Random graph models and their applications

Thomas Vallier

Lund University

Faculty of Engineering
Centre for Mathematical Sciences
Mathematics

Mathematics
Centre for Mathematical Sciences
Lund University
Box 118
SE-221 00 Lund
Sweden

http://www.maths.lth.se/

Doctoral Theses in Mathematical Sciences 2007:9
ISSN 1404-0034

ISBN 978-91-628-7378-3
LUTFMA-1028-2007

Printed in Sweden by KFS, Lund 2007

Acknowledgements

Even if only my name appears on the thesis, this work has been made with the help from several people whom I want to thank. I first want to thank my supervisor Tatyana Turova who by her work and perseverance made this thesis become true. I thank my colleagues at the mathematical centre and should have a few words for some of them. If you're not in then please forgive my mistake. Your friendly attitude was a considerable help, Azra, Sara, Anastassia, Aurélia, Henke, Sofia, Sebastian, Jimmy, Joakim, Anders, Svetlana, Klas, Mats, Johan, Johannes, Jonas, Susann, Joerg, Linda. As I promised to some of them, I thank Ica Pesto for feeding me during these four years at the department.

I thank Nadia who kindly reviewed the introduction which was terrible before her advices. I must thank Joerg, Carl-Gustav and many others for technical help. I appreciated the help from Johannes Sivén and Johan Nilsson who accepted me in their office to expose my issues when I was roaming around in the centre searching for a listener. Thank you Linda for your massive help when I needed it most and your friendly smile. I wish to give special thanks to Mona who always arranged all the administrative stuff for me and had all the answers to my non-mathematical problems. I can't write this acknowledgement without including James who manages to accumulate the functions of system engineer, psychologist and friend. He knows how much I'm in debt with him.

I'd like to thank especially two friends. Tom who translated some (a lot) part of this thesis from frenglish to a correct english. I wrote the acknowledgement on my own so that you can see his colossal task. Olivier who was always a great support during these four years and gave an unvaluable help in mathematics and moreover with LaTeX. Those who know how skilled I am with a computer can imagine how patient he was with me. I thank my parents for their support and my kids Léa and Timo who even tried to give some help by drawing balls on my notes.

The final word is for the one who never stopped believing in me and thanks to whom I made it until the end. She simply made everything that was possible to help me without asking anything in return. A lot of the acknowledgements end up with the beloved one. As I'm writing mine, I now understand why. Thank you Aurélia.

Contents

List of papers

This thesis is based on the following four papers:

A. T. Vallier: Robustness of Preferential Attachment under Deletion of Edges, *Stochastic Models*, **23** 265-276 (2007)

B. T. Turova, T. Vallier: Merging Percolation and Random Graphs: Phase transition in Dimension 1. Technical Report. arXiv:math.PR/0609594

C. T. Turova, T. Vallier: Merging Percolation on $\mathbb{Z}^d$ and Classical Random Graphs: Phase Transition. Under submission.

D. T. Vallier: Spread of Activation on Random Graphs. Technical Report.

Introduction

1 Motivation

The theory of random graphs has a wide range of applications in real life networks. It is a useful tool when studying the world wide web, neural networks or social networks. To explain what a graph is, we can take a basic example of a social network like affinities in a class. We model each student as a vertex and if the students are friends then there is an edge between the two vertices assigned to the students to show the connection. When studying such a graph, questions arise like:
Are there any isolated student?
Are they all connected to anyone else through a path of friends or are there disconnected groups?
Who's the most popular student?
Is there a big group of friends that outnumbers other groups?

These questions are part of the study of the structure of the graph. They all seem very simple but become much more complicated when one deals with large networks such as internet where there are billions of web pages linked in a large network. These questions can even become crucial in epidemiology studies, for example if one wants to know how to contain the spread of a virus with a limited stock of vaccine. Who should we vaccinate to stop the contamination?

2 Definition and terminology

Let us introduce some notations to study the structure of the graphs rigorously. We emphasise each definition with an example in real life networks to avoid a heavy accumulation of concepts without explanation of their purpose and illustrate definitions in figure 1.

Definition 1. *A graph G is an ordered pair of disjoint sets $G = (V, L)$ such that L, the set of edges is a subset of the set $V^2 = V \times V$ of unordered pairs of the set of vertices V.*

In the modelling of a network, the vertices often represent the entity and the edges are the connections or acquaintances if one thinks about a social network.

If u and v are two vertices of the graph then we denote the edge between these two vertices by (u, v). If $(u, v) \in L$ (i.e. we have an edge between the vertices u and v) then we say that the vertices u and v are *connected* through the edge (u, v). If a vertex w is

connected to no other vertex, that is

$$w \in V, \forall v \in V, v \neq w, (w,v) \notin L$$

then it is *isolated*.

In modelling a network of acquaintance by a random graph as above, we assume that the fact to know someone is mutual. This assumption is not always true. Another example of a network where the connections are oriented is the graph of the streets of a city. Some streets are two ways which corresponds to an edge and some are only one way. Therefore we must introduce the notion of orientation in a graph.

If the graph is *directed* (i.e. the edges are directed) then the pair $(\overrightarrow{u,v})$ denotes that there is an edge from the vertex u to the vertex v.

Consider that two vertices are connected through an edge if the persons modelled by those vertices are close enough to exchange germs. Someone ill might give the sickness to anyone with which it is connected. This person might in turn give the infection to its contacts and so on. If a virus appears in a very dense place then it's more likely to spread than in a scarcely inhabited one. This example allows us to introduce two important notions: the *connected component* and the *degree* of a vertex.We define the *connected components* as the sets of vertices linked by a chain of edges. An isolated vertex is the smallest possible component. The *largest connected component* is denoted C_1. If you are in a component disconnected from the original place of the infection then you can't be infected.

In a graph, the vertices can be linked to several other vertices. The vertices which are linked to u are called the neighbours of u and their set is denoted $\Gamma(u)$. The number of edges to which u is an endpoint is the *degree* of the vertex u which we denote

$$d(u) = \Big|\{(u,v), v \in V\}\Big|.$$

In a densely populated city you encounter more people and exchange more germs than in an unpopulated countryside. You are connected to more people which translates in terms of graphs as a vertex with a high degree.

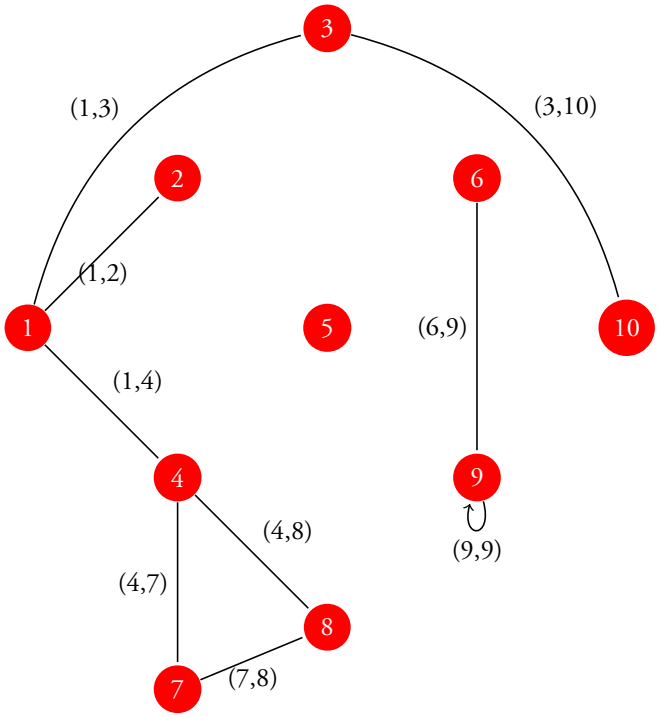
In the case of directed graphs, we make a distinction between edges $(\overrightarrow{v,u})$ and $(\overrightarrow{u,v})$. Thus we make the same distinction by considering the in-degree d_{in} and the out-degree d_{out} of the vertex where the in-degree is the number of incoming edges

$$d_{in}(u) = \Big|\{(\overrightarrow{v,u}), v \in V\}\Big|$$

and the out-degree is the number of out-coming edges. We define the degree of a vertex u as

$$d(u) = d_{in}(u) + d_{out}(u)$$

In the figure 1 we consider an example of a non-directed graph to visualise the notions introduced in this section. Circles represent vertices $V = \{1, 2, ..., 10\}$ and the couple (u,v) where $u, v \in V$ indicates the edge connecting vertices u and v.



The neighbourhood of the vertex 1 is $\Gamma(1) = \{2, 3, 4\}$.
The degree of the vertex 1 is 3, $d(1) = 3$.
The vertex 5 is isolated.
There's a loop on vertex 9.
The vertices $\{1, 2, 3, 4, 7, 8, 10\}$ form a connected component.

Figure 1: Example of graph.

3 The degree sequence

The study of the degree sequence gives indications on the homogeneity of the graph. If the degrees are concentrated around a certain value then we can consider that the vertices are hardly distinguishable. If instead we have a very wide range of degrees then we will have highly connected vertices and not highly connected ones.

3.1 The degree sequence in classical random graph

The results given in this introduction about classical random graphs are taken from the book Random Graphs by Janson, Łuczak and Ruciński (2000) (13) and from the book with the same name by Bollobás (1985) (5).

The classical model of random graph denoted $G_{n,p}$ was introduced by Gilbert (1959) in (11) but it is commonly called Erdös-Rényi random graph since they first set the basis of the probabilistic treatment of $G_{n,p}$ in a series of papers in the 60's (see (9)).

Definition 2. *The model $G_{n,p}$ consists of all graphs with vertex set $V = \{1, ..., n\}$ in which all possible edges are chosen independently and with a probability p where $p \in [0, 1]$.*

Consider a vertex v_i which is connected to any vertex independently from the other connections with probability p. Let k_i denote the degree of the vertex v_i, the probability that the vertex v_i has degree k in the graph $G_{n,p}$ is binomial $(n-1, p)$

$$\mathbb{P}\{k_i = k\} = \binom{n-1}{k} p^k (1-p)^{n-1-k} \tag{3.1}$$

where p^k is the probability that the vertex has k edges,
$(1-p)^{n-1-k}$ is the probability of absence of further edges
and $\binom{n-1}{k}$ is the number of ways of selecting the neighbouring of v_i.

The events $\{k_i = k\}$ and $\{k_j = l\}$ are not independent. It is for example impossible that a vertex has a strictly positive degree while the others have degree 0. The distribution of the number of vertices with degree k doesn't follow straightforward from (3.1).

Let X_k be the number of vertices of degree k in the graph then for a large range of probability p we have the following.

Theorem 3.1 (see (5))**.** *Let $\varepsilon > 0$ be fixed and let $\varepsilon n^{-\frac{3}{2}} \leq p = p(n) \leq 1 - \varepsilon n^{-\frac{3}{2}}$. Let $\lambda_k = n\binom{n-1}{k} p^k (1-p)^{n-1-k}$ then X_k has asymptotically a Poisson distribution with mean λ_k*

$$\mathbb{P}\{X_k = r\} \sim e^{-\lambda_k} \frac{\lambda_k^r}{r!} \tag{3.2}$$

for every fixed r, where $\sim$ means asymptotically distributed.

However, in real life networks we see a totally different type of behaviour. It has been found over the last decade that many different networks such as the citation patterns where each manuscript is a vertex and edges represent citations (i.e. there exists and edge directed from u to v if the manuscript u cites the manuscript v) have a power law distribution of the degrees. In the real life networks, the proportion of vertices with degree k is best described as a power law:

$$\mathbb{P}(k) \sim k^{-\gamma}. \tag{3.3}$$

Redner (1998) (17) showed that $\gamma = 3$ for the network of citations. In the internet where a vertex represents a web page and the edges are links pointing from one page to another, the proportion of vertices with degree k follows a power law with $\gamma = 2,1 \pm 0,1$ (3).

3.2 Preferential attachment model

Recently, many new models of random graphs have been introduced motivated by the power law sequence explained above. Barabási and Albert (1999) (3) introduced the *preferential attachment model* based on two observations:

1. Most real world networks are open and continuously incorporate new vertices to the system. An example is internet where new pages appear on the web everyday.

2. Attachment is not uniform but is preferentially to vertices that already have a large number of connections. Such as in the citation network mentioned above where a new manuscript is more likely to cite a well known and thus much cited paper. As an example, have a look at the bibliography of this thesis.

In this model, at each time step t a new vertex v_t is introduced with m edges linking v_t to the previous vertices with probabilities proportional to their degrees or attractiveness. Note that the same principle has been previously introduced in the little cited paper by Szymański (1987) (21).

Start with G^1, the graph with one vertex and one loop. Given G^{t-1} we form G^t by adding a vertex t together with a single edge $\left(\overrightarrow{t, g(t)}\right)$, where $g(t)$ is a random vertex chosen as follows. Let $D_i(t)$ denote the degree of vertex i at time t.

$$\mathbb{P}\Big(g(t) = i | D_i(t-1) = d_i\Big) = \begin{cases} \frac{d_i}{2t-1}, & \text{if } 1 \leq i \leq t-1\,, \\ \frac{1}{2t-1}, & \text{if } i = t\,. \end{cases}$$

Barabási, Albert and Jeong (1999) (4) showed by heuristic arguments supported by simulations that the degree distribution is proportional to d^{-3}. Rigorous proof of the power law distribution of degrees was given in (7). Here we give the Theorem where the number of edges introduced at each step is 1.

Theorem 3.2 (Bollobás, Riordan, Spencer, Tusnády (2001)). *Let $\sharp^n(d)$ denote the number of vertices with in-degree equal to d (i.e. with total degree $d+1$). Let*

$$\alpha(d) = \frac{4}{(d+1)(d+2)(d+3)}$$

and let $\varepsilon > 0$ be fixed. Then with probability tending to 1 as $n \to \infty$ we have

$$(1-\varepsilon)\alpha(d) \leq \frac{\sharp^n(d)}{n} \leq (1+\varepsilon)\alpha(d)$$

for every d in the range $0 \leq d \leq n^{\frac{1}{15}}$

By construction vertices with a high degree attract more edges and therefore become even more attractive. This phenomenom leads to a graph which accumulates most of the edges on the first vertices. The expected degree at time t of the vertex i introduced at time i is given by

$$\mathbb{E}\big(D_i(t)\big) = \prod_{j=i}^{t} \frac{2j}{2j-1} \sim \sqrt{t/i}.$$

This leads to a graph where the first vertices maintain the structure of the graph. Thus the model is robust against random deletion of vertices and edges but is vulnerable to an attack on the first vertices. This is explained by Bollobás and Riordan (2003) (8). The introduction of the deletion has another purpose but just checking the strength of the network. In the preferential attachment model, vertices and edges are added at each time step but never deleted. However replacement is a natural rule in a network. A social network increases by the introduction of new arrivals but also evolves in a manner that connections can disappear. Using this premise, we consider the graph where an edge is deleted a time Δ after its introduction. We prove in paper **A** that for any fixed time Δ the expected degree is uniformly bounded by a constant $e^{\frac{1}{2}}$. This shows a phase transition at $\Delta = \infty$ and proves that the preferential attachment is not robust against the ageing of edges. Moreover we prove that for any fixed Δ, the degree of any vertex goes to 0 with time.

Similar deletion was studied by Turova in (22), (23), (24) and (25). Moreover, there the life time Δ of any edge is exponentially distributed. We consider in paper **A** a life time of any edge to be a constant just for mathematical tractability. However, it is clear that our model admits generalisations. Allowing some freedom on the deletion of edges would be a closer fit to reality. If for example, the deletion happens randomly on an interval of a fixed length δ centred on Δ, then our results are still valid. Consider now a few examples of real-world networks which topological or dynamical properties are similar to our model.

Biological networks

Our model fits the description of evolution where each vertex is a species and the in-degree

quantifies the influence of the species in nature. Each species rise up to a climax and finally fade away until extinction. The growth phase refers to the period after each great extinction event when there's little or no competition between species until new species evolve better adaptations increasing come. Our model fulfils conditions of an evolution model stated by Newman (1996) in (16). "Constant change is a natural feature of evolution, on a sufficiently large scale in time, there's nothing remotely stable about evolution." It is also noted that "one billion species have inhabited the planet since Cambrian ($\sim$ -500 millions years), only a few million are still living and most species become extinct about 10 millions years after their first appearance." Our model is in good agreement with that observation. It is also assumed (16) that "the number of species the ecosystem can sustain is roughly constant over time" which is verified in our model where the number of activated vertices for $m = 1$ is between $\frac{3}{2}\Delta$ and 2Δ (making the natural assumption that old vertices are connected by only one edge).

Neural networks

We may say that the vertex "dies" as soon as its degree becomes zero, and the entire graph is being renewed and evolves over time. Hence our model is also in good agreement with the models of neural networks (see, Iglesias et al. (2005) (12) for a relevant description and citations). Roughly, it is often assumed that the synaptic connection is lost if it was not activated by an impulse for some time. Also, in a neural network a new edge (impulse) from some vertex is created if this vertex receives enough energy from other vertices. This resembles a preferential attachment: a probability of sending out a new edge is increasing with in-degree.

Social Networks

Jin, Girvan and Newman (2001) (14) consider a model of social networks where acquaintances between pairs of individuals who rarely meet decay over time. They find an upper limit on the number of friendships (links) an individual can maintain. Our model agrees with this result: the links are deleted after a while and any vertex can maintain only a finite number of connections.

Our model can be applied to other networks evolving in time in which competition occurs. The same analysis can be also done for the model of Antal and Krapivsky (2005) (1) where edges carry a weight and the vertices' weight is the sum of the weight of the adjacent edges. Assigning the same weight to every edge, we recover the Preferential Attachment model. If we introduce deletion in this model, one can show that with time the weight and therefore the degree converges to 0 when time goes to infinity. If the weight assigned to newly introduced edges converges with time to a constant different from 0 then the degree is also uniformly bounded in Δ.

In the Growing Network by Copying model of Krapivsky and Redner (2005) (15), every new vertex attaches uniformly to a previously attached one and to those, the "target vertex" already points to. If one introduces deletion then the vertices incorporated before $t - \Delta$ have out-degree 0. The probability that the new vertex is connected to a vertex with out-degree 0 goes to one when t goes to infinity and the expected number of links

goes to Δ. This incites us to add the condition that the target vertex has a strictly positive out-degree (which is not a condition with the inner model since all the vertices have out-degree strictly positive). With this assumption, the total number of links is increasing and is concentrated on the Δ last vertices. Hence, the maximal expected degree increases too in this model.

4 The largest connected component

The size of the largest connected component is a major study in the theory of random graphs. A simple application is to know whether or not a virus might spread to a large part of the population. Since the celebrated paper by Erdös and Rényi (1960) (9) the size of the largest connected component is well known in the classical random graph according to the probability p. We are mainly interested in the range when the size of the largest connected component abruptly *jumps* from a size of order $\log n$ to a positive part of the graph (order n).

4.1 The largest connected component in classical random graph

Consider a graph $G_{n,p}$. When $p = 0$ then with probability 1 there is no edge. The graph is totally disconnected and consists of isolated vertices. Conversely, if $p = 1$ then with a probability 1 any vertex is connected to all other vertices and the graph is fully connected. In the intermediary cases, we have very different structures depending on the value of p. In the study of random graphs, we consider p as a function of the number of vertices $p = p(n)$. When we increase p, the properties of the random graph change as the graph becomes denser in the sense that we have more edges. It is striking to see that the changing are sudden. The probability that a property holds can rapidly change from 0 to 1 as we increase $p(n)$. The range of the probability where this occurs is called the threshold function.

Definition 3. *(see (13)) We define the threshold function $f(n)$ of a property $\mathcal{P}$ by*

$$\mathbb{P}(\mathcal{P}) = \begin{cases} 0 \text{ if } \frac{p(n)}{f(n)} \to 0 \text{ as } n \to \infty \\ 1 \text{ if } \frac{p(n)}{f(n)} \to \infty \text{ as } n \to \infty \end{cases}$$

A monotone increasing property is a property that still holds by addition of new edges.

For every monotone property, the threshold function exists. As an example, the property that the graph is connected is an increasing property since addition of new edges cannot disconnect the graph. The size of the largest connected component can only increase by addition of new edges. Moreover, the addition of new edges can merge together two components into one big component. The increase of p implies the introduction of new edges which are more likely to link large components together than small ones. That

way, the largest connected components increase their size. For $p(n)$ sufficiently large, the largest connected components merge into a giant component. This phase is called phase transition. The threshold function for the property that the largest connected component contains a positive part of the graph which corresponds to the phase transition is $f(n) = \frac{1}{n}$. This result has been proved in a Theorem by Erdös and Rényi (1960)

Theorem 4.1 (Erdös, Rényi (1960))**.** *Let $p(n) = \frac{c}{n}$, where $c > 0$ is a constant.*

- *If $c < 1$ then*
$$\lim_{n\to\infty} \mathbb{P}\Big\{|C_1| \leq \frac{1}{1-c-\log c}\log n\Big\} = 1.$$
- *If $c > 1$, let $\beta(c) \in (0,1)$ be uniquely defined by the equation*
$$\beta(c) + e^{-c\beta(c)} = 1. \tag{4.4}$$
Then for any $\varepsilon > 0$
$$\lim_{n\to\infty} \mathbb{P}\Big\{\Big|\frac{|C_1|}{n} - \beta(c)\Big| > \varepsilon\Big\} = 0.$$

See (13) for a proof of the supercritical case using a branching process argument.

4.2 Infinite component in percolation

The results from this section can be found in the book Percolation (1999) (10) by G. Grimmett. Conversely to the classical random graph, in percolation theory, the graph has a geometric structure. In *bond percolation theory*, the vertices are located on a square lattice and there is an open edge with the nearest neighbours with a probability p. In dimension 1 the vertices are ordered on a line and can have a link with the previous and the next vertices. In dimension d the neighbourhood is composed of the $2d$ vertices surrounding the vertex. To give an example in nature of an environment that can be nicely modelled by percolation, consider a volcanic stone that you immerge into water. Water goes through the holes modelled by open edges. Will the water reach the centre of the stone? We can consider the problem the other way around, start from the centre and want to know with which probability there is a path of open edges to the borders. Considering an infinite lattice the question comes down to ask if with a positive probability there is an infinite connected cluster of open edges. In dimension 1, the critical probability above which there exists an infinite connected cluster is obviously $p_c = 1$ because if $p < 1$ then the probability that the cluster has a size larger than k decreases exponentially fast to 0. This problem is for higher dimension as simple to state as it is complex to study. Actually, only the critical probability in dimension 2 is exactly known $p_c(2) = \frac{1}{2}$ while for higher dimension, we can approximate the value of $p_c(d)$ through algorithm supported by computers. Denote the lattice in dimension d by $\mathbb{Z}^d$, we can restate the previous remark.

Theorem 4.2. *(see (10))*
The critical probability of bond percolation on $\mathbb{Z}^2$ *equals* $\frac{1}{2}$*:*

$$p_c(2) = \frac{1}{2}.$$

If $p < p_c(2) = \frac{1}{2}$ then the lattice is composed by finite open clusters separated by an infinite closed cluster.
For $p = \frac{1}{2}$ then the probability of having an open and a closed edge is equal. The graph is composed of finite open and closed clusters.
If $p > \frac{1}{2}$ then conversely to the first case there is a infinite open cluster. Moreover, this cluster is with probability 1 unique.

Theorem 4.3. *(see (10))*
If $p > p_c$*, then*

$$\mathbb{P}_p(\textit{there exists exactly one infinite open cluster}) = 1.$$

The percolation transition is markedly similar to the phase transition in random graph. They both show a phase transition where the size of the largest connected component suddenly increases from something negligible with respect to the whole graph to a unique giant connected component which contains a positive part of the graph.

However, the models are very different in nature since there is no distance between the vertices in the classical random graph model while the geometry is fundamental in percolation. By not incorporating distance between vertices, the classical model misses properties of real world networks. On the other hand, the restriction to the nearest neighbours in percolation is not in agreement with observations. In a social networks for instance we are more likely to have acquaintance with neighbours or colleagues but we do not restrict our relations to such a narrow selection. A model that would capture the feature of both classical random graph model and percolation model would better model the connections in networks.

4.3 A model merging classical random graph and percolation

The model introduced in the paper **B** incorporates the classical model of random graphs and percolation. It is highly motivated by the model designed to study biological neural networks and introduced by Turova and Villa (2007) in (26) where vertices correspond to neurons which are connected by axons represented by edges. In papers **B** and **C**, we do not concentrate on the spread of the activation in the neural networks but on the structure of the networks and more precisely on the phase transition. The exact formulation of the model is as follows.

We consider a graph on the set of vertices $V_N^d := \{1, \dots, N\}^d$ in $\mathbb{Z}^d$, where the edges between any two different vertices i and j are presented independently with probabilities

$$p_{ij} = \begin{cases} p, & \text{if } |i-j| = 1, \\ c/N^d, & \text{if } |i-j| > 1, \end{cases}$$

where $0 \leq p \leq 1$ and $0 < c < N$ are constants. This graph, call it $G_N^d(p,c)$ is a mixture of percolation model, where each pair of neighbours in $\mathbb{Z}^d$ is connected with probability p, and a random graph model, where each vertex is connected to any other vertex with probability $\frac{c}{|N^d|}$.

This model can be seen as an attempt to generalise classical models into a single model. A generalisation of the classical random graph into an inhomogeneous random graph for which the classical model is a special case has been introduced and studied by Söderberg (2002) in (18), (19) and (20). It was later extended in the manuscript by Bollobás, Janson and Riordan (2007) (6).

We prove in paper **B** that in the case when the dimension of the lattice is 1 there is a phase transition along both parameters c related to the classical random graph model and p for the percolation model. Suppose that $0 \leq p \leq 1$ is fixed then there exists a critical c denoted by

$$c^{cr}(p) = \frac{1-p}{1+p}\,.$$

such that if $c < c^{cr}(p)$, we are in the subcritical case when the largest connected component denoted by $C_1\big(G\big)$ has a size of order $\log N$ with a probability tending to 1 as $N \to \infty$. If $c > c^{cr}(p)$ then we are in the supercritical case when $C_1\big(G\big)$ spans over a positive part of the graph and

$$\frac{|C_1\big(G_N(p,c)\big)|}{N} \xrightarrow{P} \beta$$

as $N \to \infty$, with $\beta = \beta(q,c)$ defined as the maximal solution to

$$\beta = 1 - \frac{1}{\mathbb{E}X}\,\mathbb{E}\left\{Xe^{-cX\beta}\right\}. \tag{4.5}$$

In paper **C**, we extend the results above to any dimension d. Let C denote an open cluster containing the origin of $\mathbb{Z}^d$ in the bond percolation model and $B(N)$ be the box of length N then

$$c^{cr}(p) = \frac{1}{\mathbb{E}|C|}.$$

If $c < c^{cr}(p)$ then with a probability tending to 1 as $n \to \infty$ we have $|C_1\big(G\big)| \leq \alpha \log|B(N)|$ with α given in paper **C**.
If $c \geq c^{cr}(p)$ then

$$\frac{|C_1\big(G_N(p,c)\big)|}{|B(N)|} \xrightarrow{P} \beta$$

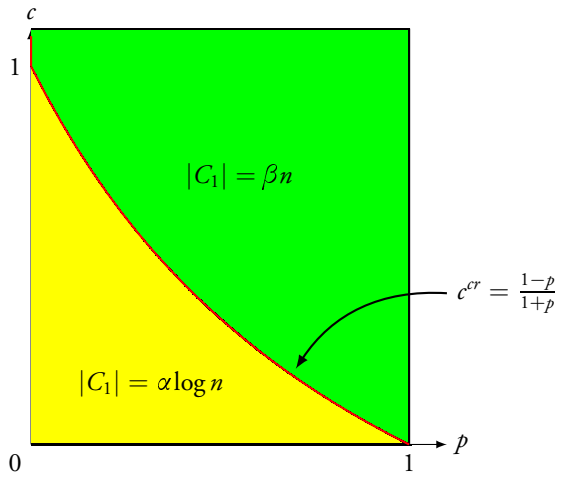


Figure 2: Phase diagram in dimension 1.

as $N \to \infty$, with $\beta = \beta(p, c)$ defined as the maximal solution to

$$\beta = 1 - \mathbb{E}\left\{e^{-c\,\beta\,|C|}\right\}. \tag{4.6}$$

Even if given in terms of c^{cr} for p fixed, it is possible to state the same result for p^{cr} with a fixed c. The phase transition happens along both parameters. The duality of the parameter p and c is manifest in dimension 1(see figure 2), where

$$c^{cr}(p) = \frac{1-p}{1+p} \text{ and } p^{cr}(c) = \frac{1-c}{1+c}$$

Notice that if we choose $p = 0$ then we do not have anymore influence from the percolation model and the model is equivalent to a classical random graph then equations (4.6) and (4.5) become (4.4).

This model is the combination of two homogeneous random graph models in the sense that the probability law of connection is the same for any vertex. It is itself an homogeneous random graph model. However, we use the theory of inhomogeneous random graph developed in (6) to tackle this problem. We consider the clusters formed by percolation (we know that for $p < p_c$ their size is finite) and consider each cluster as a macro-vertex. We build a graph on macro-vertices where each macro-vertex is of type k if the cluster contains k vertices and clusters are connected if there exists at least one edge between two vertices belonging to each macro-vertex. We derive size of the giant connected component in equations (4.5) and (4.6) directly from (6) while the subcritical case requires an entire treatment.

5 Process of activation on random graphs

The last manuscript (paper **D**) of this thesis is dedicated to the spread of activation on a classical random graph. In the paper **B** and **C**, we studied the structure of the graph proposed in (26) to model neural connections. Here we focus on the activation aspect and study the conditions under which the activation starting from a random set of activated vertices $A_n(0)$ at time 0 on a graph of n vertices spreads through the graph. In terms of biological network, it corresponds to an information that spreads in the neural network. It is known that a neuron needs several excitory impulses to become in turn excited. In paper **D**, we consider a simple process of activation where any vertex becomes activated if it has a link with at least two activated vertices. This model resembles a contact process and fits definition of a probabilistic cellular automata given in (2).

Definition 4. *Let A be a finite set of state. A probabilistic cellular automaton on $\mathbb{Z}^d$ is a stochastic process giving rise to a sequence of configurations*

$$\Phi_t : \mathbb{Z}^d \to A$$

where the state $\Phi_t(x)$ of $x \in \mathbb{Z}^d$ at time t is determined randomly with probabilities dependent on the states of the points of the neighbourhood of x at time $t-1$.

In our model, the neighbourhood as stated in the definition 4 is the entire set of vertices which can take two different states, activated or non-activated. A vertex changes from non-activated to activated with a probability depending on the probability of connection p of the vertices and the number of activated vertices.

We show that if the probability of connections is $p(n) = \frac{c}{n}$ where c is some constant and n is the number of vertices then the activation can't spread from a negligible part of the graph to a positive part of the graph. Taking $p(n)$ to be larger, we exhibit two different phases depending on the size of the original set of activated vertices $|A_n(0)|$. With probability tending to 1 as $n \to \infty$ we have the following

- If $|A_n(0)| = o(\frac{1}{np^2(n)})$ then the activation doesn't spread and the limiting set of activation has a size negligible with respect to n
- If for any $\nu > 0$, we have $|A_n(0)| > \frac{1+\nu}{np^2(n)}$ then the activation spreads through the entire graph.

Both the connections and the size of the inner set of activated vertices play a role in the transition.

This phenomenon is very natural. For instance, either you manage to contain virus or the infection expands faster and faster as the number of infected people increases. If the disease spreads to a positive part of the population, it is more likely that almost all the population will be infected leaving a few healthy people. This picture is fully described in the model of paper **D**. We summarise the results in the figure 3. This gives the size of

		$\lvert A_n(0)\rvert$			
		$o\Big(\frac{1}{np^2(n)}\Big)$	$\frac{1-\nu}{np^2(n)}$	$\frac{1+\nu}{np^2(n)}$	$\frac{\alpha(n)}{np^2(n)}$
$p(n)$	$\frac{c}{n}$	$o(n)$	†	†	$\surd$
	$\frac{c}{n} \ll \frac{\omega(n)}{n} \ll \frac{1}{\sqrt{n}}$	$o(n)$	?	$n\big(1-o(1)\big)$	$n\big(1-o(1)\big)$
	$\frac{c}{\sqrt{n}}$	$\surd$	$*n\big(1-o(1)\big)$		$n\big(1-o(1)\big)$
	$\frac{\omega(n)}{\sqrt{n}}$	$\surd$	$\surd$	$\surd$	$n\big(1-o(1)\big)$

Figure 3: The results are given with a probability tending to 1 as $n \to \infty$ except for the case marked with $*$ where the event happens with a positive probability. In the cases marked with † we already have $|A_0(n)| > \varepsilon n$ for some $\varepsilon > 0$. The case marked with $\surd$ have no sense since either $|A_0(n)| = o(1)$ or $|A_0(n)| > n$.

the limiting set of activated vertices $A_n(n)$ with respect to the probability of connections $p(n)$ and the size of the inner set of activated vertices $A_n(0)$.

Notice that in the two cases marked with a †, we can have $|A_n(0)| > n$ depending on the constant c.

5.1 Further research

The properties of the model for $|A_n(0)| = \frac{1-\nu}{np^2(n)}(1 + o(1))$ with $0 \leq \nu < 1$ are not yet known. This implies a lack of knowledge on the type of phase transition. The analysis of the model can be extended to the case when the vertices need $k \geq 2$ connections with already activated vertices to become activated. From that, it would be possible to consider a model closer to the one exposed in (26).

Each vertex is given a random potential $X_v(0) \in [0, 1]$. A vertex is activated if its potential is 1. Start with a set of activated $A_n(0)$, each activated vertex sends a potential ω through the edges to its neighbours. This increases the potential of the targeted vertex

$$X_v(t) = \min\{1, X_v(0) + k\omega\}$$

where k is the number of activated vertices at time $t - 1$ which share an edge with the vertex v. The vertex v in turn becomes activated if its potential reaches 1. The study of such a process would be of great interest in modelisation of neural networks.

Bibliography

[1] Antal T., Krapivsky P. L. (2005), Weight Driven Growing Networks *Phys. Rev E* **71** 026103.

[2] Balister P., Bollobás B. and Kozma R. (2006), Large Deviations for Mean Field Models of Probabilistic Cellular Automata. *Random Structures and Algorithms* **29** no. 3, 399-415.

[3] Barabási A. L., Albert R. (1999), Emergence of Scaling in Random Networks. *Science* **286**, 509-512.

[4] Barabási A. L., Albert R., Jeong H. (1999), Mean-Field Theory for Scale-Free Random Networks. *Physica A* **272** 173-187.

[5] Bollobás B.(1985), *Random Graphs.* Academic Press, London .

[6] Bollobás B., Janson S. and Riordan O. (2007), The phase transition in inhomogeneous random graphs. *Random Structures and Algorithms* **31** 3-122.

[7] Bollobás B., Riordan O., Spencer J., Tusnády G. (2001), The Degree Sequence of a Scale Free Random Graph Process. *Random Structures and Algorithms* **19**, 279-290.

[8] Bollobás B., Riordan O. (2003), Robustness and Vulnerability of Scale-Free Random Graphs. *Internet Mathematics* **I 1**, 1-35

[9] Erdös P. and Rényi A. (1960), On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.* **5** 17-61.

[10] Grimmett G. (1999), *Percolation* Springer-Verlag, Berlin,.

[11] Gilbert E. N. (1959), Random Graphs. *Ann. Math. Statist.* **30** 1141-1144.

[12] Iglesias J., Eriksson J., Grize F., Tomassini M., Villa A. (2005), Dynamics of Pruning in Simulated Large-Scale Spiking Neural Networks. *BioSystems* **79** 11-20.

[13] Janson S., Łuczak T. and Ruciński A. (2000), *Random Graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York,.

[14] Jin E. M., Girvan M., Newman M. E. J. (2001), The Structure of Growing Social Networks. *Phys. Rev E* **64**, 046132.

[15] Krapivsky P. L., Redner S. (2005), Network Growth by Copying *Phys. Rev E* **71** 036118.

[16] Newman M. E. J. (1996), Self Organized Criticality, Evolution and the Fossil Extinction Record, *Proceedings of the Royal Society of London* **B.263** 1605-1610.

[17] Redner S. (1998), How Popular is Your Paper? An Empirical Study of the Citation Distribution. *European Physical Journal* **B 4**, 131-134.

[18] Söderberg B. (2002), A General Formalism for Inhomogeneous Random Graphs, *Phys. Rev E* (3) **66**, 6, 066121, 6.

[19] Söderberg B. (2002), Random Graphs with Hidden Color *Phys. Rev E* **68**, 2, 026107, 12.

[20] Söderberg B. (2002), Properties of Random Graphs with Hidden Color. *Phys. Rev E* **68**, 2, 026107, 12.

[21] Szymański J. (1987), On a Nonuniform Random Recursive Tree. *Annals of Discr. Math.* **33** 297-306.

[22] Turova T. S. (2002), Dynamical Random Graphs with Memory. *Phys. Rev E* **65** 066102. (see also erratum *Phys. Rev E* **70** 059902 (2004).)

[23] Turova T. S. (2003), Long Paths and Cycles in Dynamical Random Graphs. *J. Statist. Phys.* **110**, 1-2, 385-417.

[24] Turova T. S. (2003), Phase Transitions in Dynamical Random Graphs. *J. Statist. Phys.* **123**, 5, 1007-1032.

[25] Turova T. S. (2007), Continuity of the Percolation Threshold in Randomly Grown Graphs. *Electronic Journal of Probability* **12** 1036-1047.

[26] Turova T. S. and Villa A. (2007), On a phase diagram for random neural networks with embedded spike timing dependent plasticity. *BioSystems* Volume **89**, Issues 1-3, 280-286.

A

Chapter A

Robustness of Preferential Attachment under Deletion of Edges

T. Vallier

Centre for Mathematical Sciences, Lund University, Sweden

Abstract

We study the effect of deletion of old edges in the preferential attachment model introduced by Barabási and Albert (1). We consider a model where every edge is deleted after a time Δ. The resulting graph has only Δ edges and with a high probability $(1+c)\Delta$ nodes for some positive c. However, its structure doesn't resemble the structure of the former model even for large Δ. In particular, we prove that the expected degrees of the resulting graph are uniformly bounded by a constant which does not depend on Δ. We discuss applications of our model for the evolution of networks where competition occurs.

Key words: Preferential attachment, degree sequence

1 Introduction

Recently many new random graph models have been introduced, inspired by certain features observed in large-scale real-world networks such as the world wide web, interactions between proteins, genetic networks or social and business networks (see, e.g., (2), (6) for a survey or (10)). The main observation is that in many real-world networks, the fraction $\mathbb{P}(d)$ of vertices with degree d is proportional to $d^{-\gamma}$ where γ is a constant independent of the size of the network. The study of these so-called scale-free random graphs is highly motivated by the model of preferential attachment introduced in (1). Barabási and Albert showed by approximate arguments supported by simulation that this model leads to a graph with degree distribution proportional to d^{-3}. Rigorous proof of the power law distribution of degrees was given in (3). In this model, at each time step t a new vertex v_t is introduced with m edges linking v_t to the previous vertices with probabilities proportional to their degrees or attractiveness. Note that the same principle has been previously introduced in the little cited paper by Szymański (15). This concept models real networks where the attractive components are more likely to receive new connections (as frequently visited sites are likely to get links from new sites). This process was rigorously redefined in (5).

Start with G^1, the graph with one vertex and one loop. Given G^{t-1} we form G^t by adding a vertex t together with a single edge $\big(t, g(t)\big)$, where $g(t)$ is a random vertex chosen as follows. Let $D_i(t)$ denote the degree of vertex i at time t. The degree is the number of edges incident to the vertex.

$$\mathbf{P}\Big(g(t) = i | D_i(t-1) = d_i\Big) = \begin{cases} \frac{d_i}{2t-1}, & \text{if } 1 \leq i \leq t-1\ , \\ \frac{1}{2t-1}, & \text{if } i = t\ . \end{cases}$$

Thus closed loops are allowed to occur when new vertices are nucleated, although in such cases the new sites are disconnected from the existing graph. Apart from short loops resulting from the connection of a site to itself, all connected components of the graph are essentially tree-like. However if we introduce several connections (m) from a single vertex then the graph features will differ and cycles may occur but the essential property of the graph will remain i.e. the degree distribution follows a power law (3).

In the preferential attachment model, vertices and edges are added at each time step but never deleted. However an evolving graph model incorporating deletion may model the evolution of the web graph more accurately. In this view, Cooper, Frieze and Vera introduced random deletion of vertices and edges to the process (7) to model networks such as P2P. Their model yields $\gamma \in]2; \infty[$ according to the parameters of accumulation and deletion.

The first paper on the deletion of vertices in the preferential attachment model is (4). The authors raised the following question: "How robust and vulnerable are scale free graphs ?" The robustness of a graph is defined in (4) as the stability of its properties under a *random* deletion of vertices. The vulnerability is defined as the stability of the properties

of a graph but under a *malicious attack*, i.e., when a specific part of the graph is deleted. It was proved in (4) that a scale-free random graph is more robust than a classical random graph but is also more vulnerable. Due to the definition of the preferential attachment process, vertices with a high degree attract more edges and therefore become even more attractive. This phenomenom leads to a graph which accumulates most of the edges on the first vertices. This observation incites the authors of (4) to delete the first (or the oldest) vertices in order to change the structure of the graph.

The introduction of deletion can allow the generalization of a model. Slater et al. introduced death of vertices with a rate either constant or proportionnal to their degree on a continuous time model of stochastically evolving networks (14).

The growth-deletion model introduced in (8) incorporates deletion of both vertices and edges which are chosen uniformly at random. In another model of duplication-deletion (9) describing genes, vertices are again deleted uniformly at random.

