Spread of Infection over P.A. random graphs with edge insertion

Caio Alves and Rodrigo Ribeiro

Alfréd Rényi Institute of Mathematics, Budapest.

Departament of Mathematics, University of Colorado Boulder.
E-mail address: `rodrigo.ribeiro@colorado.edu`

Abstract. In this work we investigate a bootstrap percolation process on random graphs generated by a random graph model which combines preferential attachment and edge insertion between previously existing vertices. The probabilities of adding either a new vertex or a new connection between previously added vertices are time dependent and given by a function f called the edge-step function. We show that under integrability conditions over the edge-step function the graphs are highly susceptible to the spread of infections, which requires only 3 steps to infect a positive fraction of the whole graph. To prove this result, we rely on a quantitative lower bound for the maximum degree that might be of independent interest.

1. Introduction

Investigation of spread of infections or propagation of information over networks is a crucial problem which arises naturally in many areas of science going from social sciences to economics. The recent pandemic and the dissemination of fake news in a highly interconnected society make evident the relevance of studies involving dissemination over networks.

An accurate model for spread of infections must accommodate somehow the randomness of the environment, since concrete situations such as human societies produce networks which evolve in time according to random rules. In this direction, random graph models whose rule of evolution combines the so-called *preferential attachment rule* became a natural environment to the investigation of infectious processes Berger et al. (2005); Jacob et al. (2019); Can (2015). The reason is that this mechanism of attachment which is driven by popularity, that is, individuals tends to get connected with the more popular ones, proved that it is capable of capturing network properties shared by many networks in real-file Strogatz and Watts (1998); Barabási and Albert (1999). We do not intend to cover the vast literature of preferential attachment random graphs, but we refer the reader to the books of van der Hofstad (2017) and of Durrett (2007) for a wide and rigorous introduction to many important random graph models.

In this work we investigate the bootstrap percolation process on random graphs generated by a random graph model which combines preferential attachment and edge insertion between previously existing vertices. In order to properly state and discuss the nature of our results we will define the underlying random graph model and latter the Bootstrap percolation process performed over our random graphs.

1.1. *The model.*

The model depends on a real non-negative function f, called the *edge-step function*, with domain given by the semi-line $[1, \infty)$ such that $||f||_\infty \leq 1$. Though the results here stated also apply for any finite given initial graph, we will assume henceforth that the process starts from an initial graph G_1 consisting in one vertex and one loop in order to simplify notation. As the process evolves, one of the two graph stochastic operations below may be performed on the graph G:

- *Vertex-step* - Add a new vertex v and add an edge $\{u, v\}$ by choosing $u \in G$ with probability proportional to its degree. More formally, conditionally on G, the probability of attaching v to $u \in G$ is given by

$$P(v \rightarrow u | G) = \frac{\mathrm{degree}(u)}{\sum_{w \in G} \mathrm{degree}(w)}. \tag{1.1}$$

- *Edge-step* - Add a new edge $\{u_1, u_2\}$ by independently choosing vertices $u_1, u_2 \in G$ according to the same rule described in the vertex-step. We note that both loops and parallel edges are allowed.

We let $\{Z_t\}_{t \in \mathbb{N}}$ be an independent sequence of random variables such that $Z_t \overset{\mathrm{d}}{=} \mathrm{Ber}(f(t))$. We then define a markovian random graph process $\{G_t(f)\}_{t \geq 1}$ as follows: begin with initial state G_1. Given $G_t(f)$, obtain the (multi)graph $G_{t+1}(f)$ by either performing a *vertex-step* on $G_t(f)$ when $Z_t = 1$ or performing an *edge-step* on $G_t(f)$ when $Z_t = 0$.

The above model can be seen as a modification of the one proposed by Cooper and Frieze (2001), which incorporated edge insertion into the preferential attachment rule. When modeling evolving networks for which new connections continue to be made between existing vertices, the possibility of adding new edges can be crucial. Even when the edge-step function is constant and equal to $p \in (0, 1)$, there is a qualitative change to the connectivity properties of the graph, as a large complete subgraph emerges among old vertices see Alves et al. (2017). The case when $f(t) \downarrow 0$ is also interesting from the modeling perspective, as it corresponds to networks in which new nodes appear less frequently with time, but edge-creation continues vigorously.

We will make use of natural numbers, mainly i, j and k, to denote respectively the i-th, j-th and k-th vertex added by the random graph process. Moreover, we will let $d_t(i)$ be the *degree of the i vertex at time t*. If i has not been added yet, then $d_t(i) = 0$.

1.2. *Regularity conditions.*

Again, in order to properly state and discuss our results we will need to introduce some regularity conditions for the functions f. Some of theses conditions have the objective of preventing pathological examples, for instance if one drops monotonicity it is possible to construct f such that the sequence of graphs $\{G_t(f)\}_{t \in \mathbb{N}}$ has two sub sequences of graphs: one similar to the BA random tree and the other one a quasi-complete graph, see the discussion in Section 6 of Alves et al. (2020).

Some of our results will require information about the asymptotic behavior of f and for this reason a wide class of functions will play important role: the regularly varying functions. We say that a positive function f is a *regularly varying function* (*r.v.f* for short) at infinity with *index of regular variation* $-\gamma$, for some $\gamma \geq 0$, if

$$\lim_{t \to \infty} \frac{f(at)}{f(t)} = a^{-\gamma},$$

for all $a > 0$. The special case $\gamma = 0$ is called *slowly varying function*. Below we define the conditions over f that will be useful to our purposes. For $p \in [0, 1]$, we define conditions

$$f \text{ decreases to } p. \tag{D_p}$$

We will also need to define some summability conditions:

$$\sum_{s=1}^{\infty} f(s) = \infty. \tag{V_∞}$$

The above condition relates to the expected number of vertices. Since at each step s we add a vertex with probability $f(s)$, the above condition tells us that the process keeps introducing new vertices into the graph.

$$\sum_{s=1}^{\infty} \frac{f(s)}{s} < \infty. \tag{S}$$

Whenever (S) does not hold for a specific f, we say that f satisfies (S)^c. For any $\gamma \in [0, 1]$, we let RES$(-\gamma)$ be the following class of functions

$$\text{RES}(-\gamma) := \{f : [1, \infty] \longrightarrow [0, 1] \mid f \text{ satisfies } (D_0) \text{ and is r.v.f with index } -\gamma\}. \tag{RES}$$

Due to degree of freedom one has to choose f, since it is a function, one can study this model in several different contexts. It has been studied under different regularity conditions on f and different graph properties have been investigated. In Alves et al. (2021) the authors proved that, under RES$(-\gamma)$, the empirical degree distribution of $G_t(f)$ is close to a power-law whose exponent is given by $2 - \gamma$, when $\gamma \in [0, 1)$. Whereas in Alves et al. (2020), the authors investigated how the diameter of $G_t(f)$ depends on f under different conditions on the asymptotic behavior of f. A variety of regimes for the diameter can be achieved, from $O(1)$ regimes to $\Theta(\log t)$ ones. All this variety of regimes is explained by the presence of edge-steps. For comparison, the original preferential attaching model, Barabási and Albert (1999), only generates graphs with power-law degree distribution whose exponent is greater than 2 and the diameter of the graph at time t of order $\log t$, see Bollobás et al. (2001); Bollobás and Riordan (2004).

1.3. *Main Results.* One of the central questions in the study of infectious processes is the outbreak, when the whole or a large proportion of the network becomes infected. This macro phenomena can be triggered by local decisions whose effect scales to a macro state. In a nutshell part of our results address the 'amount' of local decisions needed to trigger the outbreak and the time needed to observe it.

In order to investigate the Bootstrap percolation process we need to prove certain properties of the underlying environment. For this reason, our first result is a quantitative lower bound for the maximum degree. In order to state it, we need to introduce new notation. Given an edge-step function f, we define the normalizing function $\phi : \mathbb{N} \to (0, \infty)$ by

$$\phi(t) = \phi(t, f) := \prod_{s=1}^{t-1} \left(1 + \frac{1}{s} - \frac{f(s+1)}{2s}\right) \tag{1.2}$$

We note that condition (S) is equivalent to $\phi(t)$ growing linearly as a function of t. We will give a formal proof of this fact in Lemma 2.2. We have

Theorem 1.1 (Lower bound for maximum degree). *Let f be an edge-step function such that $f(t)$ goes to zero as t goes to infinity. For every $N \in \mathbb{N}$, there exists $C_f > 0$ depending on f only such that*

$$\mathbb{P}\left(\forall t \in \mathbb{N}, \exists i \in \{1, 2, \cdots, N\} \text{ such that } d_t(i) \geq \frac{\phi(t)}{\phi(N)}\right) \geq 1 - \exp\{-C_f N\}. \tag{1.3}$$

Furthermore, there exists almost surely a random integer $N_0 \geq 0$ such that for every $t \geq N_0$, there exists at least one vertex i in $G_t(f)$ such that

$$d_t(i) \geq \frac{\phi(t)}{\phi(N_0)}. \tag{1.4}$$

As said before, in this work we analyze the Bootstrap Percolation model over $G_t(f)$. We will consider $f \in \mathrm{RES}(-\gamma)$, with $\gamma \in [0,1)$ and satisfying condition (S). Our estimates allow us to construct structures on the graph $G_t(f)$ that make it highly susceptible to the spread of infections.

