Bootstrap Percolation in Inhomogeneous Random Graphs

H. Amini[*] N. Fountoulakis[†] K. Panagiotou[‡]

September 2021

Abstract

A bootstrap percolation process on a graph with n vertices is an "infection" process evolving in rounds. Let $r \geq 2$ be fixed. Initially, there is a subset of infected vertices. In each subsequent round every uninfected vertex that has at least r infected neighbours becomes infected as well and remains so forever.

We consider this process in the case where the underlying graph is an inhomogeneous random graph whose kernel is of rank one. Assuming that initially every vertex is infected independently with probability $p \in (0, 1]$, we provide a law of large numbers for the size of the set of vertices that are infected by the end of the process. Moreover, we investigate the case $p = p(n) = o(1)$ and we focus on the important case of inhomogeneous random graphs exhibiting a power-law degree distribution with exponent $\beta \in (2, 3)$. The first two authors have shown in this setting the existence of a critical $p_c = o(1)$ such that with high probability if $p = o(p_c)$, then the process does not evolve at all, whereas if $p = \omega(p_c)$, then the final set of infected vertices has size $\Omega(n)$. In this work we determine the asymptotic fraction of vertices that will be eventually infected and show that is also satifies a law of large numbers.

1 Introduction

A *bootstrap percolation process* with *activation threshold* an integer $r \geq 2$ on a graph $G = G(V, E)$ is a deterministic process evolving in rounds. Every vertex has two states: it is either *infected* or *uninfected* (sometimes also referred to as *active* or *inactive*, respectively). Initially, there is a subset $\mathcal{A}_0 \subseteq V$ that consists of infected vertices, whereas every other vertex is uninfected. Subsequently, in each round, if an uninfected vertex has at least r of its neighbours infected, then it also becomes infected and remains so forever. The process stops when no more vertices become infected, and we denote the final infected set by $\mathcal{A}_f$.

The bootstrap percolation process was introduced by Chalupa, Leath and Reich [16] in 1979 in the context of magnetic disordered systems. This process (as well as numerous variations of it) has been used as a model to describe several complex phenomena in diverse areas, from jamming transitions [35] and magnetic systems [31] to neuronal activity [4, 34] and spread of defaults in banking systems [5, 7]. Bootstrap percolation has also connections to the dynamics of the Ising model at zero temperature [23, 29]. A short survey regarding applications can be found in [1].

[*]Department of Risk Management and Insurance, Georgia State University, Atlanta, USA.
[†]School of Mathematics, University of Birmingham, United Kingdom.
[‡]Mathematical Institute, University of Munich, Germany.

Several qualitative characteristics of bootstrap percolation, and in particular the dependence of the initial set $\mathcal{A}_0$ on the final infected set $\mathcal{A}_f$, have been studied on a variety of graphs, such as trees [12, 22], grids [15, 24, 10], lattices on the hyperbolic plane [32], hypercubes [9], as well as on many models of random graphs [3, 13, 26]. In particular, consider the case $r = 2$ and G is the two-dimensional grid with $V = [n]^2 = \{1, \dots, n\}^2$ (i.e., a vertex becomes infected if at least two of its neighbours are already infected). Then, for $\mathcal{A}_0 \subseteq V$ whose elements are chosen independently at random, each with probability $p = p(n)$, the following sharp threshold was determined by Holroyd [24]. The probability $I(n,p)$ that the entire square is eventually infected satisfies $I(n,p) \to 1$ if $\liminf_{n\to\infty} p(n) \log n > \pi^2/18$, and $I(n,p) \to 0$ if $\limsup_{n\to\infty} p(n) \log n < \pi^2/18$. A generalization of this result to the higher dimensional case was proved by Balogh, Bollobàs and Morris [11] (when G is the 3-dimensional grid on $[n]^3$ and $r = 3$) and Balogh, Bollobàs, Duminil-Copin and Morris [10] (in general).

In this paper we study the bootstrap percolation process on inhomogeneous random graphs. Informally, these random graphs are defined through a sequence of weights that are assigned to the vertices which, in turn, determine the probability that two vertices are adjacent. More specifically, we are interested in the case where this probability is proportional to the product of the weights of these vertices. In particular, pairs of vertices where are at least one of them has a high weight are more likely to appear as edges.

A special case of our setting is the $G(n,p)$ model of random graphs, where every edge on a set of n vertices is present independently with probability p. Here every vertex has the same weight. Janson, Łuczak, Turova and Vallier [26] presented a complete analysis of the bootstrap percolation process for various ranges of p. We focus on their findings regarding the range where $p = d/n$ and $d > 0$ is fixed, as they are most relevant for the setting studied in this paper. In [26] a law of large numbers for $|\mathcal{A}_f|$ was shown when the density of $\mathcal{A}_0$ is positive, that is, when $|\mathcal{A}_0| = \theta n$, where $\theta \in (0,1)$. It was further shown that when $|\mathcal{A}_0| = o(n)$, then typically no evolution occurs. In other words, the density of the initially infected vertices must be positive in order for the density of the finally infected vertices to increase. This fact had been pointed out earlier by Balogh and Bollobás, cf. [13]. A similar behavior was observed in the case of random regular graphs [13], as well as in random graphs with given vertex degrees. These were studied by the first author in [3], when the sum of the square of degrees scales linearly with n. As we shall see shortly, the random graph model we consider here is essentially a random graph with given *expected* degrees. Finally, more recently the bootstrap process was considered in another type of inhomogeneous random graph which is the stochastic block model [36].