In this paper, we introduce another growth-deletion model where the choice of deletion is not made uniformly at random conversely to (8) and (9). Note that need for such models is discussed in (6). Instead of considering deletion of the vertices, we will consider the effect of deletion of the first edges. More precisely, we delete any edge at time Δ after its introduction in the graph. Thus each edge is given a lifetime of Δ. The choice of this attack is natural if one thinks about social, biological or in particular neural networks where the connections (impulses) are temporal in their nature (see, e.g., (11) for a relevant description and citations).

One typical property of the preferential attachment model is the accumulation of most edges on the first vertices. The deletion introduced in our model thwarts this effect. We prove that for any fixed Δ, the degree of any vertex will reach zero with time. Furthermore, when the degree of a vertex becomes zero it remains at zero forever (Theorem 3.1).

Our mechanism of deletion of edges is similar to the one introduced in (16). It was shown in (16) that if old connections are abandoned, it is still possible to keep some macro-properties of the original graph (uniformly grown in the case of (16)) almost unchanged.

Here, we observe totally different behaviour. The main result of our study shows that the expected degree of any vertex is uniformly bounded in Δ. To be more precise, no matter how large one chooses Δ the expectation of the degree in the resulting graph is bounded by a constant $me^{\frac{1}{2}}$ (Corollary 3.1). On the other hand, if $\Delta = \infty$ we recover the model of preferential attachment where the expected degree is unbounded. This yields a sharp transition at $\Delta = \infty$ proving that the preferential attachment model is not robust against the ageing of edges. All results are still valid if instead of considering the lifetime of any edge to be constant, we allow the deletion to happen randomly on an interval of fixed length centered on Δ. Here, unlike in the previous study (16), abandoning old connections leads to a loss of original properties. But as a result of the introduced dynamics our graph becomes more homogeneous and in turn less vulnerable than the

original preferential attachment model. Examples of possible applications of our model are given in the conclusion.

2 Model

We denote by $G^t = (V^t, L^t)$ the graph at time $t = \{1, 2, ...\}$ where $V^t = \{1, ..., t\}$ is the set of vertices and L^t is the set of edges. Set G^1 to be the graph with one vertex and one loop. Fix $\Delta > 0$ to be the life-time of any edge in the graph. This means that an edge introduced at time t is no longer present in the graph at time $s \geq t + \Delta$. Then as long as $t \leq \Delta$, the graph evolves as in the preferential attachment model (5). Namely, given G^{t-1} we form G^t by adding a vertex t together with a single edge $\big(t, g(t)\big)$, where $g(t)$ is a random vertex chosen as follows.

Let $D_i^{\Delta}(t)$ denote the degree of vertex i at time t. The degree is the number of edges incident to the vertex.

$$\mathbf{P}\Big(g(t) = i | D_i^{\Delta}(t-1) = d_i\Big) = \begin{cases} \frac{d_i}{2t-1}, & \text{if } 1 \leq i \leq t-1\ , \\ \frac{1}{2t-1}, & \text{if } i = t\ . \end{cases}$$

Hence $L^t = L^{t-1} \cup \{\big(t, g(t)\big)\}$.

The graph first accumulates vertices and edges (the growth phase) until it reaches its maximum capacity of Δ edges. Then the oldest edge is deleted to make way for the new one.

When $t > \Delta$, we define

$$L^t = L^{t-1} \cup \{(t, g(t))\} \setminus \{(t-\Delta, g(t-\Delta))\},$$

i.e.,

$$L^t = \{(s, g(s)), t - \Delta < s \leq t\}.$$

Clearly, if $\Delta = +\infty$, we recover the model of preferential attachment (1). From this point on, we will denote D_i^{∞} the degree of the vertex i in this case.

For all $t > \Delta$ we have $|L^t| = \Delta$ since every time one edge is introduced one edge is deleted. This implies that for all $t > \Delta$,

$$\sum_{i=1}^{t} D_i^{\Delta}(t) = 2\Delta,$$

and the transition probabilities are

$$\mathbf{P}\big(g(t) = i | D_i^{\Delta}(t-1) = d_i\big) = \begin{cases} \frac{d_i}{2\Delta+1}, & \text{if } i < t, \\ \frac{1}{2\Delta+1}, & \text{if } i = t. \end{cases}$$

We can consider an extension of the previous model where each new vertex is introduced together with m edges (i.e it has out-degree m). Each new edge attaches to previous vertices independently to the other edges introduced at the same time. Thus multiple edges are allowed and for each edge, the probability law is

$$\mathbf{P}\big(g(t)=i|D_i^{\Delta}(t-1)=d_i\big)=\begin{cases}\frac{d_i}{(2\Delta+1)m}, & \text{if } i<t,\\ \frac{1}{(2\Delta+1)m}, & \text{if } i=t.\end{cases}$$

3 Results

When $\Delta=\infty$ (model of preferential attachment (3)) the vertices with the highest degree are first. The deletion of edges in our model leads to the opposite situation. Namely, with probability one, every vertex introduced in the graph will have degree 0 after some time as it is stated in Theorem 3.1.

Theorem 3.1. *For any out-degree $m\geq 1$ and any $i\geq 1$*

$$\lim_{t\to\infty}\mathbf{P}\{D_i^{\Delta}(t)=0\}=1.$$

Recall that when $\Delta=\infty$, the expected degree has the following formula (3) for $m=1$

$$\mathbf{E}\big(D_i^{\infty}(t)\big)=\prod_{j=i}^{t}\frac{2j}{2j-1}\sim\sqrt{t/i}, \tag{3.1}$$

i.e. the expectation is unbounded in t.

In our case the expected degree of a vertex is no longer a monotone increasing function of time. In Theorem 3.2 below we give the time when the expected degree of a fixed vertex reaches its maximum.

Theorem 3.2. *Let $m\geq 1$. For any $i\geq 1$*

$$\max_{t\geq 1}\mathbf{E}\big(D_i^{\Delta}(t)\big)=\mathbf{E}\big(D_i^{\Delta}(i+\Delta-1)\big), \tag{3.2}$$

and for any $t\geq\Delta$

$$\max_{1\leq i\leq t}\mathbf{E}\big(D_i^{\Delta}(t)\big)=\mathbf{E}\big(D_{t-\Delta+1}^{\Delta}(t)\big). \tag{3.3}$$

Next corollary shows that the expectations of the degrees in our graph are uniformly bounded unlike in the corresponding model without deletion (3).

Corollary 3.1. *Let $m\geq 1$. For any $t\geq 2\Delta$*

$$\max_{1\leq i\leq t}\mathbf{E}\big(D_i^{\Delta}(t)\big)=m\big(1+\frac{1}{2\Delta+1}\big)^{\Delta}, \tag{3.4}$$

and

$$\lim_{\Delta\to\infty}\lim_{t\to\infty}\max_{1\le i\le t}\mathbf{E}\big(D_i^{\Delta}(t)\big)=me^{\frac{1}{2}}. \tag{3.5}$$

One should remark that in the model of preferential attachment,

$$\lim_{t\to\infty}\lim_{\Delta\to\infty}\max_{1\le i\le t}\mathbf{E}\big(D_i^{\Delta}(t)\big)=\lim_{t\to\infty}\max_{1\le i\le t}\mathbf{E}\big(D_i^{\infty}(t)\big)=\infty. \tag{3.6}$$

Relations (3.1), (3.5) and (3.6) show that the transition of the expected degree between the model with ephemeral edges and the model with infinite-lived connection is not continuous.

Remark 3.1. *Using the bound* (3.5) *in Corollary 3.1 and the fact that for $i>\Delta$, $t>i+\Delta$, we have $\mathbf{E}\big(D_i(t+k\Delta)\big)\le\left(\frac{1}{2}\right)^k\mathbf{E}\big(D_i(t)\big)$, one can prove that the variance is also bounded independently of Δ.*

Remark 3.2. *It is clear that our model admits generalizations. Allowing some freedom on the deletion of edges would fit reality even more. If for example the deletion happens randomly on an interval of a fixed length δ centered on Δ then our results are still valid.*

4 Proofs

4.1 Proof of Theorem 3.1

Without loss of generality, here we consider the case $m=1$.

Let us fix a vertex i and consider $D_i^{\Delta}(t),\ t\ge i$. We introduce a Markov chain $\{\mathbf{X}(t)\}_{t\ge i}$ with

$$\mathbf{X}(t)=(X_0(t),X_1(t),...,X_{\Delta}(t))\in\{0,1\}^{\Delta+1}.$$

$X_0(t)$ is deterministic

$$X_0(t)=\begin{cases}1 & \text{if } i\le t<i+\Delta,\\ 0 & \text{if } t\ge i+\Delta,\end{cases}$$

and

$$X_k(t)=\begin{cases}0 & \text{if } (i,t-k+1)\notin L^t,\\ 1 & \text{if } (i,t-k+1)\in L^t.\end{cases}$$

$X_k(t),\ 1\le k\le\Delta$ is the number of edges introduced at time $t-k+1$ incident with i. According to the definition of our model at each time $t\ge i$ a new vertex is added together with an edge $\big(t,g(t)\big)$. If $g(t)=i$ then $X_1(t)=1$ and $X_1(t)=0$ otherwise.

At time $t+1$, the value $X_1(t)$ is translated to the next coordinate : $X_2(t+1) = X_1(t)$ and $X_{k+1}(t+1) = X_k(t),\ 1 \le k \le \Delta - 1$. Then

$$\sum_{k=0}^{\Delta} X_k(t) = D_i^{\Delta}(t), \text{ for all } t \ge i.$$

Set

$$\mathbf{X}(i) = \begin{cases} (1,0,0,...,0) \text{ with probability } 1 - \frac{1}{2\Delta+1}, \\ (1,1,0,...,0) \text{ with probability } \frac{1}{2\Delta+1}. \end{cases}$$

The endpoint $g(t)$ of the edge $\big(t, g(t)\big)$ introduced at time t is chosen according to the transition probability

$$\mathbf{P}\Big(g(t) = i | D_i^{\Delta}(t-1) = d_i\Big) = \frac{d_i}{\min\{2\Delta+1, 2t-1\}}$$

$$= \mathbf{P}\Big(X_1(t) = 1 | \mathbf{X}(t-1) = \overline{\mathbf{X}}\Big) = \frac{\sum_{k=0}^{\Delta} \overline{X}_k}{\min\{2\Delta+1, 2t-1\}}. \tag{4.7}$$

To prove the statement of Theorem 3.1, we need to show that the vector (0,...,0) is absorbing and reachable from every state. Indeed, according to the transition probability (4.7) we have

$$\mathbf{P}(\mathbf{X}(t+1) = (0,...,0) | \mathbf{X}(t) = (0,...,0)) = 1,$$

i.e. $(0,...,0)$ is an absorbing state.

Now, we show that $(0,...,0)$ is reachable from any state. For any $\overline{\mathbf{X}} \in \{0,1\}^{\Delta+1}$, there exists with a positive probability a sequence of states from $\overline{\mathbf{X}}$ to (0,...,0).

Assume that $\mathbf{X}(t) = \overline{\mathbf{X}} \in \{0,1\}^{\Delta+1}$ and at each time s, $t \le s \le t+\Delta$ enforce $X_1(s) = 0$. After Δ steps, it reaches (0,...,0).

$$\mathbf{P}(\mathbf{X}(t+\Delta) = (0,...,0) | \mathbf{X}(t) = \overline{\mathbf{X}}) \ge$$

$$\frac{\prod_{i=-1}^{\Delta-1}(\Delta+i)}{(2\Delta+1)^{\Delta}} = \frac{2\Delta!}{(\Delta-1)!(2\Delta-1)(2\Delta+1)^{\Delta}} \ge \frac{1}{2^{\Delta}}.$$

This implies that the limit distribution is concentrated on the point $(0,...,0)$,

$$\lim_{t\to\infty} \mathbf{P}\big(X(t) = (0,...,0)\big) = 1.$$

Hence $\lim_{t\to\infty} \mathbf{P}\{D_i^{\Delta}(t) = 0\} = 1$ as stated in Theorem 3.1.

4.2 Recursive formula for the expectation

From now, we fix $1 \leq m$ and $0 < \Delta < \infty$ arbitrarily and write $D_i^{\Delta}(t) = D_i(t)$.

In order to prove Theorem 3.2 and Corollary 3.1, we first need the following proposition.

Proposition 4.1. *For all $i > \Delta > 0$, we have the following recursive formulas:*

$$\mathbf{E}\big(D_i(t)\big) = \begin{cases} 0, & t < i, \\ m + \frac{m}{2\Delta+1}, & t = i, \\ (1 + \frac{1}{2\Delta+1})\mathbf{E}\big(D_i(t-1)\big), & i+1 \leq t \leq i+\Delta-1, \\ (1 + \frac{1}{2\Delta+1})\mathbf{E}\big(D_i(t-1)\big) - (m + \frac{m}{2\Delta+1}), & t = i+\Delta, \\ (1 + \frac{1}{2\Delta+1})\mathbf{E}\big(D_i(t-1)\big) - \frac{1}{2\Delta+1}\mathbf{E}\big(D_i(t-\Delta-1)\big), & t > i+\Delta. \end{cases}$$

Proof of proposition 4.1. Let $t \geq i > \Delta$. Given $D_i(s) = d_i(s)$, $t - \Delta \leq s < t$ we have

$$D_i(i) = m + mI_{\{g(t)=i\}},$$

and

$$D_i(t) = D_i(t-1) + mI_{\{g(t)=i\}} - \begin{cases} 0, & \text{if } i < t < i+\Delta, \\ 1 + mI_{\{g(t-\Delta)=i\}}, & \text{if } t = i+\Delta, \\ mI_{\{g(t-\Delta)=i\}}, & \text{if } t > i+\Delta. \end{cases}$$

We prove here only the last case when $t > i + \Delta$. The rest can be found in the same way.

$$\begin{aligned} &\mathbf{E}\big(D_i(t)|D_i(t-1), D_i(t-\Delta-1)\big) \\ &= D_i(t-1) + \mathbf{E}\big(I_{\{g(t)=i\}}|D_i(t-1)\big) - \mathbf{E}\big(I_{\{g(t-\Delta)=i\}}|D_i(t-\Delta-1)\big) \\ &= (1 + \frac{1}{2\Delta+1})D_i(t-1) - \frac{D_i(t-\Delta-1)}{2\Delta+1}. \end{aligned}$$

Taking the expectation of both sides we derive

$$\mathbf{E}\big(D_i(t)\big) = (1 + \frac{1}{2\Delta+1})\mathbf{E}\big(D_i(t-1)\big) - \frac{1}{2\Delta+1}\mathbf{E}\big(D_i(t-\Delta-1)\big),$$

as required.

□

4.3 Maximum of expectation

Proof of theorem 3.2 and corollary 3.1.

Consider $\mathbf{E}\big(D_i(i+n)\big)$. We may refer to $n \geq 0$ as the age of the vertex i.

Then if $n \leq \Delta - 1$ we have

$$\mathbf{E}\big(D_i(i+n)\big) = m(1+\frac{1}{2\Delta+1})^{n+1}.$$

The right hand side is obviously increasing in n and reaches its maximum value for $n = \Delta - 1$.

To prove that $m\big(1+\frac{1}{2\Delta+1}\big)^{\Delta}$ is the maximum of the expected degree we need the following lemma.

Lemma 4.1. *Set $k \geq 1$. If $\mathbf{E}\big(D_i(i+n+1)\big) \leq \mathbf{E}\big(D_i(i+n)\big)$ for $k \leq n \leq k+\Delta$ then $\mathbf{E}\big(D_i(i+n+1)\big) \leq \mathbf{E}\big(D_i(i+n)\big)$ for $n \geq k$.*

Thanks to this lemma, we will only need to prove that the sequence $\mathbf{E}\big(D_i(i+n)\big)$ is decreasing for $\Delta - 1 \leq n \leq 2\Delta - 1$ to prove that it's decreasing for $n \geq \Delta - 1$.

Proof of lemma 4.1.

For $k \geq \Delta + 1$, the expectation follows the recursive formula

$\mathbf{E}\big(D_i(i+k+1)\big) = \mathbf{E}\big(D_i(i+k)\big) + \frac{1}{2\Delta+1}\Big(\mathbf{E}\big(D_i(i+k)\big) - \mathbf{E}\big(D_i(i+k-\Delta)\big)\Big)$.

We deduce that $\mathbf{E}\big(D_i(i+k)\big) \leq \mathbf{E}\big(D_i(i+k-\Delta)\big)$ implies $\mathbf{E}\big(D_i(i+k)\big) \leq \mathbf{E}\big(D_i(i+k)\big)$.

Finally, if

$$\mathbf{E}\big(D_i(i+k-\Delta)\big) \geq \mathbf{E}\big(D_i(i+k-\Delta+1)\big) \geq \ldots \geq \mathbf{E}\big(D_i(i+k)\big),$$

then

$$\mathbf{E}\big(D_i(i+n)\big) \geq \mathbf{E}\big(D_i(i+n+1)\big) \text{ for any } n \geq k$$

□

Let $\Delta \leq n \leq 2\Delta - 1$. We are going to show that the expectation of the degree is monotone decreasing. One can find by induction that

$$\begin{aligned}\mathbf{E}\big(D_i(i+n)\big) &= m(1+\frac{1}{2\Delta+1})^{n+1}\big(1-\frac{1}{2\Delta+1}(1+\frac{1}{2\Delta+1})^{-\Delta}(n+\Delta+2)\big)\\ &= me^{n+1\ ln(1+\frac{1}{2\Delta+1})}\big(1-C(n+\Delta+2)\big),\end{aligned}$$

with $C = \frac{1}{2\Delta+1}(1+\frac{1}{2\Delta+1})^{-\Delta}$.

Let $f(x)$, $\Delta \leq x \leq 2\Delta - 1$ be the extension to $\mathbb{R}$ of $\mathbf{E}\big(D_i(i+n)\big)$, $\Delta \leq n \leq 2\Delta - 1$.

We differentiate this function and prove that for every x, it has the same sign.

$$f'(x) = m(1+\frac{1}{2\Delta+1})^{x+1}\Big(ln(1+\frac{1}{2\Delta+1})\big(1-C(x+\Delta+2)\big)-C\Big).$$

It vanishes for $x_0 = \frac{1-C(\Delta+2)}{C} - \frac{1}{ln(1+\frac{1}{2\Delta+1})}$ but $x_0 < \Delta$ for any $\Delta > 0$ which contradicts $\Delta \leq x \leq 2\Delta - 1$.

So $f'(x) \neq 0 \ \forall \ x \in [\Delta; 2\Delta - 1]$ and is continuous which implies that all the derivatives have the same sign.

We now show that the derivative is negative for one precise n and with this result, we know that the expectation decreases for $\Delta \leq n \leq 2\Delta - 1$.

$$\begin{aligned}\operatorname{sgn}(f'(x)) &= \operatorname{sgn}(f'(x))_{|_{x=\Delta+1}} \\ &= \operatorname{sgn}\Big(\ln(1+\frac{1}{2\Delta+1})(1+\frac{1}{2\Delta+1})^{\Delta} - \frac{4}{2\Delta+1}\Big) < 0 \quad \forall \Delta > 0.\end{aligned}$$

To use the Lemma 4.1, we need a $\Delta+1^{st}$ consecutive point for which it's decreasing.

$$\begin{aligned}&\mathbf{E}\big(D_i(i+\Delta-1)\big) - \mathbf{E}\big(D_i(i+\Delta)\big) \\ &\qquad = m(1+\frac{1}{2\Delta+1})^{\Delta} - m\Big((1+\frac{1}{2\Delta+1})^{\Delta+1} - (1+\frac{1}{2\Delta+1})\Big) > 0.\end{aligned}$$

One may remark that if we consider the in-degree then the inequality is reversed and the maximal in-degree is attained at time $i+\Delta$.

As we have $\Delta+1$ consecutive points for which the expectation is monotone decreasing, we can conclude that

$$\mathbf{E}\big(D_i(i+\Delta-1)\big) \geq \mathbf{E}\big(D_i(i+n)\big) \ \forall n.$$

This proves Equation (3.2) of Theorem 3.2 .

Even if for convenience, the proof has been done for $i \geq \Delta$, it is also true for $i < \Delta$ since then edges are more intensively deleted.

The formula (3.3) of Theorem 3.2 and the Corollary 3.1 follow immediately.

$$\begin{aligned}\max_{t\geq 1}\mathbf{E}\big(D_i(t)\big) &= \mathbf{E}\big(D_i(i+\Delta-1)\big), && \forall i, \\ &= m\big(1+\frac{1}{2\Delta+1}\big)^{\Delta}, && \forall i > 2\Delta.\end{aligned}$$

As a consequence:

$$\lim_{\Delta\to\infty}\lim_{t\to\infty}\max_{1\leq i\leq t}\mathbf{E}\big(D_i^{\Delta}(t)\big) = me^{\frac{1}{2}}.$$

□

5 Conclusion

We examined the impact of deletion of edges in the Preferential Attachment model. We showed that the transition of the expected degree between the model with ephemeral edges and the model with infinite-lived connection is not continuous. Our results (Theorem 3.1) show that the deletion of edges implies that with a high probability, the vertices are renewed too. In our model, the connections are not static and the system is more applicable to the real world.

The model is separated into two phases. The first is the growth phase when the graph follows the rules of the model of preferential attachment. In this phase, the graph accumulates vertices and edges until it reaches a maximum size Δ. Once reached the graph can't continue to grow. This is followed by the renewal phase where competition occurs. In this phase, the oldest edge is deleted so as to be replaced by a new one.

The model captures essential features of social networks: people with multiple connections are more likely to acquire new frienships but there is ultimately a limit on the number of friendships that one individual can maintain (12). Our model fulfills conditions of an evolution model stated in (13) ("Constant change is a natural feature of evolution, on a sufficiently large time scale, there's nothing remotely stable about evolution.") and is in good agreement with observations such as the number of species, their lifetime and renewal. In particular, our model fits the description of neural networks

We proved that the property of accumulating edges on the first vertices is removed by the concept of lifetime for edges. However, by keeping the principle of preferential attachment for the newly introduced edges our graph is still robust to random deletion as it has a part with high degree vertices but is also much less vulnerable than the preferential attachment model as the attractive vertices are numerous. Moreover, by changing the parameter m which is the number of edges introduced into the graph with every vertex, we control the connectivity of the limiting graph.

Acknowledgements

The author thanks Tatyana Turova for helpful comments and stimulating discussions on several aspects of the problem. The author also thanks the referee for helpful remarks.

Bibliography

[1] Barabási, A.L.; Albert, R. Emergence of Scaling in Random Networks. Science **1999**, *286*, 509-512

[2] Barabási, A-L.; Albert, R. Statistical Mechanics of Complex Networks. Review of Modern Physics **2002**, *74*, 47-97

[3] Bollobás, B.; Riordan, O; Spencer, J; Tusnády, G. The Degree Sequence of a Scale Free Random Graph Process. Random Structures Algorithms **2001**,*19*, 279-290.

[4] Bollobás, B.; Riordan, O. Robustness and Vulnerability of Scale-Free Random Graphs. Internet Mathematics**2003***I 1*, 1-35

[5] Bollobás, B.; Riordan, O. The Diameter of a Scale-Free Random Graph. Combinatorica **2004**, *24 (1)*, 5-34

[6] Bonato, A. A Survey of Models of the Web Graph. The Proceedings of Combinatorial and Algorithmic Aspects of Networking **2004**

[7] Cooper, C.; Frieze, A.; Vera, J. Random Deletion in a Scale-Free Random Graph Process. Internet Mathematics **2004**, *I 4*, 463-483

[8] Chung, F.; Lu, F. Coupling Online and Offline Analyses for Random Power Law Graphs. Internet Mathematics **2004**, *I 4*, 409-461

[9] Chung, F. A Duplication-Deletion Model for Random Power Law Graphs. Models of Real-World Random Networks **2005**

[10] Dorogotsev S.N.; Mendes J.F.F *Evolution of Networks: From Biological Nets to the Internet and WWW*; Oxford University Press, **2003**

[11] Iglesias, J.; Eriksson, J.; Grize, F.; Tomassini, M.; Villa, A. Dynamics of Pruning in Simulated Large-Scale Spiking Neural Networks. BioSystems **2005**, *79*, 11-20.

[12] Jen, E.M.; Girvan, M.; Newman M.E.J. The Structure of Growing Social Networks. Phys. Rev E **2001**, *64*, 046132.

[13] Newman M.E.J.; Self Organized Criticality, Evolution and the Fossil Extinction Record,Proceedings of the Royal Society of London **1996** *B.263*, 1605-1610.

[14] Slater, J.L.; Hughes, B.D.; Landman, K.A. Evolving Mortal Networks. Phys. Rev E **2006**, *73*, 066111

[15] Szymański. J.; On a Nonuniform Random Recursive Tree. Annals of Discr. Math. **1987**, *33*, 297-306.

[16] Turova,T.S. Dynamical Random Graphs with Memory. Phys. Rev E **2002**, *65*, 066102 (see also erratum Phys. Rev E **2004**, *70*, 059902)

B

Chapter B

Merging percolation and classical random graphs: Phase transition in dimension 1

T. S. Turova, T. Vallier

Centre for Mathematical Sciences, Lund University, Sweden

Abstract

We study a random graph model which combines properties of the edge percolation model on Z^d and a classical random graph $G(n, c/n)$. We show that this model, being a *homogeneous* random graph, has a natural relation to the so-called "rank 1 case" of *inhomogeneous* random graphs. This allows us to use the newly developed theory of inhomogeneous random graphs to describe completely the phase diagram in the case $d = 1$. The phase transition is similar to the classical random graph, it is of the second order. We also find the scaled size of the largest connected component above the phase transition.

Key words: Classical random graphs, phase transition, percolation.

1 Introduction.

We consider a graph on the set of vertices $V_N^d := \{1, \dots, N\}^d$ in Z^d, where the edges between any two different vertices i and j are presented independently with probabilities

$$p_{ij} = \begin{cases} q, & \text{if } |i-j| = 1, \\ c/N^d, & \text{if } |i-j| > 1, \end{cases}$$

where $0 \le q \le 1$ and $0 < c < N$ are constants. This graph, call it $G_N^d(q,c)$ is a mixture of percolation model, where each pair of neighbours in Z^d is connected with probability q, and a random graph model, where each vertex is connected to any other vertex with probability $c/|V_N^d|$.

The introduced model is a simplification of the most common graphs designed to study natural phenomena, in particular, biological neural networks (8). Observe the difference between this and the so-called "small-world" models intensively studied after (9). In the "small-world" models where edges from the grid may be kept or removed, only finite number (often at most $2d$) of the long-range edges may come out of each vertex, and the probability of those is a fixed number.

We are interested in the limiting behaviour of the introduced graph $G_N^d(q,c)$ as $N \to \infty$. One can consider this model as a graph on Z^d or on a torus, in the limit the results are the same. The one-dimensional case which we study here, is exactly solvable. We shall write $G_N 1(q,c) = G_N(q,c)$.

Let X be a random variable with $Fs(1-q)$-distribution, i.e.,

$$\mathbf{P}\{X = k\} = (1-q)q^{k-1}, \quad k = 1, 2, \dots , \tag{1.1}$$

with

$$\mathbf{E}X = \frac{1}{1-q} .$$

Let further $C_1\Big(G\Big)$ denote the size of the largest connected component in a graph G.

Theorem 1.1. *For any* $0 \le q < 1$ *define*

$$c^{cr}(q) = \frac{\mathbf{E}X}{\mathbf{E}X2} = \frac{1-q}{1+q} . \tag{1.2}$$

i) If $c < c^{cr}(q)$ *then there exists a constant* $\alpha = \alpha(q,c) > \frac{1}{|\log q|}$ *such that*

$$\mathbf{P}\Big\{C_1\Big(G_N(q,c)\Big) > \alpha \log N\Big\} \to 0, \tag{1.3}$$

and for any $\alpha_1 < \frac{1}{|\log q|}$

$$\mathbf{P}\Big\{C_1\Big(G_N(q,c)\Big) < \alpha_1 \log N\Big\} \to 0, \tag{1.4}$$

as $N \to \infty$.

ii) If $c \geq c^{cr}(q)$ *then*

$$\frac{C_1\Big(G_N(q,c)\Big)}{N} \xrightarrow{P} \beta \tag{1.5}$$

as $N \to \infty$, *with* $\beta = \beta(q,c)$ *defined as the maximal solution to*

$$\beta = 1 - \frac{1}{\mathbf{E}X}\,\mathbf{E}\left\{Xe^{-cX\beta}\right\}. \tag{1.6}$$

Observe the following duality of this result. For any $c < 1$ we know that the subgraph induced in our model by the long-range edges may have at most $O(\log N)$ vertices in a connected component. According to Theorem 1.1, for any $c < 1$ there is

$$q^{cr}(c) = \frac{1-c}{1+c}$$

such that for all $q^{cr}(c) < q < 1$ our model will have a giant component with a size of order N, while any $q < q^{cr}(c)$ is insufficient to produce a giant component in $G_N(q,c)$. Hence, Theorem 1.1 may also tell us something about the "distances" between the components of a random graph when it is considered on the vertices of $\mathbf{Z}$.

Remark 1.1. *In the proof of (1.3) we will show how to obtain* $\alpha(q,c)$, *and will discuss how optimal this value is. Statement (1.4) is rather trivial (and far from being optimal): it follows from a simple observation (see the details below) that*

$$\mathbf{P}\left\{C_1\Big(G_N(q,c)\Big) < \frac{1}{|\log q|}\log N\right\} \leq \mathbf{P}\left\{C_1\Big(G_N(q,0)\Big) < \frac{1}{|\log q|}\log N\right\} \to 0.$$

Remark 1.2. *For any fixed c function* $\beta(q,c)$ *is continuous at* $q = 0$*: if* $q = 0$, *i.e., when our graph is merely a classical* $G_{n,c/n}$ *random graph, then* $X \equiv 1$ *and (1.6) becomes a well-known relation. Equation (1.6) can be written in an exact form:*

$$\beta = 1 - \frac{e^{c\beta}}{\big(e^{c\beta} - q\big)\,2}(1-q)2\,.$$

It is easy to check that if $c \leq c^{cr}$ then the equation (1.6) does not have a strictly positive solution, while $\beta = 0$ is always a solution to (1.6). Therefore one can derive

$$\beta_c' \mid_{\,c\downarrow c^{cr}} = 2\frac{\big(\mathbf{E}(X2)\big)3}{\big(\mathbf{E}X\big)2\mathbf{E}(X3)} = 2\frac{(1-q2)2}{q2+4q+1}. \tag{1.7}$$

This shows that the emergence of the giant component at critical parameter $c = c^{cr}$ becomes slower as q increases, but the phase transition remains of the second order (exponent 1) for any $q < 1$.

We conjecture that similar results hold in the higher dimensions if $q < Q^{cr}(d)$, where $Q^{cr}(d)$ is the percolation threshold in the dimension d. More exactly, Theorem 1.1 (as well as the first equality in (1.7)) should hold with X replaced by another random variable, which is stochastically not larger than the size of the open cluster at the origin in the edge percolation model with a probability of edge q. It is known from the percolation theory (see, e.g., (3)) that the tail of the distribution of the size of an open cluster in the subcritical phase decays exponentially. This should make possible to extend our arguments (where we use essentially the distribution of X) to the general case.

Our result in the supercritical case, namely equation (1.6) looks somewhat similar to the equation obtained in (2) for the "volume" (the sum of degrees of the involved vertices) of the giant component in the graph with a given sequence of the expected degrees. Note, however, that the model in (2) (as well as the derivations of the results) differs essentially from the one studied here. In particular, in our model the critical mean degree when $c = c^{cr}$ and $N \to \infty$ is given by

$$2q + c^{cr} = 2q + \frac{1-q}{1+q} = 1 + \frac{2q2}{1+q} \tag{1.8}$$

which is strictly greater than 1 for all positive q. This is in a contrast with the model studied in (2), where the critical expected average degree is still 1 as in the classical random graph.

Although our model (when considered on the ring or torus in higher dimensions) is a perfectly *homogeneous* random graph, in the sense that the degree distribution is the same for any vertex, we study it via *inhomogeneous* random graphs, making use of the recently developed theory from (1). The idea is the following. First, we consider the subgraph induced by the short-range edges, i.e., the edges which connect two neighbouring nodes with probability q. It is composed of the consecutive connected paths (which may consist just of one single vertex) on $V_N = \{1, \dots, N\}$. Call a *macro-vertex* each of the component of this subgraph. We say that a macro-vertex is of type k, if k is the number of vertices in it. Conditionally on the set of macro-vertices, we consider a graph on these macro-vertices induced by the long-range connections. Two macro-vertices are said to be connected if there is at least one (long-range type) edge between two vertices belonging to different macro-vertices. Thus the probability of an edge between two macro-vertices v_i and v_j of types x and y correspondingly, is

$$\widetilde{p}_{xy}(N) := 1 - \left(1 - \frac{c}{N}\right)^{xy}. \tag{1.9}$$

Below we argue that this model fits the conditions of a general inhomogeneous graph model defined in (1), find the critical parameters and characteristics for the graph on

macro-vertices, and then we turn back to the original model. We use essentially the results from (1) to derive (1.6), while in the subcritical case our approach somewhat differs from the one in (1); we discuss this in the end of Section 2.4. We shall also note that our graph on macro-vertices is similar to the model studied in (5), and our results on the critical value agree with those in (5).

Finally we comment that our result should help to study more general model for the propagation of the neuronal activity introduced in (8). Here we show that a giant component in the graph can emerge from two sources, none of which can be neglected, but each of which may be in the subcritical phase, i.e., even when both $q < 1$ and $c < 1$. In particular, for any $0 < c < 1$ we can find $q < 1$ which allows with a positive probability the propagation of impulses through the large part of the network due to the local activity.

2 Proof

2.1 Random graph on macro-vertices.

Denote X a random number of the vertices connected through short-range edges to the vertex 1 on $V_\infty^1 = \{1, 2, \ldots\}$. Clearly, X has the First success distribution defined in (1.1). Let $X_1, X_2, \ldots$, be independent copies of X, and define for any $N > 1$

$$T(N) := \min\{n \geq 0 : \sum_{i=1}^{n} X_i \leq N, \ \sum_{i=1}^{n+1} X_i > N\},$$

where we assume that a sum over an empty set equals zero.

Consider now the subgraph on $V_N = \{1, \ldots, N\}$ induced by the short-range edges. This means that any two vertices i and $i+1$ from V_N are connected with probability q independent of the rest. By the construction this subgraph, call it $G_N^{(s)}(q)$, is composed of a random number of connected paths of random sizes. We call here the size of a path the number of its vertices. Clearly, there is a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ where the number of paths in $G_N^{(s)}(q)$ equals $T(N)$ if $\sum_{i=1}^{T(N)} X_i = N$, or $T(N)+1$ if $\sum_{i=1}^{T(N)} X_i < N$. Correspondingly, the sizes of the paths follow the distribution of

$$\mathbf{X} = (X_1, X_2, \ldots, X_{T(N)}, N - \sum_{i=1}^{T(N)} X_i) \tag{2.1}$$

(where the last entry may take zero value).

On the other hand, the number of the connected components of $G_N^{(s)}(q)$ exceeds exactly by one the number of "missed" short edges on V_N. This means that on the same probability space $(\Omega, \mathcal{F}, \mathbf{P})$ there is a random variable Y_N distributed as $Bin(N-1, 1-q)$,

such that either $T(N) = Y_N + 1$ or $T(N) + 1 = Y_N + 1$, and in any case

$$0 \leq T(N) - Y_N \leq 1. \tag{2.2}$$

This together with the Strong Law of Large Numbers implies

Proposition 2.1.

$$\frac{T(N)}{N} \overset{a.s.}{\to} 1 - q = \frac{1}{\mathbf{E}X}$$

as $N \to \infty$. □

Also the relation (2.2) allows us to use the large deviation inequality from (4) (formula (2.9), p.27 in (4)) for the binomial random variables in order to obtain the following rate of convergence

$$\mathbf{P}\left\{\left|\frac{T(N)}{N} - \frac{1}{\mathbf{E}X}\right| > \delta\right\} \leq 2\exp\left(-\frac{\delta 2}{12(1-q)}N\right) \tag{2.3}$$

for all $\delta > 0$ and $N > 2/\delta$.

Define for any $k \geq 1$ an indicator function

$$I_k(x) = \begin{cases} 1, & \text{if } x = k, \\ 0, & \text{otherwise.} \end{cases}$$

As an immediate corollary of Proposition 2.1 and the Law of Large Numbers we also get the following result.

Proposition 2.2. *For any fixed* $k \geq 1$

$$\frac{1}{T(N)}\sum_{i=1}^{T(N)} I_k(X_i) \overset{P}{\to} \mathbf{P}\{X = k\} = (1-q)q^{k-1} =: \mu(k) \tag{2.4}$$

as $N \to \infty$. □

Given a vector of paths $\mathbf{X}$ defined in (2.1), we introduce another graph $\widetilde{G}_N(\mathbf{X}, q, c)$ as follows. The set of vertices of $\widetilde{G}_N(\mathbf{X}, q, c)$ we denote $\{v_1, \ldots, v_{T(N)}\}$. Each vertex v_i is said to be of type X_i, which means that v_i corresponds to the set of X_i connected vertices on V_N. We shall also call any vertex v_i of $\widetilde{G}_N(\mathbf{X}, q, c)$ a *macro-vertex*, and write

$$v_i = \begin{cases} \{1, \ldots, X_1\}, & \text{if } i = 1; \\ \{\sum_{j=1}^{i-1} X_j + 1, \ldots, \sum_{j=1}^{i-1} X_j + X_i\}, & \text{if } i > 1. \end{cases} \tag{2.5}$$

With this notation the type of a vertex v_i is simply the cardinality of set v_i. The space of the types of macro-vertices is $S = \{1, 2, \ldots\}$. According to (2.4) the distribution of type

of a (macro-)vertex in graph $\widetilde{G}_N(\mathbf{X}, q, c)$ converges to measure μ on S. The edges between the vertices of $\widetilde{G}_N(\mathbf{X}, q, c)$ are presented independently with probabilities induced by the original graph $G_N(q, c)$. More precisely, the probability of an edge between any two vertices v_i and v_j of types x and y correspondingly, is $\widetilde{p}_{xy}(N)$ introduced in (1.9). Clearly, this construction provides a one-to-one correspondence between the connected components in the graphs $\widetilde{G}_N(\mathbf{X}, q, c)$ and $G_N(q, c)$: the number of the connected components is the same for both graphs, as well as the number of the involved vertices from V_N in two corresponding components. In fact, considering conditionally on $\mathbf{X}$ graph $\widetilde{G}_N(\mathbf{X}, q, c)$ we neglect only those long-range edges from $G_N(q, c)$, which connect vertices within each v_i, i.e., the vertices which are already connected through the short-range edges.

Consider now

$$\widetilde{p}_{xy}(N) = 1 - \Big(1 - \frac{c}{N}\Big)^{xy} =: \frac{\varkappa'_N(x, y)}{N}. \tag{2.6}$$

Observe that if $x(N) \to x$ and $y(N) \to y$ then

$$\varkappa'_N(x(N), y(N)) \to cxy \tag{2.7}$$

for all $x, y \in S$. In order to place our model into the framework of the inhomogeneous random graphs from (1) let us introduce another (random) kernel

$$\varkappa_{T(N)}(x, y) = \frac{T(N)}{N}\varkappa'_N(x, y),$$

so that we can rewrite the probability $\widetilde{p}_{xy}(N)$ in a graph $\widetilde{G}_N(\mathbf{X}, q, c)$ taking into account the size of the graph:

$$\widetilde{p}_{xy}(N) = \frac{\varkappa_{T(N)}(x, y)}{T(N)}. \tag{2.8}$$

(We use notations from (1) whenever it is appropriate.) According to Proposition 2.1 and (2.7), if $x(N) \to x$ and $y(N) \to y$ then

$$\varkappa_{T(N)}(x(N), y(N)) \to \varkappa(x, y) := \frac{c}{\mathbf{E}X}xy \quad a.s. \tag{2.9}$$

as $N \to \infty$ for all $x, y \in S$.

Hence, in view of Proposition 2.2 we conclude that conditionally on $T(N) = t(N)$, where $t(N)/N \to 1/\mathbf{E}X$, our model falls into the so-called "rank 1 case" of the general inhomogeneous random graph model $G^{\mathcal{V}}(t(N), \varkappa_{t(N)})$ with a vertex space $\mathcal{V} = (S, \mu, (X_1, \dots, X_{t(N)})_{N\geq 1})$ from (1) (Chapter 16.4). Furthermore, it is not difficult to verify with a help of the Propositions 2.1 and 2.2 that

$$\varkappa \in L1(S \times S, \mu \times \mu), \tag{2.10}$$

since

$$\sum_{y=1}^{\infty}\sum_{x=1}^{\infty}(1-q)xq^{x-1}(1-q)yq^{y-1} = \Big(\frac{1}{1-q}\Big)2,$$

and for any $t(N)$ such that $t(N)/N \to 1/\mathbf{E}X$

$$\frac{1}{t(N)}\mathbf{E}\{e(\widetilde{G}_N(\mathbf{X}, q, c))|T(N) = t(N)\} \to \frac{1}{2}\sum_{y=1}^{\infty}\sum_{x=1}^{\infty}\varkappa(x,y)\mu(x)\mu(y), \tag{2.11}$$

where $e(G)$ denotes the number of edges in a graph G. According to Definition 2.7 from (1), under the conditions (2.11), (2.10) and (2.9) the sequence of kernels $\varkappa_{t(N)}$ (on the countable space $S \times S$) is called *graphical* on $\mathcal{V}$ with limit $\varkappa$.

2.2 A branching process related to $\widetilde{G}_N(\mathbf{X}, q, c)$.

Here we closely follow the approach from (1). We shall use a well-known technique of branching processes to reveal the connected component in graph $\widetilde{G}_N(\mathbf{X}, q, c)$. Recall first the usual algorithm of finding a connected component. Conditionally on the set of macro-vertices, take any vertex v_i to be the root. Find all the vertices $\{v1_{i_1}, v1_{i_2}, ..., v1_{i_u}\}$ connected to this vertex v_i in the graph $\widetilde{G}_N(\mathbf{X}, q, c)$, call them the first generation of v_i, and then mark v_i as "saturated". Then for each non-saturated but already revealed vertex, we find all the vertices connected to them but which have not been used previously. We continue this process until we end up with a tree of saturated vertices.

