A phase transition in the evolution of bootstrap percolation processes on preferential attachment graphs

Abdullah, Mohammed Amin; Fountoulakis, Nikolaos

DOI:
10.1002/rsa.20744

Document Version
Peer reviewed version

Citation for published version (Harvard):
Abdullah, MA & Fountoulakis, N 2017, 'A phase transition in the evolution of bootstrap percolation processes on preferential attachment graphs', Random Structures and Algorithms. https://doi.org/10.1002/rsa.20744

A phase transition in the evolution of bootstrap percolation processes on preferential attachment graphs

Mohammed Amin Abdullah* Nikolaos Fountoulakis†

February 6, 2017

Abstract

The theme of this paper is the analysis of bootstrap percolation processes on random graphs generated by preferential attachment. This is a class of infection processes where vertices have two states: they are either *infected* or *susceptible*. At each round every susceptible vertex which has at least $r \geq 2$ infected neighbours becomes infected and remains so forever. Assume that initially $a(t)$ vertices are randomly infected, where t is the total number of vertices of the graph. Suppose also that $r < m$, where $2m$ is the average degree. We determine a critical function $a_c(t)$ such that when $a(t) \gg a_c(t)$, complete infection occurs with high probability as $t \to \infty$, but when $a(t) \ll a_c(t)$, then with high probability the process evolves only for a bounded number of rounds and the final set of infected vertices is asymptotically equal to $a(t)$.

1 Introduction

The dissemination of contagion within a network is a fundamental problem that arises in a wide spectrum of social and economic sciences. Among the mechanisms which underlie this phenomenon is a class of dissemination processes where local decisions (or *microbehaviours*) aggregate into a large outbreak or *pandemic*. Quite frequently, these phenomena begin on a rather small scale and may end up contaminating a large part of the network. What are the particular characteristics of a network that enable or inhibit such an outbreak?

A general class of models that incorporates this kind of behaviour is what is called the *general threshold model* [31]. Here it is assumed that each vertex has one of two states: it is either *infected* or *susceptible*. Furthermore, each vertex of the underlying graph is equipped with a threshold function which depends on the states of its neighbours. This function expresses the probability that this vertex remains in a particular state given the states of its neighbours. A central problem in viral marketing is given a network, a parameter $k \geq 1$ and such a set of functions, find a set of vertices S of size k which maximizes the expected number of infected vertices at the end of the process. In [29], Kempe, Kleinberg and Tardos proved that finding such an optimal set is NP-hard. Moreover, they showed that it is NP-hard to approximate the size of the maximum expected outreach even within a polynomial factor. See also [30] for similar results.

*Mathematical and Algorithmic Sciences Lab, Huawei Technologies Ltd. mohammed.abdullah@huawei.com This work was mostly conducted whilst M. A. Abdullah was at the University of Birmingham.

†School of Mathematics, University of Birmingham, Edgbaston, B15 2TT, U.K. n.fountoulakis@bham.ac.uk
Research supported by the EPSRC Grant No. EP/K019749/1.

In this paper, we study an instance of this class of models known as *bootstrap percolation processes*. This is a threshold model that was introduced in the context of mathematical physics by Chalupa, Leath and Reich [17] in 1979 for magnetic disordered systems.

A *bootstrap percolation process* with *activation threshold* an integer $r \geq 2$ on a (multi)graph $G = G(V, E)$ is a deterministic process. Initially, there is a subset $\mathcal{I}_0 = \mathcal{I}(0) \subseteq V$ of infected vertices, whereas every other vertex is susceptible. This set can be selected either deterministically or randomly. The process evolves in rounds, where in each round, if a susceptible vertex has at least r edges connected to infected neighbours, then it also becomes infected and remains so forever. This is repeated until no more vertices become infected. We denote the final infected set by $\mathcal{I}_f$. We denote the set of susceptible (infected) vertices at round τ in the process by $\mathcal{S}(\tau)$ (respectively, $\mathcal{I}(\tau)$). Thus, $\mathcal{S}(\tau)$, $\mathcal{I}(\tau)$ form a partition of the vertex set V, and $\mathcal{I}_f = \mathcal{I}(\infty)$. Of course, the above definition makes also perfect sense when $r = 1$ – in this case $\mathcal{I}_f$ coincides with the set of vertices of the union of those components of G which contain vertices in $\mathcal{I}_0$.

Such processes (as well as several variations of them) have been used as models to describe several complex phenomena in diverse areas, from jamming transitions [36] and magnetic systems [33] to neuronal activity [5, 22]. Bootstrap percolation processes also have connections with the dynamics of the Ising model at zero temperature [23], [32]. These processes have also been studied on a variety of graphs, such as trees [10, 24], grids [16, 26, 8], lattices on the hyperbolic plane [34], hypercubes [7], as well as on several distributions of random graphs [4, 11, 28]. Ebrahimi et al. [21] showed that given a graph G, a subset of vertices V' and a positive parameter s, the question whether the final set of the bootstrap percolation process starting at V' has size at least s is P-complete. A short survey regarding applications of bootstrap percolation processes can be found in [3]. The theme of this paper is the study of bootstrap percolation processes on a preferential attachment random graph on t vertices, which we denote by $\mathrm{PA}_t(m, \delta)$.

2 Preferential attachment graphs

The preferential attachment models have their origins in the work of Yule [37], where a growing model is proposed in the context of the evolution of species. A similar model was proposed by Simon [35] in the statistics of language. The principle of these models was used by Barabási and Albert [12] to describe a random graph model where vertices arrive one by one and each of them throws a number of half-edges to the existing graph. Each half-edge is connected to a vertex with probability that is proportional to the degree of that vertex. This model was defined rigorously by Bollobás, Riordan, Spencer and Tusnády [14] (see also [13]). We will describe the most general form of the model which is essentially due to Dorogovtsev et al. [19] and Drinea et al. [20]. Our description and notation below follow that from the book of van der Hofstad [25].

The random graph $\mathrm{PA}_t(m, \delta)$ is parameterised by two constants: $m \in \mathbb{N}$, and $\delta \in \mathbb{R}$, $\delta > -m$. It gives rise to a random graph sequence (i.e., a sequence in which each member is a random graph), denoted by $(\mathrm{PA}_t(m, \delta))_{t=1}^{\infty}$. The tth term of the sequence, $\mathrm{PA}_t(m, \delta)$ is a graph with t vertices and mt edges. Further, $\mathrm{PA}_t(m, \delta)$ is a subgraph of $\mathrm{PA}_{t+1}(m, \delta)$. We define $\mathrm{PA}_t(1, \delta)$ first, then use it to define the general model $\mathrm{PA}_t(m, \delta)$ (the Barabási-Albert model corresponds to the case $\delta = 0$).

The random graph $\mathrm{PA}_1(1, \delta)$ consists of a single vertex with one self-loop. We denote the

vertices of $\mathrm{PA}_t(1,\delta)$ by $\{v_1^{(1)}, v_2^{(1)}, \ldots, v_t^{(1)}\}$. We denote the degree of vertex $v_i^{(1)}$ in $\mathrm{PA}_t(1,\delta)$ by $D_i(t)$. Then, conditionally on $\mathrm{PA}_t(1,\delta)$, the growth rule to obtain $\mathrm{PA}_{t+1}(1,\delta)$ is as follows: We add a single vertex $v_{t+1}^{(1)}$ having a single edge. The other end of the edge connects to $v_{t+1}^{(1)}$ itself with probability $\frac{1+\delta}{t(2+\delta)+(1+\delta)}$, and connects to a vertex $v_i^{(1)} \in \mathrm{PA}_t(1,\delta)$ with probability $\frac{D_i(t)+\delta}{t(2+\delta)+(1+\delta)}$ – we write $v_{t+1}^{(1)} \to v_i^{(1)}$. Any self-loop at a vertex adds 2 to the degree of that vertex. For any $t \in \mathbb{N}$, let $[t] = \{1, \ldots, t\}$. Thus,

$$\mathbb{P}\left(v_{t+1}^{(1)} \to v_i^{(1)} \mid \mathrm{PA}_t(1,\delta)\right) = \begin{cases} \frac{1+\delta}{t(2+\delta)+(1+\delta)} & \text{for } i = t+1, \\ \frac{D_i(t)+\delta}{t(2+\delta)+(1+\delta)} & \text{for } i \in [t] \end{cases}$$

The model $\mathrm{PA}_t(m,\delta)$, $m > 1$, with vertices $\{1, \ldots, t\}$ is derived from $\mathrm{PA}_{mt}(1,\delta/m)$ with vertices $\{v_1^{(1)}, v_2^{(1)}, \ldots, v_{mt}^{(1)}\}$ as follows: For each $i = 1, 2, \ldots, t$, we contract the vertices $\{v_{(i-1)+1}^{(1)}, v_{(i-1)+2}^{(1)}, \ldots, v_{(i-1)+m}^{(1)}\}$ into one super-vertex, and identify this super-vertex as i in $\mathrm{PA}_t(m,\delta)$. When a contraction takes place, all loops and multiple edges are retained. Edges shared between a set of contracted vertices become loops in the contracted super-vertex. Thus, $\mathrm{PA}_t(m,\delta)$ is a graph on $[t]$.

The above process gives a graph whose degree distribution follows a power law with exponent $3 + \delta/m$. This was suggested by the analyses in [19] and [20]. It was proved rigorously for integral δ by Buckley and Osthus [15]. For a full proof for real δ see [25]. In particular, when $-m < \delta < 0$, the exponent is between 2 and 3. Experimental evidence has shown that this is the case for several networks that emerge in applications (cf. [2]). Furthermore, when $m \geq 2$, then $\mathrm{PA}_t(m,\delta)$ is **whp** connected, but when $m = 1$ this is not the case, giving rise to a logarithmic number of components (see [25]).

We describe an alternative, though equivalent, direct construction of $(\mathrm{PA}_t(m,\delta))_{t=1}^{\infty}$. Let $\mathrm{PA}_1(m,\delta)$ be a single vertex with label 1, having m loops. Given $\mathrm{PA}_{t-1}(m,\delta)$, $t \geq 2$, the construction of $\mathrm{PA}_t(m,\delta)$ is as follows: To add vertex t to the graph, we split time step t into m sub-steps, adding one edge sequentially in each sub-step. For $j = 1, 2, \ldots, m$, denote the graph after the jth sub-step of time t by $\mathrm{PA}_{t,j}(m,\delta)$. Hence $\mathrm{PA}_t(m,\delta) \equiv \mathrm{PA}_{t,m}(m,\delta)$. For notational convenience, let $\mathrm{PA}_{t,0}(m,\delta) = \mathrm{PA}_{t-1}(m,\delta)$.

Denote the jth edge added by e_j. One end of e_j will be attached to vertex t and the other end will be attached randomly to another vertex (which may be t). Let $g(t,j)$ be the random variable representing this vertex. For $j = 1, 2, \ldots, m$, let $D_i(t,j)$ be the degree of vertex i in $\mathrm{PA}_{t,j}(m,\delta)$. That is, for $j = 1, 2, \ldots, m$, $D_i(t,j)$ is the degree of vertex i after both ends of e_j have been attached. Furthermore, for notational convenience, let $D_t(t,0) = 0$ and for $i \in [t-1]$, let $D_i(t,0) = D_i(t-1)$.

Now, for $j = 1, 2, \ldots, m$, conditionally on $\mathrm{PA}_{t,j-1}(m,\delta)$, $\mathrm{PA}_{t,j}(m,\delta)$ is generated according to the following probability rules:

$$\mathbb{P}\left(g(t,j) = i \mid \mathrm{PA}_{t,j-1}(m,\delta)\right) = \begin{cases} \frac{D_t(t,j-1)+1+j\delta/m}{(2m+\delta)(t-1)+2j-1+j\delta/m} & \text{for } i = t, \\ \frac{D_i(t,j-1)+\delta}{(2m+\delta)(t-1)+2j-1+j\delta/m} & \text{for } i \in [t-1] \end{cases}.$$

It is not difficult to see that these two constructions give rise to the same probability distribution over realisations of $(\mathrm{PA}_t(m,\delta))_{t=1}^{\infty}$. It will be sometimes convenient to refer to one form over the other.

2.1 Results

Here as well as in the rest of the paper the term *with high probability* (**whp**) means with probability $1 - o(1)$ in the space of $\mathrm{PA}_t(m, \delta)$, as $t \to \infty$. We will be using the same term for events over the product space between $\mathrm{PA}_t(m, \delta)$ and the choice of $\mathcal{I}_0$ on $[t]$.

Let X_t be a random variable on the above product space. If $a \in \mathbb{R}$, we write that $X_t \xrightarrow{p} a$ (X_t *converges to* a *in probability*) if for any $\varepsilon > 0$ we have $\mathbb{P}(|X_t - a| > \varepsilon) \to 0$ as $t \to \infty$.

The selection of $\mathcal{I}_0$ is random and each vertex is infected initially with probability $p = p(t) = a(t)/t$, independently of any other vertex. Hence, if $a(t) \to \infty$ as $t \to \infty$, the size of $\mathcal{I}_0/a(t)$ converges in probability to 1.

Recently, Ebrahimi et al. [21] investigated a threshold phenomenon that occurs in the evolution of the process on a variant of the preferential attachment model, that is very similar (though not identical) to $\mathrm{PA}_t(m, \delta)$. In our context, their results can be stated as follows. Let $\gamma = \frac{m}{2m+\delta}$. If $a(t) \gg t^{1-\gamma} \log t$, then **whp** $\mathcal{I}_f = [t]$, that is, we have complete infection. They also identified a subcritical range for $a(t)$. Assume first that $r\gamma \geq 1$. If $a(t) \ll t^{1-\gamma}$, then **whp** $\mathcal{I}_f = \mathcal{I}_0$, that is no evolution occurs. Now, if $r\gamma < 1$, then the same holds but provided that $a(t) \ll t^{1-1/r}$. Since $\gamma < 1/r$, that is, $1 - \gamma > 1 - 1/r$, it follows that this function is asymptotically smaller than the $t^{1-\gamma}$. Similar results were obtained by the two authors in [1] for $\mathrm{PA}_t(m, \delta)$.

In this paper, we complete the landscape and show that a critical phenomenon occurs "around" the function $t^{1-\gamma} =: a_c(t) = a_c$. Our results show that when $a(t) \gg a_c(t)$, there is complete infection **whp**, but if $a(t) \ll a_c(t)$ then either there is no evolution of the process or it halts in a bounded number of rounds. (In fact, for $r = 2$ we show a slightly weaker result that requires $a(t) \leq a_c(t)/\log t$.) Theorems 1 and 2(i) recover the results of Ebrahimi et al. [21], but Theorem 2(ii) covers also the case where $r\gamma \leq 1$, closing the gap between the two threshold functions that were identified by Ebrahimi et al. We should also point out that Ebrahimi et al. [21] achieve strong probability bounds, that is, the related events occur with probability that tends to 1 polynomially fast. In this case, the process accumulates only a small number of infections beyond those incurred initially, so that $\mathcal{I}_f$ is almost equal to $\mathcal{I}_0$. Inside the critical window, that is, if $a(t) = \Theta(a_c(t))$, then with probability asymptotically bounded away from zero there is complete infection, and with probability bounded away from zero we have similar behaviour as for the $a(t) \ll a_c(t)$ case.

Let $\omega = \omega(t) \to \infty$ as $t \to \infty$ arbitrarily slowly. With $\gamma = \frac{1}{2+\delta/m}$ and $a_c(t) = t^{1-\gamma}$, the above can be formalized as follows.

Theorem 1 (Supercritical case)**.** *If $r < m$ and $a(t) = \omega a_c(t)$ then all vertices in $PA_t(m, \delta)$ get infected* **whp***.*

Theorem 2 (Subcritical case)**.** *If $r \leq m$ then the following hold:*

(i) *If $a(t) = a_c(t)/\omega$ and $r\gamma > 1$, then* **whp***, $\mathcal{I}_f = \mathcal{I}_0$.*

(ii) *If $a(t) = a_c(t)/\omega$ and $r \geq 3$ then $|\mathcal{I}_f|/|\mathcal{I}_0| \xrightarrow{p} 1$ and* **whp** *the process stops in at most $\lfloor \frac{1}{\gamma} \rfloor$ rounds.*

(iii) *If $a(t) = a_c(t)/\log t$ and $r = 2$, then $|\mathcal{I}_f|/|\mathcal{I}_0| \xrightarrow{p} 1$ and* **whp** *the process stops in at most $\lfloor \frac{1}{\gamma} \rfloor + 1$ rounds.*

It should be noted that when $\delta < 0$, $r\gamma > 1$ is always satisfied, since we insist that $r \geq 2$.

Theorem 3 (Critical case)**.** *Let $r \geq 3$ and $a(t) = \lambda a_c(t)$ where λ is a constant. Then there exist constants $0 < p_1, p_2$ depending on λ such that the following hold:*

(i) *if $r \leq m$, then the following holds with probability at least p_1: vertices are infected for at most $\lfloor \frac{1}{\gamma} \rfloor$ rounds, and $|\mathcal{I}_f|/|\mathcal{I}_0| < 1 + \varepsilon$, for any $\varepsilon > 0$ and any t large enough.*

(ii) *if $r < m$, then with probability at least p_2, there is a complete infection.*

The function $a_c(t)$ was also identified by the second author and Amini [6] in the case of inhomogeneous random graphs of rank 1. However, results of Amini [4] imply that if the kernel of such a random graph gives rise to a power law degree distribution with exponent larger than 3 (corresponds to $\delta > 0$), then **whp**, a *sublinear* initial infection only results in a *sublinear* outbreak. As our results and the results in [21] show this is not the case in the preferential attachment model. In other words, a sublinear initial infection leads to an outbreak where every vertex becomes infected, provided that the amount of the initial infection is not too small. Theorems 1 and 2 identify this critical amount.

Lack of outbreak is also the case in random regular graphs of constant degree [11] as well as in binomial random graphs with constant expected degree [28]. In the latter case, the authors show that if $a(t) = o(t)$, then $|\mathcal{I}_f|/|\mathcal{I}_0| \xrightarrow{p} 1$. This behaviour is radically different from that in the preferential attachment model, where Theorem 1 implies that a sublinear initial infection may lead to pandemics.

2.1.1 The cases $r = m$ and $r > m$

It can be shown that there are a logarithmic number of self-loops in $\mathrm{PA}_t(m, \delta)$. For $r = m$, these loops make analysis of the outcome difficult. This is a rather specific artifact of the model and, is not shared with slight variations of the model, e.g., one in which self-loops are not allowed.

For $r > m$ the following "folklore" argument shows that if the number of initially infected vertices is sublinear, then the final number will be sublinear as well: Let G be the subgraph induced by all the vertices in $\mathcal{I}_f$. The number of edges in G is at least $(|\mathcal{I}_f| - |\mathcal{I}_0|)r$ but at the same time, the total number of edges in G can be at most $m|\mathcal{I}_f|$. Therefore $(|\mathcal{I}_f| - |\mathcal{I}_0|)r \leq m|\mathcal{I}_f|$ implying $|\mathcal{I}_f| \leq \frac{r}{r-m}|\mathcal{I}_0|$.

2.2 Further notation and terminology

Throughout this paper we let $\gamma = \gamma(m, \delta) = \frac{1}{2+\delta/m}$, hence $1 - \gamma = \frac{1+\delta/m}{2+\delta/m}$. Observe the condition $\delta > -m$ (which must be imposed), implies $0 < \gamma < 1$. Furthermore, $\delta < 0$ if and only if $\frac{1}{2} < \gamma < 1$.

For integers i, j with $i \leq j$, we shall sometimes write $[i, j]$ to denote the set $\{i, i+1, \ldots, j\}$. We also use $S_i(t)$ to denote the sum of degrees for vertices in the interval $[1, i]$, i.e., $S_i(t) = \sum_{j=1}^{i} D_j(t)$.

We will sometimes say a vertex j *throws* an edge e to vertex i if, in the construction of $\mathrm{PA}_j(m, \delta)$, vertex j connected edge e to vertex i. We will also say i *receives* the edge e.

Furthermore, for two non-negative functions $f(t), g(t)$ on $\mathbb{N}$ we write $f(t) \lesssim g(t)$ to denote that $f(t) = O(g(t))$. If, in addition, $g(t) = O(f(t))$, then we write $f(t) \asymp g(t)$. In this paper, the underlying asymptotic variable will always be t, the number of vertices in $\mathrm{PA}_t(m, \delta)$.

We use the notation $f(c) \overset{(m,\delta)}{\lesssim} g(c)$ to mean that there is a constant $C(m, \delta)$ such that $f(c) \leq C(m, \delta)g(c)$, and $C(m, \delta)$ depends only on m, δ.

We will begin with some general results in the next section on the concentration of the degrees, which will be used mainly in the Proof of Theorem 1.

