. Schematic plot of spreading of perturbation in giant component (GC), finite clusters (FCs) and isolated points (IPs). In GC damage spreads, in FCs it shrinks, while in IPs it does not change.

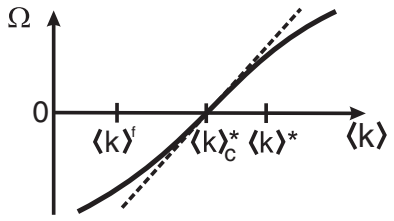

Figure 9. *Chaoticity* Ω (solid line) in the vicinity of the critical point $\langle k\rangle_c^*$. $\langle k\rangle^*$ and $\langle k\rangle^f$ are the average node degree in GC and in FCs respectively. Dashed line presents linear approximation of Ω.

To describe better the situation, let us choose the system parameters from the region that is marked by the light gray color in figure 7. Then, we introduce a quantity Ω, which measures *chaoticity* in the system as a mean damage size caused by a single node perturbation. If $\Omega > (<)0$ then mean damage size grows(shrinks) in time. Condition $\Omega = 0$ will allow us to derive the relation for the critical line in the whole network.

Let us now divide the network into three parts: giant connected component (GC), finite clusters (FCs) and isolated points (IPs). The figure 8 shows schematically how the single node perturbation evolves in time in these three parts of the network. In the studied range of parameters the giant component behaves chaotically, i.e. the mean damage size is larger than initial perturbation and $\Omega_{GC} > 0$. On the other hand, the small density of connections in finite tree-like clusters does not allow perturbation to spread out and $\Omega_{FCs} < 0$. Because the state of isolated nodes does not change in time, then $\Omega_{IPs} = 0$. Now, if one perturbs randomly a set of nodes in the whole network, fraction S of perturbations will be located in GC, fraction $(1 - S)(1 - P(k = 0))$ will be located in FCs, and the rest of them, i.e. $(1 - S)P(k = 0)$ will perturb isolated nodes. Now one can write the condition for transition from the frozen to the chaotic state in the whole network:

$$S\Omega_{GC} + (1 - S)(1 - P(k = 0))\,\Omega_{FCs} = 0, \tag{45}$$

where Ω, $P(k)$ and S depend on $\langle k\rangle$. This equation shows that the ordered behavior of small clusters can shield the chaotic behavior of the giant component. Only when *chaoticity* in GC is sufficiently developed, this shielding effect becomes neglected.

Now, expanding Ω into power series at $\langle k\rangle = \langle k\rangle_c$

$$\Omega = \Omega_0 + \frac{\partial\Omega}{\partial\langle k\rangle}(\langle k\rangle - \langle k\rangle_c), \tag{46}$$

where $\Omega_0 = 0$ in critical point (cf figure 9), one gets the final equation for the critical line:

$$S(\langle k\rangle^* - \langle k\rangle_c^*) = (S-1)(1-P(k=0))(\langle k\rangle^f - \langle k\rangle_c^*), \tag{47}$$

where $\langle k\rangle^f = \langle k\rangle u$ (cf equation (25) in [25]). The numerical solution of this implicit equation is presented in figure 7 as the solid line.

5. Conclusions

This study was done to investigate the properties of undirected KBN model in the vicinity of percolation threshold. We derived a mean field formula for the critical line characterizing the KBN model in undirected and uncorrelated random graphs with arbitrary degree distributions. We have shown that the results of classical mean field theory differ from these obtained by numerical simulations. We have shown also that, to explain the discrepancies one has to take into account the effect of correlations between adjoining nodes in the giant connected component as well as the effect of shielding by finite size clusters. As one can see, the problem is not easy even for undirected networks. As we have shown in figures 1 and 2, a directedness of the network introduces further complications in calculations. Nevertheless, we think that a similar approach can be derived even for that case. We hope that the presented work will encourage others to pursue these topics in the near future.

Acknowledgments

The work was funded in part by the European Commission Project CREEN FP6-2003-NEST-Path-012864 (P.F.), by the Project STOCHDYN (through European Science Foundation and Polish Ministry of Science and Higher Education - grant 15/ESF/2006/03) (A.F.).

References

[1] Kauffman S A 1969 *J. Theor. Biol.* **22** 437
[2] Huang S and Ingber D E 2000 *Exp. Cell Res.* **261** 91
[3] Kauffman S A and Weinberger E D 1989 *J. Theor. Biol.* **141** 211
[4] Bornholdt S and Sneppen K 2000 *Proc. R. Soc. Lond.* B **266** 2281
[5] Lambiotte R, Thurner S and Hanel R 2006 (*Preprint* physics/0612025)
[6] Wang L, Pichler E E and Ross J 1990 *Proc. Natl. Acad. Sci.* **87** 9467
[7] Baillie C F and Johnston D A 1994 *Phys. Lett.* B **326** 51
[8] Kauffman S A 1990 *Physica* D **42** 135
[9] Sole R V and Montoya J M 2001 *Proc. R. Soc. Lond.* B **268** 2039
[10] Stauffer D 1994 *J. Stat. Phys.* **74** 1293
[11] Albert R and Barabasi A L 2002 *Rev. Mod. Phys.* **74** 47
[12] Barabasi A L 2002 *Linked: The New Science of Networks* (Cambridge, MA: Perseus)
[13] Derrida B and Pomeau Y 1986 *Europhys. Lett.* **1** 45
[14] Luque B and Sole R V 1997 *Phys. Rev.* E **55** 257
[15] Aldana M 2003 *Physica* D **185** 45
[16] Aldana M and Cluzel P 2003 *Proc. Natl. Acad. Sci.* **100** 8710
[17] Lee D and Rieger H 2006 (*Preprint* cond-mat/0605730)
[18] Boguna M 2007 private communication
[19] Rohlf T, Gulbahce N and Teuscher C 2007 *Phys. Rev. Lett.* **99** 248701
[20] Ramo P, Kesseli J and Yli-Harja O 2006 *J. Theor. Biol.* **242** 164

[21] Newman M E J 2002 *Phys. Rev. Lett.* **89** 208701
[22] Noh J D 2007 (*Preprint* arXiv:0705.0087)
[23] Newman M E J, Strogatz S H and Watts D J 2001 *Phys. Rev.* E **64** 026118
[24] Barber M N 1983 *Finite-size Scaling* (*Phase Transitions and Critical Phenomena* vol 8) ed C Domb and J L Lebowitz (New York:: Academic) pp 146–268
[25] Bialas P and Oles A 2007 (*Preprint* arXiv:0710.3319)
[26] Dorogovtsev S N, Mendes J F F and Samukhin A N 2001 *Phys. Rev.* E **64** 025101(R)
[27] Schwartz N, Cohen R, ben-Avraham D, Barabasi A-L and Havlin S 2002 *Phys. Rev.* E **66** 015104(R)
[28] Dorogovtsev S N, Goltsev A V and Mendes J F F 2007 (*Preprint* arXiv:0705.0010)
[29] Albert R, Jeong H and Barabasi A-L 2000 *Nature* **406** 378
[30] Cohen R, Erez K, ben-Avraham D and Havlin S 2000 *Phys. Rev. Lett.* **85** 4626
[31] Fronczak P, Fronczak A and Holyst J A 2007 (*Preprint* arXiv:0707.1963)
[32] Molloy M and Reed B 1995 *Ran. Struct. Algor.* **6** 161
[33] Boguna M and Pastor-Satorras R 2003 *Phys. Rev.* E **68** 036112