Denote $\tau_N(x)$ the set of the macro-vertices in the tree constructed according to the above algorithm with the root at a vertex of type x.

It is plausible to think (and in our case it is correct, as will be seen below) that this algorithm with a high probability as $N \to \infty$ reveals a tree of the offspring of the following multi-type Galton-Watson process with type space $S = \{1, 2, \ldots\}$: at any step, a particle of type $x \in S$ is replaced in the next generation by a set of particles where the number of particles of type y has a Poisson distribution $Po(\varkappa(x,y)\mu(y))$. Let $\rho(\varkappa; x)$ denote the probability that a particle of type x produces an infinite population.

Proposition 2.3. *The function $\rho(\varkappa; x)$, $x \in S$, is the maximum solution to*

$$\rho(\varkappa; x) = 1 - e^{-\sum_{y=1}^{\infty}\varkappa(x,y)\mu(y)\rho(\varkappa;y)}. \tag{2.12}$$

Proof. We have

$$\sum_{y=1}^{\infty}\varkappa(x,y)\mu(y) = \frac{c}{\mathbf{E}X}\frac{x}{1-q} < \infty \text{ for any } x,$$

which together with (2.10) verifies that the conditions of Theorem 6.1 from (1) are satisfied, and the result (2.12) follows by this theorem. □

Notice that it also follows by the same Theorem 6.1 from (1) that $\rho(\varkappa; x) > 0$ for all $x \in S$ if and only if

$$\frac{c}{\mathbf{E}X}\sum_{y=1}^{\infty} y2\mu(y) = c\,\frac{\mathbf{E}X2}{\mathbf{E}X} = c\,\frac{1+q}{1-q} > 1; \tag{2.13}$$

otherwise, $\rho(\varkappa; x) = 0$ for all $x \in S$. Hence, the formula (1.2) for the critical value follows from (2.13).

As we showed above, conditionally on $T(N)$ (so that $T(N)/N \to 1/\mathbf{E}X$) the sequence $\varkappa_{T(N)}$ is graphical on $\mathcal{V}$. Hence, the conditions of Theorem 3.1 from (1) are satisfied and we derive (first, conditionally on $T(N)$, and therefore unconditionally) that

$$\frac{C_1(\widetilde{G}_N(\mathbf{X}, q, c))}{T(N)} \xrightarrow{P} \rho(\varkappa),$$

where $\rho(\varkappa) = \sum_{x=1}^{\infty} \rho(\varkappa; x)\mu(x)$. This together with Proposition 2.1 on the *a.s.* convergence of $T(N)$ implies

$$\frac{C_1(\widetilde{G}_N(\mathbf{X}, q, c))}{N} \xrightarrow{P} (1-q)\rho(\varkappa). \tag{2.14}$$

Notice that here $C_1(\widetilde{G}_N(\mathbf{X}, q, c))$ is the number of macro-vertices in $\widetilde{G}_N(\mathbf{X}, q, c)$.

2.3 On the distribution of types of vertices in $\widetilde{G}_N(\mathbf{X}, q, c)$.

Given a vector of paths $\mathbf{X}$ (see (2.1)) we define a random sequence

$$\mathcal{N} = \{\mathcal{N}_1, \ldots \mathcal{N}_N\},$$

where

$$\mathcal{N}_k = \mathcal{N}_k(\mathbf{X}) = \sum_{i=1}^{T(N)} I_k(X_i).$$

In words, $\mathcal{N}_k$ is the number of (macro-)vertices of type k in the set of vertices of graph $\widetilde{G}_N(\mathbf{X}, q, c)$. We shall prove here a useful result on the distribution of $\mathcal{N}$ (which is stronger than Proposition 2.2).

Lemma 2.1. *For any fixed* $\varepsilon > 0$

$$\mathbf{P}\{|\mathcal{N}_k/T(N) - \mu(k)| > \varepsilon k\mu(k) \quad \textit{for some } 1 \le k \le N\} = o(1) \tag{2.15}$$

as $N \to \infty$.

Proof. Let us fix $\varepsilon > 0$ arbitrarily. Observe that for any $K > 1/\varepsilon$

$$\begin{aligned} &\mathbf{P}\{|\mathcal{N}_k/T(N) - \mu(k)| > \varepsilon k\mu(k) \quad \text{for some } 1 \le k \le N\} \\ &\le \mathbf{P}\{\max_{1 \le i \le T(N)} X_i > K\} \\ &+\mathbf{P}\{|\mathcal{N}_k/T(N) - \mu(k)| > \varepsilon k\mu(k) \quad \text{for some } 1 \le k \le K\}. \end{aligned} \tag{2.16}$$

Next we shall choose an appropriate $K = K(N)$ so that we will be able to bound from above by $o(1)$ (as $N \to \infty$) each of the summands on the right in (2.16).

Let us fix $\delta > 0$ arbitrarily, and define an event

$$\mathcal{A}_{\delta,N} = \left\{ \left| \frac{T(N)}{N} - \frac{1}{\mathbf{E}X} \right| \le \delta \right\}. \tag{2.17}$$

Recall that according to (2.3)

$$\mathbf{P}(\mathcal{A}_{\delta,N}) \ge 1 - 2\exp\left(-\frac{\delta 2}{12(1-q)} N \right) = 1 - o(1) \tag{2.18}$$

as $N \to \infty$. Now we derive

$$\begin{aligned} \mathbf{P}\{ \max_{1 \le i \le T(N)} X_i > K \} &\le \mathbf{P}\{ \max_{1 \le i \le T(N)} X_i > K \mid \mathcal{A}_{\delta,N} \} \mathbf{P}(\mathcal{A}_{\delta,N}) + \mathbf{P}\{\overline{\mathcal{A}_{\delta,N}}\} \qquad (2.19) \\ &\le \left(\frac{1}{\mathbf{E}X} + \delta \right) N \, \mathbf{P}\{X_1 > K \mid \mathcal{A}_{\delta,N}\} \mathbf{P}(\mathcal{A}_{\delta,N}) + \mathbf{P}\{\overline{\mathcal{A}_{\delta,N}}\} \\ &\le \left(\frac{1}{\mathbf{E}X} + \delta \right) N \, \mathbf{P}\{X_1 > K\} + \mathbf{P}\{\overline{\mathcal{A}_{\delta,N}}\} \end{aligned}$$

as $N \to \infty$. Making use of the formula (1.1) for the distribution of X_1 we obtain from (2.19) and (2.18)

$$\mathbf{P}\{ \max_{1 \le i \le T(N)} \ge K \} \le CNq^K + 2\exp\left(-\frac{\delta 2}{12(1-q)} N \right) \tag{2.20}$$

as $N \to \infty$, where $C = C(\delta, q)$ is some finite positive constant. Let now $\omega_1(N) < N$ be any function tending to infinity with N, and set

$$K(N) = \frac{1}{|\log q|} \log N + \omega_1(N). \tag{2.21}$$

Clearly, bound (2.20) with K replaced by $K(N)$ implies

$$\mathbf{P}\{ \max_{1 \le i \le T(N)} X_i \ge K(N) \} = o(1) \tag{2.22}$$

as $N \to \infty$.

Now we consider the last term in (2.16). Let us define

$$k_0 := \max\left\{ \left[\frac{1}{\varepsilon} \right], \left[\frac{1}{|\log q|} \right] \right\} + 2. \tag{2.23}$$

Then we obtain making use of (2.18)

$$\mathbf{P}\{ |\mathcal{N}_k / T(N) - \mu(k)| > \varepsilon \, k\mu(k) \quad \text{for some } 1 \le k \le K(N) \} \tag{2.24}$$

$$\leq \sum_{k=1}^{k_0} \mathbf{P}\{|\frac{1}{T(N)}\sum_{i=1}^{T(N)} I_k(X_i) - \mu(k)| > \varepsilon\, k\mu(k)\}$$

$$+ \sum_{k=k_0+1}^{K(N)} \mathbf{P}\{|\frac{1}{T(N)}\sum_{i=1}^{T(N)} I_k(X_i) - \mu(k)| > \varepsilon\, k\mu(k) \mid \mathcal{A}_{\delta,N}\}\mathbf{P}(\mathcal{A}_{\delta,N}) + o(1)$$

$$= \sum_{k=k_0+1}^{K(N)} \mathbf{P}\{|\frac{1}{T(N)}\sum_{i=1}^{T(N)} I_k(X_i) - \mu(k)| > \varepsilon\, k\mu(k) \mid \mathcal{A}_{\delta,N}\}\mathbf{P}(\mathcal{A}_{\delta,N}) + o(1)$$

as $N \to \infty$, where the last equality is due to Proposition 2.2. Notice that for each $k > k_0$ we have $\varepsilon\, k > 1$ and therefore

$$\mathbf{P}\left\{\left|\frac{1}{T(N)}\sum_{i=1}^{T(N)} I_k(X_i) - \mu(k)\right| > \varepsilon\, k\mu(k) \mid \mathcal{A}_{\delta,N}\right\} \tag{2.25}$$

$$= \mathbf{P}\left\{\frac{1}{T(N)}\sum_{i=1}^{T(N)} I_k(X_i) - \mu(k) > \varepsilon\, k\mu(k) \mid \mathcal{A}_{\delta,N}\right\}$$

$$\leq \mathbf{P}\left\{\frac{1}{\left(\frac{1}{\mathbf{E}X} - \delta\right)N} \sum_{i=1}^{[(\frac{1}{\mathbf{E}X}+\delta)N]+1} I_k(X_i) > \mu(k) + \varepsilon\, k\mu(k) \mid \mathcal{A}_{\delta,N}\right\} =: \mathbf{P}(k).$$

Set $t(N) = \left[\left(\frac{1}{\mathbf{E}X} + \delta\right)N\right] + 1$. Then using the bound

$$\frac{\left(\frac{1}{\mathbf{E}X} - \delta\right)N}{t(N)} > 1 - \frac{5}{2}\mathbf{E}X\,\delta$$

for all $N > 2/\delta$, we derive

$$\mathbf{P}(k) \leq \mathbf{P}\left\{\frac{1}{t(N)}\sum_{i=1}^{t(N)} I_k(X_i) > \mu(k)(1+\varepsilon\, k)(1 - \frac{5\mathbf{E}X}{2}\delta) \mid \mathcal{A}_{\delta,N}\right\} \tag{2.26}$$

for all $N > 2/\delta$. Now for all $k > k_0$ and $0 < \delta < \frac{1}{10\mathbf{E}X}$ we have $(1+\varepsilon\, k)(1 - \frac{5\mathbf{E}X}{2}\delta) \geq 1 + \frac{\varepsilon}{2}\, k$, and therefore

$$\mathbf{P}(k) \leq \mathbf{P}\left\{\frac{1}{t(N)}\sum_{i=1}^{t(N)} I_k(X_i) > \mu(k)(1 + \frac{\varepsilon}{2}\, k) \mid \mathcal{A}_{\delta,N}\right\} \tag{2.27}$$

$$\leq \mathbf{P}\left\{\frac{1}{t(N)}\sum_{i=1}^{t(N)} I_k(X_i) > \mu(k)(1 + \frac{\varepsilon}{2}\, k)\right\} / \mathbf{P}(\mathcal{A}_{\delta,N}).$$

Note that $\sum_{i=1}^{t(N)} I_k(X_i)$ follows the binomial distribution $Bin(t(N), \mu(k))$. This allows us to use the large deviation inequality from (4) (see (2.5), p.26 in (4)) and derive

$$\mathbf{P}\left\{\frac{1}{t(N)}\sum_{i=1}^{t(N)} I_k(X_i) > \mu(k)(1+\frac{\varepsilon}{2}\,k)\right\} \tag{2.28}$$

$$\leq \exp\Big(-\frac{(\frac{\varepsilon}{2}k\mu(k)t(N))2}{\frac{1}{3}\varepsilon k\mu(k)t(N)+2\mu(k)t(N)}\Big) \leq \exp\big(-\frac{1}{10}\varepsilon k\mu(k)t(N)\big)$$

for all $k > k_0$. Substituting this into (2.27) we obtain

$$\mathbf{P}(k) \leq \exp\big(-\frac{1}{10}\varepsilon k\mu(k)t(N)\big)\,/\,\mathbf{P}(\mathcal{A}_{\delta,N}) \tag{2.29}$$

for all $k > k_0$. The last bound combined with (2.25) and (2.24) leads to

$$\mathbf{P}\{|\mathcal{N}_k/T(N) - \mu(k)| > \varepsilon\, k\mu(k) \quad \text{for some } 1 \leq k \leq K(N)\}$$

$$\leq \sum_{k=k_0+1}^{K(N)} \exp\big(-\frac{1}{10}\varepsilon k\mu(k)t(N)\big) + o(1),$$

as $N \to \infty$. Taking into account that function $k\mu(k)$ is decreasing for $k > k_0$ we derive from the last bound:

$$\mathbf{P}\{|\mathcal{N}_k/T(N) - \mu(k)| > \varepsilon\, k\mu(k) \quad \text{for some } 1 \leq k \leq K(N)\} \tag{2.30}$$

$$\leq \exp\big(-\frac{1}{10}\varepsilon K(N)\mu(K(N))t(N) + \log K(N)\big) + o(1),$$

as $N \to \infty$.

Setting now $\omega_1(N) = \log\log\log N$ in (2.21), it is easy to check that for

$$K(N) = \frac{1}{|\log q|}\log N + \log\log\log N$$

the entire right-hand side of the inequality (2.30) is $o(1)$ as $N \to \infty$. This together with the previous bound (2.22) and inequality (2.16) finishes the proof of lemma. □

2.4 Proof of Theorem 1.1 in the subcritical case $c < c^{cr}(q)$.

Let us fix $0 \leq q < 1$ and then $c < c^{cr}(q)$ arbitrarily. Given $\mathbf{X}$ let again v_i denote the macro-vertices with types X_i, $i = 1, 2, \ldots$, respectively, and let $\widetilde{L}$ denote a connected

component in $\widetilde{G}_N(\mathbf{X}, q, c)$. Firstly, for any $K > 0$ and $0 < \delta < 1/\mathbf{E}X$ we derive with help of (2.18)

$$\mathbf{P}\left\{C_1\Big(G_N(q,c)\Big) < K\right\} \le \mathbf{P}\left\{C_1\Big(G_N(q,0)\Big) < K\right\} = \mathbf{P}\left\{\max_{1\le i\le T(N)} X_i < K\right\} \tag{2.31}$$
$$\le \mathbf{P}\left\{\max_{1\le i\le T(N)} X_i < K \mid \mathcal{A}_{\delta,N}\right\} + o(1) \le \big(1 - \mathbf{P}\{X \ge K\}\big)^{N\left(\frac{1}{\mathbf{E}X}-\delta\right)} + o(1),$$

as $N \to \infty$, where X has the $Fs(1-q)$-distribution. Since

$$\mathbf{P}\{X \ge K\} = q^{K-1},$$

we derive from (2.31) for any $a_1 < \frac{1}{|\log q|}$ and $K = a_1 \log N$

$$\mathbf{P}\left\{C_1\Big(G_N(q,c)\Big) < a_1 \log N\right\} = o(1),$$

which proves statement (1.4).

Consider now for any positive constant a and a function $w = w(N) \ge \log N$

$$\mathbf{P}\left\{C_1\Big(G_N(q,c)\Big) > aw\right\} = \mathbf{P}\left\{\max_{\widetilde{L}} \sum_{v_i \in \widetilde{L}} X_i > aw\right\}. \tag{2.32}$$

We know already from (2.14) that in the subcritical case the size (the number of macro-vertices) of any $\widetilde{L}$ is **whp** $o(N)$. Note that when the kernel $\varkappa(x,y)$ is not bounded uniformly in both arguments, which is our case, it is not granted that the largest component in the subcritical case is at most of order $\log N$ (see, e.g., discussion of Theorem 3.1 in (1)). Therefore first we shall prove the following intermediate result.

Lemma 2.2. *If $c < c^{cr}(q)$ then*

$$\mathbf{P}\left\{C_1\Big(\widetilde{G}_N(\mathbf{X}, q, c)\Big) > N^{1/2}\right\} = o(1). \tag{2.33}$$

Proof. Let us fix $\varepsilon > 0$ and $\delta > 0$ arbitrarily and introduce the following event

$$\mathcal{B}_N := \mathcal{A}_{\delta,N} \cap \left(\max_{1\le i\le T(N)} X_i \le \frac{2}{|\log q|}\log N\right) \cap \left(\cap_{k=1}^{N}\left\{\left|\frac{\mathcal{N}_k}{T(N)} - \mu(k)\right| \le \varepsilon k \mu(k)\right\}\right). \tag{2.34}$$

According to (2.18), (2.22) and (2.15) we have

$$\mathbf{P}\{\mathcal{B}_N\} = 1 - o(1) \tag{2.35}$$

as $N \to \infty$.

Recall that $\tau_N(x)$ denotes the set of the macro-vertices in the tree constructed according to the algorithm of revealing of connected component described above. Let $|\tau_N(x)|$ denotes the number of macro-vertices in $\tau_N(x)$. Then we easily derive

$$\mathbf{P}\left\{C_1\left(\widetilde{G}_N(\mathbf{X},q,c)\right) > N^{1/2}\right\} \le \mathbf{P}\left\{\max_{1\le i\le T(N)} |\tau_N(X_i)| > N^{1/2} \mid \mathcal{B}_N\right\} + o(1) \quad (2.36)$$

$$\le N\sum_{k=1}^{N}(1+\varepsilon k)\mu(k)\left(\delta + 1/\mathbf{E}X\right)\mathbf{P}\left\{|\tau_N(k)| > N^{1/2} \mid \mathcal{B}_N\right\} + o(1)$$

as $N \to \infty$. We shall use the multi-type branching process introduced above (Section 2.2) to approximate the distribution of $|\tau_N(k)|$. Let further $\mathcal{X}^{c,q}(k)$ denote the number of the particles (including the initial one) in the branching process starting with a single particle of type k. Observe that at each step of the exploration algorithm, the number of new neighbours of x of type y has a binomial distribution $Bin(N'_y, \widetilde{p}_{xy}(N))$ where N'_y is the number of remaining vertices of type y, so that $N'_y \le \mathcal{N}_y$.

We shall explore the following obvious relation between the Poisson and the binomial distributions. Let $Y_{n,p} \in Bin(n,p)$ and $Z_a \in Po(a)$, where $0 < p < 1/4$ and $a > 0$. Then for all $k \ge 0$

$$\mathbf{P}\{Y_{n,p} = k\} \le (1+Cp2)^n\,\mathbf{P}\{Z_{n\frac{p}{1-p}} = k\}, \quad (2.37)$$

where C is some positive constant (independent of n and p). Notice that for all $x, y \le \frac{2}{|\log q|}\log N$

$$\widetilde{p}_{xy}(N) = 1 - \left(1 - \frac{c}{N}\right)^{xy} = \frac{c}{N}\,xy\,(1+o(1)), \quad (2.38)$$

and clearly, $\widetilde{p}_{xy}(N) \le 1/4$ for all large N. Therefore for any fixed positive ε_1 we can choose small ε and δ in (2.34) so that conditionally on $\mathcal{B}_N$ we have

$$N'_y\frac{\widetilde{p}_{xy}(N)}{1-\widetilde{p}_{xy}(N)} \le (1+y\varepsilon_1)\mu(y)\varkappa(x,y) \quad (2.39)$$

for all large N. Let us write further

$$\mu(y) = \mu_q(y), \quad \mu_q = \sum_{y\ge 1} y\mu_q(y)\ (= \mathbf{E}X), \quad \varkappa(x,y) = \varkappa_{c,q}(x,y)$$

emphasizing dependence on q and c. Then for any $\varepsilon_2 > 0$ and any $q' > q$ such that $\frac{1-q}{q}\frac{q'}{1-q'} < 1+\varepsilon_2$ we can choose $\varepsilon_1 < \log(q'/q)$, and derive from (2.39)

$$N'_y\frac{\widetilde{p}_{xy}(N)}{1-\widetilde{p}_{xy}(N)} \le (1+\varepsilon_2)\mu_{q'}(y)\varkappa_{c,q}(x,y) = \mu_{q'}(y)\frac{(1+\varepsilon_2)\,c}{\mu_q}xy. \quad (2.40)$$

Setting now $c' := (1+\varepsilon_2)\frac{\mu_{q'}}{\mu_q}\, c$ we rewrite (2.40) as follows

$$N'_y \frac{\widetilde{p}_{xy}(N)}{1-\widetilde{p}_{xy}(N)} \le \mu_{q'}(y)\varkappa_{c',q'}(x,y). \tag{2.41}$$

Recall that above we fixed q and $c < c^{cr}(q)$, where $c^{cr}(q)$ is decreasing and continuous in q. Hence, we can choose $q' > q$ and $c' := (1+\varepsilon_2)\frac{\mu_{q'}}{\mu_q}\, c$ so that

$$c < c' < c^{cr}(q') < c^{cr}(q), \tag{2.42}$$

and moreover c' and q' can be chosen arbitrarily close to c and q, respectively.

Now conditionally on $\mathcal{B}_N$ we can replace according to (2.37) at each (of at most N) step of the exploration algorithm the $Bin(N'_y, \widetilde{p}_{xy}(N))$ variable with $Po(N'_y \frac{\widetilde{p}_{xy}(N)}{1-\widetilde{p}_{xy}(N)})$, and further replace the last variables with the stochastically larger ones $Po(\mu_{q'}(y)\varkappa_{c',q'}(x,y))$ (recall (2.41)). As a result we get the following bound using branching process:

$$\mathbf{P}\left\{|\tau_N(k)| > N^{1/2} \mid \mathcal{B}_N\right\} \tag{2.43}$$

$$\le \left(1 + C\left(\max_{x,y\le 2\log N/|\log q|} \widetilde{p}_{xy}(N)\right) 2\right)^{N2} \mathbf{P}\left\{\mathcal{X}^{c',q'}(k) > N^{1/2}\right\}.$$

This together with (2.38) implies

$$\mathbf{P}\left\{|\tau_N(k)| > N^{1/2} \mid \mathcal{B}_N\right\} \le e^{b(\log N)4}\, \mathbf{P}\left\{\mathcal{X}^{c',q'}(k) > N^{1/2}\right\}, \tag{2.44}$$

where b is some positive constant. Substituting the last bound into (2.36) we derive

$$\mathbf{P}\left\{C_1\left(\widetilde{G}_N(\mathbf{X},q,c)\right) > N^{1/2}\right\} \tag{2.45}$$

$$\le b_2 N e^{b(\log N)4} \sum_{k=1}^{N} k\mu_{q'}(k)\mathbf{P}\left\{\mathcal{X}^{c',q'}(k) > N^{1/2}\right\} + o(1)$$

as $N \to \infty$, where b_2 is some positive constant. By the Markov's inequality

$$\mathbf{P}\{\mathcal{X}^{c',q'}(k) > N^{1/2}\} \le z^{-N^{1/2}} \mathbf{E} z^{\mathcal{X}^{c',q'}(k)} \tag{2.46}$$

for all $z \ge 1$. Denote $h_z(k) = \mathbf{E} z^{\mathcal{X}^{c',q'}(k)}$; then with a help of (2.46) we get from (2.45)

$$\mathbf{P}\left\{C_1\left(\widetilde{G}_N(\mathbf{X},q,c)\right) > N^{1/2}\right\} \le b_1 N e^{b(\log N)2} z^{-N^{1/2}} \sum_{k=1}^{N} k\mu_{q'}(k) h_z(k) + o(1). \tag{2.47}$$

Now we will show that there exists $z > 1$ such that the series

$$B_z(c', q') = \sum_{k=1}^{\infty} k\mu_{q'}(k)h_z(k)$$

converge. This together with (2.47) will clearly imply the statement of the lemma.

Note that function $h_z(k)$ (as a generating function for a branching process) satisfies the following equation

$$\begin{aligned} h_z(k) &= z \exp\left\{\sum_{x=1}^{\infty} \varkappa_{c',q'}(k,x)\mu_{q'}(x)(h_z(x)-1)\right\} \\ &= z \exp\left\{\frac{c'}{\mu_{q'}} k \left(\sum_{x=1}^{\infty} x\mu_{q'}(x)h_z(x) - \mu_{q'}\right)\right\} \\ &= z \exp\left\{\frac{c'}{\mu_{q'}} k(B_z(c', q') - \mu_{q'})\right\}. \end{aligned}$$

Multiplying both sides by $k\mu_{q'}(k)$ and summing up over k we find

$$B_z(c', q') = \sum_{k=1}^{\infty} k\mu_{q'}(k) z \exp\left\{\frac{c'}{\mu_{q'}} k(B_z(c', q') - \mu_{q'})\right\}.$$

Let us write for simplicity $B_z = B_z(c, q)$. Hence, as long as B_z is finite, it should satisfy equation

$$B_z = z\mathbf{E}X\, e^{\frac{c}{\mathbf{E}X}X(B_z - \mathbf{E}X)}, \tag{2.48}$$

which implies in turn that B_z is finite for some $z > 1$ if and only if (2.48) has at least one solution (for the same value of z). Notice that

$$B_z \geq B_1 = \mathbf{E}X \tag{2.49}$$

for $z \geq 1$. Let us fix $z > 1$ and consider equation

$$y/z = \mathbf{E}X\, e^{\frac{c}{\mathbf{E}X}X(y - \mathbf{E}X)} =: F(y) \tag{2.50}$$

for $y \geq \mathbf{E}X$. Using the properties of the distribution of X it is easy to derive that function

$$F(y) = \frac{1}{\mathbf{E}X} \frac{e^{c(\frac{y}{\mathbf{E}X} - 1)}}{\left(1 - qe^{c(\frac{y}{\mathbf{E}X} - 1)}\right)2}$$

is increasing and has positive second derivative if $\mathbf{E}X \leq y \leq y_0$, where y_0 is the root of $1 = q \exp\left\{c(\frac{y_0}{\mathbf{E}X} - 1)\right\}$. Compute now

$$\frac{\partial}{\partial y} F(y)|_{y=\mathbf{E}X} = \frac{c}{\mathbf{E}X}\mathbf{E}X2 = \frac{c}{c^{cr}}. \tag{2.51}$$

Hence, if $c < c^{cr}$ then there exists $z > 1$ such that there is a finite solution y to (2.50). Taking into account condition (2.42), we find that $B_z(c', q')$ is also finite for some $z > 1$, which finishes the proof of the lemma. □

Now we are ready to complete the proof of (1.3), following almost the same arguments as in the proof of the previous lemma. Let $S_N(x) = \sum_{v_i \in \tau_N(x)} X_i$ denote the number of vertices from V_N which compose the macro-vertices of $\tau_N(x)$. Denote

$$\mathcal{B}'_N := \mathcal{B}_N \cap \left(\max_{1 \le i \le T(N)} |\tau_N(X_i)| < N^{1/2} \right).$$

According to (2.35) and Lemma 2.2 we have

$$\mathbf{P}\{\mathcal{B}'_N\} = 1 - o(1).$$

This allows us to derive from (2.32)

$$\mathbf{P}\left\{ C_1\Big(G_N(q,c)\Big) > aw \right\} \le \mathbf{P}\left\{ \max_{1 \le i \le T(N)} S_N(X_i) > aw \mid \mathcal{B}'_N \right\} + o(1) \tag{2.52}$$

$$\le N \sum_{k=1}^{N} (1 + \varepsilon k)\mu(k)\Big(\delta + 1/\mathbf{E}X\Big)\mathbf{P}\{S_N(k) > aw \mid \mathcal{B}'_N\} + o(1).$$

Let now $S^{c,q}(y)$ denote the sum of types (including the one of the initial particle) in the total progeny of the introduced above branching process starting with initial particle of type y. Repeating the same argument which led to (2.43), we get the following bound using the introduced branching process:

$$\mathbf{P}\{S_N(k) > aw \mid \mathcal{B}'_N\} \le \left(1 + C\left(\max_{x,y \le 2\log N/|\log q|} \widetilde{p}_{xy}(N)\right) 2\right)^{b_1 N\sqrt{N}} \mathbf{P}\left\{ S^{c',q'}(k) > aw \right\}$$

as $N \to \infty$, where we take into account that we can perform at most $\sqrt{N}$ steps of exploration (the maximal possible number of macro-vertices in any $\widetilde{L}$). This together with (2.38) implies

$$\mathbf{P}\{\tau_N(k) > aw \mid \mathcal{B}'_N\} \le (1 + o(1))\mathbf{P}\left\{ S^{c',q'}(k) > aw \right\} \tag{2.53}$$

as $N \to \infty$. Substituting the last bound into (2.52) we derive

$$\mathbf{P}\left\{ C_1\Big(G_N(q,c)\Big) > aw \right\} \le Nb \sum_{k=1}^{N} k\mu_{q'}(k)\mathbf{P}\left\{ S^{c',q'}(k) > aw \right\} + o(1) \tag{2.54}$$

as $N \to \infty$, where b is some positive constant. Denote $g_z(k) = \mathbf{E}z^{S^{c',q'}(k)}$; then similar to (2.47) we derive from (2.54)

$$\mathbf{P}\left\{C_1\Big(G_N(q,c)\Big) > aw(N)\right\} \le bN\sum_{k=1}^{N} k\mu_{q'}(k)g_z(k)z^{-aw(N)} + o(1). \tag{2.55}$$

We shall search for all $z \ge 1$ for which the series

$$A_z(c',q') = \sum_{k=1}^{\infty} k\mu_{q'}(k)g_z(k)$$

converge. Function $g_z(k)$ (as a generating function for a certain branching process) satisfies the following equation

$$\begin{aligned} g_z(k) &= z^k \exp\left\{\textstyle\sum_{x=1}^{\infty} \varkappa_{c',q'}(k,x)\mu_{q'}(x)(g_z(x)-1)\right\} \\ &= z^k \exp\left\{\tfrac{c'}{\mu_{q'}} k\left(\textstyle\sum_{x=1}^{\infty} x\mu_{q'}(x)g_z(x) - \mu_{q'}\right)\right\} \\ &= z^k \exp\left\{\tfrac{c'}{\mu_{q'}} k(A_z(c',q') - \mu_{q'})\right\}. \end{aligned}$$

Multiplying both sides by $k\mu_{q'}(k)$ and summing up over k we find

$$A_z(c',q') = \sum_{k=1}^{\infty} k\mu_{q'}(k)z^k \exp\left\{\frac{c'}{\mu_{q'}} k(A_z(c',q') - \mu_{q'})\right\}. \tag{2.56}$$

It follows from here (and the fact that $A_z(c',q') \ge \mu_{q'}$ for all $z \ge 1$) that if there exists $z > 1$ for which the series $A_z(c',q')$ converge, it should satisfy

$$z < \frac{1}{q'}. \tag{2.57}$$

According to (2.56), as long as $A_z = A_z(c,q)$ is finite it satisfies the equation

$$A_z = \mathbf{E}Xz^X\, e^{\frac{c}{\mathbf{E}X}X(A_z - \mathbf{E}X)},$$

which implies that A_z is finite for some $z > 1$ if and only if the last equation has at least one solution

$$A_z \ge A_1 = \mathbf{E}X. \tag{2.58}$$

Let us fix $z > 1$ and consider equation

$$y = \mathbf{E}Xz^X\, e^{\frac{c}{\mathbf{E}X}X(y - \mathbf{E}X)} =: f(y,z). \tag{2.59}$$

Using the properties of the distribution of X it is easy to derive that

$$f(y,z) = \frac{1}{\mathbf{E}X} \frac{ze^{c(\frac{y}{\mathbf{E}X}-1)}}{\big(1-qze^{c(\frac{y}{\mathbf{E}X}-1)}\big)2}.$$

We shall consider $f(y,z)$ for $y \geq \mathbf{E}X$ and $1 \leq z < \frac{1}{q}e^{-c(\frac{y}{\mathbf{E}X}-1)}$. It is easy to check that in this area function $f(y,z)$ is increasing, it has all the derivatives of the second order, and $\frac{\partial^2}{\partial y2} f(y,z) > 0$. Compute now

$$\frac{\partial}{\partial y} f(y,z)|_{y=1,z=1} = \frac{c}{\mathbf{E}X}\mathbf{E}X2 = \frac{c}{c^{cr}}. \tag{2.60}$$

Hence, if $c > c^{cr}$ there is no solution $y \geq \mathbf{E}X$ to (2.59) for any $z > 1$. On the other hand, if $c < c^{cr}$ then there exists $1 < z_0 < 1/q$ such that for all $1 \leq z < z_0$ there is a finite solution $y \geq \mathbf{E}X$ to (2.59). One could find z_0 for example, as the (unique!) value for which function y is tangent to $f(y,z_0)$ if $y \geq \mathbf{E}X$.

Now taking into account that $c' > c$ and $q' > q$ can be chosen arbitrarily close to c and q, respectively, we derive from (2.55) that for all $1 < z < z_0$

$$\mathbf{P}\Big\{C_1\Big(G_N(q,c)\Big) > aw(N)\Big\} \leq b(z)Nz^{-aw(N)} + o(1) \tag{2.61}$$

as $N \to \infty$, where $b(z) < \infty$. This implies that for any $a > 1/\log z_0 > 1/|\log q|$

$$\mathbf{P}\Big\{C_1\Big(G_N(q,c)\Big) > a\log N\Big\} = o(1) \tag{2.62}$$

as $N \to \infty$, which proves (1.3). □

To conclude this section we comment on the methods used here. It is shown in (6) that in the subcritical case of classical random graphs the same method of generating functions combined with the Markov inequality leads to a constant which is known to be the principal term for the asymptotics of the size of the largest component (scaled to $\log N$). This gives us hope that a constant a chosen here to satisfy $a > 1/\log z_0$ is close to the minimal constant for which statement (2.62) still holds.

Similar methods were used in (7) for some class of inhomogeneous random graphs, and in (1) for a general class of models. Note, however, some difference with the approach in (1). It is assumed in (1), Section 12, that the generating function for the corresponding branching process with the initial state k (e.g., our function $g_z(k)$, $k \geq 1$) is bounded uniformly in k. As we prove here this condition is not always necessary: we need only convergence of the series A_z, while $g_z(k)$ is unbounded in k in our case. Furthermore, our approach allows one to construct constant $\alpha(q,c)$ as a function of the parameters of the model.

2.5 Proof of Theorem 1.1 in the supercritical case.

Let $\mathcal{C}_k$ denote the set of vertices in the k-th largest component in graph $G_N(q,c)$, and conditionally on $\mathbf{X}$ let $\widetilde{\mathcal{C}}_k$ denote the set of macro-vertices in the k-th largest component in graph $\widetilde{G}_N(\mathbf{X},q,c)$ (ordered in any way if there are ties). Let also C_k and $\widetilde{C}_k$ denote correspondingly, their sizes. According to our construction for any connected component $\widetilde{L}$ in $\widetilde{G}_N(\mathbf{X},q,c)$ there is a unique component L in $G_N(q,c)$ such that they are composed of the same vertices from V_N, i.e., in the notations (2.5)

$$L = \cup_{v\in\widetilde{L}} \cup_{k\in v} \{k\} =: V(\widetilde{L}).$$

Next we prove that with a high probability the largest components in both graphs consist of the same vertices.

Lemma 2.3. *For any* $0 \le q < 1$ *if* $c > c^{cr}(q)$ *then*

$$\mathbf{P}\{\mathcal{C}_1 = V(\widetilde{\mathcal{C}}_1)\} = 1 - o(1) \tag{2.63}$$

as $N \to \infty$.

Proof. In a view of the argument preceeding this lemma we have

$$\mathbf{P}\{\mathcal{C}_1 \ne V(\widetilde{\mathcal{C}}_1)\} = \mathbf{P}\{\mathcal{C}_1 = V(\widetilde{\mathcal{C}}_k) \text{ for some } k \ge 2\}.$$

According to Theorem 12.6 from (1), conditions of which are satisfied here, in the supercritical case conditionally on $T(N)$ such that $T(N)/N \to 1/\mathbf{E}X$, we have **whp** $\widetilde{C}_2 = O(\log(T(N)))$, which by Proposition 2.1 implies $\widetilde{C}_2 = O(\log N)$ **whp** . Also we know already from (2.14) that in the supercritical case $\widetilde{C}_1 = O(N)$ **whp**, and therefore $C_1 = O(N)$ **whp**. Hence, for some positive constants a and b

$$\mathbf{P}\{\mathcal{C}_1 \ne V(\widetilde{\mathcal{C}}_1)\} = \mathbf{P}\{\mathcal{C}_1 = V(\widetilde{\mathcal{C}}_k) \text{ for some } k \ge 2\} \tag{2.64}$$

$$\le \mathbf{P}\left\{\left(\max_{k\ge 2} |V(\widetilde{\mathcal{C}}_k)| > bN\right) \cap \left(\max_{k\ge 2} \widetilde{C}_k < a\log N\right)\right\} + o(1).$$

Define now for any $K = K(N)$ a set

$$B_N := \{\exists X_i \ge K \text{ for some } 1 \le i \le T(N)\}.$$

According to (2.20)

$$\mathbf{P}(B_N) \le CNq^K + o(1)$$

as $N \to \infty$ for some constant C independent of K and N. Setting from now on $K = \sqrt{N}$ we have $\mathbf{P}(B_N) = o(1)$ as $N \to \infty$. Then we derive

$$\mathbf{P}\left\{\left(\max_{k\ge 2} |V(\widetilde{\mathcal{C}}_k)| > bN\right) \cap \left(\max_{k\ge 2} \widetilde{C}_k < a\log N\right)\right\} \tag{2.65}$$

$$\leq \mathbf{P}\left\{\left(\max_{k\geq 2}|V(\widetilde{\mathcal{C}}_k)|>bN\right)\cap\left(\max_{k\geq 2}\widetilde{C}_k<a\log N\right)\cap\left(\max_{1\leq i\leq T(N)}X_i<\sqrt{N}\right)\right\}+o(1)$$
$$\leq \mathbf{P}\left\{\sqrt{N}\,a\log N>bN\right\}+o(1)=o(1).$$

Substituting this bound into (2.64) we immediately get (2.63). □

Conditionally on $\mathcal{C}_1 = V(\widetilde{\mathcal{C}}_1)$ we have

$$\begin{aligned}\tfrac{C_1}{N} &= \tfrac{1}{N}\textstyle\sum_{i=1}^{T(N)} X_i\mathbf{1}\{v_i\in\widetilde{\mathcal{C}}_1\}\\ &= \tfrac{1}{N}\textstyle\sum_{i=1}^{T(N)}\sum_{k=1}^{N} k\mathbf{1}\{X_i=k\}\mathbf{1}\{v_i\in\widetilde{\mathcal{C}}_1\} \qquad (2.66)\\ &= \tfrac{T(N)}{N}\textstyle\sum_{k=1}^{N} k\tfrac{1}{T(N)}\#\{v_i\in\widetilde{\mathcal{C}}_1 : X_i=k\}.\end{aligned}$$

Note that Theorem 9.10 from (1) (together with Proposition 2.1 in our case) implies that

$$\nu_N(k) := \frac{1}{T(N)}\#\{v_i\in\widetilde{\mathcal{C}}_1(N) : X_i=k\} \xrightarrow{P} \rho(\varkappa;k)\mu(k) \qquad (2.67)$$

for each $k\geq 1$.

We shall prove below that also

$$W_N := \sum_{k=1}^{N} k\nu_N(k) \xrightarrow{P} \sum_{k=1}^{\infty} k\rho(\varkappa;k)\mu(k) =: \beta\,\mathbf{E}X. \qquad (2.68)$$

Observe that $\rho(\varkappa;k)$ is the maximal solution to (2.12), therefore β is the maximal solution to

$$\beta = \frac{1}{\mathbf{E}X}\sum_{k=1}^{\infty} k\rho(\varkappa;k)\mu(k) = \frac{1}{\mathbf{E}X}\sum_{k=1}^{\infty} k\left(1-e^{-\sum_{y=1}^{\infty}\varkappa(k,y)\mu(y)\rho(\varkappa;y)}\right)\mu(k)$$
$$= 1-\frac{1}{\mathbf{E}X}\mathbf{E}\Big(Xe^{-cX\beta}\Big).$$

This proves that β is the maximal root of (1.6). Then (2.68) together with Proposition 2.1, which states that $T(N)/N \xrightarrow{a.s.} 1/\mathbf{E}X$, will allow us to derive from (2.66) that for any positive ε

$$\mathbf{P}\Big\{\Big|\frac{C_1\Big(G_N(c,q)\Big)}{N}-\beta\Big|>\varepsilon \mid \mathcal{C}_1 = V(\widetilde{\mathcal{C}}_1)\Big\}\to 0$$

as $N\to\infty$. This combined with Lemma 2.3 would immediately imply

$$\frac{C_1\Big(G_N(c,q)\Big)}{N} \xrightarrow{P} \beta, \qquad (2.69)$$

and hence the statement of the theorem follows.