Now, let us define the Bootstrap Percolation process. Given a finite (multi)graph $G = (V(G), E)$, a number $a \in [0, |V(G)|]$, and an integer $r \geq 2$, we define the *bootstrap percolation* measure $\mathbb{Q}_{G,a,r}$ on G with threshold r and rate of infection a in the following manner:

- Each vertex $v \in V(G)$ is *infected* at round 0 independently of the others with probability $a|V(G)|^{-1}$. The collection of all infected vertices at round 0 is denoted by $\mathcal{I}_0$.

- At round $s \in \mathbb{N}$, every vertex connected to $\mathcal{I}_{s-1}$ via at least r edges becomes infected.

- We let $\mathcal{I}_\infty$ be the set of all infected vertices when the process stabilizes, that is,

$$\mathcal{I}_\infty = \cup_{s \geq 0} \mathcal{I}_s.$$

Bootstrap percolation can be seen as a monotone cellular automata Schonmann (1992) (infected stay infected), an infection process on a graph, or a monotone version of Glauber dynamics. In the physics community, it has been proposed by Chalupa et al. (1979), where it was shown that this process exhibits first-order phase transition when the underlying graph is a tree. It has been extensively studied by the scientific community of several fields Balogh et al. (2010, 2012, 2006); Flocchini et al. (2004); Granovetter (1978).

In Amin Abdullah and Fountoulakis (2018) the authors study the bootstrap percolation process on the preferential attachment random graph where each vertex has $m \in \mathbb{N}$ outgoing edges with end vertices chosen according to an affine preferential attachment rule, that is, the PA rule in (1.1) but with a $\delta > -m$ summed in both the numerator and denominator. There they prove the existence of a *critical function* $a_t^c : \mathbb{N} \to \mathbb{R}_+$ such that the bootstrap percolation process on this random graph at time t with threshold $r \leq m$ and rate $a_t \gg a_t^c$ infects the whole graph with high probability, but the same process with rate $a_t' \ll a_t^c$ dies out without before infecting a positive proportion of the graph, also with high probability.

In our context it is not possible for the infection to spread to the whole graph due to the existence of many vertices with degree smaller than r at all times. See Theorem 1 in Alves et al. (2021), which states that, for any $d \in \mathbb{N}$, the proportion of vertices having degree d is of order $d^{-(2-\gamma)}$, w.h.p. We therefore turn the problem into asking if $\mathcal{I}_\infty$ eventually encompasses a set with positive density in the vertex set. In this sense, we show in our next result that the bootstrap percolation is *always supercritical* for the graph $G_t(f)$, with f under (S), i.e., every unbounded rate sequence gives rise to a set of infected vertices with positive density:

Theorem 1.2 (The outbreak phenomenon). *Let f be an edge-step function in $\mathrm{RES}(-\gamma)$, with $\gamma \in [0,1)$, and satisfying the summability condition (S). Then, for any sequence $(a_t)_{t \in \mathbb{N}}$ increasing to infinity and a integer number $r \geq 2$ there exists a collection of graphs $\mathcal{G}_t = \mathcal{G}_t((a_t)_{t \in \mathbb{N}}, r)$ and a (f, r)-dependent constant $c > 0$ such that,*

$$\mathbb{P}\left(G_t(f) \in \mathcal{G}_t\right) = 1 - o(1), \tag{1.5}$$

as t goes to infinity and for all $G \in \mathcal{G}_t$ the bootstrap percolation process on G with parameters a_t and $r \geq 2$ satisfies

$$\mathbb{Q}_{G,a_t,r}\left(|\mathcal{I}_\infty| \geq c|V(G)|\right) = 1 - o(1), \tag{1.6}$$

as t goes to infinity.

The above theorem illustrates how interconnected $G_t(f)$ is. From the definition of the bootstrap percolation on $G_t(f)$, it follows that the initial infected set $\mathcal{I}_0$ is essentially a_t. Thus, by Theorem 1.2 one can choose a_t of order arbitrarily smaller than the expected number of vertices in $G_t(f)$, as long as it increases to infinity, the infection manages to spread to a positive fraction of the whole graph with high probability.

In order to prove Theorem 1.2 we actually prove a stronger result, which provides an upper bound for the time it takes for the infection to spread to a positive fraction of the whole graph. Formally, for a bootstrap process on G with parameter r and a, given a positive constant c, let τ_c be the following stopping time

$$\tau_c := \inf\{s \geq 0 \ : \ |\mathcal{I}_s| \geq c|V(G)|\}. \tag{1.7}$$

Then we have the following result which is a stronger version of Theorem 1.2.

Theorem 2' (Number of steps for the outbreak). Let be $\mathcal{G}_t$ and c as in Theorem 1.2. Then, for any sequence $(a_t)_{t \in \mathbb{N}}$ increasing to infinity, $r \geq 2$ and $G \in \mathcal{G}_t$ the following holds

$$\mathbb{Q}_{G,a_t,r}\left(\tau_c \leq 3\right) = 1 - o(1).$$

Roughly speaking, the above result states that typically, a bootstrap percolation process on $G_t(f)$ needs at most 3 steps to reach a positive fraction of $G_t(f)$, making clearer the interconnectedness of the graphs $\{G_t(f)\}_{t \geq 1}$ when f is under (S) and $\text{RES}(-\gamma)$.

Remark 1.3. One could ask the question: *what happens when the process and the graph evolve simultaneously?*. Unfortunately, it does not seem to lead to new and interesting behavior, at least when the summability condition (S) is satisfied. Indeed, starting the process with any graph containing an infected vertex, eventually for large enough time this vertex will have very large degree, comparable to the maximum. When this happens, one can prove using the same construction outlined in our paper, that eventually a proportion of the vertices becomes infected. If no vertex starts infected, then no vertex can become infected later.

On the other hand, if it would be possible for vertices to either recover (as in the contact process), or become infected by themselves, this should alter the behavior of the process dramatically. We believe that these are interesting questions, which would require new techniques, so that finer connectivity properties of the model can be investigated and used.

2. Lower bound for the maximum degree

In this section, we will use martingale arguments in order to show a quantitative lower bound for the maximum degree of $G_t(f)$. This type of result is usually not as strong as its upper bound counterpart, since we may pay a relatively small price in order to make a vertex i behave like it was born much after its mean appearance time. We can however obtain stronger results when dealing with the *total degree of a collection of vertices*. This in turn gives us good lower bounds for the degree of some random vertex, by the pigeonhole Principle.

We start with an elementary stochastic recurrence identity concerning the degree of a given vertex i. Define as $\tau(i)$ the random time in which the i-th vertex is born, and for each time $s \in \mathbb{N}$, let $\Delta d_s(i)$ be

$$\Delta d_s(i) := d_{s+1}(i) - d_s(i).$$

Moreover, we will denote by $\{\mathcal{F}_t\}_{t \geq 1}$ the natural filtration for the process $\{G_t(f)\}_{t \geq 1}$.

Lemma 2.1. *Given the vertex with index i we have, in the event where $s \geq \tau(i)$,*

$$\mathbb{E}\left[\Delta d_s(i)|\mathcal{F}_s\right] = \left(\frac{1}{s} - \frac{f(s+1)}{2s}\right) d_s(i), \tag{2.1}$$

which implies that for $X_{t_i,s}$ defined as

$$X_{t_i,s} := \frac{d_s(i)}{\phi(s)} \mathbb{1}\{\tau(i) = t_i\} \tag{2.2}$$

the process $\{X_{t_i,s}\}_{s \geq t_i}$ is a martingale w.r.t $\{\mathcal{F}_s\}_{s \geq t_i}$.

Proof: We start by noticing that $\Delta d_s(i)$ is either 0, 1, or 2. In order for it to be 1, we may either add a new vertex to the graph and connect it to i, or we add a new edge and choose i to be one of its endvertices. In the case where the increment is 2, we add a new edge to the graph and choose i twice, adding a loop to it. This yields the formula, on the event where $s > \tau_i$,

$$\mathbb{E}\left[\Delta d_s(i) | \mathcal{F}_s\right] = f(s+1)\frac{d_s(i)}{2s} + 2\left(1 - f(s+1)\right)\frac{d_s(i)}{2s}\left(1 - \frac{d_s(i)}{2s}\right) + 2\left(1 - f(s+1)\right)\left(\frac{d_s(i)}{2s}\right)^2$$

$$= \left(\frac{1}{s} - \frac{f(s+1)}{2s}\right)d_s(i), \tag{2.3}$$

which in turn, by the definition of ϕ in (1.2), implies that the process in (2.2) is a martingale. $\quad\square$

Let the function $\xi : \mathbb{N} \to \mathbb{R}$ be defined by

$$\xi(s) = \xi(s, f) := \frac{\phi(s)}{s} = \prod_{r=1}^{s-1}\frac{r}{r+1}\prod_{r=1}^{s-1}\left(1 + \frac{1}{r} - \frac{f(r+1)}{2r}\right) = \prod_{r=1}^{s-1}\left(1 - \frac{f(r+1)}{2(r+1)}\right). \tag{2.4}$$