The main result of this paper provides a law of large numbers for $|\mathcal{A}_f|$ given $|\mathcal{A}_0|$ for weight sequences that satisfy fairly general and natural regularity conditions. We then consider weight sequences that follow a power law distribution, i.e., the proportion of vertices with weight w scales like $w^{-\beta}$ for some $\beta > 2$, with a particular focus on the case where $\beta \in (2,3)$. The parameter β is called the *exponent* of the power law. Note that although in this case the weight sequence has a bounded average weight, its second moment is growing with the number of vertices. Power-laws emerge in several contexts such as ranging from ecology and economics to social networks (see e.g. the survey of Mitzenmacher [28]). Already during the late 19th century Pareto observed a power law in the disrtibution of the wealth within populations [30]. In a completely different context, Lotka [27] in 1926 observed a power law distribution on the frequencies of scientists that are cited a certain number of times in Chemical Abstracts during the period 1910-1916. The article of Albert and Barabási [2] provides several examples of networks that exhibit power law degree distributions. In fact, most of these examples exhibit power laws that have exponents between 2 and 3. This range of exponents is also associated with *ultra-small* worlds. Chung and Lu [18] showed that for the model which we will consider in this paper, the average distance between two vertices in the largest (giant) component scales like $\log \log n$.

The methods of our paper have also been applied in the context of directed inhomogeneous random graphs [20]. Furthermore, they have found application in the analysis of boostrap-like processes which model cascading phenomena between financial institutions [21].

In this work we extend a theorem proved by the first two authors in [6] giving a threshold function $a_c(n) = o(n)$ such that when $a(n)$ grows slower than $a_c(n)$, then with high probability no evolution occurs, but if $a(n)$ grows faster than $a_c(n)$, then even if $a(n) = o(n)$, the final set contains a positive fraction of the vertices. Here we determine exactly this fraction and we show that as long as $a(n) = o(n)$, then it does not depend on $a(n)$ itself. In the rest of this section we proceed with the definition of the random graph model that we consider and the statement of our theorems.

Notation For non-negative sequences x_n and y_n we write $x_n = O(y_n)$ if there exist $N \in \mathbb{N}$ and $C > 0$ such that $x_n \leq C y_n$ for all $n \geq N$, and $x_n = o(y_n)$, if $x_n / y_n \to 0$, as $n \to \infty$. We sometimes also write $x_n \ll y_n$ for $x_n = o(y_n)$.

Let $\{X_n\}_{n \in \mathbb{N}}$ be a sequence of real-valued random variables on a sequence of probability spaces $\{(\Omega_n, \mathbb{P}_n)\}_{n \in \mathbb{N}, \mathcal{F}_n}$. If $c \in \mathbb{R}$ is a constant, we write $X_n \xrightarrow{p} c$ to denote that X_n *converges in probability to* c, that is, for any $\varepsilon > 0$ we have $\mathbb{P}_n(|X_n - c| > \varepsilon) \to 0$ as $n \to \infty$. Moreover, let $\{a_n\}_{n \in \mathbb{N}}$ be a sequence of real numbers that tends to infinity as $n \to \infty$. We write $X_n = o_p(a_n)$, if $|X_n|/a_n$ *converges to 0 in probability.* If $\mathcal{E}_n$ is a measurable subset of Ω_n, for any $n \in \mathbb{N}$, we say that the sequence $\{\mathcal{E}_n\}_{n \in \mathbb{N}}$ occurs *asymptotically almost surely (w.h.p.)* or *with high probability (w.h.p.)* if $\mathbb{P}_n(\mathcal{E}_n) = 1 - o(1)$ as $n \to \infty$.

2 Models and Results

The random graph model that we consider is an extension of a model considered by Chung and Lu [18], and is a special case of the so-called *inhomogeneous random graph*, which was introduced by Söderberg [33] and defined in its full generality by Bollobás, Janson and Riordan in [14].

2.1 Inhomogeneous Random Graphs with Rank-1 Kernel

Let $n \in \mathbb{N}$ and consider the vertex set $[n] := \{1, \ldots, n\}$. Each vertex i is assigned a positive weight $w_i(n)$, and we will write $\mathbf{w} = \mathbf{w}(n) = (w_1(n), \ldots, w_n(n))$. We will often suppress the dependence on n, whenever it is obvious from the context. For convenience, we will assume that $w_1 \leq w_2 \leq \cdots \leq w_n$. For any $S \subseteq [n]$, set

$$W_S(\mathbf{w}) := \sum_{i \in S} w_i.$$

In our random graph model the event of including the edge $\{i, j\}$ in the resulting graph is independent of the inclusion of any other edge, and its probability equals

$$p_{ij}(\mathbf{w}) = \min \left\{ \frac{w_i w_j}{W_{[n]}(\mathbf{w})}, 1 \right\}. \tag{1}$$

This model was studied by Chung and Lu for fairly general choices of $\mathbf{w}$, who studied in a series of papers [17, 18, 19] several typical properties of the resulting graphs, such as the average distance between two randomly chosen vertices that belong to the same component or the component size distribution. Therein, the model was defined under the additional assumption that $\max_{i\in[n]} w_i^2 < W_{[n]}$. We drop this assumption and use (1) instead. We will refer to this model as the *Chung-Lu* model, and we shall write $CL(\mathbf{w})$ for a random graph in which each possible edge $\{i,j\}$ is included independently with probability as in (1). Moreover, we will suppress the dependence on $\mathbf{w}$, if it is clear from the context which sequence of weights we refer to.