Now we are left with proving (2.68). For any $1 \le R < N$ write $W_N := W_N^R + w_N^R$, where

$$W_N^R := \sum_{k=1}^{R} k\nu_N(k), \quad w_N^R := \sum_{k=R+1}^{N} k\nu_N(k).$$

By (2.67) we have for any fixed $R \ge 1$

$$W_N^R \xrightarrow{P} \sum_{k=1}^{R} k\rho(x;k)\mu(k) \tag{2.70}$$

as $N \to \infty$. Consider w_N^R. Note that for any $k \ge 1$

$$\mathbf{E}\nu_N(k) = \mathbf{E}\mathbf{E}\Big\{ \nu_N(k) \mid T(N) \Big\} \le \mathbf{E}\frac{1}{T(N)} \sum_{i=1}^{T(N)} \mathbf{P}\{X_i = k \mid T(N)\}. \tag{2.71}$$

Using events $\mathcal{A}_{\delta,N}$ with bound (2.18) and Proposition 2.1 we obtain from (2.71) for any fixed $0 < \delta < 1/(2\mathbf{E}X)$

$$\begin{aligned}\mathbf{E}\nu_N(k) &\le \mathbf{E}\frac{1}{T(N)} \sum_{i=1}^{T(N)} \mathbf{P}\{X_i = k \mid T(N)\}\mathbf{1}\{\mathcal{A}_{\delta,N}\} + \mathbf{P}\{\overline{\mathcal{A}_{\delta,N}}\} \\ &\le \frac{(1+\delta\mathbf{E}X)}{(1-\delta\mathbf{E}X)}\,\mathbf{P}\{X_1 = k\}(1+o(1)) + \mathbf{P}\{\overline{\mathcal{A}_{\delta,N}}\}.\end{aligned}$$

Bound (2.18) allows us to derive from here that

$$\mathbf{E}\nu_N(k) \le A_1(\mu(k) + e^{-a_1 N}) \tag{2.72}$$

for some positive constants A_1 and a_1 independent of k and N. This yields

$$\mathbf{E}w_N^R = \sum_{k=R+1}^{N} k\mathbf{E}\nu_N(k) \le A_2 e^{-a_2 R} \tag{2.73}$$

for some positive constants A_2 and a_2.

Now for any $\varepsilon > 0$ we can choose R so that

$$\sum_{k=R+1}^{\infty} k\rho(x;k)\mu(k) < \varepsilon/3,$$

and then we have

$$\mathbf{P}\{|W_N - \sum_{k=1}^{\infty} k\rho(x;k)\mu(k)| > \varepsilon\} \tag{2.74}$$

$$= \mathbf{P}\{|(W_N^R - \sum_{k=1}^{R} k\rho(\varkappa;k)\mu(k)) + w_N^R - \sum_{k=R+1}^{\infty} k\rho(\varkappa;k)\mu(k)| > \varepsilon\}$$

$$\leq \mathbf{P}\{|W_N^R - \sum_{k=1}^{R} k\rho(\varkappa;k)\mu(k)| > \varepsilon/3\} + \mathbf{P}\{w_N^R > \varepsilon/3\}.$$

Markov's inequality together with bound (2.73) gives us

$$\mathbf{P}\{w_N^R > \varepsilon/3\} \leq \frac{3\mathbf{E}w_N^R}{\varepsilon} \leq \frac{3A_2 e^{-a_2 R}}{\varepsilon}. \tag{2.75}$$

Making use of (2.75) and (2.70) we immediately derive from (2.74)

$$\mathbf{P}\{|W_N - \sum_{k=1}^{\infty} k\rho(\varkappa;k)\mu(k)| > \varepsilon\} \leq o(1) + \frac{3A_2 e^{-a_2 R}}{\varepsilon} \tag{2.76}$$

as $N \to \infty$. Hence, for any given positive ε and ε_0 we can choose finite R so large that

$$\lim_{N\to\infty} \mathbf{P}\{|W_N - \sum_{k=1}^{\infty} k\rho(\varkappa;k)\mu(k)| > \varepsilon\} < \varepsilon_0. \tag{2.77}$$

This clearly proves statement (2.68), and therefore finishes the proof of the theorem. □

Bibliography

[1] B. Bollobás, S. Janson and O. Riordan, The phase transition in inhomogeneous random graphs. *Random Structures and Algorithms*, **31** 3-122.

[2] F. Chung and L. Lu , The volume of the giant component of a random graph with given expected degree. SIAM *J. Discrete Math.* **20** (2006), 395-411

[3] G. Grimmett, Percolation. Springer-Verlag, Berlin, 1999.

[4] S. Janson, T. Łuczak, and A. Ruciński, Random graphs, Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[5] M. Molloy and B. Reed, The size of the giant component of a random graph with a given degree sequence. Combin. Probab. Comput. 7 (1998), no. 3, 295–305.

[6] T.S. Turova, Note on the random graphs in the subcritical case. Dynamical systems from number theory to probability - 2, (ed. A.Yu. Khrennikov), Växjö University Press (2003), 187–192.

[7] T.S. Turova, Phase Transitions in Dynamical Random Graphs, Journal of Statistical Physics, 123 (2006), no. 5, 1007-1032.

[8] T. Turova and A. Villa, On a phase diagram for random neural networks with embedded spike timing dependent plasticity. To appear in *BioSystems*

[9] D.J. Watts and S.H. Strogatz, Collective dynamics of "small-world" networks. Nature, 393 (1998), 440–442.

C

Chapter C

Merging percolation on Z^d and classical random graphs: Phase transition

T. S. Turova, T. Vallier

Centre for Mathematical Sciences, Lund University, Sweden

Abstract

We study a random graph model which is a superposition of the bond percolation model on Z^d with probability p of an edge, and a classical random graph $G(n, c/n)$. We show that this model, being a *homogeneous* random graph, has a natural relation to the so-called "rank 1 case" of *inhomogeneous* random graphs. This allows us to use the newly developed theory of inhomogeneous random graphs to describe the phase diagram on the set of parameters $c \geq 0$ and $0 \leq p < p_c$, where $p_c = p_c(d)$ is the critical probability for the bond percolation on Z^d. The phase transition is similar to the classical random graph, it is of the second order. We also find the scaled size of the largest connected component above the phase transition.

Key words: Classical random graphs, phase transition, percolation.

1 Introduction.

We consider a graph on the set of vertices $B(N) := \{-N, \dots, N\}^d$ in Z^d, $d \geq 1$, with two types of edges: the short-range edges connect independently with probability p each pair u and v if $|u - v| = 1$, and the long-range edges connect independently any pair of two vertices with probability $c/|B(N)|$. (Here for any set A we denote $|A|$ the number of the elements in A.) This graph, call it $G_N(p, c)$ is a superposition of the bond percolation model (see, e.g., (4)), where each pair of neighbours in Z^d is connected with probability p, and a random graph model $G_{n,c/n}$ (see, e.g., (6)) on $n = |B(N)|$ vertices, where each vertex is connected to any other vertex with probability c/n; all the edges in both models are independent. By this definition there can be none, one or two edges between two vertices in graph $G_N(p, c)$, and in the last case the edges are of different types.

The introduced model is a simplification of the most common graphs designed to study natural phenomena, in particular, biological neural networks (12). Notice the difference between $G_N(p, c)$ and the so-called "small-world" models intensively studied after (13). In the "small-world" models where edges from the grid may be kept or removed, only finite number (often at most $2d$) of the long-range edges may come out of each vertex, and the probability of those is a fixed number.

We are interested in the connectivity of the introduced graph $G_N(p, c)$ as $N \to \infty$. We say that two vertices are connected, if there is a path of edges, no matter of which types, between them. Clearly, if $c = 0$, we have a purely bond percolation model on Z^d, where any edge from the grid is kept (i.e., "is open" in the terminology of percolation theory) with a probability p, or, alternatively, removed with a probability $1 - p$. Let us recall some basic facts from the percolation theory which we need here. Denote C an open cluster containing the origin of Z^d in the bond percolation model. It is known (see, e.g., (4)) that for any $d \geq 1$ there is $p_c = p_c(d)$ such that

$$\mathbf{P}\{|C| = \infty\} \begin{cases} = 0, & \text{if } p < p_c, \\ > 0, & \text{if } p > p_c, \end{cases}$$

where $0 < p_c < 1$, unless $d = 1$, in which case, obviously, $p_c = 1$. We shall assume here that $0 < p < p_c$, and thus the connected components formed by the short-range edges only are finite with probability one. Recall also that for all $0 < p < p_c$ the limit

$$\zeta(p) = lim_{n\to\infty} \left(-\frac{1}{n} \log \mathbf{P}\{|C| = n\} \right) \tag{1.1}$$

exists and satisfies $\zeta(p) > 0$ (Theorem (6.78) from (4)).

Let further $C_1\Big(G\Big)$ denote the size (the number of vertices) of the largest connected component in a graph G.

Theorem 1.1. *Assume, that $d \geq 1$ and $0 \leq p < p_c(d)$. Define*

$$c^{cr}(p) = \frac{1}{\mathbf{E}|C|}. \tag{1.2}$$

i) If $c < c^{cr}(p)$ define

$$y = \begin{cases} \text{the root of } \mathbf{E}\, c|C|e^{c|C|y} = 1, & \text{if } \mathbf{E}\, c|C|e^{|C|\zeta(p)} \geq 1, \\ \zeta(p)/c, & \text{otherwise,} \end{cases}$$

and set

$$\alpha(p,c) := \Big(c + cy - \mathbf{E}\, ce^{c|C|y}\Big)^{-1}. \tag{1.3}$$

Then for any $\alpha > \alpha(p,c)$

$$\mathbf{P}\Big\{C_1\Big(G_N(p,c)\Big) > \alpha \log|B(N)|\Big\} \to 0. \tag{1.4}$$

as $N \to \infty$.

ii) If $c \geq c^{cr}(p)$ then

$$\frac{C_1\Big(G_N(p,c)\Big)}{|B(N)|} \xrightarrow{P} \beta \tag{1.5}$$

as $N \to \infty$, with $\beta = \beta(p,c)$ defined as the maximal solution to

$$\beta = 1 - \mathbf{E}\Big\{e^{-c\,\beta\,|C|}\Big\}. \tag{1.6}$$

Remark 1.1. *Only when $d = 1$ we know the exact distribution of $|C|$ (see (1.7) below), in which case $\mathbf{E}|C|e^{|C|\zeta(p)} = \infty$. This obviously yields (for $d = 1$) that the constant y in Theorem 1.1 is defined as the root of $\mathbf{E}\, c|C|e^{c|C|y} = 1$ for any $c < c^{cr}(p)$.*

In view of (1.1) it is obvious that $\chi(p) := \mathbf{E}|C| < \infty$ for all $0 \leq p < p_c$. It is also known (see Theorem (6.108) and (6.52) in (4)) that $\chi(p)$ is analytic function of p on $[0, p_c)$ and $\chi(p) \to \infty$ as $p \uparrow p_c$. This implies that $c^{cr}(p)$ is continuous, strictly decreasing function on $[0, p_c)$ with $c^{cr}(0) = 1$ and $c^{cr}(p_c) = 0$. Hence, c^{cr} has inverse, i.e., for any $0 < c < 1$ there is a unique $0 < p^{cr}(c) < p_c = p_c(d)$ such that $c^{cr}\big(p^{cr}(c)\big) = c$. This leads to the following duality of the result in Theorem 1.1.

Corollary 1.1. *For any $0 < c < 1$ there is a unique $0 < p^{cr}(c) < p_c$ such that for any $p^{cr}(c) < p < p_c$ graph $G_N(p,c)$ has a giant component with a size $O(|B(N)|)$* **whp** *(i.e., with probability tending to one as $N \to \infty$), but for any $p < p^{cr}(c)$ the size of the largest connected component in $G_N(p,c)$* **whp** *is at most $O(\log|B(N)|)$.* □

Hence, Theorem 1.1 may also tell us something about the "distances" between the components of a random graph when it is considered on the vertices of $\mathbf{Z}^d$.

It is worth mentioning that the symmetry between c^{cr} and p^{cr} is most spectacular in the dimension one case, when $p_c(1) = 1$. Notice, that $d = 1$ case is exactly solvable, and this is the only case when we know the distribution of $|C|$:

$$\mathbf{P}\{|C| = k\} = (1-p)^2 k p^{k-1}, \quad k \geq 1. \tag{1.7}$$

Hence, if $d = 1$ we compute for all $0 \leq p < 1 = p_c(1)$

$$c^{cr}(p) = \frac{1-p}{1+p}, \tag{1.8}$$

which also yields

$$p^{cr}(c) = \frac{1-c}{1+c},$$

for all $0 \leq c < 1$. (For more details on $d = 1$ case we refer to (11).)

Remark 1.2. *For any fixed c function $\beta(p, c)$ is continuous at $p = 0$: if $p = 0$, i.e., when our graph is merely a classical $G_{n,c/n}$ random graph, then $|C| \equiv 1$ and (1.6) becomes a well-known relation.*

Furthermore, for any fixed $c < 1$, if $p = 0$, it is not difficult to derive from (1.3) that

$$\alpha(0, c) = \frac{1}{c - 1 + |\log c|}. \tag{1.9}$$

But $\log n/(c - 1 + |\log c|)$ is known (see Theorem 7a in (3)) to be the principal term in the asymptotics (in probability) of the largest connected component of $G_{n,c/n}$ when $c < 1$. This inevitably leads to the following (open) question. Will

$$\mathbf{P}\left\{C_1\Big(G_N(p, c)\Big) < \alpha_1 \log|B(N)|\right\} \to 0 \quad \text{as } N \to \infty, \tag{1.10}$$

hold also for all $\alpha_1 < \alpha(p, c)$ when $0 < p < p_c$ and $c < c^{cr}(p)$?

It is easy to check that if $c \leq c^{cr}(p)$ then the equation (1.6) does not have a strictly positive solution, while $\beta = 0$ is always a solution to (1.6). Therefore we derive for all $0 \leq p \leq p_c$

$$\frac{\partial}{\partial c}\beta(p, c) \mid_{c \downarrow c^{cr}(p)} = 2\frac{\big(\mathbf{E}|C|\big)^3}{\mathbf{E}(|C|^2)} > 0. \tag{1.11}$$

This confirms that the phase transition remains to be of the second order for any $p < p_c$, as it is for $p = 0$, i.e., in the case of classical random graph.

More similarities and differences between our model and the "mean-field" case one can see in the following example. Let $d = 1$, in which case (1.7) holds. Introducing a random variable X with the first-success distribution

$$\mathbf{P}\{X = k\} = (1-p)p^{k-1}, \quad k = 1, 2, \dots ,$$

one can rewrite (1.6) as follows (see the details in (11))

$$\beta = 1 - \frac{1}{\mathbf{E}X}\,\mathbf{E}\left\{Xe^{-cX\beta}\right\}.$$

This equation looks somewhat similar to the equation obtained in (2) for the "volume" (the sum of degrees of the involved vertices) of the giant component in the graph with a given sequence of the expected degrees. Note, however, that in our model the critical mean degree when $c = c^{cr}(p)$ and $N \to \infty$ is given according to (1.8) by

$$2p + c^{cr}(p) = 2p + \frac{1-p}{1+p} = 1 + \frac{2p^2}{1+p} \tag{1.12}$$

which is strictly greater than 1 for all positive $p < 1$. This is in a contrast with the model studied in (2), where the critical expected average degree is still 1 as in the classical random graph.

Although our model (it can be considered on a torus, in the limit the result is the same) is a perfectly *homogeneous* random graph, in the sense that the degree distribution is the same for any vertex, we study it via *inhomogeneous* random graphs, making use of the recently developed theory from (1). The idea is the following. First, we consider the subgraph induced by the short-range edges, i.e., the edges which connect two neighbouring nodes with probability p. It is composed of the connected clusters (which may consist just of one single vertex) in $B(N)$. Call a *macro-vertex* each of the connected components of this subgraph. We say that a macro-vertex is of type k, if k is the number of vertices in it. Conditionally on the set of macro-vertices, we consider a graph on these macro-vertices induced by the long-range connections. Two macro-vertices are said to be connected if there is at least one (long-range type) edge between two vertices belonging to different macro-vertices. Thus the probability of an edge between two macro-vertices v_i and v_j of types x and y correspondingly, is

$$\widetilde{p}_{xy}(N) := 1 - \left(1 - \frac{c}{|B(N)|}\right)^{xy}. \tag{1.13}$$

Below we argue that this model fits the conditions of a general inhomogeneous graph model defined in (1), find the critical parameters and characteristics for the graph on macro-vertices, and then we turn back to the original model. We use essentially the results from (1) to derive (1.6). The result in the subcritical phase (part *i*) of Theorem 1.1) does not follow by the theory in (1); we discuss this in the end of Section 2.4.

Notice also that in order to analyze the introduced model, we derive here some result on the joint distribution of the sizes of clusters in the percolation model (see Lemma 2.1 below), which may be of interest on its own.

The principle of treating some local structures in a graph as new vertices ("macro-vertices"), and then considering a graph induced by the original model on these vertices appears to be rather general. For example, in (5) a different graph model was also put into a framework of inhomogeneous graphs theory by certain restructuring. This method should be useful for analysis of a broad class of complex structures, whenever one can identify local and global connections. Some examples of such models one can find in (7).

Finally we comment that our result should help to study a model for the propagation of the neuronal activity introduced in (12). Here we show that a giant component in the graph can emerge from two sources, none of which can be neglected, but each of which may be in the subcritical phase, i.e., even when both $p < p_c$ and $c < 1$. In particular, for any $0 < c < 1$ we can find $p < p_c$ which allows with a positive probability the propagation of impulses through the large part of the network due to the local activity.

2 Proof

2.1 Random graph on macro-vertices.

Consider now the subgraph on $B(N)$ induced by the short-range edges only, which is a purely bond percolation model. By the construction this subgraph, call it $G_N^{(s)}(p)$, is composed of a random number of clusters (of connected vertices) of random sizes. We call the size of a cluster the number of its vertices (it may be just one). We recall here more results from percolation theory which we shall use later on.

Let K_N denote the number of the connected components (clusters) in $G_N^{(s)}(p)$, and let

$$\mathbf{X} = \{X_1, X_2, \ldots, X_{K_N}\} \tag{2.1}$$

denote the collection of all connected clusters X_i in $G_N^{(s)}(p)$. We shall also use X_i to denote the set of vertices in the $i-$th cluster. By this definition $\sum_{i=1}^{K_N} |X_i| = |B(N)|$.

Theorem [(4), (4.2) Theorem, p. 77]

$$\frac{K_N}{|B(N)|} \to \varkappa(p) := \mathbf{E}\frac{1}{|C|} \tag{2.2}$$

a.s. and in L^1 *as* $N \to \infty$. □

Note (see, e.g., (4)) that $\varkappa(p)$ is strictly positive and finite for all $0 < p < p_c$. Next we cite the large deviations property of K_N from (14).

Theorem [(14), Theorem 2] *Given $\varkappa(p) > \varepsilon > 0$, there exist $\sigma_j(\varepsilon, p) > 0$ for $j = 1, 2$ such that*

$$\lim_{N\to\infty} \frac{-1}{|B(N)|} \log \mathbf{P}\left(\frac{K_N}{|B(N)|} \geq \varkappa(p) + \varepsilon\right) = \sigma_1(\varepsilon, p)$$

and

$$\lim_{N\to\infty} \frac{-1}{|B(N)|} \log \mathbf{P}\left(\frac{K_N}{|B(N)|} \leq \varkappa(p) - \varepsilon\right) = \sigma_2(\varepsilon, p).$$

This theorem implies that for any $0 < \delta < \varkappa(p)$ there is a positive constant $\sigma = \sigma(\delta, p)$ such that

$$\mathbf{P}\left\{\left|\frac{K_N}{|B(N)|} - \varkappa(p)\right| > \delta\right\} \leq e^{-\sigma|B(N)|} \tag{2.3}$$

for all large N. Define for any $k \geq 1$ and $x \geq 0$ an indicator function:

$$I_k(x) = \begin{cases} 1, & \text{if } x = k, \\ 0, & \text{otherwise.} \end{cases}$$

Proposition 2.1. *For any fixed $k \geq 1$*

$$\frac{1}{K_N} \sum_{i=1}^{K_N} I_k(|X_i|) \to \frac{1}{\varkappa(p)} \frac{\mathbf{P}\{|C| = k\}}{k} =: \mu(k) \tag{2.4}$$

a.s. and in L^1 as $N \to \infty$.

Proof. Let $C_N(z)$, $z \in B(N)$, denote a connected cluster in $B(N)$ which contains vertex z, and let $C(z)$ denote an open cluster in $\mathbf{Z}^d$ which contains vertex z. Then we write

$$\frac{1}{K_N} \sum_{i=1}^{K_N} I_k(|X_i|) = \frac{|B(N)|}{K_N} \frac{1}{k} \frac{1}{|B(N)|} \sum_{z\in B(N)} I_k(|C_N(z)|). \tag{2.5}$$

Observe, that

$$\sum_{z\in B(N)} I_k(|C_N(z)|) = \sum_{z\in B(N)} I_k(|C(z)|) + \sum_{z\in B(N)} (I_k(|C_N(z)|) - I_k(|C(z)|))$$

$$= \sum_{z\in B(N)} I_k(|C(z)|) + \sum_{z\leftrightarrow \partial B(N)} (I_k(|C_N(z)|) - I_k(|C(z)|)), \tag{2.6}$$

where the last summation is over all vertices z of $B(N)$ which are connected to the surface

$$\partial B(N) = \{x \in B(N) : \max_{1\leq i\leq d} |x_i| = N\},$$

and hence the last sum in (2.6) contains at most $k|\partial B(N)|$ non-zero terms. Now we can rewrite (2.5) as follows

$$\frac{1}{K_N}\sum_{i=1}^{K_N} I_k(|X_i|) = \frac{|B(N)|}{K_N}\,\frac{1}{k}\,\frac{1}{|B(N)|}\sum_{z\in B(N)} I_k(|C(z)|) + \frac{|B(N)|}{K_N}\,\frac{\Delta_N}{|B(N)|}, \tag{2.7}$$

where

$$|\Delta_N| \le |\partial B(N)|.$$

Taking into account (2.2) it is easy to see that

$$\frac{|B(N)|}{K_N}\,\frac{\Delta_N}{|B(N)|} \to 0 \tag{2.8}$$

in L^1 and *a.s.*. By the ergodic theorem

$$\frac{1}{|B(N)|}\sum_{z\in B(N)} I_k(|C(z)|) \to \mathbf{P}\{|C| = k\} \tag{2.9}$$

a.s. as $N \to \infty$, and in L^1 as well, since

$$0 \le \frac{1}{|B(N)|}\sum_{z\in B(N)} I_k(|C(z)|) \le 1.$$

Hence, statement (2.4) follows by (2.7), (2.8) and (2.9) combined with (2.2). □

Given a collection of clusters $\mathbf{X}$ defined in (2.1), we introduce another graph $\widetilde{G}_N(\mathbf{X}, p, c)$ as follows. The set of vertices of $\widetilde{G}_N(\mathbf{X}, p, c)$ we denote $\{v_1, \ldots, v_{K_N}\}$. Each vertex v_i is said to be of type $|X_i|$, which means that v_i corresponds to the set of $|X_i|$ connected vertices in $B(N)$. We shall also call any vertex v_i of $\widetilde{G}_N(\mathbf{X}, p, c)$ a *macro-vertex*, and write sometimes

$$v_i = X_i. \tag{2.10}$$

With this notation the type of a macro-vertex v_i is simply the cardinality of set $v_i = X_i$. The space of the types of macro-vertices is $S = \{1, 2, \ldots\}$. According to (2.4) the distribution of type of a (macro-)vertex in graph $\widetilde{G}_N(\mathbf{X}, p, c)$ converges to measure μ on S. The edges between the vertices of $\widetilde{G}_N(\mathbf{X}, p, c)$ are presented independently with probabilities induced by the original graph $G_N(p, c)$. More precisely, the probability of an edge between any two vertices v_i and v_j of types x and y correspondingly, is $\widetilde{p}_{xy}(N)$ introduced in (1.13). Clearly, this construction provides a one-to-one correspondence between the connected components in the graphs $\widetilde{G}_N(\mathbf{X}, p, c)$ and $G_N(p, c)$: the number of the connected components is the same for both graphs, as well as the number of the involved vertices from $B(N)$ in two corresponding components. In fact, considering conditionally on $\mathbf{X}$ graph $\widetilde{G}_N(\mathbf{X}, p, c)$ we neglect only those long-range edges from $G_N(p, c)$, which

connect vertices within each v_i, i.e., the vertices which are already connected through the short-range edges.

Consider now

$$\widetilde{p}_{xy}(N) = 1 - \left(1 - \frac{c}{|B(N)|}\right)^{xy} =: \frac{\varkappa'_N(x,y)}{|B(N)|}. \tag{2.11}$$

Observe that if $x(N) \to x$ and $y(N) \to y$ then

$$\varkappa'_N(x(N), y(N)) \to cxy \tag{2.12}$$

for all $x, y \in S$. In order to place our model into the framework of the inhomogeneous random graphs from (1) let us introduce another (random) kernel

$$\varkappa_{K_N}(x,y) = \frac{K_N}{|B(N)|}\varkappa'_N(x,y),$$

so that we can rewrite the probability $\widetilde{p}_{xy}(N)$ in a graph $\widetilde{G}_N(\mathbf{X}, p, c)$ taking into account the size of the graph:

$$\widetilde{p}_{xy}(N) = \frac{\varkappa_{K_N}(x,y)}{K_N}. \tag{2.13}$$

(We use notations from (1) whenever it is appropriate.) According to (2.2) and (2.12), if $x(N) \to x$ and $y(N) \to y$ then

$$\varkappa_{K_N}(x(N), y(N)) \xrightarrow{a.s.} \varkappa(x,y) := c\varkappa(p)xy \tag{2.14}$$

as $N \to \infty$ for all $x, y \in S$.

Hence, in view of Proposition 2.1 we conclude that conditionally on $K_N = t(N)$, where $t(N)/|B(N)| \to \mathbf{E}(|C|^{-1})$, our model falls into the so-called "rank 1 case" of the general inhomogeneous random graph model $G^{\mathcal{V}}(t(N), \varkappa_{t(N)})$ with a vertex space

$$\mathcal{V} = (S, \mu, (v_1, \dots, v_{t(N)})_{N \geq 1})$$

(see (1), Chapter 16.4). Note, that according to (1.1) function $\mu(k)$ (defined in (2.4)) decays exponentially, which implies

$$\varkappa(x,y) \in L^1(S \times S, \mu \times \mu). \tag{2.15}$$

Furthermore, it is not difficult to verify with a help of (2.2) and Proposition 2.1 that for any $t(N)$ such that $t(N)/|B(N)| \to \mathbf{E}(|C|^{-1})$

$$\frac{1}{t(N)}\mathbf{E}\{e(\widetilde{G}_N(\mathbf{X}, p, c))|K_N = t(N)\} \to \frac{1}{2}\sum_{y=1}^{\infty}\sum_{x=1}^{\infty}\varkappa(x,y)\mu(x)\mu(y), \tag{2.16}$$

where $e(G)$ denotes the number of edges in a graph G. According to Definition 2.7 from (1), under the conditions (2.16), (2.15) and (2.14) the sequence of kernels $\varkappa_{t(N)}(x,y)$ (on the countable space $S \times S$) is called *graphical* on $\mathcal{V}$ with limit $\varkappa(x,y)$.

2.2 A branching process related to $\widetilde{G}_N(\mathbf{X}, p, c)$.

Here we closely follow the approach from (1). We shall use a well-known technique of branching processes to reveal the connected component in graph $\widetilde{G}_N(\mathbf{X}, p, c)$. Recall first the usual algorithm of finding a connected component. Conditionally on the set of macro-vertices, take any vertex v_i to be the root. Find all the vertices $\{v^1_{i_1}, v^1_{i_2}, ..., v^1_{i_n}\}$ connected to this vertex v_i in the graph $\widetilde{G}_N(\mathbf{X}, p, c)$, call them the offspring of v_i, and then mark v_i as "saturated". Then for each non-saturated but already revealed vertex, we find all the vertices connected to them but which have not been used previously. We continue this process until we end up with a tree of saturated vertices.

Denote $\tau_N(x)$ the set of the macro-vertices in the tree constructed according to the above algorithm with the root at a vertex of type x.

It is plausible to think (and in our case it is correct, as will be seen below) that this algorithm with a high probability as $N \to \infty$ reveals a tree of the offspring of the following multi-type Galton-Watson process with type space $S = \{1, 2, \ldots\}$: at any step, a particle of type $x \in S$ is replaced in the next generation by a set of particles where the number of particles of type y has a Poisson distribution $Po(\varkappa(x, y)\mu(y))$. Let $\rho(x)$ denote the probability that a particle of type x produces an infinite population.

Proposition 2.2. *The function $\rho(x)$, $x \in S$, is the maximum solution to*

$$\rho(x) = 1 - e^{-\sum_{y=1}^{\infty} \varkappa(x,y)\mu(y)\rho(y)}. \tag{2.17}$$

Proof. We have

$$\sum_{y=1}^{\infty} \varkappa(x, y)\mu(y) = c\mathbf{E}(|C|^{-1})x\frac{1}{\mathbf{E}(|C|^{-1})} = cx < \infty$$

for any x. This together with (2.15) verifies the conditions of Theorem 6.1 from (1), by which the result (2.17) follows. □

Notice that it also follows by the same Theorem 6.1 from (1) that $\rho(x) > 0$ for all $x \in S$ if and only if

$$c\mathbf{E}(|C|^{-1})\sum_{y=1}^{\infty} y^2\mu(y) = c\,\mathbf{E}(|C|^{-1})\sum_{y=1}^{\infty} y^2\frac{1}{\mathbf{E}(|C|^{-1})}\,\frac{\mathbf{P}\{|C| = y\}}{y} = c\,\mathbf{E}|C| > 1; \tag{2.18}$$

otherwise, $\rho(x) = 0$ for all $x \in S$. Hence (2.18) yields formula (1.2) for the critical value $c^{cr}(p)$.

As we showed above, conditionally on K_N so that $K_N/|B(N)| \to \mathbf{E}(|C|^{-1})$, the sequence $\varkappa_{K_N}(x, y)$ is graphical on $\mathcal{V}$. Hence, the conditions of Theorem 3.1 from (1) are

satisfied and we derive (first, conditionally on K_N, and therefore unconditionally) that

$$\frac{C_1(\widetilde{G}_N(\mathbf{X},p,c))}{K_N} \xrightarrow{P} \rho,$$

where $\rho = \sum_{x=1}^{\infty} \rho(x)\mu(x)$. This together with (2.2) implies

$$\frac{C_1(\widetilde{G}_N(\mathbf{X},p,c))}{|B(N)|} \xrightarrow{P} \mathbf{E}(|C|^{-1})\,\rho. \tag{2.19}$$

Notice that here $C_1(\widetilde{G}_N(\mathbf{X},p,c))$ is the number of *macro-vertices* in the largest connected component of $\widetilde{G}_N(\mathbf{X},p,c)$.

2.3 On the distribution of types of vertices in $\widetilde{G}_N(\mathbf{X},p,c)$.

Given a collection of clusters $\mathbf{X}$ (see (2.1)) we define for all $1 \le k \le |B(N)|$

$$\mathcal{N}_k = \mathcal{N}_k(\mathbf{X}) = \sum_{i=1}^{K_N} I_k(|X_i|).$$

In words, $\mathcal{N}_k$ is the number of (macro-)vertices of type k in the set of vertices of graph $\widetilde{G}_N(\mathbf{X},p,c)$. We shall prove here a useful result on the distribution of $\mathcal{N} = (\mathcal{N}_1,\dots,\mathcal{N}_{|B(N)|})$.

Lemma 2.1. *Set*

$$\widetilde{\mu}(k) = \sum_{n=k}^{\infty} \mathbf{P}\{|C| = n\}.$$

Then for any fixed $\nu > 2$ and $\varepsilon > 0$

$$\mathbf{P}\Big\{|\mathcal{N}_k/K_N - \mu(k)| > \varepsilon\, k^{\nu}\, \widetilde{\mu}(k) \quad \textit{for some } 1 \le k \le |B(N)|\Big\} \to 0 \tag{2.20}$$

as $N \to \infty$.

Proof. Let us fix $\nu > 2$ and $\varepsilon > 0$ arbitrarily. Fix also constant L_0 arbitrarily but so that

$$\varepsilon L_0^{\nu} \ge \mathbf{E}(|C|^{-1}). \tag{2.21}$$

Then for all $k > L_0$

$$\varepsilon\, k^{\nu}\, \widetilde{\mu}(k) > \mu(k),$$

and for any $L > L_0$

$$\mathbf{P}\{|\mathcal{N}_k/K_N - \mu(k)| > \varepsilon\, k^{\nu}\, \widetilde{\mu}(k) \quad \text{for some } 1 \le k \le |B(N)|\} \tag{2.22}$$

$$\leq \mathbf{P}\{\max_{1\leq i\leq K_N}|X_i|>L\}$$

$$+\mathbf{P}\{|\mathcal{N}_k/K_N-\mu(k)|>\varepsilon\,k^{\nu}\,\widetilde{\mu}(k)\quad\text{for some } 1\leq k\leq L\}.$$

We shall choose later on an appropriate $L=L(N)$ so that we will be able to bound from above by $o(1)$ (as $N\to\infty$) each of the summands on the right in (2.22).

First we derive

$$\mathbf{P}\{\max_{1\leq i\leq K_N}|X_i|>L\}\leq\mathbf{P}\{\max_{z\in B(N)}|C(z)|>L\}\leq|B(N)|\,\mathbf{P}\{|C|>L\},\tag{2.23}$$

where $C(z)$ is an open cluster containing z. For a further reference we note here, that according to (1.1) for any $0<\alpha<\zeta(p)$ there is constant $b>0$ such that

$$\mathbf{P}\{|C|\geq L\}\leq be^{-\alpha L}\tag{2.24}$$

for all $L\geq 1$, which together with (2.23) implies, in particular, that

$$\mathbf{P}\{\max_{1\leq i\leq K_N}|X_i|>\frac{2}{\zeta(p)}\log|B(N)|\}\to 0\tag{2.25}$$

as $N\to\infty$.

Now we consider the last term in (2.22). Let us define for any $0<\delta<\mathbf{E}(|C|^{-1})$ an event

$$\mathcal{A}_{\delta,N}=\left\{\left|\frac{K_N}{|B(N)|}-\mathbf{E}(|C|^{-1})\right|\leq\delta\right\}.\tag{2.26}$$

Recall that according to (2.3)

$$\mathbf{P}(\mathcal{A}_{\delta,N})\geq 1-e^{-\sigma|B(N)|}=1-o(1)\tag{2.27}$$

as $N\to\infty$. Then we can bound the last term in (2.22) as follows

$$\begin{aligned}
&\mathbf{P}\Big\{|\mathcal{N}_k/K_N-\mu(k)|>\varepsilon\,k^{\nu}\widetilde{\mu}(k)\quad\text{for some } 1\leq k\leq L\Big\}\\
&\qquad\leq\mathbf{P}\Big\{|\mathcal{N}_k/K_N-\mu(k)|>\varepsilon\,k^{\nu}\widetilde{\mu}(k)\quad\text{for some } 1\leq k\leq L_0\Big\}\\
&\quad+\mathbf{P}\Big\{\Big(\mathcal{N}_k/K_N>\varepsilon\,k^{\nu}\,\widetilde{\mu}(k)+\mu(k)\quad\text{for some } L_0<k\leq L\Big)\cap\mathcal{A}_{\delta,N}\Big\}+\mathbf{P}\{\overline{\mathcal{A}_{\delta,N}}\}\\
&\leq\mathbf{P}\Big\{\Big(\mathcal{N}_k/K_N>\varepsilon\,k^{\nu}\,\widetilde{\mu}(k)+\mu(k)\quad\text{for some } L_0<k\leq L\Big)\cap\mathcal{A}_{\delta,N}\Big\}+o(1),
\end{aligned}\tag{2.28}$$

as $N\to\infty$ where the last inequality follows by Proposition 2.1 and bound (2.27). Write

$$P(k):=\mathbf{P}\Big\{\Big(\frac{\mathcal{N}_k}{K_N}-\mu(k)>\varepsilon\,k^{\nu}\,\widetilde{\mu}(k)\Big)\cap\mathcal{A}_{\delta,N}\Big\}.\tag{2.29}$$

Clearly, we have by (2.28):

$$\mathbf{P}\Big\{|\mathcal{N}_k/K_N - \mu(k)| > \varepsilon\, k^\nu \widetilde{\mu}(k) \quad \text{for some } 1 \le k \le L\Big\} \le \sum_{k=L_0+1}^{L} P(k) + o(1), \quad (2.30)$$

as $N \to \infty$. Substituting now (2.30) and (2.23) into (2.22) we derive

$$\mathbf{P}\{|\mathcal{N}_k/K_N - \mu(k)| > \varepsilon\, k^\nu\, \widetilde{\mu}(k) \quad \text{for some } 1 \le k \le |B(N)|\} \quad (2.31)$$

$$\le |B(N)|\, \mathbf{P}\{|C| > L\} + \sum_{k=L_0+1}^{L} P(k) + o(1) \le |B(N)|\widetilde{\mu}(L) + \sum_{k=L_0+1}^{L} P(k) + o(1)$$

as $N \to \infty$.

Next we shall find an upper bound for $P(k)$. Due to the definition (2.26) of $\mathcal{A}_{\delta,N}$, we have

$$P(k) \le \mathbf{P}\Big\{\mathcal{N}_k > (\varkappa(p) - \delta)|B(N)|\Big(\varepsilon\, k^\nu\, \widetilde{\mu}(k) + \mu(k)\Big)\Big\}. \quad (2.32)$$

We shall use the following special case of the Talagrand's inequality.

Proposition 2.3. *For every* $0 < a \le r$

$$\mathbf{P}\Big\{\mathcal{N}_k \le a\Big\}\mathbf{P}\Big\{\mathcal{N}_k \ge r\Big\} \le \exp\left\{-\frac{(r-a)^2}{32dkr}\right\}. \quad (2.33)$$

Proof. We shall derive this result as a corollary to the Talagrand's inequality (8) which we cite here from the book (6).

Theorem. [(6), p. 40.] *Suppose that* $Z_1, ..., Z_n$ *are independent random variables taking their values in some sets* $\Lambda_1, ..., \Lambda_n$*, respectively. Suppose further that* $W = f(Z_1, ..., Z_n)$*, where* $f : \Lambda_1 \times ... \times \Lambda_n \to \mathbb{R}$ *is a function such that, for some constants* c_k*,* $k = 1, ..., n$*, and some function* Ψ*, the following two conditions hold:*

1) If $z, z' \in \Lambda = \prod_1^n \Lambda_i$ *differ only in the* $i-th$ *coordinate, then* $|f(z) - f(z')| \le c_i$*.*

2) If $z \in \Lambda$ *and* $r \in \mathbb{R}$ *with* $f(z) \ge r$*, then there exists a set* $J \subseteq \{1, ..., n\}$ *with* $\sum_{i\in J} c_i^2 \le \Psi(r)$*, such that for all* $y \in \Lambda$ *with* $y_i = z_i$ *when* $i \in J$*, we have* $f(y) \ge r$*.*

Then, for every $r \in \mathbb{R}$ *and* $t \ge 0$*,*

$$\mathbf{P}(W \le r - t)\mathbf{P}(W \ge r) \le e^{-t^2/4\,\Psi(r)}. \quad (2.34)$$

We shall show now that function $\mathcal{N}_k$ satisfies the conditions of this theorem. Let $\{e_1, \ldots, e_n\}$ be the set of all edges from the lattice $\mathbf{Z}^d$ which have both end points in $B(N)$. Define

$$Z_i = \begin{cases} 1, & \text{if } e_i \text{ is open in } G_N^s(p), \\ 0, & \text{if } e_i \text{ is closed in } G_N^s(p)\ . \end{cases}$$

According to the definition of our model, $Z_i \in Be(p)$, $i = 1, \ldots, n$, are independent random variables, and

$$\mathcal{N}_k = \mathcal{N}_k(Z_1, \ldots, Z_n)$$

since the number of the components of size k (open k-clusters) in $G_N^s(p)$ is defined completely by $Z_1, \ldots, Z_n$. Furthermore, it is clear that removing or adding just one edge in $G_N^s(p)$, may increase or decrease by at most 2 the number of k-clusters. Hence, the first condition of the Talagrand's inequality is satisfied with $c_i = 2$ for all $1 \leq i \leq n$: if configurations $z, z' \in \{0,1\}^n$ differ only in the i^{th} coordinate, then

$$|\mathcal{N}_k(z) - \mathcal{N}_k(z')| \leq 2.$$

Next we check that the second condition is fulfilled as well, and we shall determine the function Ψ. Assume, $z \in \{0,1\}^n$ corresponds such configuration of the edges in $B(N)$ that $\mathcal{N}_k(z) \geq r$, for some $r \in \{1, 2, \ldots\}$, i.e., there are at least r clusters of size k. Let $\{e_j, j \in J\} \subset \{e_1, \ldots, e_n\}$ be a set of edges which have at least one common vertex with a set of exactly r (arbitrarily chosen out of $\mathcal{N}_k(z)$) clusters of size k. Clearly, $|J| \leq 2dkr$, and for any $z' \in \{0,1\}^n$ with $z'_j = z_j$ if $j \in J$, we have

$$\mathcal{N}_k(z') \geq r.$$

Hence, the second condition of the Talagrand's inequality is satisfied with $\Psi(r) = 8dkr$ for positive integers r, since

$$\sum_{i \in J} c_i^2 = 4|J| \leq 8dkr. \tag{2.35}$$

The case when r is not integer is treated as explained in Example 2.33, page 41, in (6). Then the inequality (2.33) follows by (2.34). □

Set now

$$k_N = (\varkappa(p) - \delta)|B(N)|,$$

and consider the inequality (2.33) with

$$\begin{aligned} r &= k_N\Big(\varepsilon\, k^\nu\, \widetilde{\mu}(k) + \mu(k)\Big), \\ a &= k_N\Big(\tfrac{\varepsilon}{2}\, k^\nu\, \widetilde{\mu}(k) + \mu(k)\Big). \end{aligned} \tag{2.36}$$

By the Chebyshev's inequality

$$\mathbf{P}\Big\{\mathcal{N}_k \le a\Big\} \ge 1 - \frac{\mathbf{E}\mathcal{N}_k}{a}. \tag{2.37}$$

Consider now

$$\mathcal{N}_k = \sum_{i=1}^{K_N} I_k(|X_i|) = \frac{1}{k} \sum_{z\in B(N)} I_k(|C_N(z)|), \tag{2.38}$$

where $C_N(z)$ denote an open cluster in $B(N)$ which contains vertex z. Let further $C(z)$ denote an open cluster in $\mathbf{Z}^d$ which contains vertex z. Then by (2.6) we have

$$\sum_{z\in B(N)} I_k(|C_N(z)|) \le \sum_{z\in B(N)} I_k(|C(z)|) + \sum_{z\leftrightarrow \partial B(N)} I_k(|C_N(z)|), \tag{2.39}$$

where the last sum contains at most $k|\partial B(N)|$ non-zero terms. Note also that for any $z \leftrightarrow \partial B(N)$

$$\mathbf{P}\{|C_N(z)| = k\} \le \mathbf{P}\{|C(z)| \ge k\} = \mathbf{P}\{|C| \ge k\}.$$

Therefore we derive from (2.38) and (2.39)

$$\begin{aligned}\mathbf{E}\mathcal{N}_k &\le \frac{1}{k}|B(N)|\mathbf{P}\{|C| = k\} + |\partial B(N)|\mathbf{P}\{|C| \ge k\} \qquad (2.40)\\ &= \varkappa(p)\mu(k)|B(N)| + |\partial B(N)|\widetilde{\mu}(k).\end{aligned}$$

This yields the following bound for all $k > L_0$ (in which case $\varepsilon k^\nu \widetilde{\mu}(k) \ge \mu(k)$)

$$\begin{aligned}\frac{\mathbf{E}\mathcal{N}_k}{a} &= \frac{\mathbf{E}\mathcal{N}_k}{(\varkappa(p)-\delta)|B(N)|\Big(\frac{\varepsilon}{2}\,k^\nu\,\widetilde{\mu}(k) + \mu(k)\Big)}\\ &\le \frac{\varkappa(p)\mu(k)|B(N)| + |\partial B(N)|\widetilde{\mu}(k)}{(\varkappa(p)-\delta)|B(N)|\Big(\frac{\varepsilon}{2}\,k^\nu\,\widetilde{\mu}(k) + \mu(k)\Big)}\\ &\le \frac{\varkappa(p)}{\varkappa(p)-\delta}\left(\frac{2}{3} + 2\frac{|\partial B(N)|}{|B(N)|\varepsilon\,k^\nu}\right).\end{aligned}$$

Choosing $0 < \delta \le \varkappa(p)/18$ we derive from the last bound that for all large N such that $2\frac{|\partial B(N)|}{|B(N)|\varepsilon} < 1/24$

$$\frac{\mathbf{E}\mathcal{N}_k}{a} \le 3/4$$

uniformly in $k \ge L_0$, which together with (2.37) implies

$$\mathbf{P}\Big\{\mathcal{N}_k \le a\Big\} \ge \frac{1}{4}. \tag{2.41}$$

Using (2.41) in the Talagrand's inequality (2.33) with r and a defined in (2.36), we derive for all $k > L_0$

$$\mathbf{P}\Big\{\mathcal{N}_k \geq r\Big\} \leq \Big(\mathbf{P}\Big\{\mathcal{N}_k \leq a\Big\}\Big)^{-1} \exp\left\{-\frac{(r-a)^2}{32dkr}\right\} \tag{2.42}$$

$$\leq 4\exp\left\{-\frac{\left(\frac{\varepsilon}{2}k_N\, k^{\nu}\,\widetilde{\mu}(k)\right)^2}{32dk\left(k_N\Big(\varepsilon\, k^{\nu}\,\widetilde{\mu}(k)+\mu(k)\Big)\right)}\right\}$$

$$\leq 4\exp\left\{-\frac{\varepsilon k_N\, k^{\nu}\,\widetilde{\mu}(k)}{2^8 dk}\right\} = 4\exp\left\{-\frac{\varepsilon(\varkappa(p)-\delta)}{2^8 d}|B(N)|k^{\nu-1}\,\widetilde{\mu}(k)\right\}.$$

Substituting (2.42) into (2.32) we get

$$P(k) \leq 4\exp\left\{-b|B(N)|k^{\nu-1}\,\widetilde{\mu}(k)\right\},$$

where

$$b := \frac{\varepsilon(\varkappa(p)-\delta)}{2^8 d}.$$

The last bound combined with (2.31) yields

$$\mathbf{P}\Big\{|\mathcal{N}_k/K_N - \mu(k)| > \varepsilon\, k^{\nu}\widetilde{\mu}(k) \quad \text{for some } 1 \leq k \leq |B(N)|\Big\} \tag{2.43}$$

$$\leq |B(N)|\,\widetilde{\mu}(L) + 4\sum_{k=L_0}^{L}\exp\left\{-b|B(N)|k^{\nu-1}\,\widetilde{\mu}(k)\right\} + o(1),$$

as $N \to \infty$ for any $L \geq L_0$.