3 Vertex degrees: expectation and concentration

As we mentioned above, the degrees in $\mathrm{PA}_t(m, \delta)$ roughly follow a power-law degree distribution with exponent $3 + \delta/m$, that is, the empirical probability mass function on the degrees scales like $\frac{1}{x^{3+\delta/m}}$. In fact, many networks that emerge in applications have a degree distribution that follows a power law with exponent between 2 and 3 (cf. [2] for example), which corresponds to $\delta/m \in (-1, 0)$. The Barabasi-Albert model gives power-law with exponent 3 ($\delta = 0$). Observe that the variance on the degrees is finite if and only if the exponent is greater than 3 (corresponding to $\delta > 0$).

Consider two vertices i and j; their total weight is $D_i(t) + D_j(t) + 2\delta$, meaning probability of an edge being thrown to them is proportional to this value. Now a vertex with degree $D_i(t) + D_j(t)$ would have weight $D_i(t) + D_j(t) + \delta$. Thus, we cannot treat two separate vertices i and j as a single one of the combined degree, except when $\delta = 0$. In the special case that $\delta = 0$, the weight of a vertex is proportional to its degree, and the weight of a set of vertices is proportional to the sum of their degrees. When $\delta = 0$, we can treat a set of vertices as a bucket of *half-edges*, or *stubs*, conceptually distributing the stubs across the vertices however we like. However, when $\delta \neq 0$, the weighting is non-linear. Conceptually grouping stubs together means that one has to sum their weights not their degrees.

In summary, the probability of a vertex receiving the next edge thrown is proportional to its weight. The same holds for a set of vertices; the probability a set of vertices receiving an edge is proportional to the total weight of the set. When, and only when, $\delta = 0$, then the weight of a vertex is its degree, and the weight of a set is the total degree of the vertices in the set.

A number of results on the degree sequence are collected in van der Hofstad [25] which shows, amongst other things, that $\mathbb{E}[D_i(t)] = (1 + o(1))a\left(\frac{t}{i}\right)^{\gamma}$ where a is a constant that depends only on m and δ.

3.1 Sum of degrees

Recall that $S_i(t)$ denotes the sum of degrees for vertices in the interval $[1, i]$. We state the following without proof. It is a simple consequence of results in, e.g., [25].

Proposition 4. *There exist constants $C_\ell, C_u > 0$ that depend only on m and δ such that for each vertex $i \in [t]$,*

$$C_\ell t^\gamma i^{1-\gamma} \leq \mathbb{E}[S_i(t)] \leq C_u t^\gamma i^{1-\gamma}.$$

We next derive a concentration results for the sum of degrees. Lemma 5 is an elaboration of Lemma 2 in [18]. Its proof can be found in the appendix.

Lemma 5. *Suppose $\delta \geq 0$. There exists constants $K_0, h_0 > 0$ that depend only on m and δ, such that the following holds for all $i \in [t]$, $K > K_0$ and $h < h_0$,*

$$\mathbb{P}\left(S_i(t) < \frac{1}{K}\mathbb{E}[S_i(t)]\right) \leq e^{-hi}$$

Lemma 6. *Let $\varepsilon > 0$ be a constant. If $\delta < 0$, then there exists a positive constant $c = c(m, \delta, \varepsilon)$ that depends only on m, δ and ε, such that with probability at least $1 - e^{-ci}$,*

$$S_i(t) \geq (1 - \varepsilon)\mathbb{E}[S_i(t)] \tag{1}$$

for all $i \in [t]$.

Proof. We will use a Doob martingale in conjunction with the Azuma-Hoeffding inequality. Define $M_n^{(m,\delta)}(i,t) = \mathbb{E}[S_i(t) \mid \mathrm{PA}_n(m,\delta)]$. Observe, for $n = 1, 2, \ldots, i$, $M_n^{(m,\delta)}(i,t) = \mathbb{E}[S_i(t)]$. Now we want to bound $|M_{n+1}^{(m,\delta)}(i,t) - M_n^{(m,\delta)}(i,t)|$ for $n \geq i$. Observe that $S_i(n)$ is measurable with respect to $\mathrm{PA}_n(m,\delta)$, and $\mathbb{E}[S_i(t) \mid S_i(n), \mathrm{PA}_n(m,\delta)] = \mathbb{E}[S_i(t) \mid S_i(n)]$, i.e., that the expectation of $S_i(t)$ is independent of $\mathrm{PA}_n(m,\delta)$ given $S_i(n)$. Hence, we will instead write $M_n^{(m,\delta)}(i,t) = \mathbb{E}[S_i(t) \mid S_i(n)]$. We have, for $t > n$,

$$\begin{aligned}
\mathbb{E}[S_i(t) + \delta i \mid S_i(n)] &= \mathbb{E}[\mathbb{E}[S_i(t) + \delta i \mid S_i(t-1), S_i(n)] \mid S_i(n)] \\
&= \mathbb{E}[\mathbb{E}[S_i(t) + \delta i \mid S_i(t-1)] \mid S_i(n)].
\end{aligned}$$

We will analyse the $m = 1$ case first. Considering the inner conditional expectation,

$$\begin{aligned}
\mathbb{E}[S_i(t) + \delta i \mid S_i(t-1)] &= S_i(t-1) + \delta i + \frac{S_i(t-1) + \delta i}{(2+\delta)(t-1) + 1 + \delta} \\
&= \frac{(2+\delta)t}{(2+\delta)(t-1) + 1 + \delta}\left(S_i(t-1) + \delta i\right).
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathbb{E}[S_i(t) + \delta i \mid S_i(n)] &= \frac{t}{t - 1 + \frac{1+\delta}{2+\delta}}\mathbb{E}[S_i(t-1) + \delta i \mid S_i(n)] \\
&= (S_i(n) + \delta i)\prod_{k=n}^{t-1}\frac{k+1}{k + \frac{1+\delta}{2+\delta}} \\
&= (S_i(n) + \delta i)\frac{\Gamma(t+1)}{\Gamma(t + \frac{1+\delta}{2+\delta})}\frac{\Gamma(n + \frac{1+\delta}{2+\delta})}{\Gamma(n+1)}.
\end{aligned}$$

Consequently,

$$\begin{aligned}
&\left|M_{n+1}^{(1,\delta)}(i,t) - M_n^{(1,\delta)}(i,t)\right| = |\mathbb{E}[S_i(t) \mid S_i(n+1)] - \mathbb{E}[S_i(t) \mid S_i(n)]| \\
= \ & \frac{\Gamma(t+1)}{\Gamma(t + \frac{1+\delta}{2+\delta})}\left|(S_i(n+1) + \delta i)\frac{\Gamma(n + 1 + \frac{1+\delta}{2+\delta})}{\Gamma(n+2)} - (S_i(n) + \delta i)\frac{\Gamma(n + \frac{1+\delta}{2+\delta})}{\Gamma(n+1)}\right| \\
= \ & \frac{\Gamma(t+1)}{\Gamma(t + \frac{1+\delta}{2+\delta})}\frac{\Gamma(n + \frac{1+\delta}{2+\delta})}{\Gamma(n+1)}\left|(S_i(n+1) + \delta i)\frac{n + \frac{1+\delta}{2+\delta}}{n+1} - (S_i(n) + \delta i)\right|.
\end{aligned}$$

We have $\frac{n}{n+1} < \frac{n+\frac{1+\delta}{2+\delta}}{n+1} < 1$ and $S_i(n) \le S_i(n+1) \le S_i(n)+1$, so

$$\begin{aligned}\left|(S_i(n+1)+\delta i)\,\frac{n+\frac{1+\delta}{2+\delta}}{n+1} - (S_i(n)+\delta i)\right| &\le (S_i(n)+\delta i)\left|\frac{n+\frac{1+\delta}{2+\delta}}{n+1}-1\right| + \frac{n+\frac{1+\delta}{2+\delta}}{n+1} \\ &< \frac{S_i(n)+\delta i}{(2+\delta)(n+1)} + 1.\end{aligned}$$

Since $S_i(n) \le 2i+n-i = n+i$ and $i \le n$, the right-hand side is at most 2:

$$\frac{S_i(n)+\delta i}{(2+\delta)(n+1)} \le \frac{n+i(1+\delta)}{(2+\delta)(n+1)} \le \frac{n+n(1+\delta)}{(2+\delta)(n+1)} < 1.$$

Thus,

$$\left|M_{n+1}^{(1,\delta)}(i,t) - M_n^{(1,\delta)}(i,t)\right| < 2\frac{\Gamma(t+1)}{\Gamma(t+\frac{1+\delta}{2+\delta})}\frac{\Gamma(n+\frac{1+\delta}{2+\delta})}{\Gamma(n+1)}.$$

Recall that when $m \ge 1$ we define $\mathrm{PA}_t(m,\delta)$ in terms of $\mathrm{PA}_{mt}(1,\delta/m)$, and $S_a(b)$ in the former corresponds to $S_{ma}(mb)$ in the latter. Therefore, with $\gamma = \gamma(m,\delta) = \frac{1}{2+\delta/m}$,

$$\begin{aligned}\left|M_{n+1}^{(m,\delta)}(i,t) - M_n^{(m,\delta)}(i,t)\right| &= \left|M_{m(n+1)}^{(1,\delta/m)}(mi,mt) - M_{mn}^{(1,\delta/m)}(mi,mt)\right| \\ &= \left|\sum_{k=1}^{m}\left(M_{m(n+1)-k+1}^{(1,\delta/m)}(mi,mt) - M_{m(n+1)-k}^{(1,\delta/m)}(mi,mt)\right)\right| \\ &\le \sum_{k=1}^{m}\left|M_{m(n+1)-k+1}^{(1,\delta/m)}(mi,mt) - M_{m(n+1)-k}^{(1,\delta/m)}(mi,mt)\right| \\ &\le 2\frac{\Gamma(mt+1)}{\Gamma(mt+1-\gamma)}\sum_{k=1}^{m}\frac{\Gamma(m(n+1)-k+1-\gamma)}{\Gamma(m(n+1)-k+1)}.\end{aligned}$$

We have

$$\begin{aligned}\frac{\Gamma(mn+k-\gamma)}{\Gamma(mn+k)} &= \frac{mn+k-1-\gamma}{mn+k-1}\frac{mn+k-2-\gamma}{mn+k-2}\cdots\frac{mn+1-\gamma}{mn+1}\frac{\Gamma(mn+1-\gamma)}{\Gamma(mn+1)} \\ &\le \frac{\Gamma(mn+1-\gamma)}{\Gamma(mn+1)},\end{aligned}$$

so

$$\sum_{k=1}^{m}\frac{\Gamma(m(n+1)-k+1-\gamma)}{\Gamma(m(n+1)-k+1)} = \sum_{k=1}^{m}\frac{\Gamma(mn+k-\gamma)}{\Gamma(mn+k)} \le m\frac{\Gamma(mn+1-\gamma)}{\Gamma(mn+1)}.$$

Therefore,

$$\left|M_{n+1}^{(m,\delta)}(i,t) - M_n^{(m,\delta)}(i,t)\right| \le 2m\frac{\Gamma(mt+1)}{\Gamma(mt+1-\gamma)}\frac{\Gamma(mn+1-\gamma)}{\Gamma(mn+1)}.$$

Re-writing the above, we get

$$\begin{aligned}\left|M_{n+1}^{(m,\delta)}(i,t) - M_n^{(m,\delta)}(i,t)\right| &\le 2m\frac{\Gamma(mt+1-\gamma+\gamma)}{\Gamma(mt+1-\gamma)}\frac{\Gamma(mn+1-\gamma)}{\Gamma(mn+1-\gamma+\gamma)} \\ &\le C_{m,\delta}\left(\frac{t}{n}\right)^{\gamma},\end{aligned}$$

where $C_{m,\delta}$ is a universal constant that depends only on m and δ.

Now, applying the Hoeffding-Azuma inequality,

$$\mathbb{P}\left(S_i(t) - \mathbb{E}[S_i(t)] < -d\right) \leq \exp\left(\frac{-d^2}{C_{m,\delta}^2 \sum_{j=i+1}^{t} \left(\frac{t}{j}\right)^{2\gamma}}\right).$$

Since $\delta < 0$, we have $\sum_{j=i+1}^{t} \left(\frac{t}{j}\right)^{2\gamma} \leq K_1 t^{2\gamma} i^{1-2\gamma}$ for some constant K_1.

Hence letting $d = \varepsilon \mathbb{E}[S_i(t)] \geq \varepsilon C_\ell t^\gamma i^{1-\gamma}$ for some constant $\varepsilon > 0$,

$$\mathbb{P}\left(S_i(t) - \mathbb{E}[S_i(t)] < -d\right) \leq \exp\left(\frac{-\varepsilon^2 C_\ell^2 t^{2\gamma} i^{2(1-\gamma)}}{C_{m,\delta}^2 K_1 t^{2\gamma} i^{1-2\gamma}}\right) \leq e^{-ci}$$

for some constant $c = c(m, \delta, \varepsilon) > 0$ that depends only on m, δ and ε.

□

4 Supercritical Case: Proof of Theorem 1

The proof of this theorem relies on the fact that with high probability all of the early vertices of $\mathrm{PA}_t(m, \delta)$ become infected during the first round. Subsequently, the connectivity of the random graph is enough to spread the infection to the remaining vertices. The infection of the early vertices requires sufficiently high lower bounds on their degrees. We show these using the concentration results of the previous section together with a coupling with a Pólya urn process.

4.1 Pólya Urns

Consider the following Pólya urn process with red and black balls. Let $i \geq 2$ be an integer and let the weighting functions for the red and black balls be $W_R(k) = k + \delta$ and $W_B(k) = k + (i-1)\delta$, respectively. Under such a weighting scheme, if there are a red balls and b black balls, then the next time a ball is selected from the urn, the probability it is red is $\frac{W_R(a)}{W_R(a)+W_B(b)} = \frac{a+\delta}{a+\delta+b+(i-1)\delta} = \frac{a+\delta}{a+b+i\delta}$. Whenever a ball is picked, it is placed back in the urn with another ball of the same colour. We can ask, if there are initially a red and b black balls, and we make n selections, what is the probability that d of those selections are red?

To start with, one may calculate the probability of a particular sequence of n outcomes. If an n-sequence has d reds followed by $n - d$ blacks, then it has probability $p_{n,d,a,b}$ where

$$\begin{aligned} p_{n,d,a,b} &= \frac{a+\delta}{a+b+i\delta} \frac{a+1+\delta}{a+b+1+i\delta} \cdots \frac{a+d-1+\delta}{a+b+d-1+i\delta} \\ &\quad \times \frac{b+(i-1)\delta}{a+b+d+i\delta} \frac{b+1+(i-1)\delta}{a+b+d+1+i\delta} \cdots \frac{b+n-d-1+(i-1)\delta}{a+b+n-1+i\delta} \\ &= \frac{\Gamma(a+d+\delta)}{\Gamma(a+\delta)} \frac{\Gamma(b+n-d+(i-1)\delta)}{\Gamma(b+(i-1)\delta)} \frac{\Gamma(a+b+i\delta)}{\Gamma(a+b+n+i\delta)}. \end{aligned}$$

It is not hard to see that this is the same probability for any n-sequence with d reds and $n - d$ blacks, regardless of ordering (this is the *exchangeability* property of the Pólya urn

process). As such, letting $X_R(n,a,b)$ be the number of reds picked when n selections are made, we have

$$\mathbb{P}(X_R(n,a,b)=d)=\binom{n}{d}p_{n,d,a,b}=\binom{n}{d}\frac{\Gamma(a+d+\delta)}{\Gamma(a+\delta)}\frac{\Gamma(b+n-d+(i-1)\delta)}{\Gamma(b+(i-1)\delta)}\frac{\Gamma(a+b+i\delta)}{\Gamma(a+b+n+i\delta)}. \tag{2}$$

Now let $i\geq 2$ and consider the vertices $[1,i]$ in $(\mathrm{PA}_t(m,\delta))_{t=i}^{\infty}$. With every vertex $t=i+1,i+2,\ldots$, there are m edges created, some of which may connect to vertices in $[1,i]$. We ask, what is the probability that an edge connects to i, given that it connects to some vertex in $[1,i]$? A coupling with the above Pólya urn process is immediate: after the creation of $\mathrm{PA}_i(m,\delta)$, we create an urn with $D_i(i)$ red balls and $2mi-D_i(i)$ black balls. Every time a vertex $t>i$ connects an edge into the interval $[1,i]$, a selection is made in the urn process. A red ball is chosen if and only if the edge connects to i.

To demonstrate that the probabilities correspond, suppose in $\mathrm{PA}_{t,j-1}(m,\delta)$ we have $D_i(t,j-1)=a$. Denoting $S_{i-1}(t,j-1)=\sum_{k=1}^{i-1}D_k(t,j-1)$, suppose also $S_{i-1}(t,j-1)=b$. Then it is easily checked that $\mathbb{P}\left(g(t,j)=i\mid g(t,j)\in[1,i]\right)=\frac{a+\delta}{a+b+i\delta}$. Hence, if in $\mathrm{PA}_t(m,\delta)$ there are n edges with one end in $[1,i]$ and the other end in $[i+1,t]$, then the probability that d of those edges are attached to vertex i is given by (2). As such, we have the following proposition.

Proposition 7. *Let $m\geq 1, i\geq 2$ be integers and let $\delta>-m$ be a real number. Suppose a Pólya urn process starts with $m\leq a\leq 2m$ red and $b=2mi-a$ black balls, and has weighting functions $W_R(k)=k+\delta$ and $W_B(k)=k+(i-1)\delta$ for the red and black balls, respectively. Let the random variable $X_R(n,a)=X_R(n,a,2mi-a)$ count the total number of red choices after n selections have been made. Furthermore, consider a random graph $PA_t(m,\delta)$. If $t\geq i$, then for $0\leq d\leq n$,*

$$\mathbb{P}\left(D_i(t)=d+a\mid S_i(t)-2mi=n, D_i(i)=a\right)=\mathbb{P}(X_R(n,a)=d).$$

The following lemma will be used to bound individual vertex degrees.

Lemma 8. *Let $X_R(n,a)$ be the random variable defined in Proposition 7 and let $I=i(2m+\delta)-1$. Then for $1\leq d\leq n$,*

$$\mathbb{P}(X_R(n,a)=d)\overset{(m,\delta)}{\lesssim}\frac{1}{d}\left(\frac{Id}{I+n-d}\right)^{a+\delta}e^{-\frac{dI}{I+n}}, \tag{3}$$

and

$$\mathbb{P}(X_R(n,a)=0)\lesssim\left(\frac{I}{I+n}\right)^{a+\delta}. \tag{4}$$

Proof. As per Equation (2),

$$\mathbb{P}(X_R(n,a)=d)=\binom{n}{d}\frac{\Gamma(a+d+\delta)}{\Gamma(a+\delta)}\frac{\Gamma(b+n-d+(i-1)\delta)}{\Gamma(b+(i-1)\delta)}\frac{\Gamma(a+b+i\delta)}{\Gamma(a+b+n+i\delta)}.$$

That is, since $a+b=2mi$ and $a+b+i\delta = i(2m+\delta)$, we have

$$\mathbb{P}(X_R(n,a)=d) = \binom{n}{d}\frac{\Gamma(a+\delta+d)}{\Gamma(a+\delta)}\frac{\Gamma(i(2m+\delta)+n-(a+\delta+d))}{\Gamma(i(2m+\delta)-(a+\delta))}\frac{\Gamma(i(2m+\delta))}{\Gamma(i(2m+\delta)+n)}$$

We re-write the above as

$$\mathbb{P}(X_R(n,a)=d) = \binom{n}{d}\frac{\Gamma(a+\delta+d)}{\Gamma(a+\delta)}\frac{\Gamma(I+1+n-(a+\delta+d))}{\Gamma(I+1-(a+\delta))}\frac{\Gamma(I+1)}{\Gamma(I+1+n)}. \tag{5}$$

Suppose first that $d>0$. We can write the above as

$$\mathbb{P}(X_R(n,a)=d) = \frac{\Gamma(a+\delta+d)}{d!\Gamma(a+\delta)}\frac{\Gamma(I+1)}{\Gamma(I+1-(a+\delta))}\frac{(n)_d\Gamma(I+1+n-(a+\delta+d))}{\Gamma(I+1+n)} \tag{6}$$

($(n)_d$ denotes the falling factorial $(n)_d = n(n-1)\dots(n-d+1)$).