In order to properly calculate sums involving ϕ, the following lemma about ξ will be necessary:

Lemma 2.2. *If f satisfies (D_0), that is, $f(t)$ decreases monotonically to 0 as t increases, then ξ is a slowly varying function. Furthermore, if f satisfies condition (S), then $\phi(t)$ grows linearly in t.*

Proof: We may write ξ as

$$\xi(s) = \exp\left\{\sum_{r=1}^{s-1}\log\left(1 - \frac{f(r+1)}{2(r+1)}\right)\right\} = \exp\left\{\sum_{r=1}^{s-1}\left(-\frac{f(r+1)}{2(r+1)} + O(r^{-2})\right)\right\}. \tag{2.5}$$

Therefore ξ is a slowly varying function as long as $f(t)$ goes to zero as t goes to infinity, since for any $a \in \mathbb{R}_+$,

$$\begin{aligned}
\frac{\xi(as)}{\xi(s)} &= \exp\left\{\sum_{r=\lceil as \rceil}^{s-1}\left(\frac{f(r+1)}{2(r+1)} + O(r^{-2})\right)\right\} \\
&\leq \exp\left\{C \cdot as^{-1} + \frac{\inf_{r \geq as} f(r)}{2}(\log s - \log as)\right\} \\
&\leq \exp\left\{C \cdot as^{-1} - \frac{\inf_{r \geq as} f(r)}{2}(\log a)\right\} \\
&\xrightarrow{s \to \infty} 1.
\end{aligned} \tag{2.6}$$

From (2.5) and recalling that under condition (S), $\sum_s f(s)/s < \infty$, we have that

$$\xi(s) = \exp\left\{\sum_{r=1}^{s-1}\left(-\frac{f(r+1)}{2(r+1)} + O(r^{-2})\right)\right\} \geq \exp\left\{\sum_{r=1}^{\infty}\left(-\frac{f(r+1)}{2(r+1)} + O(r^{-2})\right)\right\} \geq C_f,$$

for some constant C_f which depends on f only. Since $\xi(s) = \phi(s)/s$, it follows that $\phi(s)$ grows linearly on s. $\quad\square$

With all the previous results at our disposal we can finally prove the main result of this section, which states a lower bound for the maximum degree of $G_t(f)$.

Proof of Theorem 1.1: Fix $N \in \mathbb{N}$ and denote by $\{W_{N,s}\}_{s \geq N}$ the process

$$W_{N,s} := \frac{\sum_{i=1}^{N} d_s(i)\mathbb{1}\{\tau(i) \leq N\}}{\phi(s)}, \tag{2.7}$$

i.e., $W_{N,s}$ denotes the sum of the degree of all vertices added by the process up to time N normalized by $\phi(s)$. Since

$$W_{N,s} = \frac{1}{\phi(s)} \sum_{i \geq 1} d_s(i)\mathbb{1}\{\tau(i) \leq N\}, \tag{2.8}$$

Lemma 2.1 implies that $\{W_{N,s}\}_{s \geq N}$ is a positive martingale w.r.t to $\{\mathcal{F}_s\}_{s \geq N}$ such that

$$\mathbb{E}W_{N,s} = W_{N,N} \equiv \frac{2N}{\phi(N)}. \tag{2.9}$$

Indeed, at time N the sum of the degree of all vertices added up to this time equals twice the number of edges, which is N. In order to shorten the notation, we define

$$D_{N,s} = \sum_{i=1}^{N} d_s(i)\mathbb{1}\{\tau(i) \leq N\}.$$

Since we can update the degree of at most two vertices by an amount of at most 2, it follows that $\{W_{N,s}\}_{s \geq N}$ has bounded increments:

$$\begin{aligned}
|\Delta W_{N,s}| \\
&= \left| \frac{1}{\phi(s+1)} D_{N,s+1} - \frac{1}{\phi(s)} D_{N,s} \right| \\
&= \frac{1}{\phi(s+1)\phi(s)} \left| \phi(s)D_{N,s+1} - \left(1 + \frac{1}{s} - \frac{f(s+1)}{2s}\right) \phi(s)D_{N,s} \right| \\
&= \left| \frac{1}{\phi(s+1)} \Delta D_{N,s} - \frac{(2 - f(s+1))}{2s\phi(s+1)} D_{N,s} \right| \\
&\leq \frac{3}{\phi(s+1)},
\end{aligned} \tag{2.10}$$

which implies $(\Delta W_{N,s})^2 \leq 9\phi(s+1)^{-2}$. We have by Lemma 2.2 that ξ^{-2} is a slowly varying function. By Karamata theory (see Theorem B.3 in the Appendix), the above inequality, the fact that f is decreasing and that $\phi(s) = \xi(s)s$, we obtain

$$\sum_{s=N}^{t} |\Delta W_{N,s}|^2 \leq 9\sum_{s=N}^{t} s^{-2}\xi(s)^{-2} \leq \frac{C}{N\xi(N)^2} = \frac{C}{\phi(N)\xi(N)}. \tag{2.11}$$

We are going to stop our martingale when it becomes unexpectedly small. For this, let η be the stopping time

$$\eta := \inf_{s \geq N} \left\{ W_{N,s} \leq \frac{N}{\phi(N)} \right\}. \tag{2.12}$$

Notice that $\{W_{N,s \wedge \eta}\}_{s \geq N}$ is a martingale having the same expected value as $\{W_{N,s}\}_{s \geq N}$, since we have that $W_{N,N \wedge \eta} \equiv 2N/\phi(N)$. Furthermore, stopping a martingale can only decrease its increments. Thus, by Azuma's inequality, we obtain, by equation (2.11) and the definition of ξ,

$$\mathbb{P}\left(W_{N,s \wedge \eta} \leq \frac{2N}{\phi(N)} - \frac{N}{\phi(N)} \right) \leq \exp\left\{ -C\frac{N^2}{\phi(N)^2}\phi(N)\xi(N) \right\} \leq \exp\{-CN\}. \tag{2.13}$$

Now, by the definition of η, we obtain

$$
\begin{aligned}
\mathbb{P}\left(\exists s \in \mathbb{N} \text{ such that } W_{N,s} \leq \frac{N}{\phi(N)}\right) &\leq \mathbb{P}\left(\eta < \infty\right) \\
&= \lim_{s \to \infty} \mathbb{P}\left(\eta \leq s\right) \\
&= \lim_{s \to \infty} \mathbb{P}\left(W_{N,s \wedge \eta} \leq \frac{N}{\phi(N)}\right) \\
&\leq \exp\left\{-CN\right\},
\end{aligned}
\tag{2.14}
$$

where in the third line we have used the identity of events below

$$
\left\{W_{N,s \wedge \eta} \leq \frac{N}{\phi(N)}\right\} = \left\{\eta \leq s\right\},
$$

which follows from the definition of η. By the pigeonhole Principle the above inequality implies (1.3) since

$$
W_{N,t} > \frac{N}{\phi(N)} \iff \sum_{i=1}^{N} d_t(i) \mathbb{1}\{\tau(i) \leq N\} > \frac{N\phi(t)}{\phi(N)},
\tag{2.15}
$$

implying the existence of at least one vertex with degree $\phi(t)/\phi(N)$ among the N first vertices. $\quad\square$

3. The outbreak phenomenon

In this section we prove Theorem 2' which is a stronger version of Theorem 1.2. The reader may find useful to recall the definition of the bootstrap percolation process at page 1224 as well as definition of τ_c in (1.7). For the sake of organization we will split the proof into several lemmas.

Throughout this section, f will be a function satisfying conditions (D$_0$), (S), and RES($-\gamma$) for $\gamma \in [0,1)$. We will also use the notation $a_n \sim b_n$ to mean $a_n/b_n \to 1$ as $n \to \infty$. Moreover, whenever we use the letter r when dealing with the graph process $\{G_t(f)\}_{t \geq 2}$, this r is the exact same r as the threshold parameter of the Bootstrap Percolation process to be performed in the graphs $G_t(f)$. We then begin by some useful properties of $G_t(f)$ and f when f satisfies the aforementioned conditions. The first lemma is a small result concerning the speed of convergence to zero of f.

Lemma 3.1. *Let f be an edge-step function satisfying conditions (S) and RES($-\gamma$) for $\gamma \in [0,1)$. Then,*

$$
\lim_{t \to \infty} f(t) \log(t) = 0
$$

Proof: In the regime $\gamma > 0$ the product $f(t) \log(t)$ converges to zero as t goes to infinity by virtue of Corollary B.2, since in this case by Theorem B.1, we have $f(t) = \ell(t) t^{-\gamma}$, where ℓ is a slowly varying function. On the other hand, if f is slowly varying ($\gamma = 0$), then for any $K \in \mathbb{N}$, the monotonicity of f implies

$$
\sum_{s \geq K} \frac{f(s)}{s} \geq \sum_{s=K}^{Kt} \frac{f(s)}{s} \geq f(Kt) \sum_{s=K}^{Kt} \frac{1}{s} \geq f(Kt) \log t(1 - o(1)) \sim f(t) \log(t)(1 - o(1)),
\tag{3.1}
$$

since f is slowly varying. Finally, using condition (S) which states that $\sum_s f(s)/s$ is finite and sending K to infinity, we obtain the desired result.