Next we shall show that for any $\Delta > 0$ one can choose a finite constant L_0 and numbers $L = L(N)$ so that

$$\sum_{k=L_0}^{L(N)}\exp\left\{-b|B(N)|k^{\nu-1}\,\widetilde{\mu}(k)\right\} < \Delta, \tag{2.44}$$

for all large N, and

$$|B(N)|\,\widetilde{\mu}(L(N)) \to 0, \tag{2.45}$$

as $N \to \infty$. This together with (2.43) will clearly imply the statement of Lemma.

We claim that both (2.44) and (2.45) hold with

$$L(N) = \min\left\{k \geq 1 : k^{\alpha}\widetilde{\mu}(k) < \frac{1}{|B(N)|}\right\}, \tag{2.46}$$

where

$$\alpha = \frac{\nu - 2}{2},$$

which is positive by the assumption of Lemma.

Recall, that along with the result (1.1) it is also proved in (4) that for all $0 < p < p_c$

$$\zeta(p) = lim_{n\to\infty}\left(-\frac{1}{n}\log \mathbf{P}\left\{|C| \geq n\right\}\right). \tag{2.47}$$

For the further reference we note here that (1.1) and (2.47) immediately imply the existence and equality of the following limits for all $0 < p < p_c$

$$\zeta(p) = lim_{n\to\infty}\left(-\frac{1}{n}\log \mu(n)\right) = lim_{n\to\infty}\left(-\frac{1}{n}\log \widetilde{\mu}(n)\right), \tag{2.48}$$

i.e., that both $\mu(n)$ and $\widetilde{\mu}(n)$ decay exponentially fast, and moreover with the same exponent in the limit.

Hence, $k^{\alpha}\widetilde{\mu}(k) \to 0$, as $k \to \infty$ for any fixed α, but $k^{\alpha}\widetilde{\mu}(k) > 0$ for all $k \geq 1$. This yields that $L(N) \to \infty$ as $N \to \infty$, which in turn implies that there exists

$$\lim_{N\to\infty} \widetilde{\mu}\big(L(N)\big)|B(N)| \leq \lim_{N\to\infty} L(N)^{-\alpha} = 0,$$

and (2.45) follows.

To prove (2.44) first we note that by the definition (2.46) of $L(N)$

$$(L(N)-1)^{\alpha}\widetilde{\mu}(L(N)-1) \geq \frac{1}{|B(N)|}. \tag{2.49}$$

Recall that according to Lemma 6.102 from (4) (p.139), for all $n, m \geq 0$

$$\frac{1}{m+n}\mathbf{P}(|C| = n+m) \geq p(1-p)^{-2}\frac{1}{m}\mathbf{P}(|C| = m)\frac{1}{n}\mathbf{P}(|C| = n). \tag{2.50}$$

When $m = 1$ the inequality (2.50) implies

$$\mathbf{P}(|C| = n+1) \geq p(1-p)^{2(d-1)}\mathbf{P}(|C| = n), \tag{2.51}$$

for all $n \geq 0$. This clearly yields

$$\widetilde{\mu}(n+1) \geq p(1-p)^{2(d-1)}\widetilde{\mu}(n)$$

for all $n \geq 0$, and in particular

$$\widetilde{\mu}(L(N)) \geq p(1-p)^{2(d-1)}\widetilde{\mu}(L(N)-1). \tag{2.52}$$

Notice that $\gamma := p(1-p)^{2(d-1)} \le p < 1$ for all $d \ge 1$. Combining (2.49) with (2.52) we immediately get

$$L(N)^{\alpha}\widetilde{\mu}(L(N)) \ge \frac{\gamma}{|B(N)|}, \tag{2.53}$$

and also by the definition (2.46) for all $k < L(N)$

$$k^{\alpha}\widetilde{\mu}(k) \ge \frac{1}{|B(N)|} \ge \frac{\gamma}{|B(N)|}. \tag{2.54}$$

Making use of (2.53) and (2.54) we derive

$$\sum_{k=L_0}^{L(N)} \exp\left\{-b|B(N)|k^{\nu-1}\,\widetilde{\mu}(k)\right\} \tag{2.55}$$

$$\le \sum_{k=L_0}^{L(N)} \exp\left\{-b\gamma k^{\nu-1-\alpha}\right\} \le a_1 \exp\left\{-b\gamma L_0^{\nu-2-\alpha}\right\},$$

where a_1 is some positive constant independent of L_0. It is clear now, that for any $\Delta > 0$ we can fix L_0 so large that (2.55) will imply (2.44), and in the same time L_0 will satisfy (2.21) and $L_0 < L(N)$. This completes the proof of the lemma. □

2.4 Proof of Theorem 1.1 in the subcritical case $c < c^{cr}(p)$.

Let us fix $0 \le p < p_c$ and then $c < c^{cr}(p)$ arbitrarily. Given $\mathbf{X}$ consider graph $\widetilde{G}_N(\mathbf{X}, p, c)$. Recall that we denote v_i the vertices of types $|X_i|$, $i \ge 1$, of this graph. Let $\widetilde{L}$ denote a connected component in $\widetilde{G}_N(\mathbf{X}, p, c)$. Clearly, for any positive constant a and a function $w = w(N)$

$$\mathbf{P}\left\{C_1\Big(G_N(p,c)\Big) > aw\right\} = \mathbf{P}\left\{\max_{\widetilde{L}} \sum_{v_i \in \widetilde{L}} |X_i| > aw\right\}. \tag{2.56}$$

We know already from (2.19) that in the subcritical case the size (the number of macro-vertices) of any $\widetilde{L}$ is $o(N)$ with probability tending to one as $N \to \infty$. Note that when the kernel $\varkappa(x,y)$ is not bounded uniformly in both arguments, which is our case, it is not granted that the largest component in the subcritical case is at most of order $\log|B(N)|$ (see, e.g., discussion of Theorem 3.1 in (1)). Therefore first we shall prove the following intermediate result.

Lemma 2.2. *If $c < c^{cr}(p)$ then*

$$\mathbf{P}\left\{C_1\Big(\widetilde{G}_N(\mathbf{X}, p, c)\Big) > |B(N)|^{1/2}\right\} = o(1), \tag{2.57}$$

as $N \to \infty$.

Proof. Let us fix $\nu > 2$, $\varepsilon > 0$ and $\delta > 0$ arbitrarily and introduce the following event

$$\mathcal{B}_N = \mathcal{A}_{\delta,N} \cap \left(\max_{1 \le i \le K_N} |X_i| \le \frac{2}{\zeta(p)} \log |B(N)| \right) \tag{2.58}$$

$$\cap \left(\cap_{k=1}^{|B(N)|} \left\{ \left| \frac{\mathcal{N}_k}{K_N} - \mu(k) \right| \le \varepsilon k^{\nu} \widetilde{\mu}(k) \right\} \right).$$

According to (2.27), (2.25) and (2.20) we have

$$\mathbf{P}\{\mathcal{B}_N\} = 1 - o(1) \tag{2.59}$$

as $N \to \infty$.

Recall that $\tau_N(x)$ denote the set of the macro-vertices in the tree constructed according to the algorithm of revealing of connected component described above. Let $|\tau_N(x)|$ denote the number of macro-vertices in $\tau_N(x)$. Then we easily derive

$$\mathbf{P}\left\{ C_1\Big(\widetilde{G}_N(\mathbf{X}, p, c)\Big) > |B(N)|^{1/2} \right\} \le \mathbf{P}\left\{ \max_{1 \le i \le K_N} |\tau_N(|X_i|)| > |B(N)|^{1/2} \mid \mathcal{B}_N \right\} + o(1) \tag{2.60}$$

$$\le |B(N)| \Big(\delta + \mathbf{E}(|C|^{-1}) \Big) \sum_{k=1}^{|B(N)|} (\mu(k) + \varepsilon k^{\nu} \widetilde{\mu}(k)) \mathbf{P}\left\{ |\tau_N(k)| > |B(N)|^{1/2} \mid \mathcal{B}_N \right\} + o(1)$$

as $N \to \infty$. We shall use the multi-type branching process introduced above (Section 2.2) to approximate the distribution of $|\tau_N(k)|$. Let further $\mathcal{X}^{c,p}(k)$ denote the total number of the particles (including the initial one) produced by the branching process starting with a single particle of type k (see the definition of this branching process in Section 2.2).

Proposition 2.4. *For any $p < p_c$ and $c < c^{cr}(p)$ one can find $p < p' < p_c$ and $c < c' < c^{cr}(p')$ arbitrarily close to p and c, correspondingly, such that for all $k \ge 1$ and all large N*

$$\mathbf{P}\left\{ |\tau_N(k)| > |B(N)|^{1/2} \mid \mathcal{B}_N \right\} \le e^{b_1 (\log |B(N)|)^4} \mathbf{P}\left\{ \mathcal{X}^{c',p'}(k) > |B(N)|^{1/2} \right\}, \tag{2.61}$$

where b_1 is some positive constant independent of k and N.

Proof. Observe that at each step of the exploration algorithm, the number of the type y offspring of a particle of type x has a binomial distribution $Bin(N'_y, \widetilde{p}_{xy}(N))$ where N'_y is the number of remaining vertices of type y, so that $N'_y \le \mathcal{N}_y$.

We shall explore the following obvious relation between the Poisson and the binomial distributions. Let $Y_{n,p} \in Bin(n,p)$ and $Z_a \in Po(a)$, where $0 < p < 1/4$ and $a > 0$. Then for all $k \ge 0$

$$\mathbf{P}\{Y_{n,p} = k\} \le (1 + \gamma p^2)^n \, \mathbf{P}\{Z_{n\frac{p}{1-p}} = k\}, \tag{2.62}$$

where γ is some positive constant (independent of n, k and p). Notice that for all $1 \le x, y \le \frac{2}{\zeta(p)} \log|B(N)|$

$$\widetilde{p}_{xy}(N) = 1 - \left(1 - \frac{c}{|B(N)|}\right)^{xy} = \frac{c}{|B(N)|}\, xy\,(1 + o(1)), \tag{2.63}$$

and clearly, $\widetilde{p}_{xy}(N) \le 1/4$ for all large N. Therefore for any fixed positive ε_1 we can choose small ε and δ in (2.58) so that conditionally on $\mathcal{B}_N$ we have

$$N'_y \frac{\widetilde{p}_{xy}(N)}{1 - \widetilde{p}_{xy}(N)} \le (\mu(y) + y^{\nu}\varepsilon_1 \widetilde{\mu}(y))\varkappa(x, y) \tag{2.64}$$

for all large N.

Let us write further

$$\mu(y) = \mu_p(y) = \frac{1}{\varkappa(p)} \frac{\mathbf{P}_p\{|C| = y\}}{y}, \quad \widetilde{\mu}(y) = \widetilde{\mu}_p(y), \quad \varkappa(x, y) = \varkappa_{c,p}(x, y) = c\varkappa(p)xy,$$

emphasizing dependence on p and c. The result (2.48) allows us to find for any $p < p' < p_c$ positive constants ε_2 and $\varepsilon_1 = \varepsilon_1(\varepsilon_2, p')$ such that

$$\mu_p(y) + y^{\nu}\varepsilon_1 \widetilde{\mu}_p(y) \le (1 + \varepsilon_2)\mu_{p'}(y), \tag{2.65}$$

and moreover $\varepsilon_2 \downarrow 0$ as $p' \downarrow p$. Setting now

$$c' = (1 + \varepsilon_2)\frac{\varkappa(p)}{\varkappa(p')}\, c$$

we derive from (2.64) with a help of (2.65) that conditionally on $\mathcal{B}_N$ with an appropriate choice of constants

$$N'_y \frac{\widetilde{p}_{xy}(N)}{1 - \widetilde{p}_{xy}(N)} \le (1 + \varepsilon_2)\mu_{p'}(y)\varkappa_{c,p}(x, y) = \mu_{p'}(y)\varkappa_{c',p'}(x, y). \tag{2.66}$$

Recall that above we fixed p and $c < c^{cr}(p)$, where $c^{cr}(p)$ is strictly decreasing and continuous in p. Furthermore, function $\varkappa(p)$ is analytic on $[0, p_c)$. Hence, we can choose $p' > p$ and $c' = (1 + \varepsilon_2)\frac{\varkappa(p)}{\varkappa(p')}c$ so that

$$c < c' < c^{cr}(p') < c^{cr}(p), \tag{2.67}$$

and moreover c' and p' can be chosen arbitrarily close to c and p, respectively. Now according to (2.62) and (2.66)

$$\mathbf{P}\{Y_{N'_y, \widetilde{p}_{xy}(N)} \ge k\} \le (1 + \gamma\widetilde{p}_{xy}(N)^2)^{N'_y}\, \mathbf{P}\{Z_{N'_y \frac{\widetilde{p}_{xy}(N)}{1 - \widetilde{p}_{xy}(N)}} \ge k\} \tag{2.68}$$

$$\leq (1+\gamma\widetilde{p}_{xy}(N)^2)^{|B(N)|}\mathbf{P}\{Z_{\mu_{p'}(y)\varkappa_{c',p'}(x,y)}\geq k\}.$$

Hence, if conditionally on $\mathcal{B}_N$ at each (of at most $|B(N)|$) step of the exploration algorithm which reveals $\tau_N(k)$, we replace the $Bin(N'_y,\widetilde{p}_{xy}(N))$ variable with the $Po\left(\mu_{p'}(y)\varkappa_{c',p'}(x,y)\right)$ one, we arrive at the following bound using branching process:

$$\mathbf{P}\left\{|\tau_N(k)|>|B(N)|^{1/2}\mid\mathcal{B}_N\right\} \tag{2.69}$$

$$\leq\left(1+C\left(\max_{x,y\leq 2\log|B(N)|/\zeta(p)}\widetilde{p}_{xy}(N)\right)^2\right)^{|B(N)|^2}\mathbf{P}\left\{\mathcal{X}^{c',p'}(k)>|B(N)|^{1/2}\right\}.$$

This together with (2.63) implies statement (2.61) of the Proposition. □

Substituting (2.61) into (2.60) we derive with a help of (2.65)

$$\mathbf{P}\left\{C_1\left(\widetilde{G}_N(\mathbf{X},p,c)\right)>|B(N)|^{1/2}\right\} \tag{2.70}$$

$$\leq b_2|B(N)|e^{b_1(\log|B(N)|)^4}\sum_{k=1}^{|B(N)|}k\mu_{p'}(k)\mathbf{P}\left\{\mathcal{X}^{c',p'}(k)>|B(N)|^{1/2}\right\}+o(1)$$

as $N\to\infty$, where b_2 is some positive constant. Using the Markov's inequality we derive from here

$$\mathbf{P}\left\{C_1\left(\widetilde{G}_N(\mathbf{X},p,c)\right)>|B(N)|^{1/2}\right\} \tag{2.71}$$

$$\leq b_2|B(N)|e^{b_1(\log|B(N)|)^4}z^{-|B(N)|^{1/2}}D_z(c',p')+o(1),$$

where

$$D_z(c,p)=\sum_{k=1}^{\infty}k\mu_p(k)\mathbf{E}z^{\mathcal{X}^{c,p}(k)}.$$

It is clear now that if we can show that for some $z>1$

$$D_z(c',p')<\infty, \tag{2.72}$$

the right-hand part of (2.71) will go to zero as $N\to\infty$, which will prove the statement of Lemma 2.2.

Set

$$h_z(k)=\mathbf{E}z^{\mathcal{X}^{c,p}(k)}$$

and consider the series

$$D_z(c,p)=\sum_{k=1}^{\infty}k\mu_p(k)h_z(k)=\varkappa(p)^{-1}\sum_{k=1}^{\infty}\mathbf{P}_p\{|C|=k\}h_z(k). \tag{2.73}$$

Note that function $h_z(k)$ (as a generating function for a branching process) satisfies the following equation

$$\begin{aligned} h_z(k) &= z \exp\left\{\sum_{x=1}^{\infty} \varkappa_{c,p}(k,x)\mu_p(x)(h_z(x)-1)\right\} \\ &= z \exp\left\{c\varkappa(p)k(D_z(c,p)-\varkappa(p)^{-1})\right\}. \end{aligned}$$

Multiplying both sides by $k\mu_p(k)$ and summing up over k we find

$$D_z(c,p) = \sum_{k=1}^{\infty} k\mu_p(k) z \exp\left\{c\varkappa(p)k(D_z(c,p)-\varkappa(p)^{-1})\right\}$$

$$= \sum_{k=1}^{\infty} \varkappa(p)^{-1}\mathbf{P}_p\{|C|=k\} z \exp\left\{c\varkappa(p)k(D_z(c,p)-\varkappa(p)^{-1})\right\}.$$

Let us write for simplicity $D_z = D_z(c,p)$. Hence, as long as D_z is finite, it should satisfy equation

$$D_z = \varkappa(p)^{-1} z\, \mathbf{E} e^{c|C|(\varkappa(p)D_z-1)}, \tag{2.74}$$

which implies in turn that D_z is finite for some $z > 1$ if and only if (2.74) has at least one solution (for the same value of z). Notice that by the definition (2.73)

$$D_z \geq D_1 = \varkappa(p)^{-1} = (\mathbf{E}(|C|^{-1}))^{-1} \tag{2.75}$$

for $z \geq 1$. Let us fix $z > 1$ and consider equation

$$y/z = \mathbf{E} e^{c|C|(y-1)} =: F(y) \tag{2.76}$$

for $y \geq 1$. Using the property (1.1) of the distribution of $|C|$ it is easy to derive that function $F(y)$ is defined on $[0, \zeta(p)/c)$ where it is finite, increasing and has positive second derivative. Compute now

$$\frac{\partial}{\partial y} F(y)|_{y=1} = c\mathbf{E}|C| = \frac{c}{c^{cr}}. \tag{2.77}$$

Hence, if $c < c^{cr}$ then there exists $z > 1$ such that there is a finite solution y to (2.76), and therefore (2.74) also has at least one solution for some $z > 1$. Taking into account condition (2.67), we find that $D_z(c',p')$ is also finite for some $z > 1$, which confirms (2.72) and therefore completes the proof of Lemma 2.2. □

Now we are ready to complete the proof of (1.4), following almost the same arguments as in the proof of the previous lemma. Let $S_N(x) = \sum_{X_i \in \tau_N(x)} |X_i|$ denote the number of vertices from $B(N)$ which compose the macro-vertices of $\tau_N(x)$. Denote

$$\mathcal{B}'_N := \mathcal{B}_N \cap \left(C_1 \left(\widetilde{G}_N(\mathbf{X}, p, c) \right) \leq |B(N)|^{1/2} \right).$$

According to (2.59) and Lemma 2.2 we have

$$\mathbf{P}\{\mathcal{B}'_N\} = 1 - o(1).$$

This allows us to derive from (2.56)

$$\mathbf{P}\left\{C_1\Big(G_N(p,c)\Big) > aw\right\} \le \mathbf{P}\left\{\max_{1\le i\le K_N} S_N(X_i) > aw \mid \mathcal{B}'_N\right\} + o(1) \tag{2.78}$$

$$\le |B(N)|\Big(\delta + \mathbf{E}(|C|^{-1})\Big)\sum_{k=1}^{|B(N)|}(\mu(k) + \varepsilon k^{\nu}\widetilde{\mu}(k))\mathbf{P}\{S_N(k) > aw \mid \mathcal{B}'_N\} + o(1).$$

Let now $S^{c,p}(y)$ denote the sum of types including the one of the initial particle, in the total progeny of the introduced above (Section 2.2) branching process starting with initial particle of type y. Repeating the same argument which led to (2.69), we get the following bound using the introduced branching process:

$$\mathbf{P}\{S_N(k) > aw \mid \mathcal{B}'_N\}$$

$$\le \left(1 + \gamma\left(\max_{x,y\le 2\log|B(N)|/\zeta(p)} \widetilde{p}_{xy}(N)\right)^2\right)^{b_1|B(N)|\sqrt{|B(N)|}} \mathbf{P}\left\{S^{c',p'}(k) > aw\right\}$$

as $N \to \infty$, where we take into account that we can perform at most $\sqrt{|B(N)|}$ steps of exploration (the maximal possible number of macro-vertices in any $\widetilde{L}$ conditioned on $\mathcal{B}'_N$). This together with (2.63) implies

$$\mathbf{P}\{\tau_N(k) > aw \mid \mathcal{B}'_N\} \le (1 + o(1))\mathbf{P}\left\{S^{c',p'}(k) > aw\right\} \tag{2.79}$$

as $N \to \infty$. Substituting the last bound into (2.78) we derive

$$\mathbf{P}\left\{C_1\Big(G_N(p,c)\Big) > aw\right\} \le b|B(N)|\sum_{k=1}^{|B(N)|} k\mu_{p'}(k)\mathbf{P}\left\{S^{c',p'}(k) > aw\right\} + o(1) \tag{2.80}$$

as $N \to \infty$, where b is some positive constant. Now similar to (2.71) we derive from (2.80) with $w = \log|B(N)|$

$$\mathbf{P}\left\{C_1\Big(G_N(p,c)\Big) > a\log|B(N)|\right\} \le b|B(N)|z^{-a\log|B(N)|}A_z(c',p') + o(1), \tag{2.81}$$

where

$$A_z(c,p) = \sum_{k=1}^{\infty} k\mu_p(k)\mathbf{E}z^{S^{c,p}(k)}.$$

It is clear that if $A_z(c',p') < \infty$ for some $z > 1$, than for any $a > \frac{1}{\log z}$ the right-hand part of (2.81) goes to zero as $N \to \infty$. Therefore we shall search for all $z \ge 1$ for which the series $A_z(c,p)$ converge.

Proposition 2.5. *Let $0 \le p < p_c$ and $c < c^{cr}(p)$. Then $A_z(c,p) < \infty$ for all*

$$0 \le z \le z_0 := \exp\{c(1 + y_0 - \mathbf{E}\, e^{c|C|y_0})\},$$

where

$$y_0 = \begin{cases} y, & \text{if } \mathbf{E}\, c|C|e^{|C|\zeta(p)} \ge 1, \\ \zeta(p)/c, & \text{otherwise,} \end{cases}$$

Proof. Denote $g_z(k) = \mathbf{E}z^{S^{c,p}(k)}$. Function $g_z(k)$ (as a generating function for a certain branching process) satisfies the following equation

$$\begin{aligned} g_z(k) &= z^k \exp\left\{\sum_{x=1}^{\infty} \varkappa_{c,p}(k,x)\mu_p(x)(g_z(x) - 1)\right\} \\ &= z^k \exp\left\{ck(\varkappa(p)A_z(c,p) - 1)\right\}. \end{aligned}$$

Multiplying both sides by $k\mu_p(k)$ and summing up over k we find

$$A_z(c,p) = \sum_{k=1}^{\infty} k\mu_p(k)z^k \exp\left\{ck(\varkappa(p)A_z(c,p) - 1)\right\}.$$

Denoting for simplicity $A_z = A_z(c,p)$, we can rewrite the last equation as follows:

$$A_z = \sum_{k=1}^{\infty} k\mu_p(k)z^k \exp\left\{ck(\varkappa(p)A_z - 1)\right\}. \tag{2.82}$$

It follows from here (and the fact that $A_z \ge 1/\varkappa(p) = A_1$ for all $z \ge 1$) that if there exists $z > 1$ for which the series A_z converge, it should satisfy by (1.1)

$$z < e^{\zeta(p)}. \tag{2.83}$$

According to (2.82), as long as A_z is finite it satisfies the equation

$$A_z = (\varkappa(p))^{-1}\mathbf{E}\left(z^{|C|}\, e^{c|C|(\varkappa(p)A_z - 1)}\right),$$

which implies that A_z is finite for some $z > 1$ if and only if the last equation has at least one solution

$$A_z \ge A_1 = 1/\varkappa(p). \tag{2.84}$$

Let us fix $z > 1$ and consider equation

$$y = \mathbf{E}\left(z^{|C|}\, e^{c|C|(y-1)}\right) \tag{2.85}$$

for $y > 1$. It is easy to check that at least for some $y > 1$ and $z > 1$ function

$$f(y,z) := \mathbf{E}\left(z^{|C|}\, e^{c|C|(y-1)}\right)$$

is increasing, it has all the derivatives of the second order, and $\frac{\partial^2}{\partial y^2} f(y,z) > 0$. Compute now

$$\frac{\partial}{\partial y} f(y,z)|_{y=1,z=1} = c\mathbf{E}|C| = \frac{c}{c^{cr}}. \tag{2.86}$$

Hence, if $c > c^{cr}$ there is no solution $y \geq 1$ to (2.85) for any $z > 1$. On the other hand, if $c < c^{cr}$ then there exists $1 < z_0 < e^{\zeta(p)}$ such that for all $1 \leq z \leq z_0$ there is a finite solution $y \geq 1$ to (2.85).

Let us rewrite (2.85) as follows. Set

$$a = \frac{1}{c} \log z,$$

then (2.85) is equivalent to

$$y = \mathbf{E}\, e^{c|C|(y-1+a)}, \tag{2.87}$$

which after the change $x = y - 1 + a$ becomes

$$x + 1 - a = \mathbf{E}\, e^{c|C|x}. \tag{2.88}$$

Here on the right we have a convex function with a positive second derivative (for all $x < \zeta(p)/c$). Hence, function

$$\frac{\partial}{\partial x} \mathbf{E}\, e^{c|C|x} = \mathbf{E}\left(c|C| e^{c|C|x}\right)$$

is strictly increasing in x when $x < \zeta(p)/c$, and continuous on the left at $x = \zeta(p)/c$. By the assumption

$$\frac{\partial}{\partial x} \mathbf{E}\, e^{c|C|x}\, |_{x=0} = \mathbf{E}\, c|C| < 1.$$

Hence, if

$$\frac{\partial}{\partial x} \mathbf{E}\, e^{c|C|x}\, |_{x=\zeta(p)/c} = \mathbf{E}\, c|C| e^{|C|\zeta(p)} \geq 1,$$

then there exists unique $y > 0$ such that

$$\frac{\partial}{\partial x} \mathbf{E}\, e^{c|C|x}\, |_{x=y} = \mathbf{E}\left(c|C| e^{c|C|y}\right) = 1. \tag{2.89}$$

Define now

$$y_0 = \begin{cases} y, & \text{if } \mathbf{E}\, c|C| e^{|C|\zeta(p)} \geq 1, \\ \zeta(p)/c, & \text{otherwise,} \end{cases}$$

and set

$$a_0 := 1 + y_0 - \mathbf{E}\, e^{c|C|y_0}, \tag{2.90}$$

which is strictly positive due to the preceeding argument. (Notice that function $x+1-a_0$ is tangent to $\mathbf{E}\, e^{c|C|_x}$.) Hence, for all $a \leq a_0$ equation (2.88) has at least one solution, which implies due to (2.87) that for all

$$z \leq z_0 := e^{ca_0} = \exp\{c(1 + y_0 - \mathbf{E}\, e^{c|C|_{y_0}})\} \tag{2.91}$$

equation (2.85) has also at least one finite solution $y > 1$. This yields in turn that A_z is finite for all $z \leq z_0$. This finishes the proof of Proposition 2.5. □

Now taking into account that $c' > c$ and $p' > p$ can be chosen arbitrarily close to c and p, respectively, the statement (1.4) of Theorem 1.1 follows by (2.81) and Proposition 2.5. □

To conclude this section we comment on the methods used here. It is shown in (9) that in the subcritical case of the classical random graph model $G_{n,c/n}$ (i.e., $p = 0$ in terms of our model) the same method of generating functions leads to a constant which is exactly $\alpha(0, c)$ (see (1.9)). The last constant is known to be the principal term for the asymptotics of the size of the largest component (scaled to $\log n$) in the subcritical case. This gives us hope that the constant $\alpha(p, c)$ is close to the optimal one also for $p > 0$.

Similar methods were used in (10) for some class of inhomogeneous random graphs, and in (1) for a general class of models. Note, however, some difference with the approach in (1). It is assumed in (1), Section 12, that the generating function for the corresponding branching process with the initial state k (e.g., our function $g_z(k)$, $k \geq 1$) is bounded uniformly in k. As we prove here this condition is not always necessary: we need only convergence of the series A_z, while $g_z(k)$ is unbounded in k in our case. Furthermore, our approach allows one to construct constant $\alpha(p, c)$ as a function of the parameters of the model.

2.5 Proof of Theorem 1.1 in the supercritical case.

Let $\mathcal{C}_k$ denote the set of vertices in the k-th largest component in graph $G_N(p, c)$, and conditionally on $\mathbf{X}$ let $\widetilde{\mathcal{C}}_k$ denote the set of macro-vertices in the k-th largest component in graph $\widetilde{G}_N(\mathbf{X}, p, c)$ (ordered in any way if there are ties). Let also C_k and $\widetilde{C}_k$ denote correspondingly, their sizes. According to our construction for any connected component $\widetilde{L}$ in $\widetilde{G}_N(\mathbf{X}, p, c)$ there is a unique component L in $G_N(p, c)$ such that they are composed of the same vertices from $B(N)$, i.e., in the notations (2.10)

$$L = \cup_{X_i \in \widetilde{L}} \cup_{z \in X_i} \{z\} =: V(\widetilde{L}).$$

Next we prove that with a high probability the largest components in both graphs consist of the same vertices.

Lemma 2.3. *For any* $0 \le p < p_c$ *if* $c > c^{cr}(p)$ *then*

$$\mathbf{P}\{\mathcal{C}_1 = V(\widetilde{\mathcal{C}}_1)\} = 1 - o(1) \tag{2.92}$$

as $N \to \infty$.

Proof. In a view of the argument preceeding this lemma we have

$$\mathbf{P}\{\mathcal{C}_1 \neq V(\widetilde{\mathcal{C}}_1)\} = \mathbf{P}\{\mathcal{C}_1 = V(\widetilde{\mathcal{C}}_k) \text{ for some } k \ge 2\}.$$

According to Theorem 12.6 from (1), conditions of which are satisfied here, in the supercritical case conditionally on K_N such that $K_N/|B(N)| \to \mathbf{E}(|C|^{-1})$, we have **whp** $\widetilde{C}_2 = O(\log(K_N))$, which by (2.2) implies $\widetilde{C}_2 = O(\log|B(N)|)$ **whp**. Also we know already from (2.19) that in the supercritical case $\widetilde{C}_1 = O(|B(N)|)$ **whp**, and therefore $C_1 = O(|B(N)|)$ **whp**. Hence, for some positive constants a and b

$$\mathbf{P}\{\mathcal{C}_1 \neq V(\widetilde{\mathcal{C}}_1)\} = \mathbf{P}\{\mathcal{C}_1 = V(\widetilde{\mathcal{C}}_k) \text{ for some } k \ge 2\} \tag{2.93}$$

$$\le \mathbf{P}\left\{\left(\max_{k\ge 2}|V(\widetilde{\mathcal{C}}_k)| > b|B(N)|\right) \cap \left(\max_{k\ge 2}\widetilde{C}_k < a\log|B(N)|\right)\right\} + o(1).$$

It follows from (2.25) that

$$\mathbf{P}\{\max_{1\le i\le K_N}|X_i| \ge \sqrt{|B(N)|}\} = o(1)$$

as $N \to \infty$. Now we derive

$$\mathbf{P}\left\{\left(\max_{k\ge 2}|V(\widetilde{\mathcal{C}}_k)| > b|B(N)|\right) \cap \left(\max_{k\ge 2}\widetilde{C}_k < a\log|B(N)|\right)\right\} \tag{2.94}$$

$$\le \mathbf{P}\left\{\left(\max_{k\ge 2}|V(\widetilde{\mathcal{C}}_k)| > b|B(N)|\right) \cap \left(\max_{k\ge 2}\widetilde{C}_k < a\log|B(N)|\right) \cap \left(\max_{1\le i\le K_N}|X_i| < \sqrt{|B(N)|}\right)\right\}$$

$$+o(1) \ \le \mathbf{P}\left\{\sqrt{|B(N)|}\, a\log|B(N)| > b|B(N)|\right\} + o(1) = o(1).$$

Substituting this bound into (2.93) we immediately get (2.92). □

Conditionally on $\mathcal{C}_1 = V(\widetilde{\mathcal{C}}_1)$ we have

$$\begin{aligned}\tfrac{C_1}{|B(N)|} &= \tfrac{1}{|B(N)|}\textstyle\sum_{i=1}^{K_N}|X_i|\mathbf{1}\{X_i \in \widetilde{\mathcal{C}}_1\} \\ &= \tfrac{1}{|B(N)|}\textstyle\sum_{i=1}^{K_N}\sum_{k=1}^{|B(N)|} k\mathbf{1}\{|X_i| = k\}\mathbf{1}\{X_i \in \widetilde{\mathcal{C}}_1\} \\ &= \tfrac{K_N}{|B(N)|}\textstyle\sum_{k=1}^{|B(N)|} k\tfrac{1}{K_N}\#\{X_i \in \widetilde{\mathcal{C}}_1 : |X_i| = k\}.\end{aligned} \tag{2.95}$$

Note that Theorem 9.10 from (1) (together with (2.2) in our case) implies that

$$\nu_N(k) := \frac{1}{K_N}\#\{X_i \in \widetilde{\mathcal{C}}_1 : |X_i| = k\} \xrightarrow{P} \rho(k)\mu(k) \tag{2.96}$$

for each $k \geq 1$ as $N \to \infty$, where $\rho(k)$ is the maximal solution to (2.17).

We shall prove below that also

$$W_N := \sum_{k=1}^{|B(N)|} k\nu_N(k) \xrightarrow{P} \sum_{k=1}^{\infty} k\rho(k)\mu(k) =: \beta\ \left(\mathbf{E}(|C|^{-1})\right)^{-1} \tag{2.97}$$

as $N \to \infty$. Observe that according to (2.17) constant β defined above is the maximal solution to

$$\beta \equiv \mathbf{E}(|C|^{-1}) \sum_{k=1}^{\infty} k\rho(k)\mu(k) = \mathbf{E}(|C|^{-1}) \sum_{k=1}^{\infty} k \left(1 - e^{-\sum_{y=1}^{\infty} \varkappa(k,y)\mu(y)\rho(y)}\right) \mu(k)$$

$$= 1 - \mathbf{E}\Big(e^{-c|C|\beta}\Big).$$

This proves that β is the maximal root of (1.6). Then (2.97) together with (2.2) will allow us to derive from (2.95) that for any positive ε

$$\mathbf{P}\Big\{\Big|\frac{C_1\Big(G_N(p,c)\Big)}{|B(N)|} - \beta\Big| > \varepsilon \mid \mathcal{C}_1 = V(\widetilde{\mathcal{C}}_1)\Big\} = \mathbf{P}\Big\{\Big|\frac{K_N}{|B(N)|} W_N - \beta\Big| > \varepsilon\Big\} \to 0$$

as $N \to \infty$. This combined with Lemma 2.3 would immediately imply

$$\frac{C_1\Big(G_N(p,c)\Big)}{|B(N)|} \xrightarrow{P} \beta, \tag{2.98}$$

and hence the statement (1.5) of the theorem follows.

Now we are left with proving (2.97). For any $1 \leq R < |B(N)|$ write $W_N := W_N^R + w_N^R$, where

$$W_N^R := \sum_{k=1}^{R} k\nu_N(k), \quad w_N^R := \sum_{k=R+1}^{|B(N)|} k\nu_N(k).$$

By (2.96) we have for any fixed $R \geq 1$

$$W_N^R \xrightarrow{P} \sum_{k=1}^{R} k\rho(k)\mu(k) \tag{2.99}$$

as $N \to \infty$. Consider w_N^R. Note that for any $k \geq 1$

$$\nu_N(k) \leq \frac{1}{K_N} \sum_{i=1}^{K_N} I_k(|X_i|) = \frac{\mathcal{N}_k}{K_N}. \tag{2.100}$$

Using events $\mathcal{A}_{\delta,N}$ defined in (2.26), we obtain from (2.100) for any fixed $0 < \delta < \mathbf{E}(|C|^{-1})/2$ and $k \geq 1$

$$\mathbf{E}\nu_N(k) \leq \mathbf{E}\left(\frac{\mathcal{N}_k}{K_N}\mathbf{1}\{\mathcal{A}_{\delta,N}\}\right) + \mathbf{E}\left(\frac{\mathcal{N}_k}{K_N}\mathbf{1}\{\overline{\mathcal{A}_{\delta,N}}\}\right)$$

$$\leq \frac{\mathbf{E}\mathcal{N}_k}{\left(\mathbf{E}(|C|^{-1}) - \delta\right)|B(N)|} + |B(N)|\mathbf{P}\{\overline{\mathcal{A}_{\delta,N}}\}.$$

Substituting bound (2.40) into the last formula we obtain

$$\mathbf{E}\nu_N(k) \leq \frac{\varkappa(p)\mu(k)|B(N)| + |\partial B(N)|\widetilde{\mu}(k)}{\left(\mathbf{E}(|C|^{-1}) - \delta\right)|B(N)|} + |B(N)|\mathbf{P}\{\overline{\mathcal{A}_{\delta,N}}\}.$$

Bound (2.27) allows us to derive from here that

$$\mathbf{E}\nu_N(k) \leq A_1(\mu(k) + \widetilde{\mu}(k) + e^{-a_1|B(N)|}) \tag{2.101}$$

for some positive constants A_1 and a_1 independent of k and N. This together with (2.48) yields

$$\mathbf{E}w_N^R = \sum_{k=R+1}^{|B(N)|} k\mathbf{E}\nu_N(k) \leq A_2 e^{-a_2 R} \tag{2.102}$$

for some positive constants A_2 and a_2.