To bound the above, we shall use the following fact: For real $x>0$,

$$\Gamma(x+1) = c_x\sqrt{2\pi}e^{-x}x^{x+\frac{1}{2}}$$

where $c_x \in [1, e^{\frac{1}{12x}}]$. Suppose now $x\to\infty$ and a is a constant. Then, the above implies that when $x+a>0$,

$$\frac{\Gamma(x+a)}{\Gamma(x)} = x^a(1+O(1/x)). \tag{7}$$

Now we bound (6): using (7), we deduce that

$$\frac{\Gamma(a+\delta+d)}{d!\Gamma(a+\delta)} \overset{(m,\delta)}{\lesssim} d^{a+\delta-1}.$$

Also by (7), $\frac{\Gamma(I+1))}{\Gamma(I+1-(a+\delta))} \overset{(m,\delta)}{\lesssim} I^{a+\delta}$, and so

$$\frac{\Gamma(a+\delta+d)}{d!\Gamma(a+\delta)}\frac{\Gamma(I+1)}{\Gamma(I+1-(a+\delta))} \overset{(m,\delta)}{\lesssim} \frac{1}{d}(Id)^{a+\delta}. \tag{8}$$

Now,

$$\frac{(n)_d\Gamma(I+1+n-(a+\delta+d))}{\Gamma(I+1+n)} = \frac{n}{I+n}\frac{n-1}{I+n-1}\cdots\frac{n-(d-1)}{I+n-(d-1)}\frac{\Gamma(I+1+n-(a+\delta+d))}{\Gamma(I+n-(d-1))}.$$

We have

$$\frac{n}{I+n}\frac{n-1}{I+n-1}\cdots\frac{n-(d-1)}{I+n-(d-1)} \le \left(\frac{n}{I+n}\right)^d \le e^{-\frac{dI}{I+n}}.$$

Furthermore, by (7)

$$\frac{\Gamma(I+1+n-(a+\delta+d))}{\Gamma(I+n-(d-1))} \overset{(m,\delta)}{\lesssim} \frac{1}{(I+n-d)^{a+\delta}}.$$

Consequently, we have the following bound:

$$\mathbb{P}(X_R(n,a)=d) \overset{(m,\delta)}{\lesssim} \frac{1}{d}\left(\frac{Id}{I+n-d}\right)^{a+\delta} e^{-\frac{dI}{I+n}}.$$

Now suppose $d=0$, then going back to (5) we have

$$\mathbb{P}(X_R(n,a)=0) = \frac{\Gamma(I+1)}{\Gamma(I+1-(a+\delta))}\frac{\Gamma(I+1+n-(a+\delta))}{\Gamma(I+1+n)} \overset{(m,\delta)}{\lesssim} \left(\frac{I}{I+n}\right)^{a+\delta}.$$

□

4.2 Proof of Theorem 1

For convenience, we rewrite as $a(t)=\omega^{10}a_c(t)$ where $\omega=\omega(t)\to\infty$ arbitrarily slowly (we can assume $\omega\le\log t$, since if not, we can just substitute $\log t$ for it and get full infection **whp**; a larger ω can only increase the probability of this happening).

Let $\kappa=\lceil\omega^{1+\delta/m}\rceil$ and choose $[\kappa]$ as a core. We wish to show all vertices in the core are infected for this $a(t)$.

For $\delta\ge 0$, we apply Lemma 5, taking h to be a sufficiently small constant such that for some constant K_ℓ, we have $S_\kappa(t)\ge K_\ell t^\gamma\kappa^{1-\gamma}$ **whp**. For $\delta<0$, we apply Lemma 6 to get the same result. We set $n=n_\kappa(t)=K_\ell t^\gamma\kappa^{1-\gamma}-2m\kappa$.

Now we wish to show that **whp**, $D_i(t)\ge\left(\frac{t}{\omega^{1+\delta/m}}\right)^\gamma\frac{1}{z}$ over all $i\in[\kappa]$, for some appropriately chosen $z=z(t)\to\infty$. Applying Lemma 8 with $I=\kappa(2m+\delta)-1$,

$$\begin{aligned}
\mathbb{P}\left(X_R(n,a)\le\frac{n}{\kappa z}\right) &= \sum_{d=0}^{n/(\kappa z)}\mathbb{P}(X_R(n,a)=d)\\
&\lesssim \left(\frac{I}{I+n}\right)^{a+\delta}+\sum_{d=1}^{n/(\kappa z)}\left(\frac{I}{I+n-d}\right)^{a+\delta}d^{a+\delta-1}e^{-\frac{dI}{I+n}}\\
&\le \left(\frac{I}{I+n}\right)^{a+\delta}+\frac{I^{a+\delta}}{(I+n-n/(\kappa z))^{a+\delta}}\sum_{d=0}^{n/(\kappa z)}d^{a+\delta-1}.
\end{aligned}$$

Since $\kappa\to\infty$ and $z\to\infty$ as $t\to\infty$, we have $n/(\kappa z)=o(n)$, so $\frac{1}{(I+n-n/(\kappa z))^{a+\delta}}\lesssim\frac{1}{(I+n)^{a+\delta}}$. Furthermore,

$$\sum_{d=0}^{n/(\kappa z)}d^{a+\delta-1}\lesssim\int_0^{n/(\kappa z)}x^{a+\delta-1}\,\mathrm{d}x\le\frac{1}{a+\delta}\left(\frac{n}{\kappa z}\right)^{a+\delta}.$$

Choosing $z=\omega^2$, we have

$$\left(\frac{I}{I+n}\right)^{a+\delta}=\left(\frac{\kappa(2m+\delta)-1}{\kappa(2m+\delta)-1+K_\ell t^\gamma\kappa^{1-\gamma}-2m\kappa}\right)^{a+\delta}\lesssim\left(\frac{\omega^{1+\delta/m}}{t}\right)^{\gamma(a+\delta)}=o\left(\frac{1}{z^{a+\delta}}\right).$$

Hence,

$$\mathbb{P}\left(X_R(n,a)\le\frac{n}{\kappa z}\right)\lesssim\left(\frac{I}{I+n}\right)^{a+\delta}+\left(\frac{I}{I+n}\right)^{a+\delta}\left(\frac{n}{\kappa z}\right)^{a+\delta}\lesssim\left(\frac{I}{I+n}\right)^{a+\delta}+\frac{1}{z^{a+\delta}}\le\frac{1}{z^{m+\delta}}.$$

Thus,

$$\mathbb{P}\left(X_R(n,a) \le \frac{n}{\kappa z}\right) \lesssim \frac{1}{z^{m+\delta}}\,.$$

Taking a union bound over all vertices in $[\kappa]$, we have a probability asymptotically bounded by $\left(\frac{\omega}{z^m}\right)^{1+\delta/m} = o(1)$.

So given $D_i(t) \ge \left(\frac{t}{\omega^{1+\delta/m}}\right)^{\gamma} \frac{1}{\omega^2}$ for each $i \in [\kappa]$, we calculate the expectation of the number of infected neighbours a vertex in the core has. This would be at least

$$\frac{a(t)}{2mt}\left(\frac{t}{\omega^{1+\delta/m}}\right)^{\gamma}\frac{1}{\omega^2} = \frac{\omega^8}{2m}\left(\frac{1}{\omega^{1+\delta/m}}\right)^{\gamma} \ge \omega^7$$

for large enough t.

To calculate the probability that at least r neighbours are infected for a fixed vertex i in the core, we bound the corresponding binomial random variable. Suppose $N = N(t) \to \infty$, $p = p(t) \to 0$ but $Np \to \infty$. Then for large enough t, $\mathbb{P}(\mathrm{Bin}(N,p) < r) \le e^{-Np/2}$.

Therefore,

$$\mathbb{P}\left(\mathrm{Bin}\left(D_i(t), \frac{a(t)}{t}\right) < r \mid D_i(t) \ge \left(\frac{t}{\omega^{1+\delta/m}}\right)^{\gamma}\frac{1}{\omega^2}\right) \le e^{-\omega^7/2}$$

and so the probability that any of the core vertices fail to be infected is at most $\omega^{1+\delta/m}e^{-\omega^7/2} \le e^{-\omega^6}$, for large enough t.

Thus, at this stage, we have proved that the core vertices, i.e., those in $[\kappa]$, all get infected **whp**. If no vertex outside the core has more than a single self-loop, then each vertex will have at least $m-1$ forward (i.e., out-going) edges. Hence, if $r \le m-1$, the entire graph will be infected if the core is. We show that no vertex outside the core has more than one self-loop.

The probability that vertex i outside the core has at least two self loops is at most $2\binom{m}{2}i^{-2}$. Summing over all $i \in [\kappa+1, t]$, this is $O\left(\int_{\omega^{1+\delta/m}}^{t} i^{-2}\,\mathrm{d}i\right) = O\left(1/\omega^{1+\delta/m}\right) = o(1)$. Hence, **whp**, no vertex outside the core has more than one self-loop. So if $r \le m-1$, the entire graph gets infected **whp**.

5 Subcritical Case: Proof of Theorem 2

The general proof strategy of Theorem 2 is based on the following argument. Suppose that a vertex i is not infected at round $\tau = 0$, but it is infected at round $\tau = 1$. Then there must be r edges connected to i that also connect to vertices infected in round $\tau = 0$. Assuming that these edges connect to different neighbours, we have a depth-1 tree. Similarly, if i gets infected in round $\tau = d$, then there must be some underlying *witness structure* which caused this. In particular, it may be that there is an r-ary tree of depth d wherein in round $\tau = 0$ all the leaves are infected and no internal vertices are. We call this a *witness tree*. More generally, such a structure may contain cycles. We shall deal with witness trees first before addressing more general witness structures. We use a first moment argument to show that witness structures of a certain depth do not exist **whp**. Before doing so, we need to develop the estimates that will allow us to bound the number of occurrences of a certain graph as a subgraph of PA_t.

We revert to the model $\mathrm{PA}_t(1,\delta)$, which, for notational convenience, we shall write as PA_t. We have $\gamma = \frac{1}{2+\delta}$.

We begin by defining a sequence of polynomials $(Q_n(x))_{n\geq 1}$ where $Q_n(x) = x(x+1)\cdots(x+n-1)$.

Lemma 9. *Let X_t be a random variable measurable with respect to PA_t. Then for $1 \leq i \leq t-1$*

$$\mathbb{E}[X_{t-1}Q_n(D_i(t)+\delta)] = \mathbb{E}[X_{t-1}Q_n(D_i(t-1)+\delta)]\frac{t+(n-1)\gamma}{t-\gamma}.$$

Proof.

$$\begin{aligned}
&\mathbb{E}[X_{t-1}Q_n(D_i(t)+\delta) \mid \mathrm{PA}_{t-1}] \\
&\quad = X_{t-1}\left[\left(1-\frac{D_i(t-1)+\delta}{(2+\delta)t-1}\right)Q_n(D_i(t-1)+\delta) + \frac{D_i(t-1)+\delta}{(2+\delta)t-1}Q_n(D_i(t-1)+\delta+1)\right] \\
&\quad = X_{t-1}\left[Q_n(D_i(t-1)+\delta)\left[1-\frac{D_i(t-1)+\delta}{(2+\delta)t-1}+\frac{D_i(t-1)+\delta+n}{(2+\delta)t-1}\right]\right].
\end{aligned}$$

Now we take expectations on both sides and the lemma follows. □

Lemma 10. *Suppose $i < j_1, j_2, \ldots, j_k$ are vertices in $PA_t(m,\delta)$. Then*

$$\mathbb{P}\left(j_1 \to i \cap j_2 \to i \cap \ldots \cap j_k \to i\right) \leq M^k \frac{1}{i^\gamma j_1^{1-\gamma}}\frac{1}{i^\gamma j_2^{1-\gamma}}\cdots\frac{1}{i^\gamma j_k^{1-\gamma}}$$

where $M = M(m,\delta)$ is a constant that depends only on m and δ.

Proof. We begin with the case $m = 1$. Assuming that $j_1 < \cdots < j_k$ we have

$$\begin{aligned}
\mathbb{E}[\mathbf{1}_{\{j_1\to i\}}\mathbf{1}_{\{j_2\to i\}}\ldots\mathbf{1}_{\{j_k\to i\}} \mid \mathrm{PA}_{j_k-1}] &= \mathbf{1}_{\{j_1\to i\}}\mathbf{1}_{\{j_2\to i\}}\ldots\mathbf{1}_{\{j_{k-1}\to i\}}\frac{D_i(j_k-1)+\delta}{(2+\delta)j_k-1} \\
&= \frac{\gamma}{j_k-\gamma}\left(\prod_{s=1}^{k-1}\mathbf{1}_{\{j_s\to i\}}\right)(D_i(j_k-1)+\delta).
\end{aligned}$$

Therefore, applying Lemma 9 repeatedly,

$$\begin{aligned}
\mathbb{E}[\mathbf{1}_{\{j_1\to i\}}\mathbf{1}_{\{j_2\to i\}}\ldots\mathbf{1}_{\{j_k\to i\}}] &= \frac{\gamma}{j_k-\gamma}\mathbb{E}\left[\left(\prod_{s=1}^{k-1}\mathbf{1}_{\{j_s\to i\}}\right)Q_1(D_i(j_k-1)+\delta)\right] \\
&= \frac{\gamma}{j_k-\gamma}\mathbb{E}\left[\left(\prod_{s=1}^{k-1}\mathbf{1}_{\{j_s\to i\}}\right)Q_1(D_i(j_k-2)+\delta)\right]\frac{j_k-1}{j_k-1-\gamma} \\
&= \vdots \\
&= \frac{\gamma}{j_k-\gamma}\mathbb{E}\left[\left(\prod_{s=1}^{k-1}\mathbf{1}_{\{j_s\to i\}}\right)Q_1(D_i(j_{k-1})+\delta)\right]\prod_{s=j_{k-1}+1}^{j_k-1}\frac{s}{s-\gamma}.
\end{aligned}$$

Now,

$$\mathbb{E}\left[\left(\prod_{s=1}^{k-1}\mathbf{1}_{\{j_s\to i\}}\right)(D_i(j_{k-1})+\delta)\mid \mathrm{PA}_{j_{k-1}-1}\right] = \left(\prod_{s=1}^{k-2}\mathbf{1}_{\{j_s\to i\}}\right)\frac{D_i(j_{k-1}-1)+\delta}{(2+\delta)j_{k-1}-1}(D_i(j_{k-1}-1)+\delta+1)$$
$$= \frac{\gamma}{j_{k-1}-\gamma}\left(\prod_{s=1}^{k-2}\mathbf{1}_{\{j_s\to i\}}\right)Q_2(D_i(j_{k-1}-1)+\delta).$$

Thus, by repeated application of Lemma 9,

$$\mathbb{E}[\mathbf{1}_{\{j_1\to i\}}\mathbf{1}_{\{j_2\to i\}}\dots\mathbf{1}_{\{j_k\to i\}}]$$
$$= \frac{\gamma}{j_k-\gamma}\frac{\gamma}{j_{k-1}-\gamma}\mathbb{E}\left[\left(\prod_{s=1}^{k-2}\mathbf{1}_{\{j_s\to i\}}\right)Q_2(D_i(j_{k-1}-1)+\delta)\right]\prod_{s=j_{k-1}+1}^{j_k-1}\frac{s}{s-\gamma}$$
$$= \frac{\gamma}{j_k-\gamma}\frac{\gamma}{j_{k-1}-\gamma}\mathbb{E}\left[\left(\prod_{s=1}^{k-2}\mathbf{1}_{\{j_s\to i\}}\right)Q_2(D_i(j_{k-2})+\delta)\right]\prod_{s=j_{k-2}+1}^{j_{k-1}-1}\frac{s+\gamma}{s-\gamma}\prod_{s=j_{k-1}+1}^{j_k-1}\frac{s}{s-\gamma}.$$

This pattern continues until we get

$$\mathbb{E}[\mathbf{1}_{\{j_1\to i\}}\mathbf{1}_{\{j_2\to i\}}\dots\mathbf{1}_{\{j_k\to i\}}]$$
$$= \frac{\gamma}{j_k-\gamma}\frac{\gamma}{j_{k-1}-\gamma}\dots\frac{\gamma}{j_1-\gamma}\prod_{s=j_1+1}^{j_2-1}\frac{s+(k-2)\gamma}{s-\gamma}\dots\prod_{s=j_{k-3}+1}^{j_{k-2}-1}\frac{s+2\gamma}{s-\gamma}\prod_{s=j_{k-2}+1}^{j_{k-1}-1}\frac{s+\gamma}{s-\gamma}\prod_{s=j_{k-1}+1}^{j_k-1}\frac{s}{s-\gamma}$$
$$\times\,\mathbb{E}[Q_k(D_i(j_1-1)+\delta)].$$

Applying Lemma 9 repeatedly,

$$\mathbb{E}[Q_k(D_i(j_1-1)+\delta)] = \mathbb{E}[Q_k(D_i(i)+\delta)]\prod_{s=i+1}^{j_1-1}\frac{s+(k-1)\gamma}{s-\gamma},$$

and

$$\mathbb{E}[Q_k(D_i(i)+\delta)] = \left(1-\frac{1+\delta}{(2+\delta)i-1}\right)Q_k(1+\delta)+\frac{1+\delta}{(2+\delta)i-1}Q_k(1+\delta+1)$$
$$= Q_k(1+\delta)\frac{i+(k-1)\gamma}{i-\gamma}.$$

Thus,

$$\mathbb{E}[\mathbf{1}_{\{j_1\to i\}}\mathbf{1}_{\{j_2\to i\}}\dots\mathbf{1}_{\{j_k\to i\}}]$$
$$= \prod_{s=i}^{j_1-1}\frac{s+(k-1)\gamma}{s-\gamma}\prod_{s=j_1+1}^{j_2-1}\frac{s+(k-2)\gamma}{s-\gamma}\dots\prod_{s=j_{k-3}+1}^{j_{k-2}-1}\frac{s+2\gamma}{s-\gamma}\prod_{s=j_{k-2}+1}^{j_{k-1}-1}\frac{s+\gamma}{s-\gamma}\prod_{s=j_{k-1}+1}^{j_k-1}\frac{s}{s-\gamma}$$
$$\times\,\frac{\gamma}{j_k-\gamma}\frac{\gamma}{j_{k-1}-\gamma}\dots\frac{\gamma}{j_1-\gamma}Q_k(1+\delta). \tag{9}$$

Observe that

$$\prod_{s=j_{k-1}+1}^{j_k-1}\frac{s}{s-\gamma} = \frac{\Gamma(j_k)}{\Gamma(j_{k-1}+1)}\frac{\Gamma(j_{k-1}+1-\gamma)}{\Gamma(j_k-\gamma)} = \frac{\Gamma(j_k)}{\Gamma(j_k-\gamma)}\frac{\Gamma(j_{k-1}+1-\gamma)}{\Gamma(j_{k-1}+1)},$$

and similarly with the other product terms. Thus, the product will give us

$$\frac{\Gamma(j_k)}{\Gamma(j_k-\gamma)}\frac{\Gamma(j_{k-1}+1-\gamma)}{\Gamma(j_{k-1}+1)}\frac{\Gamma(j_{k-1}+\gamma)}{\Gamma(j_{k-1}-\gamma)}\frac{\Gamma(j_{k-2}+1-\gamma)}{\Gamma(j_{k-2}+1+\gamma)}\frac{\Gamma(j_{k-2}+2\gamma)}{\Gamma(j_{k-2}-\gamma)}\frac{\Gamma(j_{k-3}+1-\gamma)}{\Gamma(j_{k-3}+1+2\gamma)}\cdots$$
$$\cdots\frac{\Gamma(j_2+(k-2)\gamma)}{\Gamma(j_2-\gamma)}\frac{\Gamma(j_1+1-\gamma)}{\Gamma(j_1+1+(k-2)\gamma)}\frac{\Gamma(j_1+(k-1)\gamma)}{\Gamma(j_1-\gamma)}\frac{\Gamma(i-\gamma)}{\Gamma(i+(k-1)\gamma)}.$$

Also observe that

$$\frac{\Gamma(j_{k-1}+1-\gamma)}{\Gamma(j_{k-1}+1)}\frac{\Gamma(j_{k-1}+\gamma)}{\Gamma(j_{k-1}-\gamma)}=\frac{j_{k-1}-\gamma}{j_{k-1}}\frac{\Gamma(j_{k-1}+\gamma)}{\Gamma(j_{k-1})}<\frac{\Gamma(j_{k-1}+\gamma)}{\Gamma(j_{k-1})}.$$

A similar argument bounds the other fraction pairs, thereby giving an upper bound on the product of

$$\frac{\Gamma(j_k)}{\Gamma(j_k-\gamma)}\frac{\Gamma(j_{k-1}+\gamma)}{\Gamma(j_{k-1})}\frac{\Gamma(j_{k-2}+2\gamma)}{\Gamma(j_{k-2}+\gamma)}\cdots\frac{\Gamma(j_1+(k-1)\gamma)}{\Gamma(j_1+(k-2)\gamma)}\frac{\Gamma(i-\gamma)}{\Gamma(i+(k-1)\gamma)}. \tag{10}$$

For some constant c which depends only on γ, we have $\Gamma(x+\gamma)/\Gamma(x)\le cx^\gamma$. Therefore, as the last term is less than 1, (10) is bounded by $c^k\frac{(j_k,j_{k-1}\ldots j_1)^\gamma}{i^{k\gamma}}$.

Going back to (9), we have

$$\mathbb{E}[\mathbf{1}_{\{j_1\to i\}}\mathbf{1}_{\{j_2\to i\}}\ldots\mathbf{1}_{\{j_k\to i\}}]\le Q_k(1+\delta)c^k\frac{1}{i^\gamma j_1^{1-\gamma}}\frac{1}{i^\gamma j_2^{1-\gamma}}\cdots\frac{1}{i^\gamma j_k^{1-\gamma}} \tag{11}$$

where c is a constant that depends only on γ.