$\square$

It will be useful to our purposes to know the order of magnitude of the expected number of vertices in $G_t(f)$. For this, let $V_t(f)$ be the number of vertices in $G_t(f)$, then, by the Karamata's Theorem (Theorem B) we have that

$$
\mathbb{E}V_t(f) \sim \frac{f(t)t}{1 - \gamma} \implies \mathbb{E}V_t(f) \leq \frac{5}{4(1 - \gamma)} f(t)t,
\tag{3.2}
$$

for large enough t.

In order to prove Theorem 2', we will prove all the intermediate results and the theorem itself under the assumption that the sequence $(a_t)_{t\geq 1}$ satisfies the following technical conditions

$$(i) \quad tf(t)f(a_t)\log(a_t) \to \infty; \quad (ii) \quad a_t \leq t; \quad (iii) \quad f(a_t)\log(a_t) \geq f(t)\log t, \qquad (L)$$

for sufficiently large t. The reason why we need (L) is of technical nature. Having such control over $(a_t)_{t\geq 1}$ makes some arguments easier to follow. However, the following lemma, whose proof is straightforward and will be omitted, guarantees that (L) is not restrictive, as will be proved in Lemma 3.3 below.

Lemma 3.2 (Monotonicity on $(a_t)_{t\geq 1}$). *Let $(a'_t)_{t\geq 1}$ and $(a_t)_{t\geq 1}$ be two sequence of positive real numbers such that $a'_t \leq a_t$ for all $t \geq 1$. Then, given a (multi)graph G, a positive constant c and a natural number r, the following holds*

$$\mathbb{Q}_{G,a'_t,r}\left(|\mathcal{I}_\infty| \geq c|V(G)|\right) \leq \mathbb{Q}_{G,a_t,r}\left(|\mathcal{I}_\infty| \geq c|V(G)|\right),$$

for all $t \geq 1$.

Recall that the parameter a_t is essentially the size of $\mathcal{I}_0$, which is the set of initially infected vertices. Thus, the above result states that if we increase the average size of $\mathcal{I}_0$, we increase our chances of percolating. The upshot is that we do not lose generality by imposing conditions on the growth rate of $(a_t)_{t\geq 1}$, as long as such conditions require making $(a_t)_{t\geq 1}$ go to infinity slower.

Lemma 3.3. *Given an increasing sequence of positive numbers $(a_t)_{t\geq 1}$ going to infinity, there exists a sequence $(a''_t)_{t\geq 1}$ satisfying (L) and such that $a''_t \leq a_t$ for all $t \geq 1$.*

Proof: Taking $a'_t = a_t \wedge t^\beta$ with

$$\beta = \begin{cases} \frac{1-\gamma}{2\gamma} \wedge 1, & \text{if } \gamma > 0, \\ \frac{1}{2}, & \text{if } \gamma = 0, \end{cases}$$

we obtain a sequence satisfying items (i) and (ii) of (L). Item (ii) follows by definition. By Theorem B.1 and Corollary B.2, using the decomposition of $f(t) = \ell(t)t^{-\gamma}$, where ℓ is a slowly varying function, we have

$$\begin{aligned}
tf(t)f(a'_t)\log(a'_t) &= t\frac{\ell(t)}{t^\gamma}\frac{\ell(a'_t)}{(a'_t)^\gamma}\log(a'_t) \\
&\geq t^{1-\gamma}\ell(t)\ell(a'_t)\beta\log(t)t^{-\beta\gamma} \\
&\geq t^{(1-\gamma)/2}\ell(t)\ell(a'_t)\beta\log(t) \\
&\xrightarrow{t\to\infty} \infty.
\end{aligned}$$

Now, to get a sequence satisfying condition (iii), we define a new sequence $(a''_t)_{t\geq 1}$ with $a''_t \leq a'_t$ for all $t \geq 1$ inductively. Recalling that $f(t)\log(t) \to 0$ by Lemma 3.1, we define $n_1 = 1$, and, for $k \in \mathbb{N}$,

$$n_{k+1} := \inf_{m>n_k}\left\{\sup_{n\geq m} f(n)\log(n) \leq f(a'_k)\log(a'_k)\right\}.$$

Letting $a''_t = a'_k$ for $n_k \leq t < n_{k+1}$, we obtain a sequence satisfying condition (iii) for large enough t.

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$

In general lines, in order to prove that $G_t(f)$ is supercritical for the Bootstrap percolation process, the key steps are to construct specific subgraphs in $G_t(f)$, which have good infectious properties in the sense that the infectious process reach those subgraphs more easily and they are interconnected in way that once the infection has reached them, it manages to spread to a positive proportion of the whole graph in a few steps.

In the next lemmas we will make sure $G_t(f)$ has some useful graph properties a.a.s.. We let g and h be the following functions:

$$g(t) := \frac{1-\gamma}{f(t)}; \ h(t) := \frac{C_1}{6(r+1)^2 C_2} \cdot \frac{g(t)}{\log(t)}, \tag{3.3}$$

where r is the threshold parameter of the Bootstrap Percolation process and C_1 and C_2 are positive constants such that

$$C_1 t \leq \phi(t) \leq C_2 t, \tag{3.4}$$

which exists by Lemma 2.2. Note that, by Lemma 3.1, $h(t)$ goes to infinity as t increases.

We now present a sketch of the proof of the outbreak phenomenon, proving w.h.p. the existence of a structure that is highly susceptible to infections. In what follows, everything happens w.h.p. as t goes to infinity.

(i) At time t the graph has a number of vertices of order $f(t)t$, and a distinguished vertex v^* with degree of order at least $t/h(a_t)$;

(ii) At time $(r+1)t$, the above bounds are still valid, but there exists a subgraph H whose *total degree* in $G_{(r+1)t}(f)$ has order at least t, and every vertex of H sends at least r edges to v^*;

(iii) At time $2(r+1)^2 t$, the above is still valid, but now there exists a subset $S \subset V(G_{2(r+1)^2 t}(f))$ with cardinality of the same order as $V(G_{2(r+1)^2 t}(f))$ such that each vertex of S has at least r connections with H. Other than that, the distinguished vertex v^* has order at least $f(t)t/h(a_t)$ distinct neighbors.

This is enough to prove our result for the graph $G_{2(r+1)^2 t}(f)$:

(i) The lower bound on the number of distinct neighbors of v^* and our choice of h imply that this vertex becomes infected in the first step;

(ii) Since every vertex of H has at least r connections to v^*, H as a whole becomes infected in the second step;

(iii) Now S has cardinality comparable to $V(G_{2(r+1)^2 t})$, and each vertex in S sends r connections to H, implying that S becomes infected in the third step, implying the existence of an outbreak.

We can prove the same result for the graph G_t just by adjusting the constants. The sequence of images 3.1, 3.2, and 3.3 illustrates the construction of the highly susceptible structure, while the sequence 3.4, 3.5, 3.6 shows how an infection spreads through this structure.

In accordance to the proof sketch above, for each time $t \in \mathbb{N}$, let $\mathcal{P}_t^1$ be the following collection of (multi)graphs

$$\mathcal{P}_t^1 := \left\{ G = (V, E) : |V| \leq \frac{15 f(t)t}{8(1-\gamma)} \text{ and } \exists v^* \in V, \text{with } d_G v^* \geq \frac{C_1 t}{C_2 h(a_t)} \right\}. \tag{3.5}$$

In words, $\mathcal{P}_t^1$ is the collection of multigraphs with at most $15 f(t)t/8(1-\gamma)$ vertices and containing at least one vertex v^* whose degree is at least $C_1 C_2^{-1} t/h(a_t)$, where C_1 and C_2 are the f-dependent constants in (3.4). We call the distinguished vertex v^* a *star*. The next result ensures $G_t(f)$ has property $\mathcal{P}_t^1$ a.a.s..

Lemma 3.4. *Let f be an edge-step function satisfying conditions (S) and $\text{RES}(-\gamma)$ for $\gamma \in [0, 1)$. Then, there exists a f-dependent positive constant C_f such that*

$$\mathbb{P}\left(G_t(f) \in \mathcal{P}_t^1 \right) \geq 1 - \exp\{-C_f h(a_t)\} - \exp\left\{ -\mathbb{E}V_t(f)/4 \right\}.$$

Proof: By Theorem 1.1, choosing $N = h(a_t)$, we have that

$$\mathbb{P}\left(\exists i \in \{1, 2, \cdots, h(a_t)\}, \text{ such that } d_t(i) \geq \frac{C_1}{C_2} \frac{t}{h(a_t)} \right) \geq 1 - \exp\{-C_f h(a_t)\} \tag{3.6}$$
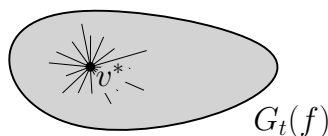
FIGURE 3.1. At time t, w.h.p. there is a star subgraph of $G_t(f)$ whose center v^* has degree at least $th(a_t)^{-1}$.
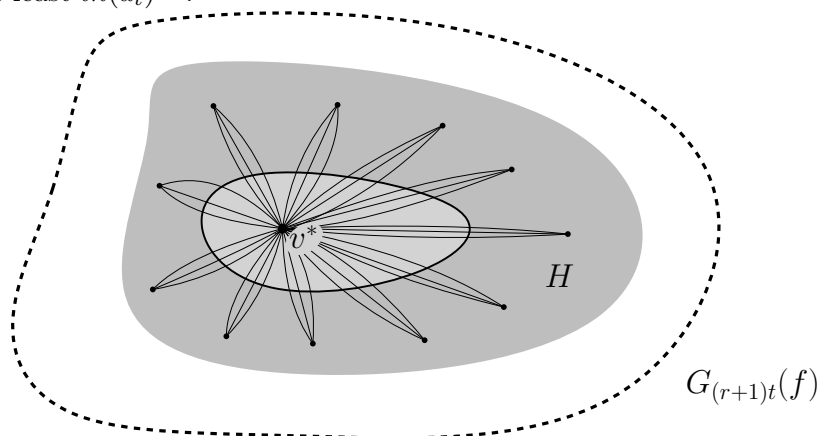