Clearly, for any $\varepsilon > 0$ we can choose R_0 so that for all $R \geq R_0$

$$\sum_{k=R+1}^{\infty} k\rho(k)\mu(k) < \varepsilon/3,$$

and then we have

$$\mathbf{P}\{|W_N - \sum_{k=1}^{\infty} k\rho(k)\mu(k)| > \varepsilon\} \tag{2.103}$$

$$= \mathbf{P}\{|(W_N^R - \sum_{k=1}^{R} k\rho(k)\mu(k)) + w_N^R - \sum_{k=R+1}^{\infty} k\rho(k)\mu(k)| > \varepsilon\}$$

$$\leq \mathbf{P}\{|W_N^R - \sum_{k=1}^{R} k\rho(k)\mu(k)| > \varepsilon/3\} + \mathbf{P}\{w_N^R > \varepsilon/3\}.$$

Markov's inequality together with bound (2.102) gives us

$$\mathbf{P}\{w_N^R > \varepsilon/3\} \le \frac{3\mathbf{E}w_N^R}{\varepsilon} \le \frac{3A_2 e^{-a_2 R}}{\varepsilon}. \tag{2.104}$$

Making use of (2.104) and (2.99) we immediately derive from (2.103)

$$\mathbf{P}\{|W_N - \sum_{k=1}^{\infty} k\rho(k)\mu(k)| > \varepsilon\} \le o(1) + \frac{3A_2 e^{-a_2 R}}{\varepsilon} \tag{2.105}$$

as $N \to \infty$. Hence, for any given positive ε and ε_0 we can choose finite R so large that

$$\lim_{N\to\infty} \mathbf{P}\{|W_N - \sum_{k=1}^{\infty} k\rho(k)\mu(k)| > \varepsilon\} < \varepsilon_0. \tag{2.106}$$

This clearly proves statement (2.97), and therefore finishes the proof of the theorem. □

Acknowledgment T.T. thanks MSRI for the hospitality at the beginning of this project.

Bibliography

[1] B. Bollobás, S. Janson and O. Riordan, The phase transition in inhomogeneous random graphs. *Random Structures and Algorithms*, **31** 3-122.

[2] F. Chung and L. Lu , The volume of the giant component of a random graph with given expected degree. SIAM *J. Discrete Math.* **20** (2006), 395-411

[3] P. Erdös and A. Rényi, On the evolution of random graphs. Publ. Math. Inst. Hung. Acad. Sci. 5 (1960), 17–61.

[4] G. Grimmett, Percolation. Springer-Verlag, Berlin, 1999.

[5] S. Janson, On a random graph related to quantum theory. Combin. Probab. Comput., to appear. arXiv:math.PR/0606454

[6] S. Janson, T. Łuczak, and A. Ruciński, Random graphs, Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[7] V.A. Malyshev, Random graphs and grammars on graphs. Discrete Math. Appl. 8 (1998), no. 3, 247–262.

[8] M. Talagrand, Concentration of measure and isoperimetric inequalities in product spaces. *Inst. Hautes Etudes Sci. Publ. Math.* **81**, 73-205.

[9] T.S. Turova, Note on the random graphs in the subcritical case. Dynamical systems from number theory to probability - 2, (ed. A.Yu. Khrennikov), Växjö University Press (2003), 187–192.

[10] T.S. Turova, Phase transitions in dynamical random graphs, Journal of Statistical Physics, 123 (2006), no. 5, 1007-1032.

[11] T.S. Turova and T. Vallier, Merging percolation and random graphs: Phase transition in dimension 1. arXiv:math.PR/0609594

[12] T. Turova and A. Villa, On a phase diagram for random neural networks with embedded spike timing dependent plasticity. *BioSystems* Volume **89**,Issues 1-3,280-286.

[13] D.J. Watts and S.H. Strogatz, Collective dynamics of "small-world" networks. Nature, 393 (1998), 440–442.

[14] Y. Zhang, A Martingale Approach in the Study of Percolation Clusters on the Z^d Lattice. *Journal of Theoretical Probability* **14** (2001) no. 1, 165-187

D

Chapter D

Spread of Activation on Random Graphs

T. Vallier

Centre for Mathematical Sciences, Lund University, Sweden

Abstract

We consider the spread of activation in the Erdös-Rényi random graph $G_{n,p}$ from a fixed set of activated vertices $A_n(0)$ with $|A_n(0)|$ depending on n. Any vertex which is linked to at least 2 activated vertices becomes activated. We describe and analyze the process of activation. In particular, when $p = \frac{c}{n}$ and $|A_n(0)| = o(n)$, the activation does not spread through a positive part of the vertices (as $n \to \infty$). When $\frac{1}{p} = o(n)$ we show that, depending on $|A_n(0)|$ the total number of activated vertices varies from $o(n)$ to $n - o(n)$. This indicates existence of a phase transition along different parameters of the model.

Key words: Classical random graphs, phase transition, contact process.

1 Introduction

In the last decade a number of models of random growth were introduced in probability to study natural phenomena. Here we study a marginal case of a network designed to study the spread of activity in a neural network and introduced in (5) . A special feature of such a network is that a neuron often needs several incoming impulses to be activated. To model this effect we introduce the following model.

Let $G_{n,p}$ denote an Erdös-Rényi random graph on the set of vertices $V_n = \{1, ..., n\}$ where the edges between any two different vertices i and j are present independently with probabilities

$$p_{ij}(n) = p(n).$$

Here we denote (i, j) as an edge between i and j in G. The set V_n is separated in two types of vertices, the activated and the non activated. Let all the vertices with indices in

$$A_n(0) \subseteq V_n,\ |A_n(0)| \geq 2.$$

be activated.

Any vertex v which is connected to at least two vertices in $A_n(0)$ becomes activated and is a vertex of the first generation which we denote $v \in B_n^1$. The vertex v remains activated and now contributes to the activation of other vertices through its edges. At each step, any non activated vertex becomes activated if it is connected to at least 2 activated vertices. Since the original set of activated vertices $A_n(0)$ is not necessarily connected, the maximal set obtained along the process of activation which we denote $A_n(n)$ does not need to be connected either. Thus $|A_n(n)|$ has no connection with the size of the largest connected component. However, any vertex of $A_n(n)$ is connected through a path to a vertex of $A_n(0)$. Consider during the process of activation the link between activated and non activated vertices as directed in the sense of activation, a vertex v is in the k^{th} generation, $v \in B^k$ if the longest path through directed edges from $A_n(0)$ to v has length k.

This model resembles contact processes (4), models of random growth (e.g., (2)). It fits a general definition of a probabilistic cellular automata (1) (which are known to be extremely difficult to analyse in a full generality), and is rather close to a "semi-totalistic" mean-field model also defined in (1). However in the "semi-totalistic" mean-field model the neighbourhood is deterministic while in this paper the neighbourhood is defined through the links in the random graph.

2 Results

Theorem 2.1. *Let $p(n) \geq \frac{c}{n}$ for some $c > 0$ arbitrarily fixed with $\lim_{n\to\infty} \sqrt{n}p(n) = 0$. If $\lim_{n\to\infty} |A_n(0)|np^2(n) = 0$ then*

$$|A_n(n)|np^2(n) \xrightarrow{P} 0. \tag{2.1}$$

Corollary 2.1. *Under conditions of Theorem 2.1*

$$\frac{|A_n(n)|}{n} \xrightarrow{P} 0. \tag{2.2}$$

Theorem 2.2. *Let* $p(n) = \frac{\omega(n)}{n}$ *for some* $\omega(n)$ *such that* $\lim_{n\to\infty} \omega(n) = \infty$ *and* $\lim_{n\to\infty} p(n)\sqrt{n} = 0$. *If* $|A(0)| \geq \frac{1+\nu}{np^2(n)}$ *where* $\nu > 0$ *then*

$$\frac{|A_n(n)|}{n} \xrightarrow{P} 1. \tag{2.3}$$

Corollary 2.1 tells us that starting with a set of vertices of size $|A_n(0)| = o\big(\frac{1}{np^2(n)}\big)$ as $n \to \infty$, with a probability tending to 1 as n goes to infinity, the activation doesn't spread through a positive part of the graph. On the other hand, if we start with a set of size at least $\frac{1+\nu}{np^2(n)}$ then the size of the set of activated vertices rescaled by $\frac{1}{n}$ converges in probability to one. (We activate almost all vertices.) The change occurs when $|A_n(0)| \sim \frac{1}{np^2(n)}$. We study the case $p(n) = \frac{c}{\sqrt{n}}$ separately where $c > 0$ for which the range of the threshold is of order constant.

Theorem 2.3. *Let* $p(n) = \frac{c}{\sqrt{n}}$.

1. *If* $|A_n(0)| = k \geq 2$ *then there exists* $\zeta(k,c) > 0$ *such that for any* $\varepsilon > 0$

$$1 > \lim_{n\to\infty} \mathbb{P}\left\{\frac{|A_n(n)|}{n} > 1-\varepsilon\right\} \geq \zeta(k,c). \tag{2.4}$$

2. *If* $|A_n(0)| = \alpha(n)$ *where* $\alpha(n)$ *is such that* $lim_{n\to\infty}\alpha(n) = \infty$ *then*

$$\frac{|A_n(n)|}{n} \xrightarrow{P} 1. \tag{2.5}$$

If we start with a finite number of vertices with a positive probability lower than 1, we activate almost all the graph. If we start with a number of vertices $\alpha(n)$ where $\lim_{n\to\infty} \alpha(n) = \infty$ then with a probability tending to 1 as n goes to infinity we activate almost all the graph.

In the case when the probability of connection is much larger than $\frac{1}{\sqrt{n}}$, even starting with a finite number of vertices, the size of the set of activated vertices rescaled by $\frac{1}{n}$ converges in probability to 1.

Theorem 2.4. *Let* $p(n) = \frac{\omega(n)}{\sqrt{n}}$ *for some* $\omega(n)$ *where* $\lim_{n\to\infty} \omega(n) = \infty$. *If* $|A(0)| = k \geq 2$ *then*

$$\frac{|A(n)|}{n} \xrightarrow{P} 1. \tag{2.6}$$

3 Basics

3.1 The algorithm

To analyse the size of $A_n(n)$ we shall introduce the following algorithm which uncovers the vertices of $A_n(n)$. This algorithm can be compared with the well-known algorithm of finding a connected component in a random graph (see, e.g., (3)).

On a graph $G \in G_{n,p}$ we construct a process $\big(A_n(t), \mathcal{L}(t)\big), t = 0, ..., n$, as follows.

The set V_n is separated in two types, the activated vertices and the non activated vertices. Let $A_n(t)$ denote the set of activated vertices at time $t \in \{0, 1, ...\}$. We start with

$$A_n(0) \subseteq V_n, \ |A_n(0)| \geq 2.$$

The spread of activation is defined according to the following algorithm.

- Step 0. Let u_0 be a vertex chosen uniformly at random among the vertices of $A_n(0)$. Denote this distribution as $u_0 \in U\big(A_n(0)\big)$.

  1. Find all the neighbours of u_0 in G (i.e. the vertices connected to u_0 in G). Call that set $N(u_0)$.
  2. Set a directed edge from the vertex u_0 to any vertex $u \in N(u_0)$ and denote it $(\overrightarrow{u_0, u})$. Call $\mathcal{L}_n(0) = \{(\overrightarrow{u_0, u}), u \in N(u_0)\}$ the set of directed edges at time 0.
  3. Mark u_0 as saturated. Denote $S_n(0) = \{u_0\}$ the set of saturated vertices at time 0. Notice that the set of saturated vertices will always be included in the set of activated vertices

  $$S_n(t) \subseteq A_n(t) \text{ for any } t \geq 0.$$

- Step 1. Let $u_1 \in U\big(A_n(0) \setminus S_n(0)\big)$.

  1. Find all the neighbours of u_1 in the graph G among the vertices $V_n \setminus A_n(0)$. Call them $N(u_1) = \{u \in V_n \setminus A_n(0), (u_1, u) \in G\}$.
  2. Set directed edges $(\overrightarrow{u_1, u})$ from u_1 to $u \in N(u_1)$ and denote $\Delta\mathcal{L}_n(1) = \{(\overrightarrow{u_1, u}), u \in N(u_1)\}$ the set of these edges. Then

  $$\mathcal{L}_n(1) = \mathcal{L}_n(0) \cup \Delta\mathcal{L}_n(1),$$

  is the set of directed edges at time 1.
  3. Mark u_1 as saturated and set

  $$S_n(1) = S_n(0) \cup \{u_1\}.$$

4. Find all the vertices among $N(u_1)$ with in-degree 2 (i.e, vertices u having two incoming edges in $\mathcal{L}_n(1)$). Call them $\Delta A_n(1)$. The set of activated vertices at the end of step 1 is
$$A_n(1) = A_n(0) \cup \Delta A_n(1).$$

Up to step $k = |A_n(0)| - 1$, we choose $u_k \in U\big(A_n(0) \setminus S_n(k-1)\big)$ and follow the same procedure as above, setting

$$\mathcal{L}_n(k) = \mathcal{L}_n(k-1) \cup \Delta\mathcal{L}_n(k),$$
$$S_n(k) = S_n(k-1) \cup \{u_k\},$$
$$A_n(k) = A_n(k-1) \cup \Delta A_n(k),$$

Then at time $k_0 = |A_n(0)|$ we have

$$S_n(k_0) = A_n(0). \tag{3.7}$$

Call
$$B_n^1 = A_n(k_0) \setminus A_n(0),$$
the first generation of activated vertices. We shall say that $A_n(0)$ is the generation 0 of activated vertices.

We now define recursively further generations as follows. Suppose we constructed the $l \geq 1$ first generations and at time $k_l = |A_n(0)| + \sum_{i=1}^{l} |B_n^i|$ we have (as in (3.7))

$$S_n(k_l) = A_n(0) \cup \bigcup_{i=1}^{l} B_n^i.$$

Call
$$B_n^{l+1} = A_n(k_l) \setminus \left\{ A_n(0) \cup \bigcup_{i=1}^{l} B_n^i \right\}$$
the generation $l+1$ of activated vertices. We can now define the algorithm at any time k.

- Step $k \geq |A_n(0)|$. Let $u_k \in U(B_n^j \setminus S_n(k))$ where $j = \min\{i \geq 1 : B_n^i \setminus S_n(k-1) \neq \emptyset\}$.

  1. Find all the neighbours of u_k in G among the vertices of $V_n \setminus A_n(k-1)$ and call this set $N(u_k)$.
  2. Set directed edges $(\overrightarrow{u_k, u})$ from u_k to $u \in N(u_k)$ and denote $\Delta\mathcal{L}_n(k) = \{(\overrightarrow{u_k, u}), u \in N(u_k)\}$ the set of these edges. Define
  $$\mathcal{L}_n(k) = \mathcal{L}_n(k-1) \cup \Delta\mathcal{L}_n(k),$$
  which is the set of directed edges at time k.

3. Mark u_k as saturated and set

$$S_n(k) = S_n(k-1) \cup \{u_k\}.$$

4. Find all the vertices among $N(u_k)$ with in-degree 2. Call them $\Delta A_n(k)$. The set of activated vertices at the end of the step k is

$$A_n(k) = A_n(k-1) \cup \Delta A_n(k).$$

We denote $A_n(n)$ the limiting set of activated vertices and $\mathcal{L}_n(n)$ the limiting set of directed between 2 vertices of $A_n(n)$. The graph of activated vertices is thus

$$G_a(n) = \big(A_n(n), \mathcal{L}_n(n)\big)$$

The algorithm stops after step $k \leq n$ in two different cases: either all the vertices are activated i.e.,

$$|A_n(k)| = n, \tag{3.8}$$

or

$$A_n(k) = S_n(k), \tag{3.9}$$

in which case we have explored the neighbourhood of all the activated vertices and $A_n(n) = A_n(k)$.

To simplify the notations, we will only write $A(k)$ for $A_n(k)$ omitting the index n.

The limiting set of activated vertices $A(n)$ is function of the given graph G and the original set of activated vertices $A(0)$ where $A(0)$ is chosen randomly. The order in which we choose the vertices in the algorithm has no incidence on $A(n)$. Similarly as in (1), the only information of relevance is the number of activated and saturated vertices at time t in the algorithm.

The following lemma implies that starting with a set $A(0)$ of activated vertices and picking the vertex u_t in any order $\mathcal{O} = (u_0, u_1, ...)$ among the activated vertices at time t, we generate the same set of activated vertices A_n.

Lemma 3.1. *Let $A(0) \subseteq V$ and let $A(n)$ and $A'(n)$ be the limiting set of activated vertices in the algorithm with an order $\mathcal{O}$ and respectively, $\mathcal{O}'$ on a graph $G \in G_{n,p}$ then*

$$A(n) = A'(n).$$

Proof of lemma 3.1. By definition a vertex $v \in A(n)$ if there exists a path of directed edges from a vertex of $A(0)$ to v:

$$A(n) = \Big\{ v \in V : \exists u \in A(0), \exists u_1, ..., u_i \in V, (\overrightarrow{u, u_i}), (\overrightarrow{u_i, u_{i-1}}), ..., (\overrightarrow{u_2, u_1}), (\overrightarrow{u_1, v}) \Big\}.$$

Suppose that there exists a vertex $v \in A(n)$ such that $v \notin A'(n)$. Since $v \in A(n)$ there exists two vertices u_1 and $u_1' \in A(n)$ such that $(\overrightarrow{u_1, v}) \in \mathcal{L}(n)$ and $(\overrightarrow{u_1', v}) \in \mathcal{L}(n)$ where $\mathcal{L}(n)$is the set of directed edges generated along the algorithm.

If $u_1 \in A'(n)$ and $u_1' \in A'(n)$ then $v \in A'(n)$. This implies that either $u_1 \notin A_1'$ or $u_1' \notin A_1'$. Suppose that $u_1 \notin A'(n)$, using the same argument as above, we show that at least one of the vertices u_2, u_2' with $(\overrightarrow{u_2, u_1}) \in \mathcal{L}(n)$ and $(\overrightarrow{u_2', u_1}) \in \mathcal{L}(n)$ doesn't belong to $A'(n)$. Say that $u_2 \notin A'(n)$. Recursively, we find a vertex $u_i \in B^1$ is such that $u_i \notin A(n)'$. If $(\overrightarrow{u_0, u_i}) \in \mathcal{L}(n)$ and $(\overrightarrow{u_0', u_i}) \in \mathcal{L}(n)$ then $u_0' \notin A'(n)$ or $u_0 \notin A'(n)$ which contradicts the fact that u_0' and $u_0 \in A(0) \subseteq A'(n)$. Finally, for any $v \in A(n)$ we have $v \in A'(n)$ from which it follows $A'(n) = A(n)$. □

Although unnecessary to find $|A(n)|$, we order the vertices in the algorithm in generations. This notion provides information on the spread of activation and allows us to give a more extensive picture of the graph of activated vertices. It is however meaningless to consider the limiting set of edges $\mathcal{L}(n)$ in the graph $G_a(n)$ as the edges are directly dependent of the order $\mathcal{O}$ in which we picked up the vertices in the algorithm. Suppose for example that a vertex of the first generation is the neighbour via undirected edges with at least 3 vertices of $A(0)$, say u_1, u_2, u_3. A vertex becomes activated if its in-degree is 2, we see that picking first u_1 and u_2, then $(\overrightarrow{u_3, u}) \notin \mathcal{L}(n)$ while if we choose u_2 and u_3 first then we will have $(\overrightarrow{u_3, u}) \in \mathcal{L}(n)$ but $(\overrightarrow{u_1, u}) \notin \mathcal{L}(n)$.

3.2 Probability of activation

In the following, we call probability of activation at step t the probability that a non activated vertex $u \notin A(t-1)$ becomes activated at step t in the algorithm, $u \in A(t)$.

Let $u \notin A(t-1)$. The probability that at step t the vertex u becomes activated is

$$\begin{aligned} p_a(t) &= \mathbb{P}\{u \in A(t) \setminus A(t-1)\} \\ &= \mathbb{P}\Big\{(\overrightarrow{u_t, u}) \cap \{\exists! u_i, i < t, (\overrightarrow{u_i, u})\}\Big\} \\ &= \mathbb{P}(\overrightarrow{u_t, u}) \sum_{i=0}^{t-1} \mathbb{P}\{(\overrightarrow{u_i, u})\} \prod_{j \neq i, j<t} \mathbb{P}\{(\overrightarrow{v_j, u})^c\} \\ &= p^2(n) t \big(1 - p(n)\big)^{t-1}. \end{aligned}$$

4 Proofs

4.1 Proof of Theorem 2.1

In this part, we prove Theorem 2.1 which states that if we start with a number of activated vertices $|A(0)| = \frac{1}{\alpha(n) n p^2(n)}$ where $\lim_{n\to\infty} \alpha(n) = \infty$ then with a probability tending to 1 as n goes to infinity, the size of the set of activated vertices remains in this region.

The following lemma states that if the size of $A(0)$ is sufficiently small then with a probability tending to 1 as n goes to infinity, no vertex is activated.

Lemma 4.1. *Let $p(n) \geq \frac{c}{n}$ for some $c > 0$ such that $\lim_{n\to\infty} p(n)\sqrt{n} = 0$. Let $\alpha(n)$ be a function such that $\lim_{n\to\infty} \alpha(n) = \infty$. If $|A(0)| \leq \frac{1}{\alpha(n)\sqrt{np(n)}}$ then*

$$\lim_{n\to\infty} \mathbb{P}\{|A(n)| - |A(0)| \geq 1\} = 0. \tag{4.10}$$

Proof of lemma 4.1. Suppose that $|A(0)| = \left\lfloor \frac{1}{\alpha(n)\sqrt{np(n)}} \right\rfloor$ The probability that a vertex u becomes activated at time s is $p_a(s)$ with

$$\begin{aligned} p_a(s) &= \mathbb{P}\{u \in A(s) \setminus A(s-1)\} \\ &= p^2(n)s\big(1-p(n)\big)^{s-1}. \end{aligned}$$

The number of activated vertices at time s follows a binomial law

$$Y(s) = |A(s) \setminus A(s-1)| \in Bin\big(n - |A(s-1)|, p_a(s)\big).$$

The fact that

$$A(t) = A(0) \cup \bigcup_{s=1}^{t} \{A(s) \setminus A(s-1)\},$$

for any $t \geq 1$ implies

$$|A(t)| = |A(0)| + \sum_{s=1}^{t} Y(s), \tag{4.11}$$

for any $t \geq 1$. The probability of activation of a vertex at time s is bounded from above by

$$p_a(s) \leq p^2(n)s.$$

The random variable $Y(s)$ is stochastically dominated by the random variable $U^+(s) \in Bin(n, p^2(n)s)$. Therefore

$$\mathbb{E}\Big(\sum_{s=1}^{|A(0)|-1} Y(s)\Big) \leq \mathbb{E}\Big(\sum_{s=1}^{|A(0)|-1} U^+(s)\Big).$$

with $\mathbb{E}\big(U^+(s)\big) = np^2(n)s$. We deduce that

$$\begin{aligned} \mathbb{E}\Big(\sum_{s=1}^{|A(0)|-1} Y(s)\Big) &\leq np^2(n) \sum_{s=1}^{|A(0)|-1} s \\ &\leq np^2(n)\frac{1}{2}\big(|A(0)|^2 - |A(0)|\big). \end{aligned} \tag{4.12}$$

By assumption $|A(0)| = \left\lfloor \frac{1}{\alpha(n)\sqrt{np}(n)} \right\rfloor$. Equation (4.12) implies

$$np^2(n)\frac{1}{2}\big(|A(0)|^2 - |A(0)|\big) \leq \frac{1}{\alpha^2(n)},$$

from which we derive

$$\mathbb{E}\bigg(\sum_{s=1}^{|A(0)|-1} Y(s)\bigg) \leq \frac{1}{\alpha^2(n)}. \tag{4.13}$$

Using Markov inequality

$$\mathbb{P}\{X \geq \varepsilon\} \leq \frac{\mathbb{E}X}{\varepsilon},$$

where X is a nonnegative random variable and equation (4.13) we find that

$$\lim_{n\to\infty} \mathbb{P}\bigg\{\bigg(\sum_{s=1}^{|A(0)|-1} Y(s)\bigg) \geq 1\bigg\} = 0.$$

which implies by (4.11) that

$$\lim_{n\to\infty} \mathbb{P}\Big\{|A\big(|A(0)|-1\big)| - |A(0)| \geq 1\Big\} = 0 \tag{4.14}$$

With a probability tending to one as n goes to infinity, the process of activation of the vertices stops at time $|A(0)|$ since all the activated vertices are saturated

$$\lim_{n\to\infty} \mathbb{P}\Big\{|A\big(|A(0)|\big)| - |S\big(|A(0)|-1\big)| = 0\Big\} = 1.$$

The condition (3.9) is fulfilled and

$$\lim_{n\to\infty} \mathbb{P}\Big\{|A\big(|A(0)|-1\big)| = |A(n)|\Big\} = 1. \tag{4.15}$$

Together, (4.14) and (4.15) imply equation (4.10)

$$\lim_{n\to\infty} \mathbb{P}\Big\{|A(n)| - |A(0)| \geq 1\Big\} = 0.$$

which proves Lemma 4.1 □

If $A(0)$ doesn't fulfil the conditions of Lemma 4.1 then vertices are activated with a positive probability. We show that if the number of activated vertices non-saturated is of order constant then there is no further activations with probability tending to 1 as $n \to \infty$.

Lemma 4.2. *Let $p(n) \geq \frac{c}{n}$ for some $c > 0$ such that $\lim_{n\to\infty} p(n)\sqrt{n} = 0$. Let $\alpha(n)$ be such that $\lim_{n\to\infty} \alpha(n) = \infty$, $t(n) = \frac{1}{\alpha(n)np^2(n)}$ and C a positive constant. Then*

$$\lim_{n\to\infty} \mathbb{P}\Big\{ |A(n)| - |A(t(n))| \geq 1 \Big| |A\big(t(n)\big)| - t(n) \leq C \Big\} = 0. \tag{4.16}$$

Proof of lemma 4.2. By the definition of saturated vertices in the algorithm, we have $|S\big(t(n)\big)| = t(n) + 1$ therefore the number of activated vertices which are non saturated is

$$|A\big(t(n)\big)| - |S\big(t(n)\big)| = |A\big(t(n)\big)| - t(n) - 1.$$

For

$$t(n) = \Big\lfloor \frac{1}{\alpha(n)np^2(n)} \Big\rfloor \tag{4.17}$$

we have in the condition in (4.16)

$$|A\big(t(n)\big)| = t(n) + C.$$

The expected number of vertices activated between time $t(n)$ and $|A\big(t(n)\big)| = t(n) + C$ is

$$\begin{aligned}
\mathbb{E}\Big(\sum_{s=t(n)+1}^{t(n)+C} Y(s) \Big) &\leq \mathbb{E}\Big(\sum_{s=t(n)+1}^{t(n)+C} U^+(s) \Big) \\
&\leq np^2(n) \sum_{s=t(n)+1}^{t(n)+C} s \\
&\leq np^2(n) \Big(Ct(n) + \frac{C^2 + C}{2} \Big).
\end{aligned}$$

By (4.17) we have

$$\mathbb{E}\Big(\sum_{s=t(n)+1}^{t(n)+C} Y(s) \Big) \leq \frac{C}{\alpha(n)} + np^2(n)\frac{C^2 + C}{2}.$$

By assumption $\lim_{n\to\infty} np^2(n) = 0$ and $\lim_{n\to\infty} \alpha(n) = \infty$ therefore

$$\lim_{n\to\infty} \mathbb{E}\Big(\sum_{s=t(n)+1}^{t(n)+C} Y(s) \Big) = 0. \tag{4.18}$$

Using Markov inequality and equation (4.18) we find

$$\lim_{n\to\infty} \mathbb{P}\Big\{ \sum_{s=t(n)+1}^{t(n)+C} Y(s) \geq 1 \Big\} = 0. \tag{4.19}$$

Similarly as in the proof of Lemma 4.1, (4.19) implies

$$\lim_{n\to\infty} \mathbb{P}\Big\{ \big|A\Big(|A\big(t(n)\big)|\Big)| - |A\big(t(n)\big)| \geq 1 \Big| t(n) = \Big\lfloor \frac{1}{\alpha(n)np^2(n)} \Big\rfloor \Big\} = 0 \tag{4.20}$$

This shows that with a probability tending to one as n goes to infinity, there is no further activation after time $|A\big(t(n)\big)|$. Using (4.20) we deduce that if $|A\big(t(n)\big)| = t(n) + C$ then

$$\lim_{n\to\infty} \mathbb{P}\Big\{ |A(n)| - |A\big(t(n)\big)| \geq 1 \Big| t(n) = \Big\lfloor \frac{1}{\alpha(n)np^2(n)} \Big\rfloor \Big\} = 0.$$

□

Now that we know under which condition the process of activation stops we can prove that the conditions of lemma 4.2 are fulfilled if $\lim_{n\to\infty} |A(0)|np^2(n) = 0$.

Proof of theorem 2.1. Write

$$|A(0)| = \Big\lfloor \frac{1}{\alpha(n)np^2(n)} \Big\rfloor \tag{4.21}$$

where $\alpha(n)$ is an arbitrary fixed function such that $\lim_{n\to\infty} \alpha(n) = \infty$.

We study separate cases.

A. In the case when $\lim_{n\to\infty} \frac{1}{\alpha(n)\sqrt{np}(n)} = 0$ which implies $\lim_{n\to\infty} |A(0)|\sqrt{np}(n) = 0$ then the condition of lemma 4.1 is fulfilled and

$$\lim_{n\to\infty} \mathbb{P}\{|A(n)| - |A(0)| \geq 1\} = 0.$$

B. Suppose now that there exists $c > 0$ such that $\alpha(n) \leq \frac{c}{\sqrt{np}(n)}$. The expected size of the first generation B^1 is by (4.12)

$$\begin{aligned} \mathbb{E}\big(|B^1|\big) &\leq np^2(n)\frac{1}{2}\big(|A(0)|^2 - |A(0)|\big) \\ &\leq \frac{1}{2}np^2(n)|A(0)|^2\big(1 - \frac{1}{|A(0)|}\big). \end{aligned}$$

Using (4.21) and the fact that $|A(0)| \geq 2$, we have

$$\begin{aligned} \mathbb{E}\big(|B^1|\big) &\leq \frac{1}{2}\frac{1}{\alpha(n)}|A(0)|\big(1 - \frac{1}{|A(0)|}\big) \\ &\leq \frac{1}{\alpha(n)}|A(0)|. \end{aligned} \tag{4.22}$$

B_i. Consider the subcase when

$$\alpha(n) = \frac{c}{\sqrt{np}(n)}, \tag{4.23}$$

for some $c > 0$. Equation (4.22) gives

$$\mathbb{E}\big(|B^1|\big) \leq C_1, \tag{4.24}$$

for some $C_1 > 0$. Let $\varepsilon > 0$ and let $M = \lfloor \frac{2C_1}{\varepsilon} \rfloor + 1$. Consider the event

$$\mathcal{D} = \{|B^1| \geq M\}.$$

Then by definition of M and using equation (4.24), we find

$$\mathbb{P}\{\mathcal{D}\} \leq \mathbb{P}\bigg\{|B^1| \geq \frac{\mathbb{E}\big(|B^1|\big)}{\varepsilon}\bigg\}.$$

Using Markov inequality, we have

$$\mathbb{P}\{\mathcal{D}\} \leq \varepsilon. \tag{4.25}$$

Under the condition that $|B^1| \leq M$ and by (4.21) we deduce

$$\lim_{n\to\infty} \mathbb{P}\{|B^2| \geq 1 | \mathcal{D}^c\} = 0,$$

where $\mathcal{D}^c = \{|B^1| \leq M\}$. We derive that for n large

$$\begin{aligned} \mathbb{P}\{|B^2| \geq 1\} &\leq \mathbb{P}\{|B^2| \geq 1 | \mathcal{D}^c\} + \mathbb{P}\{\mathcal{D}\} \\ &\leq 2\varepsilon. \end{aligned} \tag{4.26}$$

By the definition of the algorithm we have the equality of the following events

$$\{|A(n)| > |A(0)| + |B^1|\} = \{|B^2| \geq 1\}.$$

By (4.26) for n large

$$\mathbb{P}\{|A(n)| > |A(0)| + |B^1|\} \leq 2\varepsilon.$$

By the assumption (4.23) we have $\lim_{n\to\infty} |A_0| = \infty$ and using (4.25) we find

$$\lim_{n\to\infty} \mathbb{P}\{|B^1| > |A(0)|\} = 0,$$

which together with (4.26) implies

$$\lim_{n\to\infty} \mathbb{P}\{|A(n)| \geq 2|A(0)|\} = 0.$$

By assumption $\lim_{n\to\infty} |A(0)| np^2(n) = 0$, thus

$$\lim_{n\to\infty} \mathbb{P}\{|A(n)| np^2(n) = 0\} = 1.$$

This proves Theorem 2.1 in the case $\alpha(n) = \frac{c}{\sqrt{np(n)}}$.

B_{ii}. From now on, suppose that $\lim_{n\to\infty} \frac{\alpha(n)}{\sqrt{np(n)}} = 0$.

Define for any i such that $\lim_{n\to\infty} \left(\frac{2}{\alpha(n)}\right)^i |A(0)| = \infty$, the event

$$\mathcal{B}_i = \Big\{ |B^i| \leq \Big(\frac{2}{\alpha(n)}\Big)^i |A(0)| \Big\}.$$

We shall prove that for a certain k_0

$$\lim_{n\to\infty} \mathbb{P}\Big\{ \bigcap_{k=1}^{k_0-1} \mathcal{B}_k \Big\} = 1. \tag{4.27}$$

This would imply that

$$\begin{aligned}
\mathbb{P}\{|A(n)| \geq 2|A(0)|\} &\leq \mathbb{P}\Big\{ |A(n)| \geq 2|A(0)| \Big| \bigcap_{k=1}^{k_0-1} \mathcal{B}_k \Big\} + \mathbb{P}\Big(\Big\{ \bigcap_{k=1}^{k_0-1} \mathcal{B}_k \Big\}^c \Big) \\
&\leq \mathbb{P}\Big\{ |A(n)| \geq 2|A(0)| \Big| \bigcap_{k=1}^{k_0-1} \mathcal{B}_k \Big\} + o(1) \text{ as } n \to \infty.
\end{aligned}$$

To prove (4.27), consider the conditional expectation

$$\begin{aligned}
\mathbb{E}\Big(|B^k| \;\Big|\; |B^1|, ..., |B^{k-1}| \Big) &\leq \sum_{|A(0)|+\sum_{i=1}^{k-2}|B^i|+1}^{|A(0)|+\sum_{i=1}^{k-1}|B^i|} np^2(n)s \\
&\leq np^2(n)|B^{k-1}| \frac{1}{2}\big(2|A(0)| + ... + 2|B^{k-2}| + |B^{k-1}|\big).
\end{aligned} \tag{4.28}$$

The inequality (4.28) holds almost surely and for any $\mathcal{B}_i$, we have

$$\mathcal{B}_i \in \sigma(B^i),$$

where $\sigma(B^i)$ is the sigma algebra generated by the random variable B^i. Hence we find

$$\begin{aligned}
\mathbb{E}\Big(|B^k| \Big| \mathcal{B}_1, ..., \mathcal{B}_{k-1} \Big) &\leq np^2(n) \Big(\frac{2}{\alpha(n)}\Big)^{k-1} |A(0)|^2 \Big(1 + 2\sum_{i=1}^{k-1} \Big(\frac{2}{\alpha(n)}\Big)^i \Big) \\
&\leq \frac{1}{\alpha(n)} \Big(\frac{2}{\alpha(n)}\Big)^{k-1} \Big(1 + 2\sum_{i=1}^{k-1} \Big(\frac{2}{\alpha(n)}\Big)^i \Big) |A(0)|,
\end{aligned} \tag{4.29}$$

for any k such that $\lim_{n\to\infty}\left(\frac{2}{\alpha(n)}\right)^k|A(0)| = \infty$. Notice that for any $\varepsilon > 0$ there exists N such that for any $n \geq N$ we have

$$\sum_{i=1}^{k-1}\Big(\frac{2}{\alpha(n)}\Big)^i < 1+\varepsilon, \tag{4.30}$$

which implies

$$\mathbb{E}\Big(|B^k|\Big|\ \mathcal{B}_1,...,\mathcal{B}_{k-1}\Big) \leq \frac{2+\varepsilon}{\alpha(n)}\Big(\frac{2}{\alpha(n)}\Big)^{k-1}|A(0)|. \tag{4.31}$$

Define

$$k_0 = \min\Big\{k : \frac{|A(0)|}{\alpha^k(n)} < C_1\Big\}, \tag{4.32}$$

for some $C_1 > 0$. Consider the following

$$\mathbb{P}\Big\{\mathcal{B}_k^c\ \Big|\ \mathcal{B}_1,...,\mathcal{B}_{k-1}\Big\} = \mathbb{P}\Big\{|B^k| > \Big(\frac{2}{\alpha(n)}\Big)^k|A(0)|\Big|\mathcal{B}_1,...,\mathcal{B}_k\Big\}.$$

Using (4.31) and Chebyshev inequality we find

$$\mathbb{P}\Big\{|B^k| > \Big(\frac{2}{\alpha(n)}\Big)^k|A(0)|\Big|\mathcal{B}_1,...,\mathcal{B}_k\Big\} \leq 2\frac{\alpha^k(n)}{|A(0)|}.$$

We can find the probability for the event $\mathcal{B}_1 \cap ... \cap \mathcal{B}_{k_0-1}$

$$\mathbb{P}\Big\{\bigcap_{k=1}^{k_0-1}\mathcal{B}_k\Big\} \geq \prod_{k=1}^{k_0-1}\Big(1-2\frac{\alpha^k(n)}{|A(0)|}\Big),$$

which converges to 1 as n goes to infinity if

$$\lim_{n\to\infty}\sum_{k=1}^{k_0-1}\frac{\alpha^k(n)}{|A(0)|} = 0. \tag{4.33}$$

The sum in (4.33) is bounded from above as follows

$$\sum_{k=1}^{k_0-1}\frac{\alpha^k(n)}{|A(0)|} \leq \frac{1}{|A(0)|}2\alpha^{k_0-1}(n).$$

By definition of k_0 (see (4.32)) we have

$$\lim_{n\to\infty}\frac{1}{|A(0)|}2\alpha^{k_0-1}(n) = 0.$$

We find that

$$\lim_{n\to\infty} \mathbb{P}\Big(\Big\{\bigcap_{k=1}^{k_0-1} \mathcal{B}_k\Big\}^c\Big) = 0. \tag{4.34}$$

For $k = k_0$ there exists C such that

$$\mathbb{E}\Big(|B^{k_0}|\Big|\bigcap_{k=1}^{k_0-1} \mathcal{B}_k\Big) \leq C.$$

By (4.34) and using Markov inequality we find that for any $\varepsilon > 0$, there exists M such that we have

$$\mathbb{P}\Big\{|B^{k_0}| \geq M\Big|\bigcap_{k=1}^{k_0-1} \mathcal{B}_k\Big\} \leq \frac{\varepsilon}{2}. \tag{4.35}$$

Let $t(n) = |A(0)| + \frac{2}{\alpha(n)}|A(0)|\sum_{k=1}^{k_0}\Big(\frac{2}{\alpha(n)}\Big)^k$. Using Chebyshev inequality, equations (4.29) and (4.34) we find for any $\varepsilon_0 > 0$ and n large

$$\begin{aligned}\mathbb{P}\Big\{|A\big(|B^{k_0}|\big)| \leq t(n)\Big\} &\leq \mathbb{P}\Big\{|A\big(|B^{k_0}|\big)| \leq t(n)\Big|\bigcap_{k=1}^{k_0-1} \mathcal{B}_k\Big\} + \mathbb{P}\Big\{\bigcap_{k=1}^{k_0-1} \mathcal{B}_k\Big\}\\ &\leq \varepsilon_0.\end{aligned}$$

By (4.30), we have

$$t(n) \leq 2|A(0)|.$$

By assumption, $|A(0)| = \frac{1}{\alpha(n)np^2(n)}$ thus $t(n)$ fulfils the conditions of Lemma 4.2. Similarly to the case B_i. we find using (4.35) that

$$\mathbb{P}\Big\{|B^{k_0+1}| \geq 1\Big\} \leq \varepsilon.$$

With a probability tending to one as n goes to infinity the process doesn't activate further vertices after the generation k_0. Using (4.29) we find that the expected number of activated vertices is

$$\begin{aligned}\mathbb{E}\Big(|A(n)|\Big|\bigcap_{k=1}^{k_0-1} \mathcal{B}_k\Big) &\leq |A(0)| + |A(0)|\sum_{k=1}^{\infty}\Big(\frac{2}{\alpha(n)}\Big)^k\\ &\leq |A(0)|\Big(1 + \frac{2}{\alpha(n)}\frac{1}{1-\frac{2}{\alpha(n)}}\Big).\end{aligned}$$

We deduce using Chebyshev inequality that

$$\lim_{n\to\infty} \mathbb{P}\Big\{|A(n)| \geq 2|A(0)|\Big|\bigcap_{k=1}^{k_0-1} \mathcal{B}_k\Big\} = 0.$$

By (4.34) we can conclude that

$$\lim_{n\to\infty} \mathbb{P}\{|A(n)| \geq 2|A(0)|\} = 0.$$

Since $\lim_{n\to\infty} |A(0)|np^2(n) = 0$, we have for any $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}\{|A(n)|np^2(n) > \varepsilon\} = 0.$$

which implies theorem 2.1. □

4.2 Proof of Theorem 2.2

The proof of Theorem 2.2 is splitted in several propositions. We first prove that if we activate a positive part of the graph then with a probability tending to 1 as $n \to \infty$, we activate almost all the graph. Then we find under which condition we activate a positive part of the graph. In the Proposition 4.2, we show that it is sufficient that we can activate at least $\frac{c}{p(n)}$ vertices with $c > 0$. Propositions 4.2, 4.3 and 4.4 give a chain of conditions under which we activate at least $\frac{c}{p(n)}$ vertices for some $c > 0$ coming off with the condition $|A(0)| \geq \frac{1+\nu}{np^2(n)}$, $\nu > 0$.