We wish to extend the above result to $\mathrm{PA}_t(m,\delta)$ with $m>1$, that is, we wish to bound the probability of the js connecting to the same vertex i. The js do not have to be distinct. Recall that vertex i in $\mathrm{PA}_t(m,\delta)$ is created from grouping m consecutive vertices in $\mathrm{PA}_{mt}(1,\delta/m)$ and contracting them into one vertex (possibly creating loops and/or parallel edges in doing so).

Let $I=\{m(i-1)+1,m(i-1)+2,\ldots,mi\}$ be the set of vertices in $\mathrm{PA}_{mt}(1,\delta/m)$ that group to become i in $\mathrm{PA}_t(m,\delta)$. Similarly, we have sets $J_1,J_2,\ldots,J_k$ for the js.

Recall that $g(j,\ell)=i$ means that the ℓth edge of vertex j is connected to vertex $i<j$. For a vertex i, an integer N_i and a collection of pairs of indices $\{(j_n^{(i)},\ell_n^{(i)})\}_{n=1,\ldots,N_i}$ let

$$E_i=\bigcap_{n=1}^{N_i}\{g(j_n^{(i)},\ell_n^{(i)})=i\}.$$

In other words, E_i is the event that the given set of edges are thrown to vertex i. Since $\mathrm{PA}_t(m,\delta)$ is created from $\mathrm{PA}_{mt}(1,\delta/m)$, Inequality (11) implies that for some $c_1=c_1(m,\delta)>0$ we have

$$\mathbb{P}\left(E_i\right)\le c_1^{N_i}\frac{1}{i^{N_i\gamma}\prod_{n=1}^{N_i}\left(j_n^{(i)}\right)^{1-\gamma}}. \tag{12}$$

The following lemma from [25] states that E_i and $E_{i'}$ are negatively correlated for $i\neq i'$. That is, edges thrown to different vertices are negatively correlated.

Lemma 11 (Lemma 11.13 [25])**.** *For distinct vertices $i_1,i_2,\ldots,i_\ell$ in $PA_t(m,\delta)$,*

$$\mathbb{P}\left(\bigcap_{s=1}^{\ell}E_{i_s}\right)\le\prod_{s=1}^{\ell}\mathbb{P}(E_{i_s}).$$

The event $j_1 \to i$ in $\mathrm{PA}_t(m,\delta)$ occurs in $\mathrm{PA}_{mt}(1,\delta/m)$ when a vertex in J_1 throws an edge to a vertex in I. This can happen in m^2 different ways. More generally, the event $j_1 \to i \cap j_2 \to i \cap \ldots \cap j_k \to i$ in $\mathrm{PA}_t(m,\delta)$ can happen in at most m^{2k} different ways. Any such way corresponds to the realisation of a collection of events $E_{i_1},\ldots,E_{i_\ell}$ in $\mathrm{PA}_{mt}(1,\delta/m)$, for some $\ell \le m \wedge k$. So the above lemma together with (12) imply (11) (with a different multiplicative constant) for $m > 1$.

□

The following is a corollary of Lemmas 10 and 11. The js need not be distinct, and some of the js may also be is.

Corollary 12. *Suppose $i_1, j_1, i_2, j_2, \ldots, i_k, j_k$ are vertices in $PA_t(m,\delta)$ where $i_s < j_s$ for $s = 1, 2, \ldots, k$. Then*

$$\mathbb{P}(j_1 \to i_1 \cap j_2 \to i_2, \ldots, j_k \to i_k) \le Q_k(1+\delta)M^k \frac{1}{i_1^\gamma j_1^{1-\gamma}} \frac{1}{i_2^\gamma j_2^{1-\gamma}} \cdots \frac{1}{i_k^\gamma j_k^{1-\gamma}}$$

where $M = M(m,\delta,k)$ depends only on m, δ and k.

5.1 Witness structures

In order to show that a vertex i does not get infected in round $\tau = 1$ **whp**, it suffices to show that there is no depth-1 witness structure, **whp**. This can be done by showing that the expected number of such witness structures is $o(1)$. We shall deal with trees first, where every internal (non-leaf) vertex has r children.

Let $i \in \mathrm{PA}_t(m,\delta)$. A particular tree T_i, rooted at $i = \mathrm{root}(T_i)$ with leaves $L = \mathrm{leaves}(T_i)$, is a subgraph of $\mathrm{PA}_t(m,\delta)$. If $L \subseteq \mathcal{I}_0$ but no other vertex of T_i is in $\mathcal{I}_0$, then T_i is called a *witness tree*. For the sake of the analysis, in this section it will be convenient to consider edges of $\mathrm{PA}_t(m,\delta)$ to be directed, where edge (i,j) is directed from the younger to the older. Thus, given a T_i the orientations on its edges are already determined and we are not free to alter them. Suppose vertex j is a child of vertex j' in a tree T_i. If $j' < j$, then the edge $\{j', j\}$ is directed from j to j' and we call (j, j') an *up edge*; otherwise we call it a *down edge*.

A given tree T_i is a member of a rooted, directed, isomorphism class $\overrightarrow{\mathcal{T}_i}$: this consists of pairwise isomorphic rooted trees, where the root is labeled by i and the other vertices have labels in $[t] \setminus \{i\}$. Here, we assume that an isomorphism between members of this class respects edge orientations.

Alternatively, we may define $\overrightarrow{\mathcal{T}}$ to be a rooted, directed r-ary tree whose vertices are the variables $x_0, x_1, x_2, \ldots, x_N$ and x_0 is the label/variable of the root. These variables take values in $[t]$. If we set $x_0 = i$, then we denote the resulting tree (or class of trees) by $\overrightarrow{\mathcal{T}_i}$. Every assignment of the variables which respects the edge orientations gives rise to a $T_i \in \overrightarrow{\mathcal{T}_i}$.

Let $d_0 = \min\{d \in \mathbb{N} : d\gamma > 1\}$. As we shall see, we need only consider trees of depth at most d_0, which is, of course, a constant. Consequently, there is a bounded number of isomorphism classes, and since each tree is r-ary, no tree has more than r^{d_0+1} vertices.

We shall deal with the cases $r\gamma > 1$ and $r\gamma \le 1$ separately, starting with the former. There, it suffices to consider only trees of depth 1.

We require the following lemma.

Lemma 13. *For infection probability $p = O(1/t^\gamma)$, no vertex in $\mathcal{I}_0$ has parallel edges* **whp**.

Proof. Let $X_t^{\parallel}$ be a random variable that counts the number of vertices j which throw parallel edges in $\mathrm{PA}_t(m,\delta)$. Then, by Lemma 10

$$\mathbb{E}[X_t^{\parallel}] = O(1)\sum_{j=1}^{t}\sum_{i=1}^{j}\frac{1}{i^{2\gamma}j^{2(1-\gamma)}} = O(1)\sum_{j=1}^{t}\frac{1}{j^{2(1-\gamma)}}\int_1^j x^{-2\gamma}\,\mathrm{d}x = \frac{O(1)}{2\gamma-1}\sum_{j=1}^{t}\frac{1-j^{1-2\gamma}}{j^{2(1-\gamma)}}.$$

When $\delta < 0$, we have $1-2\gamma = \frac{\delta/m}{2+\delta/m} < 0$. Hence $\mathbb{E}[X_t^{\parallel}] = O(1)\int_1^t j^{-2(1-\gamma)}\,\mathrm{d}j = O(t^{2\gamma-1})$. Therefore, the expected number of vertices that are in $\mathcal{I}_0$ and throw parallel edges, or throw parallel edges to vertices in $\mathcal{I}_0$, is $O(t^{\gamma-1}) = o(1)$.

When $\delta = 0$, we have $\gamma = 1/2$ so the sum is $O((\log t)^2)$, giving probability $O((\log t)^2/t^\gamma) = o(1)$.

When $\delta > 0$, we have $0 < \gamma < 1/2$ giving probability $O(\log t/t^\gamma) = o(1)$. □

5.2 $r\gamma > 1$

In this section we prove Theorem 2**(i)**. We remind that $\delta < 0$ implies $r\gamma > 1$.

By Lemma 13, any vertex i infected in round $\tau = 1$ must be infected by a depth-1 witness tree.

Lemma 14. *Suppose $p = \frac{1}{\omega t^\gamma}$. For a vertex $i \in PA_t(m,\delta)$, the expected number of depth-1 witness trees rooted at i is $O\left(\frac{1}{\omega^r i^{r\gamma}}\right)$.*

Proof. Let T_i be such a tree with k up edges and $r-k$ down edges. Specifically, say the up leaves are vertices $j_1, j_2, \ldots, j_k$ and the down leaves are vertices $j_{k+1}, j_{k+2}, \ldots, j_r$. If $T_i \subseteq \mathrm{PA}_t(m,\delta)$ means that T_i *is a subgraph of* $\mathrm{PA}_t(m,\delta)$, then

$$\begin{aligned}\mathbb{P}\left(T_i \subseteq \mathrm{PA}_t(m,\delta)\right) &\le M^k \frac{1}{i^\gamma j_1^{1-\gamma}}\frac{1}{i^\gamma j_2^{1-\gamma}}\cdots\frac{1}{i^\gamma j_k^{1-\gamma}} \times \frac{1}{i^{1-\gamma}j_{k+1}^{\gamma}}\frac{1}{i^{1-\gamma}j_{k+1}^{\gamma}}\cdots\frac{1}{i^{1-\gamma}j_r^{\gamma}}\\ &= M^k \frac{1}{i^{\gamma k+(1-\gamma)(r-k)}}\frac{1}{(j_1 j_2\ldots j_k)^{1-\gamma}}\frac{1}{(j_{k+1}j_{k+1}\ldots j_r)^{\gamma}}.\end{aligned}$$

Therefore, the expected number of trees in the isomorphism class (i.e., those trees isomorphic to T_i, rooted at i and having the same edge orientations) is bounded from above by

$$\begin{aligned}O(1)\frac{1}{i^{\gamma k+(1-\gamma)(r-k)}}\left(\int_i^t j^{-1+\gamma}\,\mathrm{d}j\right)^k\left(\int_1^i j^{-\gamma}\,\mathrm{d}j\right)^{r-k} &= O(1)\frac{1}{i^{\gamma k+(1-\gamma)(r-k)}}t^{\gamma k}i^{(1-\gamma)(r-k)}\\ &= O(1)\left(\frac{t}{i}\right)^{\gamma k}.\end{aligned}$$

The above is therefore maximised when $k = r$, that is, when all edges to leaves are up. There are 2^r possible edge orientations, hence, multiplying by the probability that all leaves of such a tree are infected we get a bound of $O(1/\omega^r)i^{-r\gamma}$ for the expected number of depth-1 witness trees rooted at i. □

The proof of Theorem 2**(i)** is a corollary of the above: summing $O(1/\omega^r)i^{-r\gamma}$ over all i from 1 to t, the condition $r\gamma > 1$ ensures we get $o(1)$.

5.3 $r\gamma \leq 1$

In this section we prove Theorem 2**(ii)** and **(iii)**. Recall that the condition $r\gamma \leq 1$ implies that $\delta \geq 0$ and, therefore, $\gamma \leq 1/2$.

5.3.1 Witness trees

Recall that $d_0 = \min\{d \in \mathbb{N} : d\gamma > 1\}$. We shall consider witness trees of depth at most d_0. Since d_0 is a constant and each internal vertex has precisely r children, there is only a bounded number of isomorphism classes.

In round $\tau = 0$, there are $\Theta(t^{1-\gamma}/\omega)$ infected vertices in expectation. We will show that in expectation there are $o(t^{1-\gamma}/\omega)$ newly infected vertices in each of rounds $\tau = 1, 2, \ldots, d_0 - 1$, and $o(1)$ in round d_0. Consequently, the progression of the outbreak stops at or before round d_0 **whp** and, moreover, by Markov's inequality, it follows that $|\mathcal{I}_f|/|\mathcal{I}_0| \xrightarrow{p} 1$ as $t \to \infty$.

If i gets infected in round $\tau = d$, it must be the case that there is a depth-d witness structure which causes this infection. We shall bound from above the expected number of such witness structures for $d = 1, 2, \ldots, d_0$. In this section, we focus on witness structures that are trees - the general case is treated in the next section.

For the purposes of the next section, we will consider an *extended isomorphism class* $\overrightarrow{\mathcal{T}}$ which is a rooted, oriented tree, whose vetrices are variables, taking values in $[t]$, and every vertex has *at most* r children. Assuming that the tree has $N+1$ vertices, the variables that are the labels of the vertices are $x_0, x_1, \ldots, x_N$, where x_0 is the label of the root vertex. When these variables are assigned values in $[t]$ that are compatible with the direction of the edges of $\overrightarrow{\mathcal{T}}$ and the corresponding edges are present in $\mathrm{PA}_t(m, \delta)$, then we have a realisation of $\overrightarrow{\mathcal{T}}$. Thus, we may view $\overrightarrow{\mathcal{T}}$ as the set of all such realisations - we write $T \in \overrightarrow{\mathcal{T}}$. For any $i \in [t]$, we let $\overrightarrow{\mathcal{T}_i}$ denote the restriction of $\overrightarrow{\mathcal{T}}$ where the root variable x_0 has been set to i.

Let us consider, in particular, the case of an extended isomorphism class $\overrightarrow{\mathcal{T}_i}$, for some $i \in [t]$. Let $X_{\overrightarrow{\mathcal{T}_i}}$ count the number of trees $T_i \in \overrightarrow{\mathcal{T}_i}$ such that $T_i \subseteq \mathrm{PA}_t(m, \delta)$ and $L = \mathrm{leaves}(T_i) \subseteq \mathcal{I}_0$. We have

$$\mathbb{E}\left[X_{\overrightarrow{\mathcal{T}_i}}\right] = \left(\frac{1}{\omega t^{\gamma}}\right)^{|L|} \sum_{T_i \in \overrightarrow{\mathcal{T}_i}} \mathbb{P}\left(T_i \subseteq \mathrm{PA}_t(m, \delta)\right). \tag{13}$$

Since each tree has at most r^{d_0+1} edges, by Corollary 12 we have

$$\sum_{T_i \in \overrightarrow{\mathcal{T}_i}} \mathbb{P}\left(T_i \subseteq \mathrm{PA}_t(m, \delta)\right) \leq C_1(m, \delta, r) \sum_{T_i \in \overrightarrow{\mathcal{T}_i}} \prod_{(a,b) \in E(T_i)} \frac{1}{b^{\gamma} a^{1-\gamma}}. \tag{14}$$

where $C_1(m, \delta, r)$ is some constant that depends only on m, δ, r and $E(T_i)$ is the edge set of T_i. (Recall that each edge is oriented from the larger vertex to the smaller one.)

Since each tree T_i is an assignment of the variables $x_1, x_2, \ldots, x_N$ of the tree $\overrightarrow{\mathcal{T}_i}$ (recall that $x_0 = i$), to calculate the sum (13) we can perform a sum over all valid assignments. Our aim is to bound from above the above sum.

To this end, we consider a more general setting in which each vertex x_a is associated with a *valuation function* $v_a : [t] \to \mathbb{R}^+$. When the variables/vertices x_a assume some value, i.e., some assignment of a vertex in $[t]$, then the corresponding vertices get the value $v_a(x_a)$. We consider valuation functions of a certain form. Namely, if x_a is not a leaf, then

$v_a(j) = (\log j)^{\rho_a}/j^{e_a}$, where ρ_a is a non-negative integer and e_a is a non-negative real number such that $e_a = A\gamma + B(1-\gamma)$ with A, B being non-negative integers that satisfy

$$c(x_a) + A + B \geq r. \text{ Property (A)},$$

where $c(x_a)$ denotes the number of children of x_a. Furthermore, if $e_a = 0$, then $\rho_a = 0$, and we call the valuation function *trivial*. Hence, if x_a is an internal vertex with a trivial valuation function, then it has exactly r children. Now, if x_a is a leaf, then either $v_a(j) = \frac{1}{\omega t^\gamma} = p$ (for $r = 2$ we take $\omega = \log t$) or $v_a(j) = (\log j)^{\rho_a}/j^{e_a}$, where ρ_a is a non-negative integer and $e_a > 0$ satisfies Property (A). In the former case, we call the leaf *original*; otherwise, we call it a *contraction* leaf. The purpose of having a valuation function of this form will become apparent in the next section, where we consider general witness structures that are not trees. In those cases we perform a series of operations that convert a general witness structure into a tree. During these operations, we perform contractions of subtrees (hence the term *contraction leaf*). Effectively, the valuation function is (up to multiplicative constants) the probability that the vertex/root of the contracted subtree is infected through this subtree. When a leaf is original, it is meant to be externally infected, whereas a contraction leaf is infected through a certain sub-tree (that had been) rooted at it. If all the children of a vertex are original leaves, then it has exactly r children.

For a vertex x_a of $\overrightarrow{\mathcal{T}}$, we define the function $f_a : [t] \to \mathbb{R}_{\geq 0}$ recursively: If x_a is a leaf, then $f_a(j) = v_a(j)$. Otherwise, with $x_{a_1}, x_{a_2}, \ldots, x_{a_k}$ being the child variables of x_a, where $k \leq r$, with $x_{a_1}, \ldots x_{a_{k_1}}$ having up edges with x_a and the rest down, we set

$$f_a(j) := v_a(j) \left(\prod_{s=1}^{k_1} \sum_{j'=j+1}^{t} f_{a_s}(j') \frac{1}{j^\gamma j'^{1-\gamma}} \right) \left(\prod_{s=k_1+1}^{k} \sum_{j'=1}^{j-1} f_{a_s}(j') \frac{1}{j'^\gamma j^{1-\gamma}} \right). \tag{15}$$

We call f_a the *weight function* of the sub-tree that is rooted at x_a.

It is not hard to see that if the sub-tree $\overrightarrow{\mathcal{T}}_a$ that is rooted at x_a has N (non-root) vertices $x_{a_1}, \ldots, x_{a_N}$, then $f_a(j)$ is

$$f_a(j) = v_a(j) \sum_{j_1, \ldots, j_N} \prod_{\ell=1}^{N} v_{a_\ell}(j_\ell) \prod_{(x_{a_{\ell_1}}, x_{b_{\ell_2}}) \in E(\overrightarrow{\mathcal{T}}_a)} \frac{1}{j_{\ell_2}^\gamma j_{\ell_1}^{1-\gamma}} \mathbf{1}_{\{j_{\ell_1} > j_{\ell_2}\}}.$$

However, if the valuation functions of the internal vertices are trivial and all leaves are original, then simply

$$f_0(i) \geq \left(\frac{1}{\omega t^\gamma} \right)^{|L|} \sum_{T_i \in \overrightarrow{\mathcal{T}}_i} \prod_{(a,b) \in E(T_i)} \frac{1}{b^\gamma a^{1-\gamma}}.$$

Thus by (13) and (14), if we show that

$$\sum_{i=1}^{t} f_0(i) = o(1), \tag{16}$$

this will imply that **whp** there are no vertices which are infected through a tree that is isomorphic to $\overrightarrow{\mathcal{T}}$.

To this end, we will first provide an upper bound on f_0 (cf. Lemma 16 and Corollary 18 below). In fact, we will provide a more general upper bound that is applicable to a general configuration of witness trees. This general form will be useful in the next section where we analyse the expected number of occurrences of general witness structures.

Let $\overrightarrow{\mathcal{T}}(a)$ denote the subtree of $\overrightarrow{\mathcal{T}}$ rooted at x_a. Thus, in particular, $\overrightarrow{\mathcal{T}}(0) = \overrightarrow{\mathcal{T}}$. Let $\rho(a)$ be the number of down edges in $\overrightarrow{\mathcal{T}}(a)$ and let $\ell(a)$ be the number of its original leaves. We shall also be writing $a' \in \overrightarrow{\mathcal{T}}(a)$ to denote that $x_{a'}$ is a vertex of $\overrightarrow{\mathcal{T}}(a)$. For each $a' \in \overrightarrow{\mathcal{T}}$, we denote by $e_{a'}$ and $\rho_{a'}$ the exponents of the valuation function $v_{a'}$ of $x_{a'}$, as described above.

First we require the following lemma, the proof of which is in the appendix:

Lemma 15. *Assume that $x_{a'}$ is a leaf and let $f_{a'}(j') = (\log j')^{\rho_{a'}}/j'^{e_{a'}}$, if it is a contraction leaf, where $e_{a'} > 0$ satisfies Property (A). Then for any $1 \le j < t$ we have*

$$\sum_{j'=j+1}^{t} f_{a'}(j')\frac{1}{j^{\gamma}j'^{1-\gamma}} \lesssim \begin{cases} \frac{1}{\omega}\ \frac{1}{j^{\gamma}}, & \text{if } x_{a'} \text{ is an original leaf} \\ (\log j)^{\rho_{a'}} j^{-e_{a'}}, & \text{if } x_{a'} \text{ is a contraction leaf} \end{cases},$$

and

$$\sum_{j'=1}^{j-1} f_{a'}(j')\frac{1}{j'^{\gamma}j^{1-\gamma}} \lesssim \begin{cases} \frac{1}{\omega}\ \frac{1}{j^{\gamma}}, & \text{if } x_{a'} \text{ is an original leaf} \\ (\log j)^{\rho_{a'}} j^{-e_{a'}}, & \text{if } x_{a'} \text{ is a contraction leaf and } 1-\gamma > e_{a'} \\ \frac{(\log j)^{\rho_{a'}+1}}{j^{1-\gamma}}, & \text{otherwise} \end{cases}.$$

Remark The above lemma effectively shows that it is enough to consider only the case where original leaves are connected to their parent through an up edge, and so, ***we shall henceforth assume this to be the case***. We also remind that if all the children of a vertex are original leaves, then it has exactly r children.