Note that in a Chung-Lu random graph the weights (essentially) control the *expected* degrees of the vertices. Indeed, if we ignore the minimization in (1), and also allow a loop at vertex i, then the expected degree of that vertex is $\sum_{j=1}^{n} w_i w_j / W_{[n]} = w_i$.

2.2 Regular Weight Sequences

Following van der Hofstad [37], for any $n \in \mathbb{N}$ and any sequence of weights $\mathbf{w}(n)$ let

$$F_n(x) = n^{-1} \sum_{i=1}^{n} \mathbf{1}[w_i(n) \le x], \quad \forall x \in [0,\infty)$$

be the empirical distribution function of the weight of a vertex chosen uniformly at random. We will assume that F_n has a certain structure.

Definition 2.1. *We say that* $(\mathbf{w}(n))_{n\ge 1}$ *is* regular*, if it has the following properties.*

- **[Weak convergence of weight]** *There is a distribution function* $F : [0,\infty) \to [0,1]$ *such that for all* x *at which* F *is continuous* $\lim_{n\to\infty} F_n(x) = F(x)$.
- **[Convergence of average weight]** *Let* W_n *be a random variable with distribution function* F_n*, and let* W_F *be a random variable with distribution function* F*. Then* $\lim_{n\to\infty} \mathbb{E}\,(\,W_n\,) = \mathbb{E}\,(\,W_F\,) < \infty$.
- **[Non-degeneracy]** *There is a* $x_0 \in \mathbb{R}^+$ *such that* $F_n(x) = 0$ *for all* $x \in [0, x_0)$ *and* $n \in \mathbb{N}$*. (That is, the weights are bounded from below by* x_0*.)*

The regularity of $(\mathbf{w}(n))_{n\ge 1}$ guarantees two important properties. First, the weight of a random vertex is approximately distributed as a random variable that follows a certain distribution. Second, this variable has finite mean and it is easy to see that the associated Chung-Lu random graph has bounded average degree with high probability. The third property in Definition 2.1 is a minor restriction guaranteeing that no vertex has a vanishing expected degree and is added for convinience in order to simplify several of our technical considerations.

At many places in our arguments it will be important to select vertices randomly according to their weight, i.e. the probability to choose $i \in [n]$ equals $w_i / W_{[n]}(\mathbf{w})$. This is the so-called *size-biased* distribution and we denote by $W_{F_n}^*$ a random variable with this distribution. A straightforward calculation shows that for every bounded continuous function f

$$\mathbb{E}\left(\, f(W_{F_n}^*) \,\right) = \frac{\mathbb{E}\,(\, W_{F_n} f(W_{F_n}) \,)}{\mathbb{E}\,(\, W_{F_n} \,)}. \tag{2}$$

2.3 Results

The main theorem of this paper gives a law of large numbers for the size of $\mathcal{A}_f$ when $\mathcal{A}_0$ has positive density in the case where the underlying random graph is a Chung-Lu random graph with a regular weight sequence. Let $\psi_r(x)$ for $x \geq 0$ be equal to the probability that a Poisson-distributed random variable with parameter x is at least r, i.e.,

$$\psi_r(x) := \mathbb{P}(\mathsf{Po}(x) \geq r) = e^{-x} \sum_{j \geq r} x^j / j!.$$

Let X be a non-negative random variable and $p \in [0,1]$. For any $r \geq 1$ and $y \in \mathbb{R}^+$ set

$$f_r(y; X, p) = (1-p)\mathbb{E}\left[\, \psi_r(Xy) \,\right] + p - y.$$

Theorem 2.2. *Let $(\mathbf{w}(n))_{n \geq 1}$ be regular with limiting distribution function F. Consider the bootstrap percolation process on $CL(\mathbf{w})$ with activation threshold $r \geq 2$, where $\mathcal{A}_0 \subseteq [n]$ includes any vertex independently with fixed probability $p \in (0,1)$. Let $\hat{y}$ be the smallest positive solution of*

$$f_r(y; W_F^*, p) = 0. \tag{3}$$

Assume also that $f_r'(\hat{y}; W_F^*, p) < 0$. Then

$$n^{-1}|\mathcal{A}_f| \xrightarrow{p} (1-p)\mathbb{E}\left[\, \psi_r(W_F \hat{y}) \,\right] + p, \text{ as } n \to \infty. \tag{4}$$

We remark that a solution $\hat{y}$ to (3) always exists because $f_r(y; W_F^*, p)$ is continuous, $f_r(0; W_F^*, p) > 0$ and $f_r(1; W_F^*, p) \leq 0$. Note that the conclusion of our results is valid only if $f_r'(\hat{y}; W_F^*, p) < 0$. This does not happen only if

$$\mathbb{E}\left[\frac{e^{-\hat{y}W_F^*}(W_F^* \hat{y})^r}{r!}\right] = \frac{\hat{y}}{(1-p)r},$$

and for such (rather exceptional) weight sequences we expect a different behavior. Moreover, we show that (c.f. Lemma 4.12) if the weight sequence has power law distribution with exponent between 2 and 3, this case will not happen (i.e., $f_r'(\hat{y}; W_F^*, p) < 0$ always).