The following proposition implies that in a graph if the probability of connection between the vertices is $p(n) = \frac{\omega(n)}{n}$ then if a positive part of the graph is activated, this part will activate almost all the graph.

Proposition 4.1. *Let $p(n) = \frac{\omega(n)}{n}$ for some $\omega(n)$ with $\lim_{n\to\infty} \omega(n) = \infty$. Consider the graph $G_{n,p} = (V, \mathcal{L})$. Let $\vartheta > 0$ and $S_\vartheta \subset V$ with $|S_\vartheta| = \lfloor \vartheta n \rfloor$. For any set $S_{\geq\varepsilon}$ with $|S_{\geq\varepsilon}| \geq \varepsilon n$ and $S_\vartheta \cap S_{\geq\varepsilon} = \emptyset$:*

$$\lim_{n\to\infty} \mathbb{P}\Big\{\exists v \in S_{\geq\varepsilon}, \exists u_1, u_2 \in S_\vartheta : (u_1, v) \in \mathcal{L}, (u_2, v) \in \mathcal{L}\Big\} = 1$$

With this lemma, we only need to show that we activate a positive part of the graph to show that we activate almost all the graph.

Proof of proposition 4.1. Consider a set S_ϑ with $|S_\vartheta| = \lfloor \vartheta n \rfloor$, the event that a vertex v_1 has less than 2 connections (i.e. 0 or 1) with the vertices of S_ϑ is

$$F_1 = \left\{ \bigcap_{i=1}^{\lfloor \vartheta n \rfloor} (v_1, u_i)^c \cup \left\{ \bigcup_{i=1}^{\lfloor \vartheta n \rfloor} (v_1, u_i) \cap \bigcap_{l\neq i, l\leq \lfloor \vartheta n \rfloor} (v_1, u_l)^c \right\} \right\},$$

which has a probability

$$\begin{aligned}\mathbb{P}\{F_1\} &= \mathbb{P}\left\{ \bigcap_{i=1}^{\lfloor \vartheta n \rfloor} (v_1, u_i)^c \cup \left\{ \bigcup_{i=1}^{\lfloor \vartheta n \rfloor} (v_1, u_i) \cap \bigcap_{l\neq i, l\leq \vartheta n} (v_1, u_l)^c \right\} \right\} \\ &= \big(1 - p(n)\big)^{\lfloor \vartheta n \rfloor} + \lfloor \vartheta \rfloor np(n)\big(1 - p(n)\big)^{\lfloor \vartheta n \rfloor - 1}.\end{aligned}$$

By assumption $p(n) = \frac{\omega(n)}{n}$ with $\lim_{n\to\infty} \omega(n) = \infty$, then for any $\varepsilon_0 > 0$ there exists $N_0 > 0$ such that for any $n \geq N_0$

$$\mathbb{P}\{F_1\} \leq e^{-\vartheta\omega(n)} + \vartheta\omega(n)e^{-\vartheta\omega(n)}(1+\varepsilon_0).$$

The probability that any vertex $v \in S_\varepsilon$ is connected to at most 1 vertex of S_ϑ is

$$\begin{aligned} P_{\varepsilon n} &= \binom{\lfloor (1-\vartheta)n \rfloor}{\lfloor \varepsilon n \rfloor} \mathbb{P}\Big\{ \bigcap_{k=1}^{\lfloor \varepsilon n \rfloor} F_k \Big\} \\ &= \binom{\lfloor (1-\vartheta)n \rfloor}{\lfloor \varepsilon n \rfloor} \mathbb{P}\Big\{ \bigcap_{k=1}^{\lfloor \varepsilon n \rfloor} \bigcap_{i=1}^{\lfloor \vartheta n \rfloor} (v_k, u_i)^c \cup \Big\{ \bigcup_{i=1}^{\lfloor \vartheta n \rfloor} (v_k, u_i) \cap \bigcap_{l\neq i, l\leq \lfloor \vartheta n \rfloor} (v_k, u_l)^c \Big\} \Big\} \\ &\leq \binom{\lfloor (1-\vartheta)n \rfloor}{\lfloor \varepsilon n \rfloor} \Big(e^{-\vartheta\omega(n)} + \vartheta\omega(n)e^{-\vartheta\omega(n)}(1+\varepsilon_0) \Big)^{\varepsilon n} (1+\varepsilon_1). \end{aligned} \tag{4.36}$$

Using Stirling formula in (4.36), we find

$$\begin{aligned} P_{\varepsilon n} \leq \frac{1}{\sqrt{2\pi}} \frac{\sqrt{1-\vartheta}}{\sqrt{\varepsilon}\sqrt{1-\vartheta-\varepsilon}} \frac{1}{\sqrt{n}} \Big(\frac{1-\vartheta}{1-\vartheta-\varepsilon} \Big)^{(1-\vartheta)n} \\ \Big(\frac{1-\vartheta-\varepsilon}{\varepsilon} \Big)^{\varepsilon n} \Big(e^{-\vartheta\omega(n)} + \vartheta\omega(n)e^{-\vartheta\omega(n)}(1+\varepsilon_0) \Big)^{\varepsilon n} (1+\varepsilon_2) \end{aligned}$$

and

$$\lim_{n\to\infty} P_{\varepsilon n} = 0. \tag{4.37}$$

Considering S_ε as a subset of $S_{\geq\varepsilon}$ we can extend (4.37) to $S_{\geq\varepsilon}$. This proves Lemma 4.1. □

The following lemma states that at step $t(n)$ such that $\lim_{n\to\infty} t(n)p(n) = 0$, then up to time $t(n)$, for any $\varepsilon > 0$ the number of activated vertices is less than εn with a probability tending to 1 as n goes to infinity. This shows that the condition $|A\big(s(n)\big)| > \frac{c}{p(n)}$ for some $c > 0$ is the right one. More important, this implies that up to time $t(n)$ with $\lim_{n\to\infty} t(n)p(n) = 0$ then the number of vertices non activated is larger than $n(1-\varepsilon)$ for any $\varepsilon > 0$. This condition will be used along the following proofs.

Lemma 4.3. *Suppose that* $|A(0)| = \frac{1}{\alpha(n)n}$ *for some* $\alpha(n)$ *with* $\lim_{n\to\infty} \alpha(n) = \infty$. *Let* $p(n) \geq \frac{c}{n}$ *with* $c > 0$ *and* $\lim_{n\to\infty} p(n) = 0$. *For* $t(n)$ *such that* $\lim_{n\to\infty} t(n)p(n) = 0$ *consider the event*

$$\mathcal{A}_\varepsilon\big(t(n)\big) = \{|A\big(t(n)\big)| \leq \varepsilon n\},$$

then

$$\lim_{n\to\infty} \mathbb{P}\Big(\mathcal{A}_\varepsilon\big(t(n)\big)\Big) = 1.$$

Notice that we have $\mathcal{A}_\varepsilon(s+1) \subseteq \mathcal{A}_\varepsilon(s)$ which implies $\bigcap_{s=1}^{t(n)} \mathcal{A}_\varepsilon(s) = \mathcal{A}_\varepsilon\big(t(n)\big)$

Proof of lemma 4.3. Using the fact that the random variable $Y(s)$ is stochastically dominated by the random variable $U^+(s) \in Bin\big(n, p^2(n)s\big)$, we have

$$\mathbb{P}\Big(\mathcal{A}_\varepsilon^c\big(t(n)\big)\Big) \leq \mathbb{P}\Big\{\frac{|A(0)|}{n} + \frac{\sum_{s=1}^{t(n)} U^+(s)}{n} > \varepsilon\Big\}.$$

By assumption $\lim_{n\to\infty} \frac{|A(0)|}{n} = 0$. Therefore, for n sufficiently large, we have

$$\mathbb{P}\Big(\mathcal{A}_\varepsilon^c\big(t(n)\big)\Big) \leq \mathbb{P}\Big\{\frac{\sum_{s=1}^{t(n)} U^+(s)}{n} > \frac{\varepsilon}{2}\Big\}. \tag{4.38}$$

Since $U^+(s) \in Bin\big(n, p^2(n)s\big)$, we have

$$\mathbb{E}\big(U^+(s)\big) = np^2(n)s \tag{4.39}$$

and $Var\big(U^+(s)\big) \leq \mathbb{E}\big(U^+(s)\big)$. The random variables $U^+(s)$, $s = 1, \ldots, t(n)$ are independent. This gives

$$\begin{aligned} Var\Big(\sum_{s=1}^{t(n)} U^+(s)\Big) &= \sum_{s=1}^{t(n)} Var\big(U^+(s)\big) \\ &\leq \sum_{s=1}^{t(n)} \mathbb{E}\big(U^+(s)\big). \end{aligned} \tag{4.40}$$

Using the result (4.40) and Chebyshev inequality, we find

$$\begin{aligned} \mathbb{P}\Big\{\Big|\sum_{s=1}^{t(n)} U^+(s) - \sum_{s=1}^{t(n)} \mathbb{E}\big(U^+(s)\big)\Big| > \frac{\varepsilon}{2}n\Big\} &\leq \frac{Var\big(\sum_{s=1}^{t(n)} U^+(s)\big)}{(\frac{\varepsilon}{2})^2 n^2} \\ &\leq \frac{\sum_{s=1}^{t(n)} \mathbb{E}\big(U^+(s)\big)}{(\frac{\varepsilon}{2})^2 n^2}. \end{aligned} \tag{4.41}$$

By (4.39), we have

$$\sum_{s=1}^{t(n)} \mathbb{E}\big(U^+(s)\big) = np^2(n)\frac{t^2(n) + t(n)}{2}. \tag{4.42}$$

By assumption $\lim_{n\to\infty} t(n)p(n) = 0$ so for n large (4.42) implies

$$\sum_{s=1}^{t(n)} \mathbb{E}\big(U^+(s)\big) \leq \frac{\varepsilon}{4}n \tag{4.43}$$

Using (4.43) in (4.41) we find

$$\mathbb{P}\Big\{\sum_{s=1}^{t(n)}\mathbb{E}\big(U^+(s)\big)\le\frac{\varepsilon}{4}n\Big\}\le\frac{1}{\varepsilon n}$$

Hence

$$\lim_{n\to\infty}\mathbb{P}\Big\{\frac{\sum_{s=1}^{t(n)}U^+(s)}{n}>\frac{\varepsilon}{2}\Big\}=0. \tag{4.44}$$

By (4.38) and (4.44) we deduce that

$$\lim_{n\to\infty}\mathbb{P}\Big(\mathcal{A}_\varepsilon\big(t(n)\big)\Big)=1. \tag{4.45}$$

□

Lemma 4.4. *Let* $p(n)=\frac{\omega(n)}{n}$ *for some* $\omega(n)$ *such that* $\lim_{n\to\infty}\omega(n)=\infty$ *and* $\lim_{n\to\infty}p(n)=0$. *Suppose that* $\lim_{n\to\infty}\frac{|A(0)|}{n}=0$. *Let* $t(n)=\lfloor\frac{\beta}{p(n)}+\tau(n)\rfloor$ *with* $\beta>0$ *and* $\lim_{n\to\infty}p(n)\tau(n)=0$. *Then for any* $0<\delta<1$, *we have for n large*

$$\mathbb{P}\Big\{|A\big(t(n)\big)|\ge n\big(1-(\beta+1)e^{-\beta}\big)(1+\delta)\big)\Big||A\big(t(n)\big)|>t(n)\Big\}\le\frac{C(\beta)}{n}.$$

where $C(\beta)>0$ *for any* $\beta>0$.

Prof of lemma 4.4. We first bound from above the expected number of activated vertices and then using Chebyshev inequality, we find a bound on the probability.

The random variable $Y(s)$ is stochastically dominated by the random variable $U^+(s)\in Bin\big(n,p^2(n)s\big(1-p(n)\big)^{s-1}\big)$. Therefore

$$\begin{aligned}&\mathbb{P}\Big\{|A\big(t(n)\big)|\ge n\big(1-(\beta+1)e^{-\beta}\big)(1+\delta)\big)\Big||A\big(t(n)\big)|>t(n)\Big\}\\&\le\mathbb{P}\Big\{|A(0)|+\sum_{s=1}^{t(n)}U^+(s)\ge n\big(1-(\beta+1)e^{-\beta}\big)(1+\delta)\big)\Big||A\big(t(n)\big)|>t(n)\Big\}.\end{aligned} \tag{4.46}$$

We have

$$\mathbb{E}\big(U^+(s)\big)=np^2(n)s\big(1-p(n)\big)^{s-1}.$$

We bound the sum in (4.46) using integrals. Consider the function $g(s)=np^2(n)s\big(1-p(n)\big)^{s-1}$ for $s\in\mathbb{R}$. It has a single maximum at $s_0=\frac{1}{p(n)}+o\Big(\frac{1}{p(n)}\Big)$ with

$$g(s_0)=np(n)e^{-1}+O\big(np^2(n)\big). \tag{4.47}$$

We have

$$\int_1^{t(n)-1} g(s)ds - g(s_0) \leq \sum_{s=1}^{t(n)} \mathbb{E}\big(U^+(s)\big) \leq \int_1^{t(n)+1} g(s)ds + g(s_0). \tag{4.48}$$

Suppose that $t(n) = \frac{\beta}{p(n)} + \tau(n)$, with $\lim_{n\to\infty} p(n)\tau(n) = 0$. We change the variable into $s = \frac{x}{p(n)}$, $ds = \frac{1}{p(n)}dx$. The upper bound in (4.48) becomes

$$\int_0^{t(n)+1} g(s)ds = n\int_0^{\beta+(\tau(n)+1)p(n)} x\Big((1-p(n)\big)^{\frac{1}{p(n)}}\Big)^{x-p(n)}dx.$$

Since $\big(1-p(n)\big)^{\frac{1}{p(n)}} \leq e^{-1}$ we have

$$\begin{aligned}\int_0^{t(n)+1} g(s)ds &\leq n\int_0^{\beta+(\tau(n)+1)p(n)} xe^{-x}dx \\ &\leq n\Big(-1-\big(\beta+(\tau(n)+1)p(n)\big)e^{-\beta+(\tau(n)+1)p(n)}\Big)+1.\end{aligned}$$

By assumption $\lim_{n\to\infty} p(n) = 0$ and $\lim_{n\to\infty} p(n)\tau(n) = 0$ so for any $\varepsilon_0 > 0$, there exists N_0 such that for any $n > N_0$ we have

$$\begin{aligned}\sum_{s=1}^{t(n)} \mathbb{E}U^+(s) &\leq \int_0^{t(n)+1} g(s)ds + f(s_0) \\ &\leq (1+\varepsilon_0)n\big(1-(\beta+1)e^{-\beta}\big).\end{aligned}$$

For the lower bound, we use the fact that for any ε_1, there exists N_1 such that for any $n > N_1$ we have $e^{-1}(1-\varepsilon_1) < \big(1-p(n)\big)^{\frac{1}{p(n)}}$ which gives

$$\begin{aligned}\int_1^{t(n)-1} g(s)ds &= \int_{p(n)}^{\beta+(\tau(n)-1)p(n)} x\Big((1-p(n)\big)^{\frac{1}{p(n)}}\Big)^{x-p(n)}dx \\ &\geq \int_{x=p(n)}^{\beta+(\tau(n)-1)p(n)} x\Big(e^{-1}(1-\varepsilon_1)\Big)^{x}dx.\end{aligned}$$

There exists N_2 such that for any $n > N_2$

$$\int_1^{t(n)-1} g(s)ds - f(s_0) \geq (1-3\varepsilon_1)\Big(1-(1+2\varepsilon_1)(\beta+1)e^{-\beta}\Big). \tag{4.49}$$

Similarly to the proof of Lemma 4.3 equation (4.40), we use the fact that the random variables $U^+(s)$ are independent, have a binomial distribution

$U^+(s) \in Bin\Big(n, p^2(n)s\big(1-p(n)\big)^{s-1}\Big)$ and $Var\big(U^+(s)\big) \leq \mathbb{E}\big(U^+(s)\big)$. Thus for any $\varepsilon > 0$ we have

$$\mathbb{P}\Big\{\sum_{s=1}^{t(n)} U^+(s) \geq (1+\varepsilon)n\big(1-(\beta+1)e^{-\beta}\big)\Big\} \leq \frac{1}{\varepsilon^2}\frac{1}{n}\frac{1}{(1-3\varepsilon_1)\Big(1-(1+2\varepsilon_1)(\beta+1)e^{-\beta}\Big)}.$$

By assumption $\lim_{n\to\infty}\frac{|A(0)|}{n} = 0$, we have for any $\beta > 0$

$$\mathbb{P}\Big\{|A\big(t(n)\big)| \geq n\big(1-(\beta+1)e^{-\beta}\big)(1+\varepsilon)\Big||A\big(t(n)\big)| > t(n)\Big\} \leq \frac{C(\beta)}{n}, \tag{4.50}$$

where $C(\beta) > 0$ for any $\beta > 0$. Notice that $C(\beta)$ is a decreasing function of β and

$$\lim_{\beta\to 0} C(\beta) = \infty. \tag{4.51}$$

□

Lemma 4.4 gives us a lower bound on the number of non-activated vertices. Thanks to this bound, we can find a lower bound on the number of activated vertices and prove in the following proposition that if we have at least $\frac{\beta}{p(n)}$ activated vertices with $\beta > 0$ then we activate a positive part of the graph.

Proposition 4.2. *Let $p(n) = \frac{\omega(n)}{n}$ for some $\omega(n)$ with $\lim_{n\to\infty}\omega(n) = \infty$ and $\lim_{n\to\infty}\frac{|A(0)|}{n} = 0$. Let $t(n) = \lfloor\frac{\beta}{p(n)+\tau(n)}\rfloor$ with $\beta > 0$ and $\lim_{n\to\infty} p(n)\tau(n) = 0$ then for any $\varepsilon > 0$*

$$\lim_{n\to\infty}\mathbb{P}\Big\{|A\big(t(n)\big)| < n\Big(1-\Big(\frac{1}{2}\beta^2+\beta+1\Big)e^{-2\beta}\Big)(1-\varepsilon)\Big||A\big(t(n)\big)| > t(n)\Big\} = 0. \tag{4.52}$$

Proof of proposition 4.2. We suppose that $\Big\{|A\big(t(n)\big)| > t(n)\Big\}$ is fulfilled. We first notice that for any $s_1 < s_2$ we have $\{|A(s_1)| > s_1\} \subseteq \{|A(s_2)| > s_2\}$ since if we have $\{|A(s_1)| = s_1\}$ then the process of activation stops and $\{|A(s_2)| = s_1\}$. Thus $\{|A(s_2)| > s_2\}^c \subseteq \{|A(s_1)| > s_1\}$ and we have the required inclusion for $s_1 < s_2$

$$\{|A(s_1)| > s_1\} \subseteq \{|A(s_2)| > s_2\}.$$

Conditioning on $\{|A\big(t(n)\big)| > t(n)\}$, we have $\{|A(s)| > s\}$ for any $s \leq t(n)$. Denote $f(\beta) = 1-\Big(\frac{1}{2}\beta^2+\beta+1\Big)e^{-2\beta}$. Let $\delta > 0$, consider the following events.

- If $\lfloor s = \frac{\alpha}{p(n)} + \tau(n)\rfloor$ with $\alpha < \delta$ and $\lim_{n\to\infty}\tau(n)p(n) = 0$, define

$$\mathcal{A}_\varepsilon(s) = \Big\{|A(s)| \leq n\Big(1-(\delta+1)e^{-\delta}\Big)(1+\frac{\varepsilon}{2})\Big||A(s)| > s\Big\}. \tag{4.53}$$

- If $s = \lfloor \frac{\alpha}{p(n)} + \tau(n) \rfloor$ with $\alpha \geq \delta$ define

$$\mathcal{A}_\varepsilon(s) = \Big\{ |A(s)| \leq n\Big(1 - (\alpha+1)e^{-\alpha}\Big)(1 + \frac{\varepsilon}{2}) \Big| |A(s)| > s \Big\}. \tag{4.54}$$

We fix $\delta > 0$ to avoid the problem of control on $C(\alpha)$ since by equation (4.51) we have $\lim_{\alpha \to 0} C(\alpha) = \infty$.

We consider the following probability measure

$$\mathbb{P}_{t(n)}\{X\} = \mathbb{P}\Big\{ X \Big| |A\big(t(n)\big)| > t(n) \Big\}.$$

We shall prove that

$$\lim_{n \to \infty} \mathbb{P}_{t(n)} \Big\{ \bigcap_{s=1}^{\lfloor \frac{\beta}{p(n)} \rfloor} \mathcal{A}_\varepsilon(s) \Big\} = 0. \tag{4.55}$$

This would imply that

$$\begin{aligned}
&\mathbb{P}\Big\{ \Big|A\big(t(n)\big)\Big| < nf(\beta)(1-\varepsilon) \Big| |A\big(t(n)\big)| > t(n) \Big\} \\
&\leq \mathbb{P}_{t(n)} \Big\{ \Big|A\Big(\lfloor \frac{\beta}{p(n)} \rfloor\Big)\Big| < nf(\beta)(1-\varepsilon) \Big| \bigcap_{s=1}^{\lfloor \frac{\beta}{p(n)} \rfloor} \mathcal{A}_\varepsilon(s) \Big\} + \mathbb{P}_{t(n)} \Big(\Big\{ \bigcap_{s=1}^{\lfloor \frac{\beta}{p(n)} \rfloor} \mathcal{A}_\varepsilon(s) \Big\}^c \Big) \\
&\leq \mathbb{P}_{t(n)} \Big\{ \frac{\Big|A\Big(\lfloor \frac{\beta}{p(n)} \rfloor\Big)\Big|}{n} < nf(\beta)(1-\varepsilon) \Big| \bigcap_{s=1}^{\lfloor \frac{\beta}{p(n)} \rfloor} \mathcal{A}_\varepsilon(s) \Big\} + o(1) \text{ as } n \to \infty.
\end{aligned} \tag{4.56}$$

Let $s = \lfloor \frac{\alpha}{p(n)} + \tau(n) \rfloor$ with $\alpha < \delta$. The number of activated vertices is a monotone increasing function of time thus

$$\begin{aligned}
\mathbb{P}_{t(n)}\big(\mathcal{A}_\varepsilon^c(s)\big) &= \mathbb{P}_{t(n)} \Big\{ |A(s)| > n\Big(1 - (\delta+1)e^{-\delta}\Big)(1+\varepsilon) \Big\} \\
&\leq \mathbb{P}_{t(n)} \Big\{ \Big|A\Big(\lfloor \frac{\delta}{p(n)} \rfloor\Big)\Big| > n\Big(1 - (\delta+1)e^{-\delta}\Big)(1+\varepsilon) \Big\}.
\end{aligned}$$

By (4.50) we have for any $\varepsilon_3 > 0$ and for n large

$$\mathbb{P}_{t(n)}\big(\mathcal{A}_\varepsilon^c(s)\big) \leq \frac{C(\delta)}{n}(1+\varepsilon_3).$$

Let $s = \lfloor \frac{\alpha}{p(n)} + \tau(n) \rfloor$ with $\alpha \geq \delta$. By the same reasoning as above, we find

$$\mathbb{P}_{t(n)}\big(\mathcal{A}_\varepsilon^c(s)\big) \leq \frac{C(\alpha)}{n}(1+\varepsilon_3).$$

Since $C(\alpha)$ is a decreasing function of α, we have

$$\mathbb{P}_{t(n)}\big(\mathcal{A}_\varepsilon^c(s)\big) \leq \frac{C(\delta)}{n}(1+\varepsilon_3).$$

Summing up over the $\frac{\beta}{p(n)}$ first steps, we find

$$\sum_{s=1}^{\lfloor \frac{\beta}{p(n)} \rfloor} \mathbb{P}_{t(n)}\Big(\mathcal{A}_\varepsilon^c(s)\Big) \leq 2\frac{\beta C(\delta)}{np(n)}.$$

Hence

$$\mathbb{P}_{t(n)}\left(\left\{\bigcap_{s=1}^{\lfloor \frac{\beta}{p(n)} \rfloor} \mathcal{A}_\varepsilon(s)\right\}^c\right) \leq 2\frac{\beta C(\delta)}{np(n)}.$$

By assumption $\lim_{n\to\infty} \frac{1}{np(n)} = 0$. We can conclude that

$$\lim_{n\to\infty} \mathbb{P}_{t(n)}\left(\left\{\bigcap_{s=1}^{\lfloor \frac{\beta}{p(n)} \rfloor} \mathcal{A}_\varepsilon(s)\right\}^c\right) = 0.$$

In the case (4.53) where $\lfloor s = \frac{\alpha}{p(n)} + \tau(n) \rfloor$ with $\alpha < \delta$ the random variable $\Big(Y(s)|\mathcal{A}_\varepsilon(s)\Big)$ stochastically dominates the independent random variable $U_1^-(s)$ defined as

$$U_1^-(s) \in Bin\left(\left\lfloor n\Big(1-\frac{\varepsilon}{2}\Big)(\delta+1)e^{-\delta}\right\rfloor, p_a(s)\right).$$

In the case (4.54) where $s = \lfloor \frac{\alpha}{p(n)} + \tau(n) \rfloor$ with $\alpha \geq \delta$ the random variable $\Big(Y(s)|\mathcal{A}_\varepsilon(s)\Big)$ stochastically dominates the random variable $U_2^-(s)$ defined as

$$U_2^-(s) \in Bin\left(\left\lfloor n\Big(1-\frac{\varepsilon}{2}\Big)(\alpha+1)e^{-\alpha}\right\rfloor, p_a(s)\right).$$

Define

$$U^-(s) = \begin{cases} U_1^-(s) & \text{if } s = \lfloor \frac{\alpha}{p(n)} + \tau(n) \rfloor \text{ with } \alpha < \delta, \\ U_2^-(s) & \text{if } s = \lfloor \frac{\alpha}{p(n)} + \tau(n) \rfloor \text{ with } \alpha \geq \delta. \end{cases}$$

Suppose that $t(n) = \lfloor \frac{\alpha}{p(n)} + \tau(n) \rfloor$. Changing the variable into $s = \frac{x}{p(n)}$, we find for any $\varepsilon_1 > 0$

$$\sum_{s=1}^{t(n)} \mathbb{E}\big(U^-(s)\big) \geq (1-\varepsilon)\left(n(\delta+1)e^{-\delta}\int_{p(n)}^{\delta} xe^{-x}dx + n\int_{\delta}^{\beta} x(x+1)e^{-2x}dx\right) - g(s_0),$$

where $g(s_0)$ is given by (4.47). Notice that this inequality holds because $x(x+1)e^{-2x} \leq xe^{-x}$. Computations give

$$\sum_{s=1}^{t(n)} \mathbb{E}\big(U^-(s)\big) \geq (1-\varepsilon)n\bigg((\delta+1)e^{-\delta}\Big(\big(1-p(n)\big)e^{-p(n)} - (1+\delta)e^{-\delta}\Big) + \Big(\frac{1}{2}\delta^2+\delta+1\Big)e^{-2\delta} - \Big(\frac{1}{2}\beta^2+\beta+1\Big)e^{-2\beta}\bigg) - g(s_0). \quad (4.57)$$

We can bound from below the right hand side of (4.57) as follows. For any $\varepsilon_0 > 0$ then there exists δ_0 small enough such that for any $\delta < \delta_0$ and n large we can bound from below the right hand side of (4.57) by $f(\beta)(1-\frac{\varepsilon_0}{2})$. This gives that for n large

$$\sum_{s=1}^{t(n)} \mathbb{E}\big(U^-(s)\big) \geq nf(\beta)(1-\varepsilon_0).$$

The random variables $U^-(s), s = 1, \ldots t$ are binomial and independent. Therefore

$$Var\bigg(\sum_{s=1}^{t(n)} U^-(s)\bigg) \leq \sum_{s=1}^{t(n)} \mathbb{E}\big(U^-(s)\big).$$

Let $|A_*\big(t(n)\big)| = |A(0)| + \sum_{s=1}^{t(n)} U^-(s)$. Using Chebyshev inequality, we find for $t(n) = \lfloor \frac{\beta}{p(n)} + \tau(n) \rfloor$

$$\lim_{n\to\infty} \mathbb{P}\bigg\{ |A_*\big(t(n)\big)| \leq nf(\beta)(1-2\varepsilon_0) \bigg| |A\big(t(n)\big)| > t(n) \bigg\} = 0. \quad (4.58)$$

By equations (4.56) and (4.58) we have

$$\lim_{n\to\infty} \mathbb{P}\bigg\{ |A\big(t(n)\big)| \leq n\bigg(1 - \Big(\frac{1}{2}\beta^2+\beta+1\Big)e^{-2\beta}\bigg)(1-2\varepsilon_0) \bigg| |A\big(t(n)\big)| > t(n) \bigg\} = 0.$$

which proves Proposition 4.2. □

This result implies that conditioned on the fact that we activate at least $\lfloor \frac{\beta}{p(n)} \rfloor$ vertices then when we reach the step $\lfloor \frac{\beta}{p(n)} \rfloor$ in the process of activation, we have already activated a positive part of the graph. So for any $x > 0$ we will continue the process of activation for any step $\frac{x}{p(n)}$ and we will activate at least $\frac{1}{2}n$ vertices. The following part is dedicated to the investigation of the weakest conditions under which with a probability tending to 1 as $n \to \infty$ we activate at least $\lfloor \frac{\beta}{p(n)} \rfloor$ vertices for some $\beta > 0$.

We first prove that if we have $|B^i| = \frac{1}{n^{1-\frac{1}{2^k}} p^{2-\frac{1}{2^k}}(n)}$ then with a probability tending to 1 as n goes to infinity $|B^{i+k}| = \frac{\beta}{p(n)}$. In a second part, we prove that if we have $t(n)$ activated vertices such that $t(n)\, np^2(n) \to \infty$ as n goes to infinity which means that

$$t(n) = \frac{\alpha(n)}{np^2(n)},$$

where $\lim_{n\to\infty} \alpha(n) = \infty$, then with a probability tending to 1 as n goes to infinity we activate at least $\frac{1}{n^{1-\frac{1}{2^k}} p^{2-\frac{1}{2^k}}(n)}$ vertices. Finally, we prove using a branching process argument that if we start with $|A(0)| > \frac{1}{np^2(n)}(1+\delta)$ vertices then we activate $t(n)$ vertices with $t(n)$ fulfilling the condition (4.2).

The following lemma implies that in our case, when the number of activated vertices is $|A\big(t(n)\big)| = \frac{\alpha(n)}{np^2(n)}$ then with probability tending to 1 as $n \to \infty$, the size of the new generation is much larger than the set of previously activated vertices.

Lemma 4.5. *Let* $p(n) = \frac{\omega(n)}{n}$ *with* $\lim_{n\to\infty} \omega(n) = \infty$. *Suppose that* $\lim_{n\to\infty} t_0(n) np^2(n) = +\infty$ *and* $\lim_{n\to\infty} t_0(n) p(n) = 0$ *then*

$$\frac{t_0(n)}{\sum_{s=1}^{t_0(n)} Y(s)} \xrightarrow{P} 0. \tag{4.59}$$

Proof of Lemma 4.5. Suppose that $t_0(n) = \lfloor \frac{\alpha(n)}{np^2(n)} \rfloor$ with $\lim_{n\to\infty} \alpha(n) = \infty$. Suppose that $\lim_{n\to\infty} \frac{\alpha(n)}{\omega(n)} = 0$ so that $\lim_{n\to\infty} p(n) t_0(n) = 0$. The random variable $\Big(Y(s) | \mathcal{A}_\varepsilon\big(t_0(n)\big)\Big)$ stochastically dominates the random variable $Y_*(s) \in Bin(n(1-\varepsilon), p_a(s))$ with $\varepsilon > 0$. Write

$$\mathbb{P}\left\{ \frac{t_0(n)}{\sum_{s=1}^{t_0(n)} Y(s)} > \varepsilon_1 \right\} \le \mathbb{P}\left\{ \frac{t_0(n)}{\sum_{s=1}^{t_0(n)} Y(s)} > \varepsilon_1 \middle| \mathcal{A}_\varepsilon\big(t_0(n)\big) \right\} + \mathbb{P}\left\{ \mathcal{A}_\varepsilon^c\big(t_0(n)\big) \right\}.$$

By equation (4.45) this gives

$$\begin{aligned}
\mathbb{P}\left\{ \frac{t_0(n)}{\sum_{s=1}^{t_0(n)} Y(s)} > \varepsilon_1 \right\} &\le \mathbb{P}\left\{ \frac{t_0(n)}{\sum_{s=1}^{t_0(n)} Y(s)} \varepsilon_1 \middle| \mathcal{A}_\varepsilon\big(t_0(n)\big) \right\} + o(1) \text{ as } n \to \infty \\
&\le \mathbb{P}\left\{ \frac{t_0(n)}{\sum_{s=1}^{t_0(n)} Y_*(s)} \varepsilon_1 \middle| \mathcal{A}_\varepsilon\big(t_0(n)\big) \right\} + o(1) \text{ as } n \to \infty. \qquad (4.60)
\end{aligned}$$

For n large, we have

$$\mathbb{E}\Big(Y_*(s)\Big) \ge n(1-2\varepsilon)p^2(n)s.$$

This implies for n large

$$\mathbb{E}\Big(\sum_{s=1}^{t_0(n)} Y_*(s)\Big) \geq n(1-2\varepsilon)p^2(n)\frac{1}{2}t_0^2(n)\Big(1+\frac{1}{t_0(n)}\Big)$$
$$\geq \frac{1}{2}n(1-3\varepsilon)p^2(n)t_0^2(n). \tag{4.61}$$

By assumption $t_0(n) = \lfloor \frac{\alpha(n)}{np^2(n)} \rfloor$ which gives

$$\mathbb{E}\Big(\sum_{s=1}^{t_0(n)} Y_*(s)\Big) \geq \frac{1}{2}n(1-3\varepsilon)\alpha(n)t_0(n).$$

We deduce

$$\frac{t_0(n)}{\mathbb{E}\Big(\sum_{s=1}^{t_0(n)} Y_*(s)\Big)} \leq \frac{2}{\alpha(n)}.$$

By definition of $\alpha(n)$, we have $\lim_{n\to\infty} \alpha(n) = \infty$. We find

$$\lim_{n\to\infty} \frac{t_0(n)}{\mathbb{E}\Big(\sum_{s=1}^{t_0(n)} Y_*(s)\Big)} = 0.$$

Using Markov inequality, we find that

$$\mathbb{P}\Big\{\frac{t_0(n)}{\sum_{s=1}^{t_0(n)} Y_*(s)} > \varepsilon_1\Big\} \leq \frac{1}{\varepsilon_1}\mathbb{E}\Big(\frac{t_0}{\sum_{s=1}^{t_0(n)} Y_*(s)}\Big). \tag{4.62}$$

Using Jensen's inequality in (4.62) we find

$$\mathbb{P}\Big\{\frac{t_0(n)}{\sum_{s=1}^{t_0(n)} Y_*(s)} > \varepsilon_1\Big\} \leq \frac{1}{\varepsilon_1}\frac{t_0(n)}{\mathbb{E}\Big(\sum_{s=1}^{t_0(n)} Y_*(s)\Big)}$$
$$\leq \frac{1}{\varepsilon_1}\frac{2}{\alpha(n)}.$$

By (4.60), we can conclude

$$\lim_{n\to\infty} \mathbb{P}\Big\{\frac{t_0(n)}{\sum_{s=1}^{t_0(n)} Y(s)} > \varepsilon_1\Big\} = 0$$

which is equivalent to equation (4.59) from Lemma 4.5 □

The Lemma 4.5 implies that when we compute the size of the next generation then if we approximate it with the total number of activated vertices, we have a good approximation. We use this fact in the following proofs to simplify the computations. Moreover by Lemma 4.5 we have for any $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}\Big\{ \frac{|A(0)| + |B^1| + ... + |B^i|}{|B^1| + ... + |B^i| + |B^{i+1}|} > \varepsilon \Big\} = 0,$$

or simply

$$\lim_{n\to\infty} \mathbb{P}\Big\{ \frac{|A(0)| + |B^1| + ... + |B^i|}{|B^{i+1}|} > \varepsilon \Big\} = 0.$$

Proposition 4.3. *Suppose that $p(n) = \frac{\omega(n)}{n}$ and there exists $x > 0$ such that $p(n) \leq \frac{x}{\sqrt{n}}$ then there exists $C > 0$ such that for any $i \geq 0$*

$$\lim_{n\to\infty} \mathbb{P}\Big\{ |B^{i+k}| > \frac{C}{p(n)} \Big| |B^i| \geq \frac{1}{n^{1-\frac{1}{2^k}} p^{2-\frac{1}{2^k}}(n)} \Big\} = 1. \tag{4.63}$$

Proof of proposition 4.3. Suppose that $|B^i| = t_0(n) = \left\lfloor \frac{1}{n^{1-\frac{1}{2^k}} p^{2-\frac{1}{2^k}}(n)} \right\rfloor$. If we prove for such $|B^i|$ that we have $\lim_{n\to\infty} \mathbb{P}\Big\{ |B^{i+k}| > \frac{C}{p(n)} \Big| |B^i| = t_0(n) \Big\} = 1$ then proposition 4.3 is proved. If $|B^i| \geq \frac{1}{n^{1-\frac{1}{2^k}} p^{2-\frac{1}{2^k}}(n)}$ then the subset B'^i of size $\left\lfloor \frac{1}{n^{1-\frac{1}{2^k}} p^{2-\frac{1}{2^k}}(n)} \right\rfloor$ activates vertices such that we have $\lim_{n\to\infty} \mathbb{P}\Big\{ |B'^{i+k}| > \frac{C}{p(n)} \Big| |B^i| = t_0(n) \Big\} = 1$. Thus we have it for $|B^{i+k}|$ too.

We consider the following process which underlies the process of activation. We start with a set of activated vertices of size $|G^0| = |B^i| = t_0 = \left\lfloor \frac{1}{n^{1-\frac{1}{2^k}} p^{2-\frac{1}{2^k}}(n)} \right\rfloor$. We explore as in the algorithm the neighbourhood of the vertices, comparing it with the neighbourhood of the previously activated vertices. When all the vertices of G^0 are saturated, we have the first generation of vertices G^1. We do not consider anymore the vertices of G^0. We compare the neighbourhood of the vertices only among the vertices of G^1 during the process of activation. This process is dominated by the process of activation because we do not consider the vertices activated by ancestors being from two different generations. However, it is very close to the inner model because by Lemma 4.5 the size of the previous generations is with a probability tending to 1 as $n \to \infty$ negligible with respect to the size of the last generation. The main contribution to the process of activation comes from the last generation.