Lemma 16. *Suppose that $\overrightarrow{\mathcal{T}}$ is as above. Let x_a be an internal vertex of $\overrightarrow{\mathcal{T}}$ or a contraction leaf. Then uniformly for all $j \in [t]$,*

$$f_a(j) \lesssim \frac{1}{\omega^{\ell(a)}}\ \frac{(1 \vee (\log j)^{\rho(a)'})}{j^{y_a}}, \tag{17}$$

where $y_a = A\gamma + B(1-\gamma)$ with non-negative integers A, B satisfying $A + B \ge r$. Furthermore, $B > 0$ implies $y_a \ge \ell(a)\gamma + \sum_{a' \in \overrightarrow{\mathcal{T}}(a)} e_{a'}$, and $\rho(a)' = \rho(a) + \sum_{a' \in \overrightarrow{\mathcal{T}}(a)} \rho_{a'}$.

Remark The hidden constant factor in (17) depends only on m, δ and r.

Proof. If x_a is a contraction leaf, then $f_a = v_a$ by definition, and the lemma is immediate. Hence, we assume below that x_a is an internal vertex.

For the sake of notational convenience we shall omit the '$1\vee$' component of terms of the form $(1 \vee (\log j)^{\rho})$. We shall give a proof by induction starting from the bottom and going "up" the tree. The reader is reminded that we are ultimately aiming to bound the right-hand side of (15)

Assume that $x_a = j$ and suppose x_{a_1} is a child of x_a. We shall consider a number of cases regarding x_{a_1} and combine them in the end. Consider first of all that x_{a_1} is a leaf. Recalling

that for a leaf x_{a_1} the weight function f_{a_1} is defined to be equal to the valuation function v_{a_1}, we apply Lemma 15. Thus, if x_{a_1} connects to x_a by an up edge, we have

$$\sum_{j'=j+1}^{t} f_{a'}(j')\frac{1}{j^{\gamma}j'^{1-\gamma}} \lesssim \begin{cases} \frac{1}{\omega}\ \frac{1}{j^{\gamma}}, & \text{if } x_{a'} \text{ is an original leaf} \\ (\log j)^{\rho_{a'}} j^{-e_{a'}}, & \text{if } x_{a'} \text{ is a contraction leaf} \end{cases},$$

and if x_{a_1} connects to x_a by a down edge, we have

$$\sum_{j'=1}^{j-1} f_{a'}(j')\frac{1}{j'^{\gamma}j^{1-\gamma}} \lesssim \begin{cases} (\log j)^{\rho_{a'}} j^{-e_{a'}}, & \text{if } x_{a'} \text{ is a contraction leaf and } 1-\gamma > e_{a'} \\ \frac{(\log j)^{\rho_{a'}+1}}{j^{1-\gamma}}, & \text{otherwise} \end{cases}.$$

(Recall, as per the remark after the statement of Lemma 15, we have made the assumption that original leaves are only connected through up edges).

Now suppose x_{a_1} is an internal vertex and a child of x_a. Letting $\ell_1 = \ell(a_1)$, the number of original leaves in the subtree rooted at x_{a_1}, we have by the induction hypothesis, $f_{a_1}(j_1) \lesssim \left(\frac{1}{\omega}\right)^{\ell_1} \frac{(\log j_1)^{\rho_1}}{j_1^{A_1\gamma+B_1(1-\gamma)}}$ for some appropriate A_1, B_1 and ρ_1 as in the statement of the lemma. In particular, these are such that $A_1 + B_1 \geq r$.

Consider now the case where $x_{a_1} = j_1$ is connected by an up edge. We have

$$\begin{aligned} \sum_{j_1=j+1}^{t} f_{a_1}(j_1)\frac{1}{j^{\gamma}j_1^{1-\gamma}} &\lesssim \left(\frac{1}{\omega}\right)^{\ell_1} \sum_{j_1=j+1}^{t} \frac{1}{j^{\gamma}j_1^{1-\gamma}} \frac{(\log j_1)^{\rho_1}}{j_1^{A_1\gamma+B_1(1-\gamma)}} \\ &\lesssim \left(\frac{1}{\omega}\right)^{\ell_1} \frac{1}{j^{\gamma}} \int_{j}^{t} x_1^{-1+\gamma-(A_1\gamma+B_1(1-\gamma))} (\log x_1)^{\rho_1}\, \mathrm{d}x_1. \end{aligned}$$

Observe $A_1\gamma + B_1(1-\gamma) > \gamma$. Indeed, since $\delta \geq 0$ we have $\gamma \leq 1-\gamma$. Thereby, $A_1\gamma + B_1(1-\gamma) \geq (A_1+B_1)\gamma \geq r\gamma > \gamma$. Therefore, we can apply the following lemma to bound the above integral. Its proof can be found in Section 8.2 (appendix).

Lemma 17. *Let $k \geq 0$ be an integer, let $\alpha > 0$ be a real number and let*

$$I_{k,\alpha}(j) := \int_{j}^{t} (\log x)^k x^{-1-\alpha}\, \mathrm{d}x.$$

Then uniformly for $j \geq 1$ we have

$$I_{k,\alpha}(j) \lesssim \frac{(\log j \vee 1)^k}{j^{\alpha}}.$$

This gives

$$\sum_{j_1=j+1}^{t} f_{a_1}(j_1)\frac{1}{j^{\gamma}j_1^{1-\gamma}} \lesssim \left(\frac{1}{\omega}\right)^{\ell_1} \frac{(\log j)^{\rho_1}}{j^{A_1\gamma+B_1(1-\gamma)}}. \tag{18}$$

Note that due to the fact that we consider trees of bounded degree and depth, terms such as ρ and $A\gamma + B(1-\gamma)$ will always be bounded from above and below by constants that depend only on m, δ and r. Therefore, the constant factor incurred by the above integration is always bounded by some constant that only depends on these parameters.

Observe that (18) is the same (up to multiplicative constants) as the expression for $f_{a_1}(j_1)$ except that j has replaced j_1. In this sense, we see that an up edge causes the parent vertex to "reverse inherit" the exponent of the child, in this case, that exponent being $A_1\gamma + B_1(1-\gamma)$.

Now we will consider what happens if it is a down edge, where, by assumption, x_{a_1} is an internal vertex. We have

$$\begin{aligned}
\sum_{j_1=1}^{j-1} f_{a_1}(j_1)\frac{1}{j^{\gamma}j_1^{1-\gamma}} &\lesssim \left(\frac{1}{\omega}\right)^{\ell_1} \sum_{j_1=1}^{j-1} \frac{1}{j_1^{\gamma}j^{1-\gamma}} \frac{(\log j_1)^{\rho_1}}{j_1^{A_1\gamma+B_1(1-\gamma)}} \\
&\lesssim \left(\frac{1}{\omega}\right)^{\ell_1} \frac{(\log j)^{\rho_1}}{j^{1-\gamma}} \int_1^j x_1^{-\gamma-(A_1\gamma+B_1(1-\gamma))}\,\mathrm{d}x_1 \\
&\lesssim \begin{cases} \left(\frac{1}{\omega}\right)^{\ell_1} \frac{(\log j)^{\rho_1+1}}{j^{1-\gamma}} & \text{if } 1-\gamma-(\gamma A_1+(1-\gamma)B_1) \le 0 \\ \left(\frac{1}{\omega}\right)^{\ell_1} \frac{(\log j)^{\rho_1}}{j^{A_1\gamma+B_1(1-\gamma)}} & \text{if } 1-\gamma-(\gamma A_1+(1-\gamma)B_1) > 0 \end{cases}
\end{aligned} \tag{19}$$

Once again, we emphasise that the integration incurs a constant factor that is bounded by a constant that depends only on m, r and δ.

Let $\mathcal{C}(a)$ denote the set of indices of the children of x_a. For any $a' \in \mathcal{C}(a)$, let $r_{a'}$ denote the exponent of j in the expression of the upper bound of $f_{a'}(j)$. We let $\mathcal{C}_1 \subseteq \mathcal{C}(a)$ denote the set of indices of the original leaves among the members of $\mathcal{C}(a)$. Also, we let $\mathcal{C}_2 \subseteq \mathcal{C}(a)$ denote the set of the indices of those children of x_a that are not original leaves but are connected to x_a through up edges. Let $\mathcal{C}_2'$ denote the set of the indices of those children that are not original leaves, are connected to x_a through down edges and $1-\gamma > r_{a'}$, for $a' \in \mathcal{C}_2'$. Similarly, we define as $\mathcal{C}_2''$ the set of the indices of those children that are not original leaves, are connected to x_a through down edges but $1-\gamma \le r_{a'}$, for $a' \in \mathcal{C}_2''$.

By the above and the the definition in (15),

$$f_a(j) \lesssim v_a(j)\left(\frac{1}{\omega}\right)^{\sum_{a'\in\mathcal{C}(a)}\ell(a')} \frac{(\log j)^{|\mathcal{C}_2''|+\sum_{a'\in\mathcal{C}_2\cup\mathcal{C}_2'\cup\mathcal{C}_2''}\rho(a')}}{j^{|\mathcal{C}_1|\gamma+\sum_{a'\in\mathcal{C}_2\cup\mathcal{C}_2'} r_{a'}+(1-\gamma)|\mathcal{C}_2''|}} \tag{20}$$

Let y_a denote the exponent of j. Firstly, note that $\ell(a) = \sum_{a'\in\mathcal{C}(a)}\ell(a')$.

Assume that $\mathcal{C}_2 \cup \mathcal{C}_2' = \emptyset$. Then $y_a = |\mathcal{C}_1|\gamma + (1-\gamma)|\mathcal{C}_2''| + e_a$. But as $|\mathcal{C}_1| + |\mathcal{C}_2''| = c(x_a)$ and e_a satisfies Property (A), it follows that $|\mathcal{C}_1|\gamma + (1-\gamma)|\mathcal{C}_2''| + e_a = A\gamma + B(1-\gamma)$, where A, B are non-negative integers that satisfy $A + B \ge r$.

If $\mathcal{C}_2 \cup \mathcal{C}_2' \neq \emptyset$, then $r_{a'} > 0$, for some $a' \in \mathcal{C}_2 \cup \mathcal{C}_2'$, which has the form $A\gamma + B(1-\gamma)$, for some A, B that are non-negative integers satisfying $A + B \ge r$. Hence, the exponent of j satisfies this as well.

Assume now that y_a cannot be written in the form $A\gamma + B(1-\gamma)$ with A, B non-negative integers and $B > 0$. Then this is the case for $r_{a'}$ for any $a' \in \mathcal{C}_2 \cup \mathcal{C}_2'$. Hence, by the induction hypothesis $\sum_{a'\in\mathcal{C}_2\cup\mathcal{C}_2'} r_{a'}$ is equal to the number of original leaves that are contained in the sub-tree that is rooted at those $x_{a'}$ together with the sum of the exponents $e_{a'}$ of the valuation functions of the vertices of these sub-trees. Moreover, $|\mathcal{C}_2''| = 0$ and recall that $|\mathcal{C}_1|$ is the number of original leaves that are directly connected to x_a. Thereby,

$$y_a = \ell(a)\gamma + \sum_{a'\in\overrightarrow{\mathcal{T}}(a)} e_{a'}.$$

The base case, where x_a has only leaves for children, is immediate from the above.

□

The above lemma now implies the following.

Corollary 18. *If the valuation functions of the internal vertices of $\overrightarrow{\mathcal{T}}$ are trivial and all leaves are original, then*

$$f_0(i) \lesssim \left(\frac{1}{\omega}\right)^{\ell} \left(\frac{(1 \vee (\log i)^{\rho})}{i^{y_0}}\right) \tag{21}$$

where $\ell = \ell(0)$ and $\rho = \rho(0)$ and either $y_0 = A\gamma + B(1-\gamma)$ where A, B are non-negative integers that satisfy $A + B \geq r$ and $B > 0$ or $y_0 = \ell\gamma$.

We conclude with the proof of (16) for depth d_0. Consider the expression on the right-hand side of (21). If $0 < B < r$ and $r \geq 3$, then $A \geq r - B$ and it is easy to check that $A\gamma + B(1-\gamma) > 1$ (it is a convex combination of two positive numbers that are at least 1, one of which is bigger than one, where $\gamma \neq 0, 1$). If $B \geq r$, then $A\gamma + B(1-\gamma) \geq r(1-\gamma) \geq 3(1-\gamma)$. But $\gamma \leq 1/2$, whereby $3(1-\gamma) \geq 3/2 > 1$. If $y_0 = \ell\gamma$, then $\ell \geq d_0$ implies $y_0 \geq \gamma d_0 > 1$.

If $r = 2$ then we are not necessarily guaranteed $A\gamma + B(1-\gamma) > 1$ since, for example, $B = 1$ and $r = 2$ only assures $\gamma A + (1-\gamma)B \geq 1$.

If $\gamma A + (1-\gamma)B > 1$, then

$$\sum_{i=1}^{t} f_0(i) \lesssim \left(\frac{1}{\omega}\right)^{\ell} \sum_{i=1}^{t} \frac{(1 \vee (\log i)^{\rho})}{i^{\gamma A + (1-\gamma)B}} \lesssim \left(\frac{1}{\omega}\right)^{\ell} \int_1^t \frac{(\log x)^{\rho}}{x^{\gamma A + (1-\gamma)B}} \, \mathrm{d}x = O\left(\left(\frac{1}{\omega}\right)^{\ell}\right). \tag{22}$$

In this case, the expected number of witness trees of this isomorphism class, over all i, is $o(1)$.

In the case that $y_0 = \ell\gamma \leq 1$, the sum is $\left(\frac{1}{\omega}\right)^{\ell} (\log t)^{\rho+1} t^{1-\ell\gamma}$ and the expected number of witness trees of this isomorphism class, over all i, is $O((\log t)^{\rho+1} t^{1-\ell\gamma}/\omega^{\ell}) = o(t^{1-\gamma}/\omega^{\ell})$ since $\ell \geq 2$. In other words, the expected number of witness trees of depth less than d_0 is $o(t^{1-\gamma}/\omega^{\ell})$.

As stated above, there are only a bounded number of isomorphism classes that we need to consider, hence the relevant constant factors are absorbed into the $O(.)$ terms above.

We would like to extend this to include $r = 2$, wherein if we take the depth of the tree to be equal to d_0, then it may be the case that the exponent of i in $f_0(i)$ is 1 (which is the minimum it can be when $\ell \geq d_0$). In that case, the integral in (22) would grow like $(\log t)^{\rho+1}$. To bypass this difficulty, when $r = 2$ we consider witness trees that have depth equal to $d_0 + 1$. Recall that in this case we assume that $p_0 = \frac{1}{\log t} \, \frac{1}{t^{\gamma}}$. Also, as we have already commented in the remark after Lemma 15, we may assume that the witness trees we consider are such that all their leaves are connected to the rest of the tree through up edges.

Let x_1, x_2 be the children of x_0 and assume without loss of generality that the subtree that is rooted at x_1 has depth d_0. Suppose the exponent of j_1 is 1. That is, recalling that this subtree has $\ell(1)$ original leaves, by Lemma 16 we have $f_1(j_1) \lesssim \frac{1}{(\log t)^{\ell(1)}} \frac{(\log j_1)^{\rho(1)}}{j_1}$. Thus, if x_1 is connected by an up edge with x_0, by (18) the exponent transfers, and we get a factor $\frac{1}{(\log t)^{\ell(1)}} \frac{(\log i)^{\rho(1)}}{i}$ in $f_0(i)$. If it is connected through a down edge, by (19), we get the factor

$$\frac{1}{(\log t)^{\ell(1)}} \, \frac{1}{i^{1-\gamma}} \int_1^i x_1^{-\gamma-1} (\log x_1)^{\rho(1)} \, \mathrm{d}x_1 \lesssim \frac{1}{(\log t)^{\ell(1)}} \, \frac{1}{i^{1-\gamma}}.$$

If x_2 is an original leaf, then by Lemma 15 it contributes a factor that is at most (up to a multiplicative constant) $\frac{1}{(\log t)}\frac{1}{i^\gamma}$, thus giving in total $\frac{1}{(\log t)^{\ell(1)+1}}\frac{1}{i^{1+\gamma}}$ or $\frac{1}{(\log t)^{\ell(1)+1}}\frac{1}{i}$. In any case,

$$f_0(i) \lesssim \frac{1}{(\log t)^{\ell(1)+1}}\ \frac{1}{i}. \tag{23}$$

If x_2 is not an original leaf, then by Lemma 16 $f_2(j_2) \lesssim \frac{1}{(\log t)^{\ell(2)}}\frac{(\log j_2)^{\rho(2)}}{j_2^{y_2}}$, where either y_2 can be written as $A\gamma + B(1-\gamma)$ for some non-negative integers A, B that satisfy $B \geq 1$ and $A + B \geq 2$, or $y_2 \geq \ell(2)\gamma$, where in this case $\ell(2) \geq 2$.

If x_2 is joined to x_0 by an up edge, then by (18) it contributes a factor that is at most (up to a constant) $\frac{1}{(\log t)^{\ell(2)}}\ \frac{(\log i)^{\rho(2)}}{i^{y_2}} \lesssim \frac{1}{(\log t)^{\ell(2)}}\ \frac{(\log i)^{\rho(2)}}{i^{2\gamma}}$, giving a total

$$f_0(i) \lesssim \frac{1}{(\log t)^{\ell(1)}}\frac{(\log i)^{\rho(1)}}{i^{1-\gamma}}\frac{1}{(\log t)^{\ell(2)}}\frac{(\log i)^{\rho(2)}}{i^{2\gamma}} = \frac{1}{(\log t)^{\ell(0)}}\frac{(\log i)^{\rho(0)}}{i^{1+\gamma}}. \tag{24}$$

If x_2 is not an original leaf and is connected to x_0 by a down edge, the possibilities are

$$f_0(i) \lesssim \frac{1}{(\log t)^{\ell(1)}}\frac{(\log i)^{\rho(1)}}{i}\frac{1}{(\log t)^{\ell(2)}}\frac{(\log i)^{\rho(2)+1}}{i},$$

or

$$f_0(i) \lesssim \frac{1}{(\log t)^{\ell(1)}}\frac{(\log i)^{\rho(1)}}{i}\frac{1}{(\log t)^{\ell(2)}}\frac{(\log i)^{\rho(2)+1}}{i^{2\gamma}},$$

or

$$f_0(i) \lesssim \frac{1}{(\log t)^{\ell(1)}}\ \frac{1}{i^{1-\gamma}}\frac{1}{(\log t)^{\ell(2)}}\frac{(\log i)^{\rho(2)+1}}{i},$$

or

$$f_0(i) \lesssim \frac{1}{(\log t)^{\ell(1)}}\ \frac{1}{i^{1-\gamma}}\frac{1}{(\log t)^{\ell(2)}}\frac{(\log i)^{\rho(2)+1}}{i^{2\gamma}}.$$

In all cases,

$$f_0(i) \lesssim \frac{1}{(\log t)^{\ell(0)}}\frac{(\log i)^{\rho(0)}}{i^{1+\gamma}}. \tag{25}$$

Summing (23), (24) or (25) over $i = 1, \ldots, t$ gives $o(1)$.

Consequently, the expected number of witness trees of depth $d_0 + 1$ when the initial infection probability is $p_0 = \frac{1}{(\log t)t^\gamma}$ is $o(1)$ as well.

We have shown that if $r \geq 3$, then **whp** the process stops in *less than* d_0 rounds, whereas for $r = 2$ (with the appropriate choice of p_0) it stops in *at most* d_0 rounds. Note that $d_0 = \lfloor \frac{1}{\gamma} \rfloor + 1$. To be more precise, we have shown the bounds of Theorem 2 for witness structures that are trees. We need to argue about general witness structures that may contain cycles. In this case, we show that the expected number of occurrences of such a structure is bounded by the expected number of occurrences of a tree that is appropriately constructed and has depth either d_0 or $d_0 + 1$, depending on the value of r.

5.3.2 General witness structures

We now consider general witness structures that may have cycles. Recall that we are only assuming that $\delta \geq 0$, since $\delta < 0 \Rightarrow r\gamma > 1$. Recall that in this case $\gamma \leq 1/2$.

Firstly, the following lemma allows us to consider witness structures where the initially infected vertices are vertices which do not belong to cycles.