Intuitively, the quantity $\hat{y}$ represents the limit of the probability that infection is passed through a random neighbour of a vertex. The fixed-point equation $f_r(y; W_F^*, p) = 0$, whose solution is $\hat{y}$, effectively says that a vertex is infected if either it is initially infected (which occurs with probability p) or (if not, which occurs with probability $1-p$) it has at least r infected neighbours. The latter is a Poisson-distributed random variable with parameter equal to $W_F^* \hat{y}$. The first factor essentially states the fact that a vertex becomes some other vertex's neighbour with probability proportional to the latter's weight, whereas it is infected with probability approximately $\hat{y}$.

We will now see an extension of the above theorem to the case where p is not anymore bounded away from 0. Under certain conditions the above theorem can be transferred to this case simply by setting $p = 0$. These conditions ensure that a positive but rather small fraction of the vertices become infected and this effectively corresponds to taking a p that is in fact bounded away from 0 but small.

2.4 Power-law Weight Sequences

Our second result focuses on an important special case of weight sequences, namely those following a power law distribution. This is described by the following condition.

Definition 2.3. *We say that a regular sequence $(\mathbf{w}(n))_{n\geq 1}$ follows* a power law with exponent β, *if there are $0 < c_1 < c_2$, $c_3, x_0 > 0$ and $0 < \zeta \leq 1/(\beta-1)$ such that for all $x_0 \leq x < c_3 \cdot n^\zeta$*

$$c_1 x^{-\beta+1} \leq 1 - F_n(x) \leq c_2 x^{-\beta+1},$$

and $F_n(x) = 0$ for $x < x_0$ and $F_n(x) = 1$ for $x \geq c_3 \cdot n^\zeta$. Moreover, for any $x > x_0$ we have for some $c > 0$

$$\lim_{n\to\infty} F_n(x) = F(x) = 1 - cx^{-\beta+1}.$$

We say that such a sequence belongs to the class $PL(\beta, \zeta)$.

In the above definition the maximum weight of a vertex is close to $c_3 \cdot n^\zeta$, for any n sufficiently large. Furthermore, if $\zeta = 1/(\beta-1)$, then $c_3 \leq c_2^{1/(\beta-1)}$.

A particular example of a power-law weight sequence is given in [18], where the authors choose $w_i = d(n/(i+i_0))^{1/(\beta-1)}$ for some $d > 0$. This results typically in a graph with a power law degree sequence with exponent β, average degree $O(d)$, and maximum degree proportional to $(n/i_0)^{1/(\beta-1)}$, see also [37] for a detailed discussion. When $\beta \in (2,3)$, these random graphs are also characterized as *ultra-small worlds*, due to the fact that the typical distance of two vertices that belong to the same component is $O(\log\log n)$, see [18, 37].

Theorem 2.2 addresses the case where the initial set $\mathcal{A}_0$ has positive density. Our second result is complementary and considers the setting where $p = p(n) = o(1)$, with a particular focus on the case where the exponent of the power law is in $(2,3)$ Assume that $\mathcal{A}_0$ has density $a(n)/n$. In [6] the first two authors determined a function $a_c(n)$ (which we also give in the statement of the next theorem) such that, for ζ satisfying $\frac{r-1}{2r-\beta+1} < \zeta \leq \frac{1}{\beta-1}$, when $a(n) = o(a_c(n))$, then w.h.p. $|\mathcal{A}_0| = |\mathcal{A}_f|$, whereas if $a(n) = \omega(a_c(n))$ but $a(n) = o(n)$, then w.h.p. $|\mathcal{A}_f| > \varepsilon n$, for some $\varepsilon > 0$. However, for $\zeta \leq \frac{r-1}{2r-\beta+1}$, they showed a weaker result and identified two functions $a_c^-(n) \ll a_c^+(n) = o(n)$ such that if $a(n) \gg a_c^+(n)$, then $|\mathcal{A}_f| > \varepsilon n$, for some $\varepsilon > 0$, but if $a(n) \ll a_c^-(n)$, then w.h.p. $|\mathcal{A}_0| = |\mathcal{A}_f|$. (In particular, $a_c^-(n) = a_c(n)$ and $a_c^+(n) = n^{1-\zeta\frac{r-\beta+2}{r-1}}$.) We refine this result using the proof of Theorem 2.2 and determine the fraction of vertices that belong to $\mathcal{A}_f$.

Theorem 2.4. *Let $(\mathbf{w}(n))_{n\geq 1} \in PL(\beta,\zeta)$ for some $\beta \in (2,3)$. Consider the bootstrap percolation process on $CL(\mathbf{w})$ with activation threshold $r \geq 2$. Let*

$$a_c(n) = n^{(r(1-\zeta)+\zeta(\beta-1)-1)/r}.$$

and

$$a_c^+(n) = n^{1-\zeta\frac{r-\beta+2}{r-1}}.$$

Assume that $\mathcal{A}_0$ is a random subset of $[n]$ where each vertex is included independently with probability $a(n)/n$. If $a(n) = o(n)$ and $a(n) = \omega(a_c(n))$ (for $\frac{r-1}{2r-\beta+1} < \zeta \leq \frac{1}{\beta-1}$) and $a(n) = \omega(a_c^+(n))$ (for $\zeta \leq \frac{r-1}{2r-\beta+1}$), then

$$n^{-1}|\mathcal{A}_f| \xrightarrow{p} \mathbb{E}\left[\psi_r(W_F\hat{y})\right] \text{ as } n\to\infty,$$

where $\hat{y}$ is the smallest positive solution of

$$y = \mathbb{E}\left[\psi_r(W_F^* y)\right].$$

When $\beta > 3$ the regularity assumptions of Theorem 2.2 are satisfied and the asymptotics of the size of the final set is given by this. When $\beta = 3$ these assumptions are no longer satisfied. However, the techniques that are used for proof of Theorem 2.4 in Section 3.2 do not apply immediately but need a significant refinement.