Define for any $j \leq k$

$$t_{j+1}(n) = \frac{1}{2} n p^2(n) t_j^2(n).$$

Hence

$$t_j(n) = \left(\frac{1}{2}\right)^{2^j-1} \frac{1}{n^{1-\frac{1}{2^{k-j}}} p^{2-\frac{1}{2^{k-j}}}(n)}. \tag{4.64}$$

We notice that $t_{k-1}(n) = \left(\frac{1}{2}\right)^{2^k-1} \frac{1}{n^{\frac{1}{2}} p^{\frac{3}{2}}(n)}$ so we are in the case when $\lim_{n\to\infty} p(n)t_{k-1}(n) = 0$. Consider the event

$$\mathcal{A}_\varepsilon\big(2t_{k-1}(n)\big) = \Big\{A\big(2t_{k-1}(n)\big) < \varepsilon n\Big\}.$$

We consider the following probability measure

$$\mathbb{P}_{t_0}\Big\{X\Big\} = \mathbb{P}\Big\{X\Big||G^0| = t_0\Big\}.$$

We write

$$\mathbb{P}_{t_0}\Big\{|G^k| < \frac{C}{p(n)}\Big\} \leq \mathbb{P}_{t_0}\Big\{|G^k| < \frac{C}{p(n)}\Big|\mathcal{A}_\varepsilon\big(2t_{k-1}(n)\big)\Big\} + \mathbb{P}_{t_0}\Big\{\mathcal{A}^c_\varepsilon\big(2t_{k-1}(n)\big)\Big\}.$$

By (4.45) this gives

$$\mathbb{P}_{t_0}\Big\{|G^k| < \frac{C}{p(n)}\Big\} \leq \mathbb{P}_{t_0}\Big\{|G^k| < \frac{C}{p(n)}\Big|\mathcal{A}_\varepsilon\big(2t_{k-1}(n)\big)\Big\} + o(1) \text{ as } n \to \infty. \tag{4.65}$$

Consider the random variables $Y_*(s) \in Bin\Big(\lfloor n(1-\varepsilon)\rfloor, p_a(s)\Big)$ which are stochastically dominated by the random variable $\Big(Y(s)|\mathcal{A}_\varepsilon\big(2t_{k-1}(n)\big)\Big)$ and call $|G^l_*|$ the l^{th} generation where in the process explained above where here the vertices are activated at time s with a distribution $Y_*(s)$. Suppose $|G^0_*| = |G^0|$, using equation (4.65), we have

$$\mathbb{P}_{t_0}\Big\{|G^k| < \frac{C}{p(n)}\Big\} \leq \mathbb{P}_{t_0}\Big\{|G^k_*| < \frac{C}{p(n)}\Big\} + o(1) \text{ as } n \to \infty.$$

Consider the events

$$\mathcal{C}_j = \Big\{\sum_{i=1}^{|G^j_*|} Y_*(s) > (1-4\varepsilon)t_{j+1}\Big\}.$$

We write

$$\mathbb{P}_{t_0}\Big\{|G^k_*| < \frac{c}{p(n)}\Big\} \leq \mathbb{P}_{t_0}\Big\{|G^k_*| < \frac{c}{p(n)}\Big|\mathcal{C}_{k-1}\Big\} + \mathbb{P}_{t_0}\Big(\mathcal{C}^c_{k-1}\Big). \tag{4.66}$$

We shall prove that

$$\lim_{n\to\infty} \mathbb{P}_{t_0}\Big(\mathcal{C}^c_{k-1}\Big) = 0, \tag{4.67}$$

which combined with (4.66) will give

$$\mathbb{P}_{t_0}\Big\{|G_*^k| < \frac{C}{p(n)}\Big\} \le \mathbb{P}_{t_0}\Big\{|G_*^k| < \frac{C}{p(n)}\Big|\mathcal{C}_{k-1}\Big\} + o(1) \text{ as } n \to \infty.$$

To prove (4.67), we consider the development of (4.67) into conditional probabilities in the following way

$$\mathbb{P}_{t_0}\Big(\mathcal{C}_{k-1}^c\Big) \le \sum_{j=1}^{k-2} \mathbb{P}_{t_0}\Big(\mathcal{C}_{j+1}^c\Big|\mathcal{C}_j\Big) + \mathbb{P}_{t_0}\Big(\mathcal{C}_1^c\Big). \tag{4.68}$$

We first compute the probability of $\mathcal{C}_1^c$

$$\mathbb{P}_{t_0}\Big(\mathcal{C}_1^c\Big) = \mathbb{P}\Big\{\sum_{i=1}^{t_0(n)} Y_*(s) \le (1-4\varepsilon)t_1\Big\}.$$

where the condition $G^0 = t_0(n)$ is inserted in the sum. By (4.61) we have

$$(1-3\varepsilon)t_1(n) \le \sum_{s=1}^{t_0(n)} \mathbb{E}\big(Y_*(s)\big) \le t_1(n)\Big(1+\frac{1}{t_0(n)}\Big).$$

The random variables $Y_*(s)$ are binomial and independent. Thus

$$\begin{aligned} Var\Big(\sum_{s=1}^{t_0(n)} Y_*(s)\Big) &\le \sum_{s=1}^{t_0(n)} \mathbb{E}\big(Y_*(s)\big) \\ &\le t_1(n)\Big(1+\frac{1}{t_0(n)}\Big). \end{aligned} \tag{4.69}$$

We have

$$\mathbb{P}\Big\{\sum_{i=1}^{t_0(n)} Y_*(s) \le (1-4\varepsilon)t_1\Big\} \le \mathbb{P}\Big\{\Big|\sum_{i=1}^{t_0(n)} Y_*(s) - \sum_{i=1}^{t_0(n)} \mathbb{E}Y_*(s)\Big| \le \varepsilon t_1(n)\Big\} \tag{4.70}$$

We use Chebyshev inequality in (4.70). With the help of equation (4.69) we find

$$\mathbb{P}\Big\{\sum_{i=1}^{t_0(n)} Y_*(s) \le (1-4\varepsilon)t_1\Big\} \le \frac{1+\frac{1}{t_0(n)}}{\varepsilon^2 t_1(n)}.$$

Hence

$$\lim_{n\to\infty} \mathbb{P}_{t_0}\Big(\mathcal{C}_1^c\Big) = 0. \tag{4.71}$$

The conditional probabilities are

$$\mathbb{P}_{t_0}\Big\{\mathcal{C}_{j+1}^c\Big|\mathcal{C}_j\Big\} = \mathbb{P}\Big\{\sum_{s=1}^{|G_*^{j+1}|} Y_*(s) \le (1-4\varepsilon)^{j+2}t_{j+2}(n)\Big|\sum_{s=1}^{|G_*^j|} Y_*(s) > (1-4\varepsilon)^{j+1}t_{j+1}(n)\Big\}$$
$$\le \mathbb{P}\Big\{\sum_{s=1}^{(1-4\varepsilon)^{j+1}t_{j+1}(n)} Y_*(s) \le (1-4\varepsilon)^{j+2}t_{j+2}(n)\Big\}. \tag{4.72}$$

The event conditioned on $\sum_{s=1}^{|G_*^j|} Y_*(s) > (1-4\varepsilon)^{j+1}t_{j+1}(n)$ is independent to $|G^0| = t_0(n)$. By (4.61) we have

$$(1-3\varepsilon)\frac{1}{2}np^2(n)t_l^2(n) \le \sum_{s=1}^{t_l(n)}\mathbb{E}\big(Y_*(s)\big) \le \frac{1}{2}np^2(n)t_l^2(n)\Big(1+\frac{1}{t_l(n)}\Big).$$

By definition of $t_l(n)$ in (4.64) we have

$$(1-3\varepsilon)t_{l+1}(n) \le \sum_{s=1}^{t_l(n)}\mathbb{E}\big(Y_*(s)\big) \le t_{l+1}(n)\Big(1+\frac{1}{t_l(n)}\Big).$$

This gives

$$(1-3\varepsilon)(1-4\varepsilon)^{j+1}t_{j+2}(n) \le \sum_{s=1}^{(1-4\varepsilon)^{j+1}t_{j+1}(n)}\mathbb{E}\big(Y_*(s)\big)$$
$$\le (1-4\varepsilon)^{j+2}t_{j+2}(n)\Big(1+\frac{1}{(1-4\varepsilon)^{j+1}t_{j+1}}\Big). \tag{4.73}$$

We find with the help of the first inequality of (4.73) in (4.72)

$$\mathbb{P}_{t_0}\Big\{\mathcal{C}_{j+1}^c\Big|\mathcal{C}_j\Big\} \le \mathbb{P}\Big\{\Big|\sum_{s=1}^{(1-4\varepsilon)^{j+1}t_{j+1}(n)} Y_*(s) - \sum_{s=1}^{(1-4\varepsilon)^{j+1}t_{j+1}(n)}\mathbb{E}Y_*(s)\Big| \ge \varepsilon(1-4\varepsilon)^{j+1}t_{j+2}\Big\}. \tag{4.74}$$

The random variables $Y_*(s)$ are binomial and independent. From (4.73) we have

$$Var\Big(\sum_{s=1}^{(1-4\varepsilon)^{j+1}t_{j+1}(n)} Y_*(s)\Big) \le \sum_{s=1}^{(1-4\varepsilon)^{j+1}t_{j+1}(n)}\mathbb{E}\big(Y_*(s)\big)$$
$$\le (1-4\varepsilon)^{j+2}t_{j+2}(n)\Big(1+\frac{1}{(1-4\varepsilon)^{j+1}t_{j+1}}\Big). \tag{4.75}$$

Using Chebyshev inequality in the equation (4.74), we find with the help of (4.75)

$$\lim_{n\to\infty} \mathbb{P}_{t_0}\left\{\mathcal{C}_{j+1}\middle|\mathcal{C}_j\right\} = 0. \tag{4.76}$$

We have a finite number of generations k, therefore we deduce by (4.68) that

$$\lim_{n\to\infty} \mathbb{P}_{t_0}\left(\mathcal{C}_{k-1}^c\right) = 0.$$

For the k^{th} generation we have

$$\begin{aligned}\mathbb{P}_{t_0}\left\{|G_*^k| \le (1-4\varepsilon)^{k+1}\left(\frac{1}{2}\right)^{2^k-1}\frac{1}{p(n)}\ \middle|\mathcal{C}_{k-1}\right\} &= \mathbb{P}_{t_0}\left\{\sum_{s=1}^{|G_*^{k-1}|} Y_*(s) \le (1-4\varepsilon)^{k-1}t_{k-1}(n)\middle|\mathcal{C}_{k-1}\right\}\\ &= \mathbb{P}_{t_0}\left\{\mathcal{C}_k\middle|\mathcal{C}_{k-1}\right\}. \end{aligned}\tag{4.77}$$

Using equation (4.76) in (4.77), we find

$$\lim_{n\to\infty} \mathbb{P}_{t_0}\left\{|G_*^k| \le (1-4\varepsilon)^{k+1}\left(\frac{1}{2}\right)^{2^k-1}\frac{1}{p(n)}\ \middle|\mathcal{C}_{k-1}\right\} = 0. \tag{4.78}$$

If we combine (4.78) with (4.68) and (4.71), we find

$$\lim_{n\to\infty} \mathbb{P}_{t_0}\left\{|G_*^k| \le (1-4\varepsilon)^{k+1}\left(\frac{1}{2}\right)^{2^k-1}\frac{1}{p(n)}\right\} = 0 \tag{4.79}$$

which proves equation (4.63) with $C = (1-4\varepsilon)^{k+1}\left(\frac{1}{2}\right)^{2^k-1} > 0$ in the case $|B^i| = \left\lfloor \frac{1}{n^{1-\frac{1}{2^k}}p^{2-\frac{1}{2^k}}(n)} \right\rfloor$. Equation (4.79) still holds if we consider $|B^i| \ge \left\lfloor \frac{1}{n^{1-\frac{1}{2^k}}p^{2-\frac{1}{2^k}}(n)} \right\rfloor$ instead,

$$\lim_{n\to\infty} \mathbb{P}\left\{|B^{i+k}| > \frac{C}{p(n)}\middle||B^i| \ge \frac{1}{n^{1-\frac{1}{2^k}}p^{2-\frac{1}{2^k}}(n)}\right\} = 1.$$

This proves Proposition 4.3.

□

Proposition 4.4. *Let $p(n) = \frac{\omega(n)}{n}$ for some $\omega(n)$ with $\lim_{n\to\infty}\omega(n) = \infty$ and suppose there exists $c > 0$ such that $p(n) \le \frac{c}{\sqrt{n}}$. Let $\alpha(n)$ be a function with $\lim_{n\to\infty}\alpha(n) = \infty$ and let $\varepsilon > 0$. Then there exists $k > 0$ and $C > 0$ such that*

$$\lim_{n\to\infty} \mathbb{P}\left\{|B^{i+k}| > \frac{C}{n^{1-\varepsilon}p^{2-\varepsilon}(n)}\middle||B^i| \ge \frac{\alpha(n)}{np^2(n)}\right\} = 1. \tag{4.80}$$

Proof of proposition 4.4. We consider the process of activation described in the proposition 4.3. We start with $G^0 = t_0(n) = \lfloor \frac{\alpha(n)}{np^2(n)} \rfloor$ and follow the process until we reach a generation G^k such that with a probability tending to 1 as $n \to \infty$ the size of that generation fulfils $|G^k| > \frac{C}{n^{1-\varepsilon}p^{2-\varepsilon}(n)}$. By Proposition 4.3, we know that with a probability tending to 1 as $n \to \infty$, if $|G^k| > \frac{C_1}{n^{1-\varepsilon}p^{2-\varepsilon}(n)}$ we need several generations k_ε to reach the size $|G^{k+k_\varepsilon}| > \frac{C_2}{p(n)}$. Thus as we consider the process of activation until the new generation of activated vertices reaches a size $\frac{C_1}{n^{1-\varepsilon}p^{2-\varepsilon}(n)} < |G^k| < \frac{C_2}{n^{1-\varepsilon}p^{2-\varepsilon}(n)}$, by Lemma 4.5 this implies that the number of activated vertices has with probability 1 the same property. Thus we have the following event with a probability tending to 1 as $n \to \infty$

$$\mathcal{A}\Big(\frac{1}{n^{\frac{3}{4}}p^{\frac{7}{4}}(n)}\Big) = \Big\{ |A\Big(\frac{1}{n^{\frac{3}{4}}p^{\frac{7}{4}}(n)}\Big)| < \frac{1}{n^{\frac{1}{2}-\varepsilon}p^{\frac{3}{2}-\varepsilon}(n)} \Big\}.$$

We denote

$$\mathbb{P}_{t_0}(X) = \mathbb{P}\Big\{ X \Big| |G^0_*| = t_0 \Big\}.$$

We have

$$\mathbb{P}_{t_0}\Big\{ |G^k_*| < \frac{C}{n^{1-\varepsilon}p^{2-\varepsilon}(n)} \Big\} \leq \mathbb{P}_{t_0}\Big\{ |G^k_*| < \frac{C}{n^{1-\varepsilon}p^{2-\varepsilon}(n)} \Big| \mathcal{A}\Big(\frac{1}{n^{\frac{3}{4}}p^{\frac{7}{4}}(n)}\Big) \Big\} + \mathbb{P}\Big(\mathcal{A}^c\Big(\frac{1}{n^{\frac{3}{4}}p^{\frac{7}{4}}(n)}\Big)\Big).$$

The random variable $\Big(Y(s) | \mathcal{A}(s) \Big)$ stochastically dominates the random variable

$$Y_*(s) \in Bin\Big(n - \frac{1}{n^{\frac{1}{2}-\varepsilon}p^{\frac{3}{2}-\varepsilon}(n)}, \frac{p^2(n)}{1-p(n)} s\big(1 - sp(n)\big)\Big).$$

Consider the process where the number of vertices activated at time s is distributed as $Y_*(s)$ then if $|G^0_*| = t_0(n)$ we have for any ε, there exist n sufficiently large such that

$$\frac{1}{2}np^2(n)t^2(1-\varepsilon) \leq \mathbb{E}\Big(G^1_*\Big) \leq \frac{1}{2}np^2(n)t^2.$$

Therefore we introduce the sequence $t_{k+1}(n) = \frac{1}{2}np^2(n)t^2_k(n)$ and starting with $t_0(n) = \frac{\alpha(n)}{np^2(n)}$, we have that

$$t_k(n) = \frac{1}{2}\big(\frac{1}{2}\alpha(n)\big)^{2^k} \frac{1}{np^2(n)}. \tag{4.81}$$

This sequence reaches the required size $\frac{C}{n^{1-\varepsilon}p^{2-\varepsilon}(n)}$ for

$$k > \frac{\log(\log(np(n)))}{\log 2}. \tag{4.82}$$

We have

$$\begin{aligned}\Big(\frac{1}{2}\alpha(n)\Big)^{2^{\frac{\log(\log(n))}{\log 2}}} &= e^{e^{\log(\log(n))}\log(\frac{1}{2}\alpha(n))}\\ &= e^{\log(np(n))\log(\frac{1}{2}\alpha(n))}\\ &= (np(n))^{\log(\frac{1}{2}\alpha(n))} > (np(n))^{\varepsilon} \text{ for any } \varepsilon > 0. \end{aligned} \tag{4.83}$$

It means that if we can show that the number of vertices in each generation is close to the sequence of t_k then we need at most $\frac{\log(\log(np(n)))}{\log 2}$ generations so as to reach the size $\frac{C}{n^{1-\varepsilon}p^{2-\varepsilon}(n)}$. We can see in the equation (4.83) that the exponent $\log(\frac{1}{2}\alpha(n))$ is very large which is not needed. It comes from the fact that we need about $\frac{\log(\log(np(n)))}{\log 2}$ steps so as to reach a size $\frac{\big(np(n)\big)^{\varepsilon}}{np^2(n)}$ and then, the size increases to $\frac{1}{p(n)}$ in a finite number of steps.

Let $|G_*^0| = t_0(n) = |G^0|$. Consider the events

$$\mathcal{D}(i) = \Big\{|G_*^i| \leq \Big(\frac{1}{2}\Big)^{2^i-1} t_i\Big\}$$

We prove that for k as in (4.82)

$$\lim_{n\to\infty} \mathbb{P}_{t_0}\Big(\mathcal{D}(k)\Big) = 0$$

which implies that $|G_*^k| > \frac{1}{n^{1-\varepsilon}p^{2-\varepsilon}(n)}$ for some $\varepsilon > 0$. We write

$$\mathbb{P}_{t_0}\Big(\mathcal{D}(k)\Big) \leq \sum_{i=1}^{k} \mathbb{P}_{t_0}\Big(\mathcal{D}(k)\Big|\mathcal{D}^c(k-1)\Big) + \mathbb{P}_{t_0}\big(\mathcal{D}^c(1)\big). \tag{4.84}$$

We first treat the last term of the right hand side of (4.84). By definition of the event, we have

$$\mathbb{P}_{t_0}\big(\mathcal{D}^c(1)\big) = \mathbb{P}_{t_0}\Big\{|G_*^1| \leq \frac{1}{2}t_1\Big\}.$$

We compute this probability using Chebyshev inequality and the fact that $Var\big(|G_*^i|\big) \leq \mathbb{E}\big(|G_*^i|\big)$. This gives

$$\begin{aligned}\mathbb{P}_{t_0}\Big\{|G_*^1| \leq \frac{1}{2}t_1(n)\Big\} &\leq \frac{\mathbb{E}\big(|G_*^1|\big)}{\frac{1}{4}t_1^2(n)}\\ &\leq \frac{4}{t_1(n)}.\end{aligned}$$

We immediately notice that

$$\lim_{n\to\infty} \mathbb{P}_{t_0}\Big\{|G_*^1| \leq \frac{1}{2}t_1(n)\Big\} = 0. \tag{4.85}$$

We turn to the conditional probabilities which appear in (4.84). By definition of the event, we have

$$\mathbb{P}_{t_0}\Big(\mathcal{D}(k)\Big|\mathcal{D}^c(k-1)\Big) = \mathbb{P}\Big\{|G_*^k| \le \Big(\frac{1}{2}\Big)^{2^k-1} t_k \Big| |G_*^{k-1}| \ge \Big(\frac{1}{2}\Big)^{2^{k-1}-1} t_{k-1}\Big\},$$

where the event conditioned on $|G_*^{k-1}| \ge \left(\frac{1}{2}\right)^{2^{k-1}-1} t_{k-1}$is independent from $|G_*^0| = t_0(n)$. Considering the fact that

$$\begin{aligned}\mathbb{P}\Big\{|G_*^i| \le \Big(\frac{1}{2}\Big)^{2^i-1} t_k \Big| |G_*^{i-1}| \ge \Big(\frac{1}{2}\Big)^{2^{i-1}-1} t_{i-1}\Big\}& \\ \le \mathbb{P}\Big\{|G_*^i| \le \Big(\frac{1}{2}\Big)^{2^i-1} t_k \Big| |G_*^{i-1}| = \Big\lfloor \Big(\frac{1}{2}\Big)^{2^{i-1}-1} t_{i-1}\Big\rfloor\Big\}&\end{aligned}$$

then one can find by use of the Chebyshev inequality similarly as it is done for (4.85) that

$$\mathbb{P}\Big\{|B_*^i| \le \Big(\frac{1}{2}\Big)^{2^i-1} t_i \Big| |B_*^{i-1}| = \Big\lfloor \Big(\frac{1}{2}\Big)^{2^{i-1}-1} t_{i-1}\Big\rfloor\Big\} \le \frac{2^{2^i}}{t_i}.$$

Hence we can compute the sum of the conditional probabilities and prove that it converges to 0. We have

$$\begin{aligned}\sum_{i=1}^k \mathbb{P}_{t_0}\Big(\mathcal{D}(i)\Big|\mathcal{D}^c(i-1)\Big) &= \sum_{i=1}^k \mathbb{P}\Big\{|B_*^i| \le \Big(\frac{1}{2}\Big)^{2^i-1} t_i \Big| |B_*^{i-1}| = \Big(\frac{1}{2}\Big)^{2^{i-1}-1} t_{i-1}\Big\} \\ &\le \sum_{i=1}^k \frac{2^{2^i}}{t_i}.\end{aligned}$$

By definition of t_i given in (4.81) this gives

$$\begin{aligned}\sum_{i=1}^k \mathbb{P}_{t_0}\Big(\mathcal{D}(i)\Big|\mathcal{D}^c(i-1)\Big) &\le 2\sum_{i=1}^k \frac{2^{2^i}}{(\frac{1}{2}\alpha(n))^{2^i}\frac{1}{np^2(n)}} \\ &\le 2np^2(n)\sum_{i=1}^k \Big(\frac{4}{\alpha(n)}\Big)^{2^i}. \qquad (4.86)\end{aligned}$$

Since $\lim_{n\to\infty}\alpha(n) = \infty$, for n large, we have $\frac{4}{\alpha(n)} < 1$ and we can bound (4.86) the following way

$$\begin{aligned}\sum_{i=1}^k \mathbb{P}_{t_0}\Big(\mathcal{D}(i)\Big|\mathcal{D}^c(i-1)\Big) &\le 2np^2(n)\sum_{i=1}^k \Big(\frac{4}{\alpha(n)}\Big)^{i} \\ &\le 2np^2(n)\frac{1}{\alpha(n)} \qquad (4.87)\end{aligned}$$

By assumption $p(n) \leq \frac{c}{\sqrt{n}}$ so we have

$$np^2(n) \leq C^2 \text{ and } lim_{n\to\infty}\alpha(n) = \infty,$$

so

$$\lim_{n\to\infty} \sum_{i=1}^{k} \mathbb{P}_{t_0}\Big(\mathcal{D}(i)\Big|\mathcal{D}^c(i-1)\Big) = 0. \tag{4.88}$$

We conclude by (4.85) and (4.88) that

$$\lim_{n\to\infty} \mathbb{P}_{t_0}\Big(\mathcal{D}(k)\Big) = 0. \tag{4.89}$$

This implies that with a probability tending to 1 as $n \to \infty$, the process underlying our process of activation activates at least $\frac{1}{n^{1-\varepsilon}p^{2-\varepsilon}}$ vertices for some $\varepsilon > 0$ so

$$\lim_{n\to\infty} \mathbb{P}\Big\{|B^{i+k}| > \frac{C}{n^{1-\varepsilon}p^{2-\varepsilon}(n)}\Big||B^i| \geq \frac{\alpha(n)}{np^2(n)}\Big\} = 1.$$

which proves proposition 4.4. □

We now turn to the last proposition needed to prove Theorem 2.2. We use branching process argument to show that starting from $\lfloor\frac{1+\varepsilon}{np^2(n)}\rfloor$ vertices, with probability tending to 1 as $n \to \infty$ we activate at least $\frac{\alpha(n)}{np^2(n)}$ vertices with $\lim_{n\to\infty}\alpha(n) = \infty$.

Proposition 4.5. *Under conditions of Theorem 2.2, i.e. suppose* $p(n) = \frac{\omega(n)}{n}$ *for some* $\omega(n)$ *with* $\lim_{n\to\infty}\omega(n) = \infty$ *and* $\lim_{n\to\infty}\sqrt{n}p(n) = 0$. *Suppose that* $|A(0)| = \lfloor\frac{1+\nu}{np^2(n)}\rfloor$ *with* $\nu > 0$ *then there exists* $\alpha(n)$ *with* $\lim_{n\to\infty}\alpha(n) = \infty$ *and* $t(n)$ *such that*

$$\lim_{n\to\infty} \mathbb{P}\Big\{|A(t(n))| \geq \frac{\alpha(n)}{np^2(n)}\Big\} = 1.$$

Proof of proposition 4.5. To prove Proposition 4.5, we are going to prove that with a probability tending to 1 as n goes to infinity for a non negligible part of the $\lfloor\frac{1+\nu}{np^2(n)}\rfloor$ vertices, the process of activation survives. Denote by $D \subset A(0)$ the set of vertices which have infinitely many descendant. Let $\gamma(n)$ be any function such that $\lim_{n\to\infty}\gamma(n)np^2(n) = 0$ and let $\beta > 1$ then for $t = \lfloor\frac{\beta}{np^2(n)} + \gamma(n)\rfloor > \frac{1}{np^2(n)}$ the probability of activation is $p_a(t) = \frac{\beta}{n}\big(1-p(n)\big)^{t-1}\Big(1+\frac{\gamma(n)}{np^2(n)}\Big)$. For n sufficiently large we have

$$p_a(t) > \begin{cases} \frac{1}{n}\beta\Big(1-\frac{\beta-1}{2}\Big) & \text{if } 1<\beta<2 \\ \frac{2-\varepsilon}{n} & \text{if } \beta = 2 \\ \frac{2}{n} & \text{if } \beta > 2. \end{cases} \tag{4.90}$$

We treat here the first case in (4.90). The other cases can be proved the same way. The process of activation can be bounded from below for any $\varepsilon > 0$ by a branching process with distribution $X_n(t) \in Bin\Big(n(1-\varepsilon), \frac{1}{n}\beta\Big(1-\frac{\beta-1}{2}\Big)\Big)$ where $\beta\Big(1-\frac{\beta-1}{2}\Big) > 1$. For such a process, the probability of survival $1 - x_\beta(n)$ is strictly positive where $x_\beta(n)$ is the probability of ultimate extinction. Thus, the probability of survival for any process starting from any of the vertices $v \in \left[\frac{1+\frac{\nu}{2}}{np^2(n)}; \frac{1+\nu}{np^2(n)}\right]$ is strictly positive and bounded from below by the probability of survival of the branching process with distribution $X_n(t) \in Bin\Big(n(1-\varepsilon), \frac{1}{n}\big(1+\frac{\nu}{2}\big)\big(1-\frac{\nu}{4}\big)\Big)$.

We notice that the probability of survival for each of the vertices is bounded away from 1 as there is a positive probability that they have no descendant at the first generation.

If we can prove that among the vertices $v \in \left[\frac{1+\frac{\nu}{2}}{np^2(n)}; \frac{1+\nu}{np^2(n)}\right]$, the probability that at most $\gamma(n)$ survive with $\lim_{n\to\infty} \gamma(n)np^2(n) = 0$ converges to 0 then we have our result.

Let H denote the set of vertices with a branching process which survives. The probability that exactly h branching processes with distribution $X_n(\lfloor\frac{1+\frac{\nu}{2}}{np^2(n)}\rfloor)$ survive is

$$\mathbb{P}\Big\{|H| = h\Big\} = \binom{\lfloor\frac{\nu}{2}\frac{1}{np^2(n)}\rfloor}{h}(1-x_\beta(n))^h x_\beta(n)^{\frac{\frac{\nu}{2}}{np^2(n)}-s},$$

where $1-x_\beta(n)$ is the probability of survival, we have h branching processes which survive and $\Big\lfloor\frac{\frac{\nu}{2}}{np^2(n)}\Big\rfloor - h$ which don't and $\binom{\lfloor\frac{\nu}{2}\frac{1}{np^2(n)}\rfloor}{h}$ way to choose the h vertices

The union is disjoint so the probability that at most $\gamma(n)$ vertices are activated is

$$\mathbb{P}\Big\{\bigcup_{h=0}^{\lfloor\gamma(n)\rfloor}|H| = h\Big\}) = \sum_{h=0}^{\lfloor\gamma(n)\rfloor}\binom{\lfloor\frac{\nu}{2}\frac{1}{np^2(n)}\rfloor}{h}(1-x_\beta(n))^h x_\beta(n)^{\frac{\frac{\nu}{2}}{np^2(n)}-h}. \tag{4.91}$$

Suppose that there exists $\gamma_1(n)$ with $\lim_{n\to\infty}\gamma_1(n)np^2(n) = 0$, such that there exists $\vartheta > 0$ and for any $N_0 > 0$, there exists $n > N_0$ with

$$\mathbb{P}\Big\{\bigcup_{h=0}^{\lfloor\gamma_1(n)\rfloor}|H| = h\Big\} > \vartheta. \tag{4.92}$$

or in other words $\overline{\lim}_{n\to\infty}\mathbb{P}\Big\{\cup_{h=0}^{\lfloor\gamma_1(n)\rfloor}|H| = h\Big\} > \vartheta$. Equation (4.92) implies using (4.91) that

$$\sum_{h=0}^{\lfloor\gamma_1\rfloor(n)}\binom{\lfloor\frac{\nu}{2}\frac{1}{np^2(n)}\rfloor}{h}(1-x_\beta(n))^h x_\beta(n)^{\frac{\frac{\nu}{2}}{np^2(n)}-h} > \vartheta. \tag{4.93}$$

By assumption $\lim_{n\to\infty} \gamma_1(n)np^2(n) = 0$ then there exists δ and for any $h < \gamma_1(n)$, we have

$$\binom{\lfloor \frac{\nu}{2}\frac{1}{np^2(n)}\rfloor}{h} \geq \frac{\left(\lfloor \frac{\nu}{2}\frac{1}{np^2(n)}\rfloor\right)^h}{h!}(1-\delta).$$

This gives for (4.93)

$$a^{\frac{\nu}{2}\frac{1}{np^2(n)}} \sum_{h=0}^{\lfloor \gamma_1\rfloor(n)} \frac{1}{h!}\Big(\Big\lfloor \frac{\nu}{2}\frac{1}{np^2(n)}\Big\rfloor \frac{1-x_\beta(n)}{x_\beta(n)}\Big)^h (1-\delta) > \vartheta.$$

By assumption $\lim_{n\to\infty} hnp^2(n) = 0$, therefore the function

$$f(h) = \frac{1}{h!}\Big(\Big\lfloor \frac{\nu}{2}\frac{1}{np^2(n)}\Big\rfloor \frac{1-x_\beta(n)}{x_\beta(n)}\Big)^h$$

is increasing with h.

Now, consider the sum from 0 to $\eta(n)\gamma_1(n)$ where $\eta(n)$ is such that $\lim_{n\to\infty} \eta(n)\gamma_1(n)np^2(n) = 0$ and $\lim_{n\to\infty} \eta(n) = \infty$, we have

$$\begin{aligned}\sum_{h=0}^{\lfloor \eta(n)\gamma_1(n)\rfloor} \frac{1}{h!}\Big(\Big\lfloor \frac{\nu}{2}\frac{1}{np^2(n)}\Big\rfloor \frac{1-x_\beta(n)}{x_\beta(n)}\Big)^h &\geq \eta(n) \sum_{h=0}^{\lfloor \gamma_1(n)\rfloor} \frac{1}{h!}\Big(\Big\lfloor \frac{\nu}{2}\frac{1}{np^2(n)}\Big\rfloor \frac{1-x_\beta(n)}{x_\beta(n)}\Big)^h \\ &\geq \eta(n)\vartheta \to \infty \text{ as } n \to \infty,\end{aligned}$$

which is impossible as it is the lower bound on the probability that at most $\eta(n)\gamma_1(n)$ vertices survive.

Then for any $\gamma_1(n)$ such that $\lim_{n\to\infty} \gamma_1(n)np^2(n) = 0$

$$\lim_{n\to\infty} \mathbb{P}\Big(\bigcup_{h=0}^{\lfloor \gamma_1(n)\rfloor} \{|H| = h\}\Big) = 0,$$

This proves that there exists $\varepsilon_0 > 0$ such that

$$\lim_{n\to\infty} \mathbb{P}\Big\{|H| > \frac{\varepsilon_0}{np^2(n)}\Big\} = 1.$$

This in turns implies that

$$\lim_{n\to\infty} \mathbb{P}\Big\{|A\big(t(n)\big)| \geq \frac{\alpha(n)}{np^2(n)}\Big\} = 1$$

for some function $\alpha(n)$ with $\lim_{n\to\infty} \alpha(n) = \infty$, which proves Proposition 4.5. □

We now have all elements to prove Theorem 2.2. We described the conditions to activate almost all the graph trying to find weaker and weaker conditions. To prove Theorem 2.2 , we start from the last proposition and take them in reverse order.

Proof of Theorem 2.2. Let $p(n) = \frac{\omega(n)}{n}$ for some $\omega(n)$ with $\lim_{n\to\infty} \omega(n) = \infty$ and $\lim_{n\to\infty} \sqrt{n}p(n) = 0$. Suppose that $|A(0)| \geq \frac{1+\nu}{np^2(n)}$ with $\nu > 0$.

By Proposition 4.5, there exists a time $t(n)$ and $\alpha(n)$ with $\lim_{n\to\infty} \alpha(n) = \infty$ such that

$$\lim_{n\to\infty} \mathbb{P}\Big\{A\big(t(n)\big) \geq \frac{\alpha(n)}{np^2(n)}\Big\} = 1.$$

Lemma 4.5 implies that for any $\varepsilon > 0$ then for n large we have

$$\lim_{n\to\infty} \mathbb{P}\Big\{\frac{|A(0)| + \sum_{l=1}^{i} |B^l|}{|B^{i+1}|} < \varepsilon\Big\} = 1$$

and we derive that there exists i such that

$$\lim_{n\to\infty} \mathbb{P}\Big\{|B^i| \geq C_1 \frac{\alpha(n)}{np^2(n)}\Big\} = 1. \tag{4.94}$$

By Proposition 4.4, we have there exists $k > 0$ such that

$$\lim_{n\to\infty} \mathbb{P}\Big\{|B^{i+k}| > \frac{C_2}{n^{1-\varepsilon}p^{2-\varepsilon}}(n)\Big||B^i| \geq C_1 \frac{\alpha(n)}{np^2(n)}\Big\} = 1. \tag{4.95}$$

Combine (4.95) with (4.94), we find that if $|A(0)| \geq \frac{1+\nu}{np^2(n)}$ then there exists $l > 0$ such that

$$\lim_{n\to\infty} \mathbb{P}\Big\{|B^l| > \frac{C_2}{n^{1-\varepsilon}p^{2-\varepsilon}(n)}\Big\} = 1. \tag{4.96}$$

By Proposition 4.3 we have that for $\varepsilon > \frac{1}{2^k}$

$$\lim_{n\to\infty} \mathbb{P}\Big\{|B^{l+k}| > \frac{C_3}{p(n)}\Big||B^l| \geq \frac{C_2}{n^{1-\varepsilon}p^{2-\varepsilon}(n)}\Big\} = 1. \tag{4.97}$$

Combine (4.97) with (4.96), we find that if $|A(0)| \geq \frac{1+\nu}{np^2(n)}$ then there exists $m > 0$ such that

$$\lim_{n\to\infty} \mathbb{P}\Big\{|B^m| > \frac{C_3}{p(n)}\Big\} = 1. \tag{4.98}$$

By Proposition 4.2, we have that there exists $C_4 > 0$ such that

$$\lim_{n\to\infty} \mathbb{P}\Big\{\frac{|A(\lfloor \frac{C_3}{p(n)} \rfloor)|}{n} > C_4 n\Big|\Big|A\Big(\Big\lfloor \frac{C_3}{p(n)} \Big\rfloor\Big)\Big| > \frac{C_3}{p(n)}\Big\} = 1. \tag{4.99}$$

Combine (4.98) with (4.99), we find that if $|A(0)| \geq \frac{1+\nu}{np^2(n)}$ then there exists a step $t(n) = \lfloor \frac{C_3}{p(n)} \rfloor$ such that

$$\lim_{n\to\infty} \mathbb{P}\left\{ \frac{|A(t(n))|}{n} > C_4 n \right\} = 1. \tag{4.100}$$

By Proposition 4.1, we know that if we activate a positive part of the graph then with a probability tending to 1 as $n \to \infty$, we activate almost all the graph. For any $\varepsilon > 0$ then

$$\lim_{n\to\infty} \mathbb{P}\left\{ \frac{|A(n)|}{n} < 1 - \varepsilon \middle| \frac{|A(t(n))|}{n} > C_4 \right\} = 0 \tag{4.101}$$

Combine (4.100) with (4.101), we find that if $|A(0)| \geq \frac{1+\nu}{np^2(n)}$ then for any $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}\left\{ \frac{|A(n)|}{n} < 1 - \varepsilon \right\} = 0$$

which proves Theorem 2.2 □

4.3 Proof of Theorem 2.3

We consider the case when the probability of having an edge between two vertices is $p(n) = \frac{c}{\sqrt{n}}$. In such a case $\frac{1}{np^2(n)} = \frac{1}{c^2}$ and the critical size is of order constant. We prove in the following that starting with a constant number of vertices greater than two then with a positive probability, we activate almost all the graph. Notice that Propositions 4.1, 4.4, 4.3, 4.2 and Lemma 4.5 apply under conditions of Theorem 2.3, i.e. with $p(n) = \frac{c}{\sqrt{n}}$.

Proof of Theorem 2.3. Let $p(n) = \frac{c}{\sqrt{n}}$ where $c > 0$ is a constant. Let $|A(0)| = k \geq 2$, suppose that t is a constant then the probability of activation is

$$\begin{aligned} p_a(t) &= \frac{c^2}{n} t \Big(1 - \frac{c}{\sqrt{n}}\Big)^{t-1} \\ &\geq \frac{c^2}{n} t \Big(1 - (t-1)\frac{c}{\sqrt{n}}\Big). \end{aligned} \tag{4.102}$$

Let $\varepsilon > 0$ then there exists $N > 0$ such that for $n > N$ we have

$$p_a(t) \geq \frac{c^2}{n} t(1 - \varepsilon). \tag{4.103}$$

In the case when $t = o\big(\sqrt{n}\big)$, the probability of activation increases with the number of activations. We can bound from below the process of activation after time t by a

branching process where the number of offspring from a point is distributed as $Bin(n(1-\varepsilon), \frac{c^2 t}{n}(1-\varepsilon))$.

If $c^2 t(1-\varepsilon)^2 > 1$ then with a positive probability, the branching process survives and hence the process of activation too.

If we start with t vertices and $c^2 t \leq 1$ then with a positive probability, we can activate $\lfloor \frac{1}{c^2} \rfloor + 1 - t$ vertices. At step $t_0 = \lfloor \frac{1}{c^2} \rfloor + 1$, we have $c^2\Big(\lfloor \frac{1}{c^2} \rfloor + 1\Big) > 1$ so for $\varepsilon > 0$ small enough this gives $c^2 t_0(1-\varepsilon)^2 > 1$ and we have

$$p_a(t_0) > \frac{1}{n}$$

and then as before, the process of activation survives with a positive probability. We know that there exists $\alpha(n)$ with $\lim_{n\to\infty} \alpha(n) = \infty$, a time $t(n)$ and a positive constant ξ such that

$$\lim_{n\to\infty} \mathbb{P}\Big\{ |A\big(t(n)\big)| > \alpha(n) \Big\} = \xi > 0. \tag{4.104}$$

Moreover Propositions 4.1, 4.2, 4.3 and 4.4 apply in the case of $p(n) = \frac{c}{\sqrt{n}}$. Using the same chain of arguments as in the proof of Theorem 2.2, this gives

$$\lim_{n\to\infty} \mathbb{P}\left\{ \frac{|A(n)|}{n} < 1 - \varepsilon \middle| |A\big(t(n)\big)| > \alpha(n) \right\} = 0, \tag{4.105}$$

which implies equation (2.5) of Theorem 2.3. Combine (4.104) with (4.105), we can conclude that if $|A(0)| = k \geq 2$, with a positive probability, we activate almost all the graph.

The probability that the activation spreads through a positive part of the graph is bounded away from 1 as the probability that we have no activation after $|A(0)|$ steps is positive. We have equation (2.4) of Theorem 2.3

$$1 > \lim_{n\to\infty} \mathbb{P}\left\{ \frac{|A_n(n)|}{n} > 1 - \varepsilon \right\} \geq \zeta(k, c).$$

□

4.4 Proof of Theorem 2.4

We consider the last case when the probability of having an edge between two vertices is much larger than $\frac{1}{\sqrt{n}}$. Suppose that $p(n) = \frac{\omega(n)}{\sqrt{n}}$ where $\lim_{n\to\infty} \omega(n) = \infty$ and that $|A(0)| = k \geq 2$ then with a probability tending to 1 as n goes to infinity the process of activation survives and so we activate almost all the graph.

Proof of Theorem 2.4 . Let $p(n) = \frac{\omega(n)}{\sqrt{n}}$ where $\lim_{n\to\infty} \omega(n) = \infty$ and $\lim_{n\to\infty} p(n) = 0$. Suppose $|A(0)| = k \geq 2$ is a constant then the probability of activation is $p_a(k) =$

$\frac{\omega^2(n)k}{n}\Big(1-\frac{\omega(n)}{\sqrt{n}}\Big)^{k-1}$. The probability of activation $p_a(t)$ increases with t. We can bound from below the process of activation after time k by a branching process where the number of offspring from a point is distributed as $X \in Bin\Big(n(1-\varepsilon), \frac{\omega^2(n)k}{n}(1-\varepsilon_0)\Big)$.

For such a process, the probability of ultimate extinction is the smallest nonnegative root of the equation.

$$tu = g_X(u), \tag{4.106}$$

where $g_X(u) = \mathbb{E}u^X$ is the probability generating function of the random variable X. By assumption $X \in Bin(n(1-\varepsilon), \frac{\omega^2(n)k}{n})$, (4.106) becomes

$$\Big(1-\frac{\omega^2(n)}{n}k(1-\varepsilon_0)(1-u)\Big)^{n(1-\varepsilon)} = u. \tag{4.107}$$

We know that

$$\Big(1-\frac{1}{n}\omega^2(n)k(1-\varepsilon)(1-u)\Big)^{n(1-\varepsilon)} \leq e^{-\omega^2(n)k(1-\varepsilon)(1-\varepsilon_0)(1-u)}.$$

and $u_0 = 1$ is always solution. For any $u_1 > 0$ there exists N_0 such that for any $n > N_0$ we have $e^{-\omega(n)k(1-\varepsilon)(1-\varepsilon_0)(1-u_1)} < u_1$ and so $u_1 > 0$ is not a solution.

We know that $\lim_{n\to\infty} e^{-\omega^2(n)k(1-\varepsilon)(1-\varepsilon_0)(1-u)} = 0$ then $\lim_{n\to\infty}\Big(1+\frac{1}{n}\omega^2(n)k(1-\varepsilon)(u-1)\Big)^n = 0$ and $u = 0$ is the limiting solution to the equation (4.107). So the probability of ultimate extinction of the process is 0 and the survival probability of the branching process bounding from below our process converges to 1 as $n \to \infty$.

Hence the process survives with probability tending to 1 as $n \to \infty$ and reaches a size $\alpha(n)$. From that point we can bound from below our process with the process where the probability of a link between two vertices is $p(n) = \frac{c}{\sqrt{n}}$ then using Theorem 2.3 we find that

$$\lim_{n\to\infty} \mathbb{P}\left\{\frac{|A(n)|}{n} < 1-\varepsilon\right\} = 0.$$

which proves Theorem 2.4.

□

Bibliography

[1] P. Balister, B. Bollobás and R. Kozma, Large Deviations for Mean Field Models of Probabilistic Cellular Automata. Random Structures and Algorithms 29 (2006), no. 3, 399–415.

[2] M. Deijfen and O. Häggström, Nonmonotonic coexistence regions for the two-type Richardson model on graphs. Electron. J. Probab. 11 (2006), no. 13, 331–344)

[3] S. Janson, T. Łuczak, and A. Ruciński, Random graphs, Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[4] T. M. Liggett, Stochastic interacting systems: contact, voter and exclusion processes. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 324. Springer-Verlag, Berlin, 1999.

[5] T. Turova and A. Villa, On a phase diagram for random neural networks with embedded spike timing dependent plasticity. BioSystems (2007) Volume 89, Issues 1-3, 280-286.