Lemma 19. *Let K be positive constant. If $p = O(1/t^\gamma)$, then with high probability, no initially infected vertex lies on a cycle of size at most K*

Proof. For a cycle $C = (a_1, a_2, \ldots, a_k)$ of size $k \leq K$, we apply Corollary 12: for some constant $M = M(K, \delta, m)$ we have

$$\mathbb{P}(C \subseteq \mathrm{PA}_t(m, \delta)) \leq M^{2k} \prod_{i=1}^{k} \frac{1}{(a_i \wedge a_{i+1})^\gamma (a_i \vee a_{i+1})^{1-\gamma}} \leq \frac{M^{2k}}{a_1 \ldots a_k},$$

where we have used the fact that for $i < j$, $\frac{1}{i^\gamma j^{1-\gamma}} \leq \frac{1}{(ij)^{1/2}}$ when $\gamma \leq \frac{1}{2}$.

Thus, the expected number of cycles in $\mathrm{PA}_t(m, \delta)$ of size at most K is bounded from above by

$$\sum_{3 \leq k \leq K} \sum_{a_1, \ldots, a_k} \frac{M^{2k}}{a_1 \ldots a_k} = O((\log t)^K) \tag{26}$$

and so the number of initially infected vertices on such cycles is $O((\log t)^{K+1}/t^\gamma) = o(1)$. □

However, it is not true that initially infected vertices lie far from cycles. Hence, when one is supposed to trace the infection history, one has to take into account those local structures that are not trees. Furthermore, a vertex may get its first infection through such a structure. We will now formalise this notion.

Recall that if a vertex i becomes infected in round τ, then it must have been infected by some neighbours, at least one of which got infected in round $\tau - 1$. Iterating this argument, there must be a chain of infections of length τ that started in a set of initially infected vertices. This is witnessed by a rooted subgraph, whose root is vertex i and whose other vertices can be classified according to their *depths*. Let us consider this notion more precisely. Suppose x and y are neighbours in $\mathrm{PA}_t(m, \delta)$, $x \in \mathcal{I}(\tau) \cap \mathcal{S}(\tau - 1)$, and $y \in \mathcal{I}(\tau - 1)$. Then we say x is a *parent* of y and y a *child* of x. If x is a parent of y and $x < y$ then $\{x, y\}$ is an up edge. If $x > y$, then it is a down edge. The notion of *parent-child* gives rise to the *depth* of a vertex. Let $\mathrm{depth}(i) = 0$ and $\mathrm{depth}(y) = 1 + \max\{\mathrm{depth}(x) \ : \ y \text{ is a child of } x\}$. We shall use this notion later in our proof.

Suppose a vertex $i \in \mathcal{I}(\tau)$ for some $\tau > 0$. Then there must exist a subgraph $S_i \subseteq \mathrm{PA}_t(m, \delta)$ such that the following hold:

(1) every vertex in S_i except i has a parent in S_i;

(2) the set $L = S_i \cap \mathcal{I}(0)$, which we call the set of *leaves*, is non-empty;

(3) every parent in S_i has exactly r edges in S_i which go to children in S_i.

If, furthermore, $i \in \mathcal{I}(\tau) \cap \mathcal{S}(\tau - 1)$, that is, i got infected in round τ, then we also have:

(4) $\text{depth}(S_i) = \max_{j \in S_i} \text{depth}(j) = \tau$, where $\text{depth}(i) = 0$.

We call such an S_i a *witness structure rooted at* i. Observe that **(1)** forces S_i to be connected. Of course, leaves cannot be parents. Additionally, recall that we only need to analyse bounded size structures and, therefore, only have bounded size cycles. Hence, by Lemmas 13 and 19, any leaf will, with high probability, have degree 1 in S_i. We will assume this to be the case.

Condition **(3)** implies that a parent has at most r children in S_i, and the witness structure is a witness tree as per the previous definition, if and only if every parent has exactly r children and every vertex except i has exactly one parent. For a tree, it is also the case that the depth as defined here in terms of infections coincides with the standard meaning of depth – the graph distance from the root i to a vertex.

Our aim is to bound from above the expected number of witness structures that are rooted at i. To this end, we will bound this expected value by the expected number of occurrences of a tree which is produced from this witness structure through a bounded number of transformations. Informally, during each transformation we "destroy" vertices which belong to cycles in this witness structure. Eventually, having destroyed all such vertices we will obtain a tree whose vertices are equipped with certain valuation functions. We finally bound the expected number of occurrences of this tree using Lemma 16.

As with trees, we let $\overrightarrow{\mathcal{S}}$ denote an isomorphism class of a witness structure. This can be viewed as a directed graph whose vertices $x_0, \dots, x_N$ are variables taking values in $[t]$, that satisfies Conditions **(1)** and **(3)**. We assume that its root is x_0. If S is a witness structure on $[t]$ that is isomorphic to $\overrightarrow{\mathcal{S}}$, where adjacent vertices are compatible with the directions of the corresponding edges of $\overrightarrow{\mathcal{S}}$, then we write $S \in \overrightarrow{\mathcal{S}}$. We let $\overrightarrow{\mathcal{S}_i}$ denote the subset of the isomorphism class $\overrightarrow{\mathcal{S}}$, where the root is vertex i. That is, $x_0 = i$.

Let $X_{\overrightarrow{\mathcal{S}_i}}$ count the number of copies $S_i \in \overrightarrow{\mathcal{S}_i}$ such that $S_i \subseteq \text{PA}_t(m, \delta)$ and $L = \text{leaves}(S_i) \subseteq \mathcal{I}_0$. We have

$$\mathbb{E}\left[X_{\overrightarrow{\mathcal{S}_i}}\right] = \left(\frac{1}{\omega t^\gamma}\right)^{|L|} \sum_{S_i \in \overrightarrow{\mathcal{S}_i}} \mathbb{P}\left(S_i \subseteq \text{PA}_t(m, \delta)\right). \tag{27}$$

Using Corollary 12 we have

$$\sum_{S_i \in \overrightarrow{\mathcal{S}_i}} \mathbb{P}\left(S_i \subseteq \text{PA}_t(m, \delta)\right) \le C_2(m, \delta, r) \sum_{S_i \in \overrightarrow{\mathcal{S}_i}} \prod_{(a,b) \in E(S_i)} \frac{1}{b^\gamma a^{1-\gamma}}.$$

where $C_2(m, \delta, r)$ is some constant that depends only on m, δ, r and $E(S_i)$ denotes the edge set of S_i.

As in the case of trees, we will consider the notion of a *generalised witness structure*, where each vertex x_a is associated with a *valuation function* $v_a : [t] \to \mathbb{R}^+$. The valuation functions we consider are as those we considered in the previous section.

Given such a witness structure $\overrightarrow{\mathcal{S}}$, we will define a function $f_{\overrightarrow{\mathcal{S}}} : [t] \to \mathbb{R}^+$, which generalises the weight function of a tree that was defined in the previous sub-section. When the valuation functions are trivial, then $f_{\overrightarrow{\mathcal{S}}}(i)$ is (up to multiplicative constants) the expected number of occurrences of $\overrightarrow{\mathcal{S}}$ rooted at i, in the product space of $\text{PA}_t(m, \delta)$ and the set of initially infected vertices. Assume that the vertices of $\overrightarrow{\mathcal{S}}$ are $x_0, , \dots, x_N$, where x_0 is the root. We will be associating the index j_a with the variable x_a. Also, recall that the edges of

$\overrightarrow{\mathcal{S}}$ are directed and therefore the edges are *ordered* pairs. Letting $j_0 = i$, we set

$$f_{\overrightarrow{\mathcal{S}}}(i) = v_0(i) \sum_{j_1,\ldots,j_N} \prod_{a=1}^{N} v_a(j_a) \prod_{(x_a,x_b)\in E(\overrightarrow{\mathcal{S}_i})} \frac{1}{j_b^{\gamma} j_a^{1-\gamma}} \mathbf{1}_{\{j_a > j_b\}}.$$

This should not be confused with notation introduced in (15). However, it is not hard to see that if $\overrightarrow{\mathcal{S}}$ is a tree, then the above function coincides with the function $f_0(i)$.

Fix a directed isomorphism class $\overrightarrow{\mathcal{S}}$. We demonstrate how a sequence of transformations can transform $\overrightarrow{\mathcal{S}}$ into a tree isomorphism class $\overrightarrow{\mathcal{T}}$, such that each class in the sequence is an upper bound (in terms of expectation of witness structures) for the previous. Note that by Lemmas 13 and 19, it suffices to consider witness structures of bounded depth where all initially infected vertices have degree 1.

Let x_a be a vertex on a cycle such that it has maximum depth (as defined above in terms of the parent-child relation) among all vertices on cycles. Let $\overrightarrow{\mathcal{T}}(a)$ be the sub-tree rooted at vertex x_a. We apply Lemma 16 to $\overrightarrow{\mathcal{T}}(a)$ and obtain

$$f_{\overrightarrow{\mathcal{T}}(a)}(j_a) \lesssim \frac{1}{\omega^{\ell(a)}} \frac{(\log j_a)^{\rho(a)}}{j_a^{y_a}},$$

where $\rho(a)$ and y_a are as in Lemma 16. In particular, $y_a = A\gamma + B(1-\gamma)$, where A, B are non-negative integers that satisfy $A + B \geq r$.

We are now ready to define the witness structure $T\overrightarrow{\mathcal{S}}$. Assume that x_a has $k > 1$ parents $x_{a_1}, \ldots, x_{a_k}$ (not necessarily distinct). Also assume that x_a is connected to $x_{a_1}, \ldots, x_{a_h}$ through up edges and to $x_{a_{h+1}}, \ldots x_{a_k}$ through down edges, where $0 \leq h \leq k$. Let Δ now be the index of a parent of the highest depth among $x_{a_1}, \ldots, x_{a_h}$, if $h > 0$. To construct $T\overrightarrow{\mathcal{S}}$

1. remove $\overrightarrow{\mathcal{T}}(a)$ together with x_a;
2. multiply $v_{a_\Delta}(j_{a_\Delta})$ by $(\log j_{a_\Delta})^{\rho(a)} j_{a_\Delta}^{-y_a}$;
3. multiply $v_{a_i}(j_{a_i})$ by $j_{a_i}^{-\gamma}$, for all $i \neq \Delta$ and $i \leq h$;
4. multiply $v_{a_i}(j_{a_i})$ by $(\log j_{a_i})^{\rho(a)+1} / j_{a_i}^{(1-\gamma)\wedge y_a}$, for all $i = h+1, \ldots, k$.

If one of the x_{a_i}s is connected to x_a through parallel edges, then the appropriate step from the above is applied once for each edge. For the particular case of x_{a_Δ}, Step 2 is applied once for one of the parallel edges, whereas for the others we apply Step 3. If the parallel edges are down edges, then we apply Step 4 once for each of them.

Note that if the valuation functions v_{a_i} which are modified have exponents e_{a_i} satisfying Property (A), then the modifications incurred by Steps 2-4 preserve this property. Steps 3 and 4 simply remove a child of x_{a_i} and add to the exponent e_{a_i} a γ or a $y_a \wedge 1-\gamma$, thus preserving Property (A). Step 2 removes a child of x_{a_Δ} and adds y_a to e_{a_Δ}. But $y_a = A\gamma + B(1-\gamma)$, for some non-negative integers A, B that satisfy $A + B \geq r$. Hence, Property (A) is also preserved for this exponent.

Steps 2-4 yield

$$\sum_{a' \in T\overrightarrow{\mathcal{S}}} e_{a'} = \sum_{a' \in \overrightarrow{\mathcal{S}} \setminus \overrightarrow{\mathcal{T}}(a)} e_{a'} + \begin{cases} k \min\{1-\gamma, y_a\} & \text{if } h = 0 \\ (h-1)\gamma + y_a + (k-h)\min\{1-\gamma, y_a\} & \text{if } h > 0 \end{cases}. \tag{28}$$

As we shall see in the proof of the next lemma, Steps 2-4 essentially correspond to a step among a sequence of steps that transform $\overrightarrow{\mathcal{S}}$ into a tree. In each step, we have the creation of copies of x_a, which we denote by $x_{a^{(1)}}, \ldots, x_{a^{(k)}}$, where $x_{a^{(i)}}$ is attached to x_{a_i} through an up edge if $i \leq h$ or through an down edge if $i > h$. Thereafter, $x_{a^{(\Delta)}}$ as well as $x_{a^{(i)}}$, for $i > h$, each becomes the root of a copy of $\overrightarrow{\mathcal{T}}(a)$ (cf. the remark below), whereas for the remaining is, the vertices $x_{a^{(i)}}$ become original leaves (cf. Figure 1). We denote the resulting directed graph by $\hat{T}\overrightarrow{\mathcal{S}}$. Note that this is not the directed graph $T\overrightarrow{\mathcal{S}}$. The latter may be thought as coming from $\hat{T}\overrightarrow{\mathcal{S}}$ with the subtrees rooted at each $x_{a^{(i)}}$ *contracted* into x_{a_i}, multiplying the corresponding valuation functions of x_{a_i} by certain factors, as in Steps 2-4. These factors are upper bounds on the probability that $x_{a^{(i)}}$ will be infected through the sub-tree that is rooted at it.

Remark Note also that the depth of $\hat{T}\overrightarrow{\mathcal{S}}$ is equal to the depth of $\overrightarrow{\mathcal{S}}$. This is the case as all $x_{a^{(i)}}$, for $i > h$, are the roots of a copy of $\overrightarrow{\mathcal{T}}(a)$ as well as $x_{a^{(\Delta)}}$. The latter is adjacent to the deepest parent x_{a_Δ} among the x_{a_i}s, for $i \leq h$. This is the reason for which we "assign" a rooted copy of $\overrightarrow{\mathcal{T}}(a)$ to this particular vertex and we do not treat it as an original leaf, as we did in the case of the other vertices that are connected to x_a through up-edges.

Because of (27), we are interested in the case where the initial witness structure $\overrightarrow{\mathcal{S}}$ has only trivial valuation functions. In this case, $\sum_{a' \in \overrightarrow{\mathcal{S}}} e_{a'} = 0$. Now make the conservative[1] assumption that each time we apply T, we have $1 - \gamma \geq y_a$, which means each y_a will have the form $A\gamma$ for some integer $A \geq r$. It then follows from (28) that during the jth transformation the sum of the exponents of the valuation functions increases by $\gamma\ell_j$, where ℓ_j is the number of leaves that are added during the transition from $\hat{T}^{(j-1)}\overrightarrow{\mathcal{S}}$ to $\hat{T}^{(j)}\overrightarrow{\mathcal{S}}$. This can be seen inductively: original leaves with up edges contribute γ each and each $y_a = A\gamma$ represents the accumulation of γ terms from leaves in the subtree rooted at x_a.

Assume that the process stops after step j_0. Thus, $\hat{T}^{(j_0)}\overrightarrow{\mathcal{S}}$ is an r-ary tree where all its leaves are original and, by the above remark, has depth that is equal to the maximum depth in $\overrightarrow{\mathcal{S}}$. If L_{j_0} is the number of leaves of this tree and $\ell(T^{(j_0)}\overrightarrow{\mathcal{S}})$ is the number of original leaves of $T^{(j_0)}\overrightarrow{\mathcal{S}}$, then

$$L_{j_0}\gamma = \sum_{a' \in T^{(j_0)}\overrightarrow{\mathcal{S}}} e_{a'} + \ell(T^{(j_0)}\overrightarrow{\mathcal{S}})\gamma. \tag{29}$$

We will use this fact towards the end of our analysis. We now proceed with our basic inductive step which will allow us to bound $f_{\overrightarrow{\mathcal{S}}}$ after the application of a sequence of transformations T.

Lemma 20. *Let $\overrightarrow{\mathcal{S}}$ be a witness structure and x_a be a vertex of maximum depth on which we perform the above transformation. If $\ell(a)$ denotes the number of original leaves in $\overrightarrow{\mathcal{T}}(a)$, then uniformly for all $i \in [t]$*

$$f_{\overrightarrow{\mathcal{S}}}(i) \lesssim \frac{1}{\omega^{\ell(a)}}\ f_{T\overrightarrow{\mathcal{S}}}(i).$$

Proof. Let x_a be a vertex of $\overrightarrow{\mathcal{S}}$ of maximum depth as in the statement of the lemma. Also, we denote the set of indices of the vertices of $\overrightarrow{\mathcal{T}}(a)$ by $\mathcal{T}(a)$. Let $x_{a_1}, \ldots, x_{a_k}$ be the parents of x_a. Denote by $\mathcal{P}(a)$ the set of indices of $\overrightarrow{\mathcal{S}}$ not in $\mathcal{T}(a)$. Note that $a_j \in \mathcal{P}(a)$, for $j = 1, \ldots, k$.

[1]The assumption is conservative because a larger exponent can only make the final upper bound smaller.

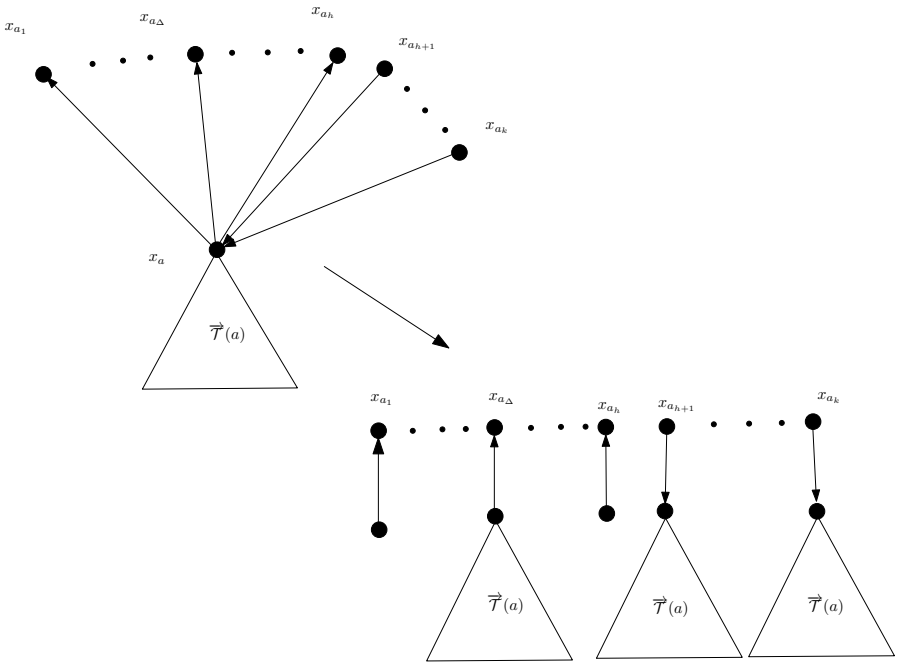

Figure 1: Transformation $\hat{T}$

Finally assume that the edges $(x_a, x_{a_1}), \ldots, (x_a, x_{a_h})$, where $0 \le h \le k$ are all up edges and $(x_a, x_{a_{h+1}}), \ldots, (x_a, x_{a_k})$ are all down edges.

Now for a set of indices $\mathcal{S}$ we define the function

$$f_{\mathcal{S}}(j_a \ : \ a \in \mathcal{S}) = \prod_{a \in \mathcal{S}} v_a(j_a) \prod_{(x_a, x_b) \in E(\mathcal{S})} \frac{1}{j_b^{\gamma} j_a^{1-\gamma}} \mathbf{1}_{\{j_a > j_b\}},$$

where $E(\mathcal{S})$ denotes the set of directed edges that is induced by $\mathcal{S}$. Using this, we write

$$f_{\overrightarrow{\mathcal{S}}}(i) =$$

$$\sum_{j_{a'} \ : \ a' \in \mathcal{P}(a) \setminus \{a_1, \ldots, a_k\}} \ \sum_{j_{a_1}, \ldots, j_{a_k}} f_{\mathcal{P}(a)}(j_{a'} \ : \ a' \in \mathcal{P}(a)) \sum_{j_a > j_{a_1} \vee \cdots \vee j_{a_h}}^{j_{a_{h+1}} \wedge \cdots \wedge j_{a_k} \wedge t} f_a(j_a) \prod_{i=1}^{h} \frac{1}{j_{a_i}^{\gamma} j_a^{1-\gamma}} \prod_{i=h+1}^{k} \frac{1}{j_a^{\gamma} j_{a_i}^{1-\gamma}}, \tag{30}$$

where f_a is the weight function of $\overrightarrow{\mathcal{T}}(a)$. Note that the index i in the LHS defines the index a in the RHS.

We will take an upper bound for each case of the definition of T with the use of Lemma 16. In particular, using Lemma 16 we will obtain an upper bound on

$$\sum_{j_a > j_{a_1} \vee \cdots \vee j_{a_h}}^{j_{a_{h+1}} \wedge \cdots \wedge j_{a_k} \wedge t} f_a(j_a) \prod_{i=1}^{h} \frac{1}{j_{a_i}^{\gamma} j_a^{1-\gamma}} \prod_{i=h+1}^{k} \frac{1}{j_a^{\gamma} j_{a_i}^{1-\gamma}}. \tag{31}$$

(If $h = 0$ then we let $j_{a_1} \vee \cdots \vee j_{a_h} = 0$, and if $h = k$, then we let $j_{a_{h+1}} \wedge \cdots \wedge j_{a_k} \wedge t = t$.)