FIGURE 3.2. At time $(r+1)t$, w.h.p. there is a subgraph H with total degree of order t, where every vertex in H has at least r connections with v^*. In the figure, we represent $r = 3$.
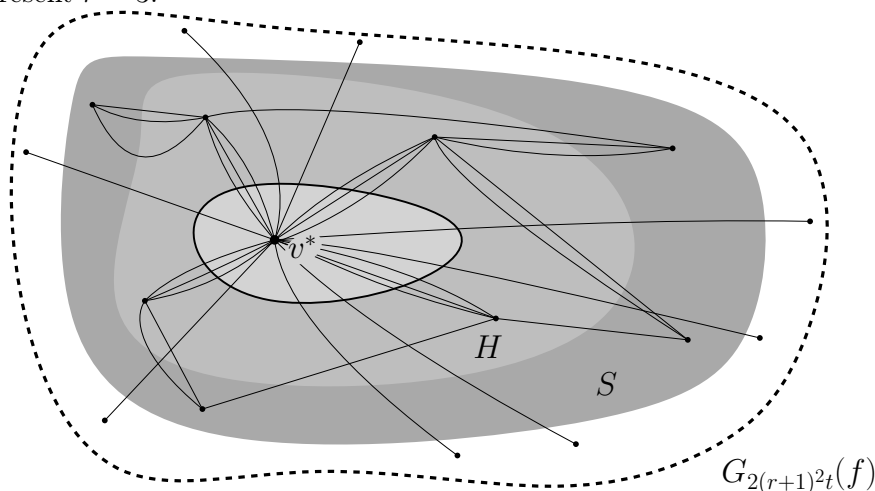

FIGURE 3.3. At time $2(r+1)^2 t$, w.h.p. there is a subset S comprising a positive proportion of vertices of the graph, where every vertex of S sends r connections to H. Furthermore, the number of distinct neighbors of v^* has a proportion of order at least $h(a_t)^{-1}$ of the total number of vertices.

for some f-dependent constant C_f. Notice that for this choice of N, we do have that

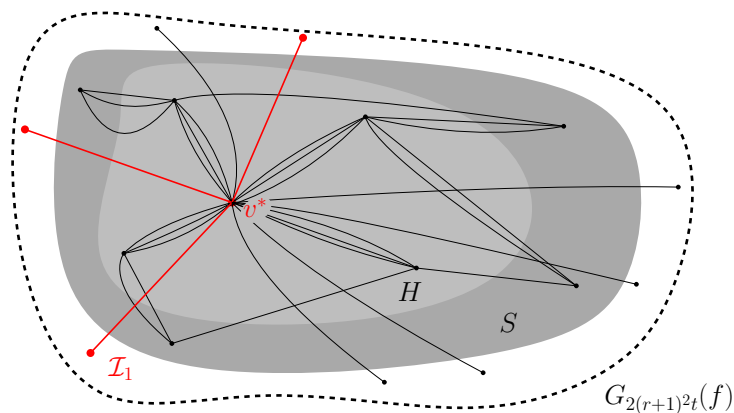$$\frac{\phi(t)}{\phi(N)} \geq \frac{C_1 t}{C_2 h(a_t)}.$$

FIGURE 3.4. The process begins, infecting with high probability a large number of neighbors of v^*. In the first step, v^* becomes infected.
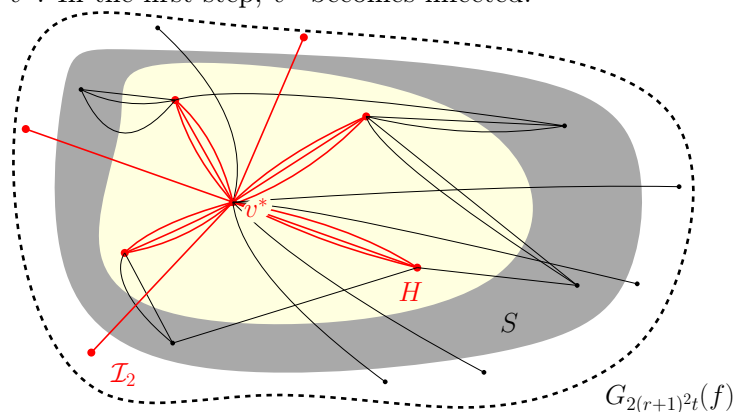

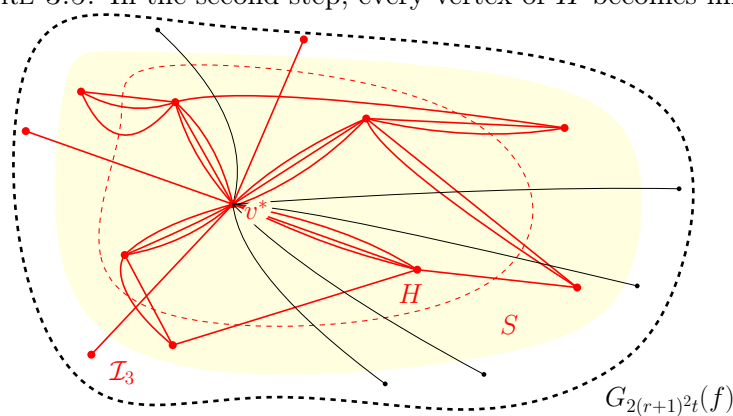FIGURE 3.5. In the second step, every vertex of H becomes infected.

FIGURE 3.6. In the third step, every neighbor of S becomes infected, giving rise to an outbreak phenomenon in 3 steps.

Since $V_t(f)$ is the sum of independent Bernoulli random variables and $\mathbb{E}V_t(f) \leq 5f(t)t/4(1-\gamma)$, a Chernoff bound yields

$$\mathbb{P}\left(V_t(f) \geq \frac{15f(t)t}{8(1-\gamma)}\right) \leq \exp\left\{-\mathbb{E}V_t(f)/4\right\}, \tag{3.7}$$

which proves that $G_t(f)$ has property $\mathcal{P}_t^1$ a.a.s. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

Now, if H is a subset of vertices of a (multi)graph G, then $d_G(H)$, the degree of H, is simply the sum of the degree of the vertices in H. With this terminology in mind, let $\mathcal{P}_{(r+1)t}^2$ be the property below

$$\mathcal{P}_{(r+1)t}^2 := \left\{ \begin{array}{c} G = (V, E) : |V| \leq \frac{15f((r+1)t)(r+1)t}{8(1-\gamma)}, \exists v^* \in V, \text{ s.t } d_G(v^*) \geq \frac{C_1 t}{C_2 h(a_t)}, \\ \exists H \subset G \text{ s.t. } d_G(H) \geq t/8 \\ \text{and } \forall v \in H \text{ there are } r \text{ edges between } v \text{ and } v^* \end{array} \right\}. \tag{3.8}$$

Now we will prove that $G_t(f)$ has property $\mathcal{P}_{(r+1)t}^2$ a.a.s.

Lemma 3.5. *Let f be an edge-step function satisfying conditions (S) and $\mathrm{RES}(-\gamma)$ for $\gamma \in [0, 1)$. Then,*

$$\mathbb{P}\left(G_t(f) \in \mathcal{P}_{(r+1)t}^2\right) = 1 - o(1).$$

Proof: Since f will be fixed the entire proof, to avoid clutter, we will drop the reference on f in the notation $G_t(f)$, writing simply G_t. We will prove first that conditioned on $G_t \in \mathcal{P}_t^1$, the graph $G_{(r+1)t}$ has property $\mathcal{P}_{(r+1)t}^2$ a.a.s. Thus, suppose that G_t has property 1 and let v^* be its star. Let j be a vertex whose degree is at least $g(t)$ at time t. Notice that on the event $V_t(f) \leq 15f(t)t/8(1-\gamma)$, there exists at least one such vertex, otherwise we would have

$$\sum_{v \in G_t} d_t(v) \leq \frac{15g(t)f(t)t}{8(1-\gamma)} = \frac{15t}{8} < 2t,$$

which is a contradiction since G_t is a graph with t edges. Now, fix $k \in \{2, 3, \cdots, r+1\}$ and note that, for any $s \in [(k-1)t, kt]$, we have

$$\mathbb{P}\left(j \text{ connects to } v^* \text{ at time } s+1 \mid G_s, d_t(j) \geq g(t), G_t \in \mathcal{P}_t^1\right)$$