Let us remark here that the (rescaled) size of the final set does not depend on $|\mathcal{A}_0|$.

More generally, the above theorem holds as long as the initial density is such that a.a.s most vertices of weight that is larger some big constant become infected.

2.5 Outline

The proofs of Theorems 2.2 and 2.4 are based on a *finitary* approximation of the weight sequence $\mathbf{w}(n)$. In the following section we construct a sequence of weight sequences having only a finite number of weights and that "approximate" the initial sequence in a certain well-defined sense. Thereafter, we show the analogue of Theorem 2.2 for finitary sequences; this is Theorem 3.4 stated below. The proof of Theorem 3.4 is based on the so-called *differential equation* method, which was developed by Wormald [38, 39], and is used to keep track of the evolution of the bootstrap percolation process through the exposure of the neighbours of each infected vertex. Such an exposure algorithm was also applied in the homogeneous setting [26]. Of course, the inhomogeneous setting imposes significant obstacles. We close the paper with the proof of some rather technical results, which transfer the condition on the derivative that appears in the statement of Theorem 2.2 in the finitary setting.

3 Finitary Weight Sequences

In this section we will consider so-called *finitary* weight sequences on $[n]$ that are suitable approximations of an arbitrary weight sequence $\mathbf{w}(n)$. As a first step we are going to "remove" all weights from $\mathbf{w}$ that are too large in the following sense. Suppose that $\mathbf{w}(n)$ is regular and that the corresponding sequence of empirical distributions converges to F. Let $(c_j)_{j\in\mathbb{N}}$ be an increasing sequence of points of continuity of F with the properties that

1. $\lim_{j\to\infty} c_j = \infty$;
2. $2c_j$ is also a point of continuity.

For $\gamma > 0$ let

$$C_\gamma = C_\gamma(F) := \min\{c_j : F(c_j) \geq 1 - \gamma\}.$$

Then, as $n \to \infty$, the following facts are immediate consequences. Let $\mathsf{C}_\gamma = \mathsf{C}_\gamma(n, F)$ be the set of vertices in $[n]$ with weight at least $C_\gamma(F)$. Then:

1. With $h_F(\gamma) := 1 - F(C_\gamma) \leq \gamma$, we have $|\mathsf{C}_\gamma(n, F)|/n \to h_F(\gamma)$.
2. $n^{-1} W_{\mathsf{C}_\gamma(n,F)}(\mathbf{w}(n)) \to \int_{C_\gamma}^{\infty} x dF(x) =: W_\gamma(F)$, where the latter is the Lebesque-Stieltjes integral with respect to F.
3. The assumption $\mathbb{E}\left[W_F \right] = d < \infty$ implies that $\mathbb{P}\left[W_F > x \right] = o(1/x)$ as $x \to \infty$. Thus

$$C_\gamma(F)\mathbb{P}\left[W_F > C_\gamma(F) \right] \to 0, \text{ as } \gamma \downarrow 0. \tag{5}$$

   Also, $W_\gamma(F)/C_\gamma(F) \to 0$ as $\gamma \downarrow 0$. We will be using this observation in several places in our proofs.

We will approximate a regular $(\mathbf{w}(n))_{n\geq 1}$ by a sequence where most vertices have their weights within a finite set of values and moreover the weights are bounded by $2C_\gamma(F)$ (cf. [37] where a similar approach is followed in a different context).

Definition 3.1. *Let $\ell \in \mathbb{N}$ and $\gamma \in (0,1)$.*

For a function $n' = n'(n) \in \mathbb{N}$ with $n' \geq n - |\mathsf{C}_\gamma(F)|$, we say that a regular weight sequence

$$(\mathbf{W}^{(\ell,\gamma)}(n))_{n\geq 1} = \left(W_1^{(\ell,\gamma)}(n), \dots, W_{n'}^{(\ell,\gamma)}(n)\right)_{n\geq 1}$$

is a (ℓ,γ)-discretisation *of a regular weight sequence $(\mathbf{w}(n))_{n\geq 1}$ with limiting distribution function F, if the following conditions are satisfied. There are an increasing sequence of natural numbers $(p_\ell)_{\ell\in\mathbb{N}}$ and positive constants $\gamma_1, \dots, \gamma_{p(\ell)} \in (0,1)$ such that $\sum_{i=1}^{p_\ell} \gamma_i = 1 - h_F(\gamma)$ and real weights $0 < W_0 < W_1 < \cdots < W_{p_\ell} \leq C_\gamma(F)$ which satisfy the following properties. There is a partition of $[n] \setminus \mathsf{C}_\gamma(F)$ into p_ℓ parts, denoted by $\mathsf{C}_1(n), \dots, \mathsf{C}_{p_\ell}(n)$ such that:*