Applying Lemma 16 to $f_a(j_a)$, we obtain

$$f_a(j_a) \lesssim \frac{1}{\omega^{\ell(a)}} \frac{(\log j_a)^{\rho(a)}}{j_a^{y_a}}, \tag{32}$$

where y_a is as in Lemma 16.

Assume first that $h > 0$. Then with Δ as above, we have

$$\begin{aligned}
&\sum_{j_a = j_{a_1} \vee \cdots \vee j_{a_h}}^{j_{a_{h+1}} \wedge \cdots \wedge j_{a_k} \wedge t} \prod_{i=1}^{h} \frac{1}{j_{a_i}^{\gamma} j_a^{1-\gamma}} \prod_{i=h+1}^{k} \frac{1}{j_a^{\gamma} j_{a_i}^{1-\gamma}} f_a(j_a) \lesssim \\
&\frac{1}{\omega^{\ell(a)}} \sum_{j_a = j_{a_1} \vee \cdots \vee j_{a_h}}^{j_{a_{h+1}} \wedge \cdots \wedge j_{a_k} \wedge t} \frac{1}{j_{a_1}^{\gamma} j_a^{1-\gamma}} \cdots \frac{1}{j_{a_h}^{\gamma} j_a^{1-\gamma}} \frac{1}{j_a^{\gamma} j_{a_{h+1}}^{1-\gamma}} \cdots \frac{1}{j_a^{\gamma} j_{a_k}^{1-\gamma}} \frac{(\log j_a)^{\rho(a)}}{j^{y_a}} \\
&\quad \lesssim \frac{1}{\omega^{\ell(a)}} \frac{1}{j_{a_1}^{\gamma}} \cdots \frac{1}{j_{a_h}^{\gamma}} \frac{1}{j_{a_{h+1}}^{1-\gamma}} \cdots \frac{1}{j_{a_k}^{1-\gamma}} \int_{j_a = j_{a_1} \vee \cdots \vee j_{a_h}}^{j_{a_{h+1}} \wedge \cdots \wedge j_{a_k} \wedge t} (\log z)^{\rho(a)} z^{-(k-h)\gamma - h(1-\gamma) - y_a} \, \mathrm{d}z \\
&\quad \le \frac{1}{\omega^{\ell(a)}} \frac{1}{j_{a_1}^{\gamma}} \cdots \frac{1}{j_{a_h}^{\gamma}} \frac{1}{j_{a_{h+1}}^{1-\gamma}} \cdots \frac{1}{j_{a_k}^{1-\gamma}} \int_{j_{a_\Delta}}^{t} (\log z)^{\rho(a)} z^{-(k-h)\gamma - h(1-\gamma) - y_a} \, \mathrm{d}z.
\end{aligned} \tag{33}$$

But by Lemma 17, we have

$$\begin{aligned}
\frac{1}{j_{a_\Delta}^{\gamma}} \int_{j_{a_\Delta}}^{t} (\log z)^{\rho(a)} z^{-(k-h)\gamma - h(1-\gamma) - y_a} &\lesssim \frac{1}{j_{a_\Delta}^{\gamma}} \frac{(\log j_{a_\Delta})^{\rho(a)}}{j_{a_\Delta}^{-1+(k-h)\gamma + h(1-\gamma) + y_a}} \\
&= \frac{(\log j_{a_\Delta})^{\rho(a)}}{j_{a_\Delta}^{(k-h)\gamma + (h-1)(1-\gamma) + y_a}} \le \frac{(\log j_{a_\Delta})^{\rho(a)}}{j_{a_\Delta}^{y_a}}.
\end{aligned}$$

Thereby, (33) becomes

$$\sum_{j_a = j_{a_1} \vee \cdots \vee j_{a_h}}^{j_{a_{h+1}} \wedge \cdots \wedge j_{a_k} \wedge t} \prod_{i=1}^{h} \frac{1}{j_a^{\gamma} j_{a_i}^{1-\gamma}} \prod_{i=h+1}^{k} \frac{1}{j_{a_i}^{\gamma} j_a^{1-\gamma}} f_a(j_a) \lesssim \frac{1}{\omega^{\ell(a)}} \frac{1}{j_{a_1}^{\gamma}} \cdots \frac{(\log j_{a_\Delta})^{\rho(a)}}{j_{a_\Delta}^{y_a}} \frac{1}{j_{a_{\Delta+1}}^{\gamma}} \cdots \frac{1}{j_{a_h}^{\gamma}} \frac{1}{j_{a_{h+1}}^{1-\gamma}} \cdots \frac{1}{j_{a_k}^{1-\gamma}}. \tag{34}$$

Now, if $h = 0$, then (31) yields

$$\begin{aligned}
\sum_{j_a=1}^{j_{a_1} \wedge \cdots \wedge j_{a_k}} \frac{1}{j_a^{\gamma} j_{a_1}^{1-\gamma}} \cdots \frac{1}{j_a^{\gamma} j_{a_k}^{1-\gamma}} \frac{(\log j_a)^{\rho(a)}}{j_a^{y_a}} &\lesssim \frac{(\log j_{a_1})^{\rho(a)}}{j_{a_1}^{1-\gamma}} \cdots \frac{(\log j_{a_k})^{\rho(a)}}{j_{a_k}^{1-\gamma}} \int_1^{j_{a_k}} z^{-k\gamma - y_a} \, \mathrm{d}z \\
&\lesssim \begin{cases}
\frac{(\log j_{a_1})^{\rho(a)}}{j_{a_1}^{1-\gamma}} \cdots \frac{(\log j_{a_k})^{\rho(a)}}{j_{a_k}^{1-\gamma}} j_{a_k}^{1-k\gamma-y_a} & \text{if } 1 - k\gamma > y_a \\
\frac{(\log j_{a_1})^{\rho(a)+1}}{j_{a_1}^{1-\gamma}} \cdots \frac{(\log j_{a_k})^{\rho(a)+1}}{j_{a_k}^{1-\gamma}} & \text{if } 1 - k\gamma = y_a \\
\frac{(\log j_{a_1})^{\rho(a)+1}}{j_{a_1}^{1-\gamma}} \cdots \frac{(\log j_{a_k})^{\rho(a)+1}}{j_{a_k}^{1-\gamma}} & \text{if } 1 - k\gamma < y_a
\end{cases}
\end{aligned} \tag{35}$$

In the first case, we have $y_a < 1 - \gamma$, since $y_a < 1 - k\gamma$. Thus the last factor is

$$\frac{1}{j_{a_k}^{(k-1)\gamma + y_a}} \le \frac{1}{j_{a_k}^{y_a}} = \frac{1}{j_{a_k}^{(1-\gamma) \wedge y_a}}.$$

Hence, in any case (35) is bounded by

$$\sum_{j_a=1}^{j_{a_1}\wedge\cdots\wedge j_{a_k}} \frac{1}{j_a^{\gamma} j_{a_1}^{1-\gamma}} \cdots \frac{1}{j_a^{\gamma} j_{a_k}^{1-\gamma}} \frac{(\log j_a)^{\rho(a)}}{j_a^{y_a}} \lesssim \prod_{i=1}^{k} \frac{(\log j_{a_i})^{\rho(a)+1}}{j_{a_i}^{(1-\gamma)\wedge y_a}}. \tag{36}$$

Setting

$$\hat{v}_{a_i}(j_{a_i}) := \begin{cases} \frac{1}{j_{a_i}^{\gamma}} & \text{if } i \le h \text{ and } i \neq \Delta, \\ (\log j_{a_i})^{\rho(a)} \cdot \frac{1}{j_{a_i}^{y_a}}, & \text{if } i = \Delta\ , \\ (\log j_{a_i})^{\rho(a)+1} \cdot \frac{1}{j_{a_i}^{(1-\gamma)\wedge y_a}} & \text{if } i > h, \end{cases}$$

now (34) and (36) yield

$$\sum_{j_a > j_{a_1}\vee\cdots\vee j_{a_h}}^{j_{a_{h+1}}\wedge\cdots\wedge j_{a_k}\wedge t} \prod_{i=1}^{h} \frac{1}{j_a^{\gamma} j_{a_i}^{1-\gamma}} \prod_{i=h+1}^{k} \frac{1}{j_{a_i}^{\gamma} j_a^{1-\gamma}} f_a(j_a) \lesssim \ \tfrac{1}{\omega^{\ell(a)}} \textstyle\prod_{i=1}^{k} \hat{v}_{a_i}(j_{a_i}). \tag{37}$$

So, substituting the bound of (37) into (30) we obtain

$$\begin{aligned} f_{\overrightarrow{\mathcal{S}}}(i) &\lesssim \frac{1}{\omega^{\ell(a)}} \sum_{j_{a'}\ :\ a'\in\mathcal{P}(a)\setminus\{a_1,\ldots,a_k\}} \sum_{j_{a_1},\ldots,j_{a_k}} f_{\mathcal{P}(a)}(j_{a'}\ :\ a'\in\mathcal{P}(a)) \prod_{i=1}^{k} \hat{v}_{a_i}(j_{a_i}) \\ &= \frac{1}{\omega^{\ell(a)}} \sum_{j_{a'}\ :\ a'\in\mathcal{P}(a)\setminus\{a_1,\ldots,a_k\}} \sum_{j_{a_1},\ldots,j_{a_k}} \prod_{a'\in\mathcal{P}(a)\setminus\{a_1,\ldots,a_k\}} v_{a'}(j_{a'})\times \\ &\qquad \prod_{i=1}^{k} v_{a_i}(j_{a_i})\hat{v}_{a_i}(j_{a_i}) \prod_{(x_{a'},x_{b'})\in E(\mathcal{P}(a))} \frac{1}{j_{b'}^{\gamma} j_{a'}^{1-\gamma}} \mathbf{1}_{\{j_{a'}>j_{b'}\}} \\ &= \frac{1}{\omega^{\ell(a)}} f_{T\overrightarrow{\mathcal{S}}_i}(i). \end{aligned} \tag{38}$$

Note that the upper bounds in Lemma 15 imply that $\prod_{i=1}^{k} \hat{v}_{a_i}(j_{a_i})$ is the bound we would get if x_a is replicated k times into $x_{a^{(1)}},\ldots,x_{a^{(k)}}$ and $x_{a^{(i)}}$ is attached to x_{a_i} through an up edge if $i \le h$ or through an down edge if $i > h$. Thereafter, $x_{a^{(\Delta)}}$ as well as $x_{a^{(i)}}$, for $i > h$, each becomes the root of a copy of $\overrightarrow{\mathcal{T}}(a)$, whereas for the remaining is, the vertices $x_{a^{(i)}}$ become leaves with valuation functions that are equal to $1/t^{\gamma}$. Note that the latter is ωp_0 - so essentially these become original leaves. □

Starting with the original witness structure $\overrightarrow{\mathcal{S}}$, we get a sequence of structures $T\overrightarrow{\mathcal{S}}, T^{(2)}\overrightarrow{\mathcal{S}}, \ldots, T^{(j)}\overrightarrow{\mathcal{S}}$ by applying the transformation T in the following way: If $T^{(j-1)}\overrightarrow{\mathcal{S}}$ is a tree, we are done; otherwise choose a vertex x_a of $T^{(j-1)}\overrightarrow{\mathcal{S}}$ such that x_a is on a cycle and has maximum depth among such vertices in $T^{(j-1)}\overrightarrow{\mathcal{S}}$. Now apply to x_a and its parents the transformation T, as appropriate, to get $T^{(j)}\overrightarrow{\mathcal{S}}$. Note that $T^{(j)}\overrightarrow{\mathcal{S}}$ has at least one less vertex that lies on a cycle. In general, the number of vertices lying on cycles reduces by at least one each time we apply the transformation. Hence, there exists a $j_0 \ge 0$ such that $T^{(j_0)}\overrightarrow{\mathcal{S}}$ is a generalised witness tree. Moreover, the depth of this tree is no more than the depth of $\overrightarrow{\mathcal{S}}$.

If $x_{a_1}, \dots, x_{a_{j_0}}$ denote the vertices that were split in each transformation, the repeated application of Lemma 20 yields

$$f_{\overrightarrow{\mathcal{S}}}(i) \lesssim \frac{1}{\omega^{\sum_{j=1}^{j_0} \ell(a_j)}} f_{T^{(j_0)}\overrightarrow{\mathcal{S}}}(i),$$

where $\ell(a_j)$ is the number of original leaves of $\overrightarrow{\mathcal{T}}(a_j)$ in $T^{(j-1)}\overrightarrow{\mathcal{S}}$.

Since $T^{(j_0)}\overrightarrow{\mathcal{S}}$ is a generalised witness tree, we can apply Lemma 16 and deduce that for some $\rho \geq 0$ and with ℓ being the number of original leaves in $T^{(j_0)}\overrightarrow{\mathcal{S}}$ we have

$$f_{\overrightarrow{\mathcal{S}}}(i) \lesssim \frac{1}{\omega^{\ell+\sum_{j=1}^{j_0} \ell(a_j)}} \frac{(\log i)^\rho}{i^y},$$

and either y can be expressed as $A\gamma + B(1-\gamma)$ where A, B are non-negative integers such that $A + B \geq r$ and $B > 0$ or

$$y = \sum_{a' \in T^{(j_0)}\overrightarrow{\mathcal{S}}} e_{a'} + \ell\gamma.$$

Note that $\ell + \sum_{j=1}^{j_0} \ell(a_j) \geq r$.

By (29) the latter is equal to $L_{j_0}\gamma$, where L_{j_0} is the number of leaves of $\hat{T}^{(j_0)}\overrightarrow{\mathcal{S}}$. But $\hat{T}^{(j_0)}\overrightarrow{\mathcal{S}}$ is an r-ary tree of depth d_0 and therefore $L_{j_0} \geq d_0$. Thus, $y \geq \gamma d_0 > 1$. Hence (22) also holds in this case, implying that the right-hand side of (27) is $o(1)$. Now, if the depth of this particular witness structure is less than d_0, then by the same principles as in the case of trees we obtain an upper bound on the number of vertices that can be infected through a witness structure isomorphic to $\overrightarrow{\mathcal{S}}$ which is $o(t^{1-\gamma})$.

This completes the proof of Theorem 2**(ii)** and 2**(iii)**.

6 Critical case

*Proof of Theorem 3**(i)**.* Let G be a realisation of $\mathrm{PA}_t(m, \delta)$. Let $\mathcal{T}(G, d)$ be the set of trees in G which have depth d and for which every internal vertex has r children. For a tree $T \in \mathcal{T}(G, d)$ let A_T be the event that all leaves in T are initially infected. Note that this is an event on the product space of initial infection, where every vertex is infected independently with probability p. Also, note that this is a *non-decreasing* event: if we infect more vertices, then A_T will not stop holding.

We wish to show $\mathbb{P}\left(\bigcap_{T\in\mathcal{T}(G,d)} A_T^c\right) > 0$. To this end, we apply the FKG inequality (see for example Theorem 6.3.2 in [27]):

$$\begin{aligned} \mathbb{P}\Big(\bigcap_{T\in\mathcal{T}(G,d)} A_T^c\Big) &\geq \prod_{T\in\mathcal{T}(G,d)} (1 - \mathbb{P}(A_T)) \\ &\geq \exp\left(-2 \sum_{T\in\mathcal{T}(G,d)} \mathbb{P}(A_T)\right), \end{aligned} \tag{39}$$

where the last inequality follows as $1 - x \geq e^{-2x}$ when x is small enough. In this case, it will be small enough provided that t is large, since $\mathbb{P}(A_T) = p^\ell$ where ℓ is the number of leaves in T and $p = o(1)$.

Let $\mathcal{T}(G,d,\ell) \subseteq \mathcal{T}(G,d)$ be those depth-d trees in G with ℓ leaves. We have

$$\sum_{T\in\mathcal{T}(G,d)} \mathbb{P}(A_T) = \sum_{\ell\geq d} \sum_{T\in\mathcal{T}(G,d,\ell)} p^\ell = \sum_{\ell\geq d} p^\ell \, |\mathcal{T}(G,d,\ell)|$$

Let C be some large constant, let $\sigma(d,\ell) = \{G \in \mathrm{PA}_t(m,\delta) : |\mathcal{T}(d,\ell)| \leq C\mathbb{E}[|\mathcal{T}(d,\ell)|]\}$ where $|\mathcal{T}(d,\ell)|$ is the random variable on $\mathrm{PA}_t(m,\delta)$ that counts the number of depth-d trees with ℓ leaves and each internal vertex having r children. Let $\sigma(d) = \bigcap_{\ell\geq d} \sigma(d,\ell)$. Then if $G \in \sigma(d)$

$$\sum_{T\in\mathcal{T}(G,d)} \mathbb{P}(A_T) \leq C \sum_{\ell\geq d} p^\ell \, \mathbb{E}[|\mathcal{T}(d,\ell)|] = O(1),$$

where the last equality follows from (22) when $d = d_0$, replacing ω in $p = p_c/\omega$ (which gave us $o(1)$) with $1/\lambda$.

Now

$$\mathbb{P}(|\mathcal{T}(\mathrm{PA}_t(m,\delta),d,\ell)| > C\mathbb{E}[|\mathcal{T}(d,\ell)|]) \leq \frac{1}{C}.$$

Hence $\mathbb{P}(\mathrm{PA}_t(m,\delta) \notin \sigma(d)) \leq r^{d+1}/C$ since $d \leq \ell \leq r^{d+1}$. Of course, we choose $C > r^{d+1}$.

Getting back to (39), we see that with probability at least $1 - r^{d+1}/C$,

$$\mathbb{P}\Big(\bigcap_{T\in\mathcal{T}(G,d)} A_T^c | G \in \sigma(d)\Big) \geq \exp\left(-2 \sum_{T\in\mathcal{T}(G,d)} \mathbb{P}(A_T)\right) = \Omega(1).$$

Consequently, with probability at least $p_1 > 0$, there is no witness tree of depth $d = d_0$, meaning no infection occurs in this round or thereafter.

The same argument applies to witness structures which are not trees. As per above, their expected number of occurrences is bounded from above by that of witness trees.

When $d < d_0$, the results of the previous section show that the expected number of infected vertices in round $d > 0$ is $o(t^{1-\gamma})$. Hence, the above analysis together with Markov's inequality yields $|\mathcal{I}_f|/|\mathcal{I}_0| < 1 + \varepsilon$, for $\varepsilon > 0$, with probability at least $p_1 > 0$, for any t large enough. □

*Proof of Theorem 3**(ii)**.* We wish to show there is a full outbreak. This will happen if, for some $k \geq 1$, the first k vertices $[k]$ get infected, and additionally, no vertex has more than one self-loop. We will show that this happens with some probability bounded away from zero.

Consider vertex 1. The argument is along the following lines: The expected degree of vertex 1 is about t^γ. Suppose that the actual degree of vertex 1 is roughly its expected degree. When the infection probability is $p = \lambda/t^\gamma$ where λ is a constant, then the probability of vertex 1 getting infected in round $\tau = 1$ is $\mathbb{P}(\mathrm{Bin}(t^\gamma, \lambda/t^\gamma) \geq r)$, which is asymptotically bounded away from 0.

For $\delta \geq 0$, we can use Lemma 5. Setting $h > 0$ to be a sufficiently small constant, we get $\mathbb{P}(S_1(t) < \mathbb{E}[S_1(t)]/K) < 1/e^h < 1$. Since $S_1(t) = D_1(t)$ and so $\mathbb{P}(D_1(t) \geq \epsilon t^\gamma) \geq \epsilon_1$ for some constants $\epsilon, \epsilon_1 > 0$. For $\delta < 0$ we apply Lemma 6 with $i = 1$ to get the same result.

Let $\mathcal{E}_i$ be the event that vertex i has at most one self-loop and let $\mathcal{E} = \bigcap_{i>1} \mathcal{E}_i$. Let $\mathcal{A}_\epsilon$ be the event $D_1(t) \geq \epsilon t^\gamma$. It is clear that $\mathbb{P}(\mathcal{A}_\epsilon \cap \mathcal{E}) \geq \mathbb{P}(\mathcal{A}_\epsilon)\mathbb{P}(\mathcal{E})$.

As per the previous sections, for $i > 1$, $\mathbb{P}(\mathcal{E}_i^c) = O(1/i^2)$ and so $\liminf_{t\to\infty} \mathbb{P}(\mathcal{E}) > 0$. Therefore, with some probability bounded away from zero, no vertex has more than one self loop, and vertex 1 is infected in round $\tau = 1$. Consequently, all vertices become infected eventually. □

7 Conclusions - open questions

This paper studies the evolution of a bootstrap percolation process on random graphs that have been generated through preferential attachment and generalise the classical Barabási-Albert model. For $r < m$, where $2m$ is the average degree, we determine a critical function $a_c(t)$ such that when the size $a(t)$ of the initial set "crosses" $a_c(t)$ the evolution of the bootstrap percolation process with activation threshold r changes abruptly from *almost* no evolution to full infection. The critical function satisfies $a_c(t) = o(t)$, which implies that a sublinear initial infection leads to full infection.