$$= (1 - f(s+1))\frac{d_s(j)d_s(v^*)}{2s^2} \tag{3.9}$$

$$\geq \frac{C_1}{3(r+1)^2 C_2} \frac{g(t)t}{t^2 h(a_t)},$$

since the degree of a vertex is increasing in s, $s \leq (r+1)t$ and $f \searrow 0$ implies that $1 - f(t) \geq 2/3$ for sufficiently large t. Thus, using condition (L), our choice of h leads to

$$\mathbb{P}\left(j \text{ is not connected to } v^* \text{ in } G_{kt} \mid G_{(k-1)t}, d_{(k-1)t}(j) \geq g(t), G_t \in \mathcal{P}_t^1\right)$$

$$\leq \left(1 - \frac{C_1}{3(r+1)^2 C_2}\frac{g(t)}{th(a_t)}\right)^t$$

$$\leq \exp\left\{-2\frac{f(a_t)\log(a_t)}{f(t)}\right\} \tag{3.10}$$

$$\leq \exp\left\{-2\frac{f(a_t)\log(a_t)}{f(t)\log t}\log t\right\}$$

$$\leq e^{-2\log t}.$$

Now, let $M_t(g(t))$ be the set of vertices in G_t whose degree is at least $g(t)$. Observe that $|M_t(g(t))| \leq t$ almost surely. Then, for each $k \in \{2, \ldots, r\}$ we have

$$\mathbb{P}\left(\exists j \in M_t(g(t)) \text{ that does not connect to } v^* \text{ in the interval } [(k-1)t, kt] \mid G_t \in \mathcal{P}_t^1\right) \leq \frac{1}{t}.$$

Consequently, we have

$$\mathbb{P}\left(\exists j \in M_t(g(t)) \text{ that does not send } r \text{ edges to } v^* \text{ in } G_{(r+1)t} \mid G_t \in \mathcal{P}_t^1\right) \leq \frac{r}{t}. \tag{3.11}$$

Finally, let H be the subgraph of G_t composed by all its vertices whose degree is at least $g(t)$. On the event where $V_t(f) \leq 15 f(t)t/8(1-\gamma)$ the sum of the degree of all vertices not in H satisfies

$$\sum_{v \notin H} d_t(v) < \frac{15}{8(1-\gamma)} g(t) f(t) t = \frac{15t}{8}. \tag{3.12}$$

Since the sum of all degrees equals in G_t equals $2t$, by the above inequality,

$$d_t(H) := \sum_{v \in H} d_t(v) \geq t/8,$$

whenever $V_t(f) \leq 15 f(t)t/8$, which is satisfied when $G_t \in \mathcal{P}_t^1$. This shows that whenever G_t is in $\mathcal{P}_t^1$, then $w.h.p$ $G_{(r+1)t}$ has a subgraph H with total degree in $G_{(r+1)t}$ larger than $t/8$, such that each vertex in H shares at least r edges with v^*. Moreover, when v^* is seen as a vertex of $G_{(r+1)t}$ it is a vertex of degree at least $C_1 t / C_2 h(a_t)$ as well, since the degree of a vertex is non-decreasing in time. Therefore, using Lemma 3.4, we obtain

$$\mathbb{P}\left(G_{(r+1)t} \notin \mathcal{P}_{(r+1)t}^2\right) \leq \mathbb{P}\left(G_{(r+1)t} \notin \mathcal{P}_{(r+1)t}^2, G_t \in \mathcal{P}_t^1\right) + \mathbb{P}\left(G_t \notin \mathcal{P}_t^1\right)$$

$$\leq \frac{r}{t}\mathbb{P}\left(G_t \in \mathcal{P}_t^1\right) + \mathbb{P}\left(G_{(r+1)t}^1 \notin \mathcal{P}_{(r+1)t}^1\right) + \mathbb{P}\left(G_t \notin \mathcal{P}_t^1\right) = o(1), \tag{3.13}$$

which completes the proof. $\qquad\square$

We move to our third required property. To avoid clutter in its definition we define the properties we are interested in below. Fix a (multi)graph $G = (V, E)$ and a number $\zeta > 0$ and consider the graph properties:

(1) $|V| \leq \frac{15 f((r+1)^2 t)(r+1)^2 t}{8(1-\gamma)}$;

(2) there exists $v^* \in V$ whose degree is bounded from bellow by

$$d_G(v^*) \geq \frac{C_1 t}{C_2 h(a_t)}$$

(3) there exists $H \subset V$ such that $d_G(H) \geq t/8$ and $\forall v \in H$ there are r edges between v and v^*;

(4) there exists $S \subset V$ such that $|S| \geq \zeta f(t)t$ and each $v \in S$ sends at least r edges to H.

For a fixed number $\zeta > 0$, we let $\mathcal{P}_{(r+1)^2 t}^3(\zeta)$ be the (multi)graph family defined below

$$\mathcal{P}_{(r+1)^2 t}^3(\zeta) := \{G = (V, E) : G \text{ satisfies all the properties (1)-(4) with parameter } \zeta\}.$$

Lemma 3.6. *Let f be an edge-step function satisfying conditions (S) and $\mathrm{RES}(-\gamma)$ for $\gamma \in [0,1)$. Then, there exists a number $\zeta_0 > 0$ depending on r and f only such that*

$$\mathbb{P}\left(G_{(r+1)^2 t}(f) \in \mathcal{P}_{(r+1)^2 t}^3(\zeta_0)\right) = 1 - o(1).$$

Proof: Observe that (1) is a consequence of Lemma 3.4 applied to $G_{(r+1)^2 t}(f)$. Whereas (2) and (3) are implied by Lemma 3.5 applied to $G_{(r+1)^2 t}(f)$.

Thus, we are left to prove that (4) holds a.a.s. The way we prove this is by showing that conditioned on $G_{(r+1)t}(f) \in \mathcal{P}_{(r+1)t}^2$, a positive proportion of vertices of degree 1 in $G_{(r+1)t}(f)$ connects to H at time $2(r+1)t$. Then we iterate this argument r times, similarly to what we have done to show the existence of H. For this, we need some definitions. Denote by $S_{(r+1)t}(1)$ the random set consisting of vertices of degree exactly 1 at time $(r+1)t$ and let $N_{(r+1)t}(1)$ be its cardinality. Also, let A_s and C_s be the random sets

$$A_s := \left\{v \in S_{(r+1)t}(1); \ v \text{ does not connect to } H \text{ up to time } s\right\};$$

$$C_s := \left\{v \in S_{(r+1)t}(1); \ v \text{ connects to } H \text{ between times } (r+1)t \text{ and } s\right\}. \tag{3.14}$$

We say a vertex in A_s is *available* to connect to H. At each step, we try to connect an available vertex to H using an edge-step. We are going to prove that *w.h.p.* $|C_{2(r+1)t}|$ is a positive proportion of $S_{(r+1)t}(1)$. So, let $\{Y_s\}_{s \geq t}$ be the following process

$$Y_s := \mathbb{1}\left\{\text{some vertex } v \text{ in } A_s \text{ connects to } H \text{ at time } s\right\}. \tag{3.15}$$

Note that $|C_{2(r+1)t}| = \sum_{s=t}^{2(r+1)t} Y_s$. Consider the following stopping time

$$\eta := \inf\left\{s \geq t : |A_s| \leq N_{(r+1)t}(1)/2\right\}, \tag{3.16}$$

To simplify our writing, we will denote by $\mathbb{P}_{G_{(r+1)t}}$ the law of $\{G_s(f)\}_{s \geq (r+1)t}$ conditioned on the event where graph $G_{(r+1)t}(f) = G_{(r+1)t}$ for some given admissible $G_{(r+1)t} \in \mathcal{P}^2_{(r+1)t}$. With this notation in mind, observe that, for $s \in [(r+1)t, 2(r+1)t]$ and large enough t

$$\mathbb{P}_{G_{(r+1)t}}\left(Y_{s+1} = 1 \mid G_s(f)\right) \geq (1 - f(s+1))\frac{|A_s|d_s(H)}{2s^2} \geq \frac{|A_s|}{65(r+1)^2 t}, \tag{3.17}$$

since $d_s(H) \geq t/8$ and $1 - f(s+1) \geq 64/65$ for large t. Then, by the definition of η, it also follows that

$$\mathbb{P}_{G_{(r+1)t}}\left(Y_{s+1} = 1, \eta > s \mid G_s(f)\right) \geq \frac{N_{(r+1)t}(1)}{130(r+1)^2 t}\mathbb{1}_{\{\eta > s\}} =: q_{(r+1)t}\mathbb{1}_{\{\eta > s\}}. \tag{3.18}$$