1. *For all $1 \leq i \leq p_\ell$ and for all $j \in \mathsf{C}_i(n)$ we have $W_j^{(\ell,\gamma)}(n') = W_i$.*
2. *Let $\mathsf{C}'_\gamma(n) := [n'] \setminus \cup_{i=1}^{p_\ell} \mathsf{C}_i(n)$. Then $C_\gamma(F) \leq W_j^{(\ell,\gamma)}(n') \leq 2C_\gamma(F)$ for all $j \in \mathsf{C}'_\gamma(n)$.*

Moreover, as $n \to \infty$:

3. *For all $1 \leq i \leq p_\ell$, $n^{-1}|\mathsf{C}_i(n)| \to \gamma_i$.*
4. *There is a $h_F(\gamma) \leq \gamma' < h_F(\gamma) + 2W_\gamma(F)/C_\gamma(F)$ such that $n^{-1}|\mathsf{C}'_\gamma(n)| \to \gamma'$.*
5. *There is a $0 \leq W'_\gamma \leq 4W_\gamma(F)$ such that $n^{-1}W_{\mathsf{C}'_\gamma(n)}(\mathbf{W}^{(\ell,\gamma)}(n)) \to W'_\gamma$.*
6. *The weight sequence $\mathbf{W}^{(\ell,\gamma)}(n)$ gives rise to a sequence of the corresponding empirical distributions which we denote by $F_n^{(\ell,\gamma)}$ and we assume that they converge weakly to a limiting distribution $F^{(\ell,\gamma)}$.*

The upper bounds in *4.* and *5.* are tailored to the proof of Theorem 2.2. Note that in the previous definition no assumption is made on the W_is, and thus $\mathbf{W}^{(\ell,\gamma)}$ might look very different from $\mathbf{w}$. The next definition quantifies when a (ℓ,γ)-discretisation is "close" to a given regular $(\mathbf{w}(n))_{n\geq 1}$ with limiting distribution function F. For a cumulative distribution function G, let G^* denote the distribution function of the size-biased version of an G-distributed random variable.

Definition 3.2. *Let $(\mathbf{w}(n))_{n\geq 1}$ be regular and let F be its limiting distribution function. A family $((\mathbf{W}^{(\ell,\gamma)}(n))_{n\geq 1})_{\ell\in\mathbb{N},\gamma\in(0,1)}$ of (ℓ,γ)-discretisations of $(\mathbf{w}(n))_{n\geq 1}$ with limiting distribution functions $F^{(\ell,\gamma)}$ is called F*-convergent *if*

1. *for every $x \in \mathbb{R}$ that is a point of continuity of F we have*

$$\lim_{\gamma\downarrow 0}\lim_{\ell\to\infty} F^{(\ell,\gamma)}(x) = F(x),\ \lim_{\gamma\downarrow 0}\lim_{\ell\to\infty} F^{*(\ell,\gamma)}(x) = F^*(x),$$

and,

2.

$$\lim_{\gamma\downarrow 0}\lim_{\ell\to\infty}\left|\int_0^\infty x dF^{(\ell,\gamma)}(x)-\mathbb{E}(W_F)\right|=0,\ \lim_{\gamma\downarrow 0}\lim_{\ell\to\infty}\left|\int_0^{C_\gamma} x dF^{(\ell,\gamma)}(x)-\mathbb{E}(W_F)\right|=0.$$

Let $U^{(\ell,\gamma)}$ ($U^{*(\ell,\gamma)}$, respectively) be a random variable whose distribution function is $F^{(\ell,\gamma)}$ ($F^{*(\ell,\gamma)}$, resp.). Let us observe that

$$\mathbb{P}\left[U^{*(\ell,\gamma)}>C_\gamma\right]=\frac{\mathbb{E}\left[U^{(\ell,\gamma)}\mathbf{1}_{U^{(\ell,\gamma)}>C_\gamma}\right]}{\mathbb{E}\left[U^{(\ell,\gamma)}\right]}\leq 2/W_0\cdot\mathbb{E}\left[U^{(\ell,\gamma)}\mathbf{1}_{U^{(\ell,\gamma)}>C_\gamma}\right],$$

since $\mathbb{E}\left[U^{(\ell,\gamma)}\right]\geq W_0/2$ for any γ and any ℓ sufficiently large. By Part 2 of Definition 3.2, we have $\lim_{\gamma\downarrow 0}\lim_{\ell\to\infty}\mathbb{E}\left[U^{(\ell,\gamma)}\mathbf{1}_{U^{(\ell,\gamma)}>C_\gamma}\right]=0$. We can thus deduce the following lemma, which will be used later.

Lemma 3.3. *If $f:\mathbb{R}\to\mathbb{R}$ is a bounded function, then*

$$\lim_{\gamma\downarrow 0}\lim_{\ell\to\infty}\left|\mathbb{E}\left(f(U^{*(\ell,\gamma)})\mathbf{1}_{U^{*(\ell,\gamma)}>C_\gamma}\right)\right|=0.$$

For technical reasons we consider a slightly different definition of the random graph model that we denote by $CL'(\mathbf{W}^{(\ell,\gamma)})$. In this modified model the edge probabilities are proportional to the product of the weights of the vertices, except that the normalizing factor is not equal to the sum of the weights in $\mathbf{W}^{(\ell,\gamma)}$, but it is equal to $W_{[n]}(\mathbf{w}(n))$, that is, the edge $\{i,j\}$ is contained in $CL'(\mathbf{W}^{(\ell,\gamma)})$ with probability

$$p_{ij}(\mathbf{W}^{(\ell,\gamma)}(n),\mathbf{w}(n))=\min\left\{\frac{w_i^{(\ell,\gamma)}w_j^{(\ell,\gamma)}}{W_{[n]}(\mathbf{w})},1\right\}.$$

The next theorem quantifies the number of the finally infected vertices when the weight sequence is a discretisation of a given regular $(\mathbf{w}(n))_{n\geq 1}$. It is general enough so that it can be used in the proof of Theorem 2.4 as well.