Our results are somewhat less tight for $r = 2$. It would be interesting to find out whether the sharpness of the threshold that we deduced for $r \geq 3$ also holds in this case. Also, the critical window itself for the case $r = 2$ has not been explored in the present work. Furthermore, it would be interesting to determine the number of rounds until the complete infection of all vertices in the supercritical case.

References

[1] M.A. Abdullah and N. Fountoulakis, Bootstrap percolation processes on preferential attachment graphs: from sparse contagion to pandemics, preprint available at `http://arxiv.org/abs/1404.4070`, 36 pages.

[2] R. Albert and A.-L. Barabási, Statistical mechanics of complex networks, *Reviews of Modern Physics*, 74:47–97, 2002.

[3] J. Adler and U. Lev. Bootstrap percolation: visualizations and applications. *Brazilian Journal of Physics*, 33(3):641–644, 2003.

[4] H. Amini. Bootstrap percolation and diffusion in random graphs with given vertex degrees. *Electronic Journal of Combinatorics, 17: R25*, 2010.

[5] H. Amini. Bootstrap percolation in living neural networks. *Journal of Statistical Physics*, 141:459–475, 2010.

[6] H. Amini and N. Fountoulakis. Bootstrap percolation in power-law random graphs, *Journal of Statistical Physics* 155: 72–92, 2014.

[7] J. Balogh and B. Bollobás. Bootstrap percolation on the hypercube. *Probability Theory and Related Fields*, 134(4):624–648, 2006.

[8] J. Balogh, B. Bollobás, H. Duminil-Copin, and R. Morris. The sharp threshold for bootstrap percolation in all dimensions. *Trans. Amer. Math. Soc.*, 36:2667 – 2701, 2012.

[9] J. Balogh, B. Bollobás, and R. Morris. Bootstrap percolation in three dimensions. *Annals of Probability*, 37:1329–1380, 2009.

[10] J. Balogh, Y. Peres, and G. Pete. Bootstrap percolation on infinite trees and non-amenable groups. *Combinatorics, Probability and Computing*, 15(5):715–730, 2006.

[11] J. Balogh and B. G. Pittel. Bootstrap percolation on the random regular graph. *Random Structures Algorithms*, 30(1-2):257–286, 2007.

[12] A.L. Barabási and R. Albert, Emergence of scaling in random networks, *Science* 286 (5439):509-?512, 1999.

[13] B. Bollobás and O. Riordan, The diameter of a scale-free random graph, *Combinatorica* 24:5–34, 2004.

[14] B. Bollobás, O. Riordan, J. Spencer and G. Tusnády, The degree sequence of a scale-free random graph process, *Random Structures and Algorithms* 18: 279-290, 2001.

[15] P.G. Buckley and D. Osthus, Popularity based random graph models leading to a scale-free degree sequence, *Discrete Math.* 282: 53–68, 2004.

[16] R. Cerf and F. Manzo. The threshold regime of finite volume bootstrap percolation. *Stochastic Processes and their Applications*, 101(1):69–82, 2002.

[17] J. Chalupa, P. L. Leath, and G. R. Reich. Bootstrap percolation on a Bethe lattice. *Journal of Physics C: Solid State Physics*, 12:L31–L35, 1979.

[18] C. Cooper, T. Radzik and Y. Siantos, A fast algorithm to find all high degree vertices in graphs with a power law degree sequence, In *Proceedings of 9th Workshop on Algorithms and Models for the Web Graph (WAW)*, pp. 165–178, 2012.

[19] S.N. Dorogovtsev, J.F.F. Mendes and A.N. Samukhin, Structure of growing networks with preferential linking. *Physical Review Letters* 85(21): 4633–4636, 2000.

[20] E. Drinea, M. Enachescu and M. Mitzenmacher, Variations on random graph models for the web. Technical report TR-06-01, Harvard University, Department of Computer Science, 2001, 3 pages.

[21] R. Ebrahimi, J. Gao, G. Ghasemiesfeh and G. Schoenenbeck, How complex contagions spread quickly in the preferential attachment model and other time-evolving networks, preprint available at `http://arxiv.org/pdf/1404.2668v1.pdf`, 23 pages.

[22] J.P. Eckmann and T. Tlusty. Remarks on bootstrap percolation in metric networks. *Journal of Physics A: Mathematical and Theoretical*, 42:205004, 2009.

[23] L. R. Fontes, R. H. Schonmann, and V. Sidoravicius. Stretched exponential fixation in stochastic Ising models at zero temperature. *Communications in Mathematical Physics*, 228:495–518, 2002.

[24] L. Fontes and R. Schonmann. Bootstrap percolation on homogeneous trees has 2 phase transitions. *Journal of Statistical Physics*, 132:839–861, 2008.

[25] R. van der Hofstad, *Random Graphs and Complex Networks*, 2013 (book available at `http://www.win.tue.nl/ rhofstad/NotesRGCN.pdf`).

[26] A. E. Holroyd. Sharp metastability threshold for two-dimensional bootstrap percolation. *Probability Theory and Related Fields*, 125(2):195–224, 2003.

[27] N. Alon and J. Spencer, *The probabilistic method*, John Wiley and Sons, 2008, xv+352 pages.

[28] S. Janson, T. Łuczak, T. Turova, and T. Vallier. Bootstrap percolation on the random graph $G_{n,p}$. *The Annals of Applied Probability*, 22(5):1989–2047, 2012.

[29] D. Kempe, J. Kleinberg and É. Tardos, Maximizing the spread of influence in a social network, In *Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discorvery and Data Mining*, pp. 137–146, 2003.

[30] D. Kempe, J. Kleinberg and É. Tardos, Influential nodes in a diffusion model for social networks, In *Proceedings of the 32nd International Colloquium on Languages, Automata and Programming (ICALP)*, pp. 1127–1138, 2005.

[31] J. Kleinberg, Cascading Behavior in Networks: Algorithmic and Economic Issues. In *Algorithmic Game Theory (N. Nisan, T. Roughgarden, E. Tardos, V. Vazirani, eds.)*, Cambridge University Press, 2007.

[32] R. Morris. Zero-temperature Glauber dynamics on $\mathbb{Z}^d$. *Probability Theory and Related Fields*, 149:417–434, 2009.

[33] S. Sabhapandit, D. Dhar, and P. Shukla. Hysteresis in the random-field Ising model and bootstrap percolation. *Physical Review Letters*, 88(19):197202, 2002.

[34] F. Sausset, C. Toninelli, G. Biroli, and G. Tarjus. Bootstrap percolation and kinetically constrained models on hyperbolic lattices. *Journal of Statistical Physics*, 138:411–430, 2010.

[35] H.A. Simon. On a class of skew distribution functions. *Biometrika*, 42: 425–440, 1955.

[36] C. Toninelli, G. Biroli, and D. S. Fisher. Jamming percolation and glass transitions in lattice models. *Physical Review Letters*, 96(3):035702, 2006.

[37] G.U. Yule. A mathematical theory of evolution, based on the conclusions of Dr. J.G. Willis F.R.S. *Phil. Trans. Roy. Soc. London, B,* 213: 21–87, 1925.

8 Appendix

8.1 Proofs

Proof of Lemma 5. Assume $h, c_t, A > 0$. Recall also that here we assume that $\delta \geq 0$. Let $Z_t = S_i(t)$.

$$\mathbb{P}\left(Z_t < A\right) = \mathbb{P}\left(e^{\frac{-hZ_t}{c_t}} > e^{\frac{-hA}{c_t}}\right).$$

$Z_t = Z_{t-1} + Y_t$. Then $Y_t \succeq X_t \sim \text{Bin}\left(m, \frac{Z_{t-1}}{mt(2+\delta/m)}\right)$.

$$\mathbb{E}\left[e^{\frac{-hX_t}{c_t}} \mid Z_{t-1}\right] = \left(1 - p + pe^{\frac{-h}{c_t}}\right)^m$$

where $p = \frac{Z_{t-1}}{mt(2+\delta/m)}$.

Using $e^{-x} \leq 1 - x + x^2$,

$$\begin{aligned}
\left(1 - p + pe^{\frac{-h}{c_t}}\right)^m &\leq \left(1 - p + p - p\frac{h}{c_t} + p\left(\frac{h}{c_t}\right)^2\right)^m \\
&= \left(1 - p\frac{h}{c_t}\left(1 - \frac{h}{c_t}\right)\right)^m \\
&\leq \exp\left(-\frac{mph}{c_t}\left(1 - \frac{h}{c_t}\right)\right) \\
&= \exp\left(-\frac{hZ_{t-1}}{c_t\,(2+\delta/m)\,t}\left(1 - \frac{h}{c_t}\right)\right)
\end{aligned}$$

Then

$$\mathbb{E}\left[e^{\frac{-hZ_{t-1}}{c_t}} e^{\frac{-hY_t}{c_t}} \mid Z_{t-1}\right] \leq \exp\left(-\frac{hZ_{t-1}}{c_t\,(2+\delta/m)\,t}\left(1 - \frac{h}{c_t}\right) - \frac{hZ_{t-1}}{c_t}\right).$$

Taking expectations on both sides,

$$\mathbb{E}\left[\exp\left(\frac{-hZ_t}{c_t}\right)\right] \leq \mathbb{E}\left[\exp\left(-\frac{hZ_{t-1}}{c_t}\left(1 + \frac{1 - h/c_t}{(2+\delta/m)t}\right)\right)\right].$$

Let $c_i = 1$ and $c_t = \left(1 + \frac{\gamma}{t}\right) c_{t-1} = \left(1 + \frac{1}{(2+\delta/m)t}\right) c_{t-1}$ for $t > i$, and note that $c_t \sim \left(\frac{t}{i}\right)^\gamma$. We have,

$$\begin{aligned}
\mathbb{E}\left[\exp\left(\frac{-hZ_t}{c_t}\right)\right] &\leq \mathbb{E}\left[\exp\left(-\frac{hZ_{t-1}}{c_{t-1}}\frac{1 + \frac{1-h/c_t}{(2+\delta/m)t}}{1 + \frac{1}{(2+\delta/m)t}}\right)\right] \\
&\leq \mathbb{E}\left[\exp\left(-\frac{hZ_{t-1}}{c_{t-1}}\left(1 - \frac{h}{(2+\delta/m)c_t t}\right)\right)\right].
\end{aligned}$$

Iterating,

$$\begin{aligned}
\mathbb{E}\left[\exp\left(\frac{-hZ_t}{c_t}\right)\right] &\leq \mathbb{E}\left[\exp\left(-\frac{hZ_{t-1}}{c_{t-1}}\left(1 - \frac{h\gamma}{c_t t}\right)\right)\right] \\
&\leq \mathbb{E}\left[\exp\left(-\frac{hZ_{t-2}}{c_{t-2}}\left(1 - \frac{h\gamma}{c_t t}\right)\left(1 - \frac{h\gamma}{c_{t-1}(t-1)}\right)\right)\right] \\
&\vdots \\
&\leq \mathbb{E}\left[\exp\left(-\frac{hZ_i}{c_i}\prod_{j=i}^{t}\left(1 - \frac{h\gamma}{c_j j}\right)\right)\right] \\
&= \mathbb{E}\left[\exp\left(-2hmi\prod_{j=i}^{t}\left(1 - \frac{h\gamma}{c_j j}\right)\right)\right].
\end{aligned}$$

$$\begin{aligned}\prod_{j=i}^{t}\left(1-\frac{h\gamma}{c_j j}\right) &\geq 1-h\gamma\sum_{j=i}^{t}\frac{1}{jc_j}\\ &= 1-O\left(h\sum_{j=i}^{t}\frac{1}{j\left(\frac{j}{i}\right)^{\gamma}}\right)\\ &= 1-O\left(hi^{\gamma}i^{-\gamma}\right)\\ &= 1-O\left(h\right).\end{aligned}$$

So

$$\mathbb{E}\left[\exp\left(\frac{-hZ_t}{c_t}\right)\right]\leq \mathbb{E}\left[\exp\left(-2hmi\left(1-O\left(h\right)\right)\right)\right]=\exp\left(-2hmi\left(1-O\left(h\right)\right)\right).$$

Hence using Markov's inequality,

$$\mathbb{P}\left(e^{\frac{-hZ_t}{c_t}}>e^{\frac{-hA}{c_t}}\right)\leq\frac{e^{-2hmi(1-O(h))}}{e^{\frac{-hA}{c_t}}}.$$

Recalling that $ic_t\sim i\left(\frac{t}{i}\right)^{\gamma}=t^{\gamma}i^{1-\gamma}$ and $\mathbb{E}[S_i(t)]\geq C_\ell t^{\gamma}i^{1-\gamma}$ (cf. Proposition 4), choose a sufficiently large constant constant K such that $\mathbb{E}[S_i(t)]/K<C_\ell ic_t/\sqrt{K}$ and let $A=C_\ell ic_t/\sqrt{K}$. Then,

$$\begin{aligned}\mathbb{P}\left(S_i(t)\leq\frac{1}{K}\mathbb{E}[S_i(t)]\right) &\leq \mathbb{P}\left(e^{\frac{-hZ_t}{c_t}}>e^{\frac{-hA}{c_t}}\right)\\ &\leq \exp\left(-2hmi\left(1-O\left(h\right)\right)+hiC_\ell/\sqrt{K}\right)\\ &= \exp\left(-hi\left(2m-O(h)-C_\ell/\sqrt{K}\right)\right)\\ &\leq \exp\left(-hi\right),\end{aligned}$$

where the last inequality follows if $K>K_0$ where $K_0>0$ is a sufficiently large constant that need only depend on m,δ, and if h is small enough. □

Proof of Lemma 15. The first sum is bounded from above by an integral:

$$\sum_{j'=j+1}^{t}f_{a'}(j')\frac{1}{j^{\gamma}j'^{1-\gamma}}\leq\frac{1}{j^{\gamma}}\int_{j}^{t}v_{a'}(x)\frac{1}{x^{1-\gamma}}dx.$$

Assume that $x_{a'}$ is a contraction leaf. In this case, the above integral becomes

$$\int_{j}^{t}v_{a'}(x)\frac{1}{x^{1-\gamma}}dx=\int_{j}^{t}\frac{(\log x)^{\rho_{a'}}}{x^{1-\gamma+e_{a'}}}dx.$$

The value of this integral now depends on the sign of $-\gamma+e_{a'}$. Recall that $e_{a'}>0$ and it satisfies Property (A). Assume that $e_{a'}=A\gamma+B(1-\gamma)$. Since $1-\gamma\geq\gamma$, the latter is $e_{a'}\geq(A+B)\gamma\geq r\gamma$. Since $r>1$, it follows that $-\gamma+e_{a'}>(r-1)\gamma>0$. Hence, by Lemma 17

$$\int_{j}^{t}v_{a'}(x)\frac{1}{x^{1-\gamma}}dx\lesssim(\log j)^{\rho_{a'}}j^{\gamma-e_{a'}},$$

and therefore

$$\sum_{j'=j+1}^{t} f_{a'}(j')\frac{1}{j^{\gamma}j'^{1-\gamma}} \lesssim (\log j)^{\rho_{a'}} j^{-e_{a'}}.$$

Assume now that $x_{a'}$ is an original leaf. In this case,

$$\sum_{j'=j+1}^{t} f_{a'}(j')\frac{1}{j^{\gamma}j'^{1-\gamma}} \lesssim \frac{1}{j^{\gamma}}\int_{j}^{t} v_{a'}(x)\frac{1}{x^{1-\gamma}}dx = p\frac{1}{j^{\gamma}}\int_{j}^{t}\frac{1}{x^{1-\gamma}}dx \lesssim p\frac{1}{j^{\gamma}}t^{\gamma} = \frac{1}{\omega}\ \frac{1}{j^{\gamma}}.$$

Consider now the second sum in the statement of the Lemma. If $x_{a'}$ is an original leaf, then

$$\sum_{j'=1}^{j-1} f_{a'}(j')\frac{1}{j'^{\gamma}j^{1-\gamma}} \lesssim \frac{1}{j^{1-\gamma}}\int_{1}^{j} v_{a'}(x)\frac{1}{x^{\gamma}}dx = p\ \frac{1}{j^{1-\gamma}}\int_{1}^{j}\frac{1}{x^{\gamma}}dx \lesssim \frac{1}{\omega t^{\gamma}}\ \frac{1}{j^{1-\gamma}}\ j^{1-\gamma} \overset{j\leq t}{\leq} \frac{1}{\omega}\ \frac{1}{j^{\gamma}}.$$

If $x_{a'}$ is a contraction leaf, then we have

$$\sum_{j'=1}^{j-1} f_{a'}(j')\frac{1}{j'^{\gamma}j^{1-\gamma}} \lesssim \frac{1}{j^{1-\gamma}}\int_{1}^{j} v_{a'}(x)\frac{1}{x^{\gamma}}dx = \frac{1}{j^{1-\gamma}}\int_{1}^{j}\frac{(\log x)^{\rho_{a'}}}{x^{\gamma+e_{a'}}}dx.$$

If $1-\gamma > e_{a'}$, then the above becomes

$$\sum_{j'=1}^{j-1} f_{a'}(j')\frac{1}{j'^{\gamma}j^{1-\gamma}} \lesssim \frac{1}{j^{1-\gamma}}(\log j)^{\rho_{a'}} j^{1-\gamma-e_{a'}} = (\log j)^{\rho_{a'}} j^{-e_{a'}}.$$

If $1-\gamma \leq e_{a'}$, then we will get

$$\sum_{j'=1}^{j-1} f_{a'}(j')\frac{1}{j'^{\gamma}j^{1-\gamma}} \lesssim \frac{(\log j)^{\rho_{a'}+1}}{j^{1-\gamma}}.$$

□

8.2 Proof of Lemma 17

In this section, we prove the following lemma 17, which has been useful for our calculations. We restate it here for completeness.

Lemma 21. *Let $k \geq 0$ be an integer, let $\alpha > 0$ be a real number and let*

$$I_{k,\alpha}(j) := \int_{j}^{t} (\log x)^{k} x^{-1-\alpha}\,\mathrm{d}x.$$

Then uniformly for $j \geq 1$ we have

$$I_{k,\alpha}(j) \lesssim \frac{(\log j \vee 1)^{k}}{j^{\alpha}}.$$

Proof. Let $v = (\log x)^k$, meaning $\frac{dv}{dx} = \frac{k(\log x)^{k-1}}{x}$. Let $\frac{du}{dx} = x^{-1-\alpha}$, meaning $u = -\frac{x^{-\alpha}}{\alpha}$. Integration by parts gives

$$\begin{aligned}
I_{k,\alpha}(j) &= \left[-(\log x)^k \frac{x^{-\alpha}}{\alpha}\right]_j^t + \frac{k}{\alpha}\int_j^t (\log x)^{k-1} x^{-1-\alpha}\,\mathrm{d}x \\
&\leq \frac{1}{\alpha}\frac{(\log j \vee 1)^k}{j^\alpha} + \frac{k}{\alpha} I_{k-1,\alpha} \\
&\leq \frac{1}{\alpha}\frac{(\log j \vee 1)^k}{j^\alpha} + \frac{k}{\alpha}\left[\frac{1}{\alpha}\frac{(\log j \vee 1)^{k-1}}{j^\alpha} + \frac{k-1}{\alpha} I_{k-2,\alpha}\right] \\
&= \frac{1}{\alpha}\frac{(\log j \vee 1)^k}{j^\alpha} + \frac{k}{\alpha^2}\frac{(\log j \vee 1)^{k-1}}{j^\alpha} + \frac{k(k-1)}{\alpha^2} I_{k-2,\alpha} \\
&\leq \frac{1}{\alpha}\frac{(\log j \vee 1)^k}{j^\alpha} + \frac{k}{\alpha^2}\frac{(\log j \vee 1)^{k-1}}{j^\alpha} + \frac{k(k-1)}{\alpha^3}\frac{(\log j \vee 1)^{k-2}}{j^\alpha} + \\
&\quad \ldots + \frac{k(k-1)\ldots 2}{\alpha^k}\,\frac{\log j \vee 1}{j^\alpha} + \frac{k(k-1)\ldots 2}{\alpha^k}\, I_{0,\alpha}(j).
\end{aligned}$$

Now

$$I_{0,\alpha}(j) = \int_j^t x^{-1-\alpha}\,\mathrm{d}x \leq \frac{1}{\alpha j^\alpha}.$$

Thus, we get

$$\begin{aligned}
I_{k,\alpha}(j) &\leq \frac{1}{\alpha}\frac{(\log j \vee 1)^k}{j^\alpha} + \frac{k}{\alpha^2}\frac{(\log j \vee 1)^{k-1}}{j^\alpha} + \ldots + \frac{k(k-1)\ldots 2}{\alpha^k}\frac{\log j \vee 1}{j^\alpha} + \frac{k!}{\alpha^{k+1}}\frac{1}{j^\alpha} \\
&\lesssim \frac{(\log j \vee 1)^k}{j^\alpha},
\end{aligned}$$

uniformly over all $j \geq 1$. □