The above inequality implies that, on the event $\{\eta > 2(r+1)t\}$, the sum $\sum_{s=(r+1)t}^{2(r+1)t} Y_s$ dominates a r.v. following a binomial distribution of parameters $(r+1)t$ and $q_{(r+1)t}$. Then,

$$\mathbb{P}_{G_{(r+1)t}}\left(|C_{2(r+1)t}| \leq \frac{N_{(r+1)t}(1)}{260(r+1)^2}\right) \leq \mathbb{P}_{G_{(r+1)t}}\left(\text{bin}(t, q_{(r+1)t}) \leq \frac{N_{(r+1)t}(1)}{260(r+1)^2}, \eta > 2(r+1)t\right)$$
$$+ \mathbb{P}_{G_{(r+1)t}}\left(|C_{2(r+1)t}| \leq \frac{N_{(r+1)t}(1)}{260(r+1)^2}, \eta \leq 2(r+1)t\right)$$
$$= \mathbb{P}_{G_{(r+1)t}}\left(\text{bin}(t, q_{(r+1)t}) \leq \frac{N_{(r+1)t}(1)}{260(r+1)^2}, \eta > 2(r+1)t\right), \tag{3.19}$$

since on $\{\eta \leq 2(r+1)t\}$ we have that $|C_{2(r+1)t}| \geq N_{(r+1)t}(1)/2$. Thus, by Chernoff bounds

$$\mathbb{P}_{G_{(r+1)t}}\left(|C_{2(r+1)t}| \leq \frac{N_{(r+1)t}(1)}{260(r+1)^2}\right) \leq \mathbb{P}_{G_{(r+1)t}}\left(\text{bin}(t, q_{(r+1)t}) \leq \frac{N_{(r+1)t}(1)}{260(r+1)^2}\right)$$
$$\leq e^{-N_{(r+1)t}(1)/2^5 \cdot 5 \cdot 13 \cdot (r+1)^2}, \tag{3.20}$$

since $\text{bin}(t, q_{(r+1)t})$ has mean $N_{(r+1)t}(1)/130(r+1)^2$ (which is measurable with respect to $G_{(r+1)t}$). As stated in the introduction, Theorem 1 of Alves et al. (2021) states that w.h.p the proportion of vertices having degree d has order $d^{-(2-\gamma)}$. That result together with a Chernoff bound imply that, w.h.p. $N_{(r+1)t}(1) \geq c_f f(t)t$ for an f-dependent constant $c_f > 0$ (indeed, just take $A = o(\mathbb{E}V_{(r+1)t}(f))$ in the Theorem mentioned, and use a Chernoff bound to show that $V_{(r+1)t}(f)$ concentrates). The above upper bound combined with Lemma 3.5 gives us that the probability $\mathbb{P}\left(C_{2(r+1)t} \leq (1/65(r+1)^2) \cdot c_f f(t)t\right)$ is bounded from above by

$$\mathbb{P}\left(|C_{2(r+1)t}| \leq N_{(r+1)t}(1)/260(r+1)^2, N_{(r+1)t}(1) \geq c_f f(t)t, G_{(r+1)t}(f) \in \mathcal{P}^2_{(r+1)t}\right)$$
$$+ \mathbb{P}\left(G_{(r+1)t}(f) \notin \mathcal{P}^2_{(r+1)t}\right) + \mathbb{P}\left(N_{(r+1)t}(1) \leq ctf(t)\right)$$
$$= o(1),$$

Now we can iterate this argument with the set $C_{2(r+1)t}$ instead of $N_{(r+1)t}(1)$ to guarantee *w.h.p* that there exist at least $c'f(t)t$, for some constant c' depending on r and f, vertices sending at least

two edges to H in $G_{3(r+1)t}(f)$ and so on. Choosing ζ_0 appropriately, and recalling that, by the fact f is a regularly varying function, we have that

$$\frac{f(t)t}{(r+1)f((r+1)t)t} \to (r+1)^{\gamma-1},$$

as t goes to infinity, is enough to conclude the proof.

$\square$

We define the last property we want our graphs to satisfy a.a.s. Given a vertex v in a (multi)graph G, we let $\Gamma(v)$ be the number of neighbors of v in G. For the particular case of v in $G_t(f)$, we will write $\Gamma_t(v)$ to stress out the dependence on time of the number of neighbors of v. Moreover, due to the presence of edge-steps, $G_t(f)$ might be a multigraph which implies that $\Gamma_t(v) \leq d_t(v)$, with the strict inequality happening when v has multiple edges connecting it to one of its neighbors. For two fixed positive numbers ζ and κ, consider the following list of graph properties

(1)' $|V| \leq \frac{15f(2(r+1)^2t)2(r+1)^2t}{8(1-\gamma)}$;

(2)' there exists $v^* \in V$ whose degree and number of neighbors are bounded from bellow by

$$d_G(v^*) \geq \frac{C_1t}{C_2h\,(a_t)}; \quad \Gamma(v^*) \geq \kappa\frac{f(t)t}{h(a_t)};$$

(3)' there exists $H \subset V$ such that $d_G(H) \geq t/16$ and $\forall v \in H$ there are r edges between v and v^*;

(4)' there exists $S \subset V$ such that $|S| \geq \zeta f(t)t$ and each $v \in S$ sends at least r edges to H.

We then let $\mathcal{P}^4_{2(r+1)^2t}(\zeta, \kappa)$ be

$$\mathcal{P}^4_{2(r+1)^2t}(\zeta, \kappa) := \left\{G = (V, E) : G \text{ satisfies all properties } (1)' - (4)' \text{ with parameters } \zeta \text{ and } \kappa\right\}.$$

Next result also guarantees that the above property holds a.a.s..

Lemma 3.7. *Let f be an edge-step function satisfying conditions (S) and $\mathrm{RES}(-\gamma)$ for $\gamma \in [0, 1)$. Then, there exist positive parameters ζ_0 and κ_0 depending on f and r only such that*

$$\mathbb{P}\left(G_t(f) \in \mathcal{P}^4_{2(r+1)^2t}(\zeta_0, \kappa_0)\right) = 1 - o(1).$$

Proof: As usual we will prove that $G_{2(r+1)^2t}(f)$ satisfies $(1)' - (4)'$ for some parameters ζ_0 and κ_0. The property $(1)'$ is consequence of Lemma 3.4 applied directly to $G_{2(r+1)^2t}(f)$. Whereas $(3)'$ and $(4)'$ follow by conditioning on $G_{(r+1)^2t}(f) \in \mathcal{P}^3_{(r+1)^2t}(\zeta)$ and seeing it as a subgraph of $G_{2(r+1)^2t}(f)$.

Thus we are left to prove $(2)'$. The key step is to show $(2)'$ is seeing that, conditioned on $G_{(r+1)^2t}(f) \in \mathcal{P}^3_{(r+1)^2t}(\zeta_0)$, the graph $G_{2(r+1)^2t}(f)$ has w.h.p a vertex v^* such that $\Gamma_{2(r+1)^2t}(v^*) \geq \kappa f(t)t/h(a_t)$, for some positive number κ w.h.p.. For this purpose, observe that if v^* has degree at least $C_1C_2^{-1}t/h(a_t)$ at time $(r+1)^2t$, then for any $s \in ((r+1)^2t, 2(r+1)^2t)$, we have that

$$\mathbb{P}\left(\Gamma_{s+1}(v^*) - \Gamma_s(v^*) = 1 \,\Big|\, \mathcal{F}_s, d_t(v^*) \geq \frac{C_1t}{C_2h(a_t)}\right) \geq \frac{C_1}{C_2}f(s)\frac{t}{2sh(a_t)} \geq \frac{C_1}{C_2}\frac{f(2(r+1)^2t)}{4(r+1)^2h(a_t)}.$$

Thus, using the fact that f is a regularly varying function with index $-\gamma$, setting

$$\kappa_1 = \frac{C_1}{4(r+1)^{2+\gamma}C_2},$$

and conditioning on $G_{(r+1)^2t}(f) \in \mathcal{P}^3_{(r+1)^2t}(\zeta)$, the number of neighbors v^* has at time $2(r+1)^2t$ dominates a binomial random variable with parameters $(r+1)^2t$ and $\kappa_1f(t)/2h(a_t)$ for large t. Recalling the definition of functions h and g at (3.3) it follows that

$$\frac{\kappa_1(r+1)^2tf(t)}{h(a_t)} = \kappa_1(r+1)^2\frac{6}{C_1(1-\gamma)}C_2(r+1)^2tf(t)f(a_t)\log(a_t) \to \infty,$$

by condition (L). Thus, conditioned on $G_{(r+1)^2t}(f) \in \mathcal{P}^3_{(r+1)^2t}$, by Chernoff bounds on the binomial

$$\mathrm{bin}((r+1)^2t, \kappa_1 f(t)/2h(a_t)),$$

it follows that the (multi)graph $G_{2(r+1)^2t}(f)$ has a vertex v^* with at least $2\kappa_1 tf(t)/3h(a_t)$ neighbors, w.h.p. Using that $G_{(r+1)^2t}(f) \in \mathcal{P}^3_{(r+1)^2t}(\zeta)$ a.a.s., and defining $\kappa_0 = (2/3)\kappa_1$, we conclude the proof. $\qquad\square$

Finally we are able to prove Theorem 2' which implies Theorem 1.2 immediately.

Proof of Theorem 2': We first fix the parameter r of the bootstrap percolation process and define $\mathcal{G}_t$ as $\mathcal{P}^4_t(\zeta_0, \kappa_0)$, where ζ_0 and κ_0 are the positive numbers given by Lemma 3.7. Now we are left to prove that the bootstrap percolation process on a (multi)graph in the collection $\mathcal{G}_t$ and parameters a_t and r w.h.p has an outbreak of infection in at most 3 steps.