Theorem 3.4. *Let $(\mathbf{w}(n))_{n\geq 1}$ be regular and let F be its limiting distribution function. Let $((\mathbf{W}^{(\ell,\gamma)}(n))_{n\geq 1})_{\ell\in\mathbb{N},\gamma\in(0,1)}$ be a family of (ℓ,γ)-discretisations of $(\mathbf{w}(n))_{n\geq 1}$ which is F-convergent. Moreover, assume that $f_r'(\hat{y};W_F^*,p)<0$ (cf. Theorem 2.2).*

Let $r\geq 2$. Assume that initially all vertices of $CL'(\mathbf{W}^{(\ell,\gamma)})$ that belong to $\mathsf{C}_\gamma'(n)$ are infected, whereas each vertex in $\mathsf{C}_i(n)$ is infected independently with probability $p\in[0,1)$, for each $i=1,\dots,p_\ell$. Let $\mathcal{A}_f^{(\ell,\gamma)}$ denote the set of vetices in $[n']\setminus\mathsf{C}_\gamma'(n)$ that become eventually infected during a bootstrap percolation process with activation threshold r. There exists $c>0$ for which the following holds: for $\gamma\in(0,c)$ and for any $\delta\in(0,1)$ there is a subsequence $\mathcal{S}:=\{\ell_k\}_{k\in\mathbb{N}}$ such that for any $\ell\in\mathcal{S}$ with probability at least $1-o(1)$

$$n^{-1}|\mathcal{A}_f^{(\ell,\gamma)}|=(1\pm\delta)\left((1-p)\mathbb{E}\left[\psi_r(W_F\hat{y})\right]+p\right).$$

3.1 Proof of Theorem 2.2

Given a regular $(\mathbf{w}(n))_{n\geq 1}$, Theorem 2.2 follows from Theorem 3.4 by constructing an F-convergent family $((\mathbf{W}^{(\ell,\gamma)}(n))_{n\geq 1})_{\ell\in\mathbb{N},\gamma\in(0,1)}$. We first describe our construction and prove some properties of it, and then proceed with the proofs of our main results.

3.1.1 The construction of approximating weight sequences

Let $(\mathbf{w}(n))_{n\geq 1}$ be regular and consider the limiting distribution function F. For $\gamma \in (0,1)$, recall that $F(C_\gamma) \geq 1-\gamma$. Recall also that from Definition 2.1 there is a positive real number x_0 such that $F(x) = 0$ for $x < x_0$.

We define a set of intervals $\mathcal{P}_\ell$ whose union is a superset of $[x_0, C_\gamma)$ as follows. Let $\varepsilon_\ell = 1/\ell$. Firstly, for $i \geq 0$, we set

$$x_{i+1} = \sup\{x \in (x_i, C_\gamma) \ : \ F(x) - F(x_i) < \varepsilon_\ell\}.$$

Set $t_\ell = \min\{i \ : \ x_i = C_\gamma\}$ and $x_{-1} = 0$. For each $i = 0, \ldots, t_\ell$, let y_{2i}, y_{2i+1} be such that

1. $\max\{\frac{1}{2}(x_{i-1} + x_i), x_i - \varepsilon_\ell\} < y_{2i} < x_i$;
2. $x_i < y_{2i+1} < \min\{\frac{1}{2}(x_i + x_{i+1}), x_i + \varepsilon_\ell\}$ or $y_{2i+1} = C_\gamma$, if $i = t_\ell$;
3. y_{2i}, y_{2i+1} are points of continuity of F.

Now, we set $\mathcal{P}_\ell := \{[y_0, y_1), \ldots, [y_{2t_\ell}, C_\gamma)\}$. With $p_\ell = 2t_\ell$, for $i = 0, \ldots, p_\ell$, we set $I_i = [y_i, y_{i+1})$.

Given this partition and the weight sequence $\mathbf{w}(n)$, for each $n \geq 1$ we define two finitary weight sequences $\mathbf{W}^{(\ell,\gamma)+}(n')$ and $\mathbf{W}^{(\ell,\gamma)-}(n'')$ on the sets $[n']$ and $[n'']$, respectively, as follows. The partition $\mathcal{P}_\ell$ gives rise to a partition of $[n] \setminus \mathsf{C}_\gamma$, where for each $i = 1, \ldots, p_\ell$ we have $\mathsf{C}_i = \{j \ : \ w_j(n) \in I_i\}$. We denote this partition by $\mathcal{P}_{n,\ell,\gamma}$ and we let this be the associated partition of $\mathbf{W}^{(\ell,\gamma)+}(n')$ and $\mathbf{W}^{(\ell,\gamma)-}(n'')$. In particular,

– consider the random subset of C_γ, in which every element of C_γ is included independently with probability p. An application of the Chernoff bounds implies that w.h.p. this has size at least $\lfloor p|\mathsf{C}_\gamma| - n^{2/3}\rfloor =: k_-$. Consider a set of vertices $\mathsf{C}_\gamma^- = \{v_1, \ldots, v_{k_-}\}$ which is disjoint from $[n]$. We identify with $[n'']$ the set $(\cup_{i=1}^{p_\ell}\mathsf{C}_i)\bigcup \mathsf{C}_\gamma^-$, through a bijective mapping $\varphi^- : (\cup_{i=1}^{p_\ell}\mathsf{C}_i)\bigcup \mathsf{C}_\gamma^- \to [n'']$. It follows that $n'' = (1 - h_F(\gamma) + ph_F(\gamma))n(1 + o(1))$.