Summarizing all we have proven about $G_t(f)$ so far, a (multi)graph in $\mathcal{G}_t$ has the following structure:

(1) $G_t(f)$ has at most $15f(t)t/8(1-\gamma)$ vertices;
(2) There exists a star $v^* \in G_t(f)$ such that

$$\Gamma_t(v^*) \geq \frac{\kappa_0 f(t)t}{h(a_{t/2(r+1)^2})}; \qquad\qquad (3.21)$$

(3) There exists a subgraph H of $G_t(f)$ such that $\forall v \in H$ there are r edges between v and v^*; Moreover, recall from Lemma 3.5 that H is formed by those vertices of degree at least $g(t)$ at time $t/(2(r+1)^2)$;
(4) There exists a subset $S \subset G_t(f)$ containing at least $\zeta_0 f(t)t$ vertices whose elements are connected to H at least r times, that is, each vertex in S sends at least r edges to H.

Having all the properties above in mind, we start the bootstrap percolation process. In the round zero we construct $\mathcal{I}_0$. We add each vertex in $G_t(f)$ to $\mathcal{I}_0$ independently and with probability $a_t/|V(G_t(f))|$. Observe that by (1), it follows that

$$\frac{a_t}{|V(G_t(f))|} \geq \frac{c_1 a_t}{f(t)t},$$

where $c_1 = 8(1-\gamma)/15$.

Now, observe that the number of neighbors of v^* infected at round zero dominates a binomial random variable with parameters $\kappa_0 f(t)t/h(a_{t/2(r+1)^2})$ and $c_1 a_t/f(t)t$. Since f is a regularly varying function with index $-\gamma \in (-1,0]$, and a_t is increasing to infinity, Corollary B.2 implies that $a_t/h(a_{t/2(r+1)^2})$ goes to infinity as t goes to infinity. This and Chernoff bounds implies that

$$\mathbb{Q}_{G_t(f),a_t,r}\left(\mathcal{I}_0 \text{ contains at least } r \text{ neighbors of } v^*\right) = 1 - o(1). \qquad (3.22)$$

In words, at round zero v^* has at least r infected neighbors *w.h.p.* Now observe that given that v^* has at least r neighbors in $\mathcal{I}_0$, it follows that at round one v^* gets infected. Then, at the next round all the subgraph H gets infected since all of its vertices send at least r edges to v^*. Finally at round three, S gets infected since all of its vertices send at least r edges to H. This proves that $\mathcal{I}_3$ is a positive proportion of $G_t(f)$ and proves our theorem. $\qquad\square$

Appendix A. Martingale concentration inequalities

For the sake of completeness we state here two useful concentration inequalities for martingales which are used throughout the paper.

Theorem A.1 (Azuma-Höeffding Inequality - see Chung and Lu (2006))**.** *Let $(M_n, \mathcal{F})_{n \geq 1}$ be a martingale satisfying*

$$|M_{i+1} - M_i| \leq a_i$$

Then, for all $\lambda > 0$ we have

$$\mathbb{P}\left(|M_n - M_0| > \lambda\right) \leq \exp\left(-\frac{\lambda^2}{\sum_{i=1}^n a_i^2}\right).$$

Appendix B. Karamata theory

The three following results are used throughout the paper.

Theorem B.1 (Representation theorem - Theorem 1.4.1 of Bingham et al. (1987))**.** *Let f be a continuous regularly varying function with index of regular variation γ. Then, there exists a slowly varying function ℓ such that*

$$f(t) = t^\gamma \ell(t),\tag{B.1}$$

for all t in the domain of f.

Corollary B.2. *Let f be a continuous regularly varying function with index of regular variation $\gamma < 0$. Then,*

$$f(x) \to 0,\tag{B.2}$$

as x tends to infinity. Moreover, if ℓ is a slowly varying function, then for every $\varepsilon > 0$

$$x^{-\varepsilon}\ell(x) \to 0 \text{ and } x^\varepsilon \ell(x) \to \infty \tag{B.3}$$

Proof: Comes as a straightforward application of Theorem 1.3.1 of Bingham et al. (1987) and Corollary B.1. □

Theorem B.3 (Karamata's theorem - Proposition 1.5.8 of Bingham et al. (1987))**.** *Let ℓ be a continuous slowly varying function and locally bounded in $[x_0, \infty)$ for some $x_0 \geq 0$. Then*

(a) *for $\alpha > -1$*

$$\int_{x_0}^x t^\alpha \ell(t)dt \sim \frac{x^{1+\alpha}\ell(x)}{1+\alpha}.\tag{B.4}$$

(b) *for $\alpha < -1$*

$$\int_x^\infty t^\alpha \ell(t)dt \sim \frac{x^{1+\alpha}\ell(x)}{1+\alpha}.\tag{B.5}$$

References

Alves, C., Ribeiro, R., and Sanchis, R. Large communities in a scale-free network. *J. Stat. Phys.*, **166** (1), 137–149 (2017). MR3592854.

Alves, C., Ribeiro, R., and Sanchis, R. Diameter of P.A. random graphs with edge-step functions. *Random Structures Algorithms*, **57** (3), 612–636 (2020). MR4144077.

Alves, C., Ribeiro, R., and Sanchis, R. Preferential attachment random graphs with edge-step functions. *J. Theoret. Probab.*, **34** (1), 438–476 (2021). MR4207498.

Amin Abdullah, M. and Fountoulakis, N. A phase transition in the evolution of bootstrap percolation processes on preferential attachment graphs. *Random Structures Algorithms*, **52** (3), 379–418 (2018). MR3783205.

Balogh, J., Bollobás, B., Duminil-Copin, H., and Morris, R. The sharp threshold for bootstrap percolation in all dimensions. *Trans. Amer. Math. Soc.*, **364** (5), 2667–2701 (2012). MR2888224.

Balogh, J., Bollobás, B., and Morris, R. Bootstrap percolation in high dimensions. *Combin. Probab. Comput.*, **19** (5-6), 643–692 (2010). MR2726074.

Balogh, J., Peres, Y., and Pete, G. Bootstrap percolation on infinite trees and non-amenable groups. *Combin. Probab. Comput.*, **15** (5), 715–730 (2006). MR2248323.

Barabási, A.-L. and Albert, R. Emergence of scaling in random networks. *Science*, **286** (5439), 509–512 (1999). MR2091634.

Berger, N., Borgs, C., Chayes, J. T., and Saberi, A. On the spread of viruses on the internet. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 301–310. ACM, New York (2005). MR2298278.

Bingham, N. H., Goldie, C. M., and Teugels, J. L. *Regular variation*, volume 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge (1987). ISBN 0-521-30787-2. MR898871.

Bollobás, B. and Riordan, O. The diameter of a scale-free random graph. *Combinatorica*, **24** (1), 5–34 (2004). MR2057681.

Bollobás, B., Riordan, O., Spencer, J., and Tusnády, G. The degree sequence of a scale-free random graph process. *Random Structures Algorithms*, **18** (3), 279–290 (2001). MR1824277.

Can, V. H. Metastability for the contact process on the preferential attachment graph. *ArXiv Mathematics e-prints* (2015). arXiv: 1502.05633.

Chalupa, J., Leath, P. L., and Reich, G. R. Bootstrap percolation on a Bethe lattice. *J. Phys. C: Solid State Phys.*, **12** (1), L31–L35 (1979). DOI: 10.1088/0022-3719/12/1/008.

Chung, F. and Lu, L. *Complex graphs and networks*, volume 107 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI (2006). ISBN 978-0-8218-3657-6; 0-8218-3657-9. MR2248695.

Cooper, C. and Frieze, A. M. A general model of undirected web graphs. In *Algorithms—ESA 2001 (Århus)*, volume 2161 of *Lecture Notes in Comput. Sci.*, pp. 500–511. Springer, Berlin (2001). MR1914060.

Durrett, R. *Random graph dynamics*, volume 20 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge (2007). ISBN 978-0-521-86656-9; 0-521-86656-1. MR2271734.

Flocchini, P., Lodi, E., Luccio, F., Pagli, L., and Santoro, N. Dynamic monopolies in tori. *Discrete Appl. Math.*, **137** (2), 197–212 (2004). MR2048030.

Granovetter, M. Threshold Models of Collective Behavior. *American Journal of Sociology*, **83** (6), 1420–1443 (1978). DOI: 10.2307/2778111.

Jacob, E., Linker, A., and Mörters, P. Metastability of the contact process on fast evolving scale-free networks. *Ann. Appl. Probab.*, **29** (5), 2654–2699 (2019). MR4019872.

Schonmann, R. H. On the behavior of some cellular automata related to bootstrap percolation. *Ann. Probab.*, **20** (1), 174–193 (1992). MR1143417.

Strogatz, S. and Watts, D. J. Collective dynamics of 'small-world' networks. *Nature*, **393**, 440–442 (1998). DOI: 10.1038/30918.

van der Hofstad, R. *Random graphs and complex networks. Vol. 1*. Cambridge Series in Statistical and Probabilistic Mathematics, [43]. Cambridge University Press, Cambridge (2017). ISBN 978-1-107-17287-6. MR3617364.