– for any vertex $j \in \mathsf{C}_\gamma$ such that $w_j(n) \geq 2C_\gamma$ we consider $c_j := 2\lfloor \frac{w_j(n)}{C_\gamma}\rfloor$ copies of this vertex each having weight $2C_\gamma$, which we label as $v_{j1}, \ldots, v_{jc_j}$. For each such j we let $\varepsilon_j(n) = \frac{w_j(n)}{C_\gamma} - \lfloor \frac{w_j(n)}{C_\gamma}\rfloor$ and we set $R = \lceil 2\sum_{j \ : \ w_j(n)\geq 2C_\gamma} \varepsilon_j(n)\rceil$. If $j \in \mathsf{C}_\gamma$ is such that $C_\gamma \leq w_j(n) < 2C_\gamma$, then we introduce a single copy v_{j1} having weight equal to w_j (in other words $c_j = 1$).

We let C_γ^+ be the set that is the union of these copies together with a set of R vertices which we denote by $\mathcal{R}$ (disjoint from the aforementioned sets) each having weight $2C_\gamma$:

$$\mathsf{C}_\gamma^+ := \mathcal{R} \cup \bigcup_{j\in \mathsf{C}_\gamma} \{v_{j1}, \ldots, v_{jc_j}\}.$$

Let $n' = \left|(\cup_{i=1}^{p_\ell}\mathsf{C}_i)\bigcup \mathsf{C}_\gamma^+\right|$ and identify the set $[n']$ with the vertices in $(\cup_{i=1}^{p_\ell}\mathsf{C}_i)\bigcup \mathsf{C}_\gamma^+$, through a bijection $\varphi^+ : (\cup_{i=1}^{p_\ell}\mathsf{C}_i)\bigcup \mathsf{C}_\gamma^+ \to [n']$. We will use the symbol C_γ^+ to denote the set $[n'] \setminus \varphi^+\left((\cup_{i=1}^{p_\ell}\mathsf{C}_i)\right)$. In other words, the set C_γ^+ consists of the replicas of the vertices in C_γ, as these were defined above, together with the set of vertices corresponding to $\mathcal{R}$. This completes the definition of $\mathbf{W}^{(\ell,\gamma)+}(n')$.

– for each $i = 1, \ldots, p_\ell$, we set $W_i^- = y_{2i}$ and $W_i^+ = y_{2i+1}$; for each $j \in \mathsf{C}_i$, we set

$$W_{\varphi^-(j)}^{(\ell,\gamma)-}(n'') := W_i^-, \text{ and } W_{\varphi^+(j)}^{(\ell,\gamma)+}(n') := W_i^+.$$

For any $j \in [n''] \setminus \varphi^{-}(\cup_{i=1}^{p_\ell} \mathsf{C}_i)$ we set $W_j^{(\ell,\gamma)-}(n) := C_\gamma$. Note that

$$\lim_{n\to\infty} \frac{|\mathsf{C}_\gamma^-|}{n} = p h_F(\gamma),$$

and if $W_{\mathsf{C}_\gamma^-}(\mathbf{W}^{(\ell,\gamma)-})$ denotes the total weight of these vertices, then this satisfies

$$\lim_{n\to\infty} \frac{W_{\mathsf{C}_\gamma^-}(\mathbf{W}^{(\ell,\gamma)-})}{n} = p h_F(\gamma) C_\gamma =: W_\gamma^- < W_\gamma < 4W_\gamma.$$

Furthermore,

$$\begin{aligned} |\mathsf{C}_\gamma^+| &= \sum_{j\,:\,C_\gamma \le w_j < 2C_\gamma} 1 + \sum_{j\,:\,w_j \ge 2C_\gamma} 2\lfloor \frac{w_j}{C_\gamma} \rfloor + R \\ &= \sum_{j\,:\,C_\gamma \le w_j < 2C_\gamma} 1 + 2\sum_{j\,:\,w_j \ge 2C_\gamma} \frac{w_j}{C_\gamma} + e(n), \end{aligned}$$

with $0 \le e(n) < 1$. By the weak convergence of F_n to F and since $\mathbb{E}[W_n] \to \mathbb{E}[W_F] < \infty$, it follows that

$$\lim_{n\to\infty} \frac{|\mathsf{C}_\gamma^+|}{n} = \mathbb{P}[C_\gamma \le W_F \le 2C_\gamma] + 2\frac{\mathbb{E}\left[\mathbf{1}_{\{W_F \ge 2C_\gamma\}} W_F\right]}{C_\gamma} =: \gamma^+, \tag{6}$$

where $\gamma^+ \downarrow 0$ as $\gamma \downarrow 0$. So $n'/n \to 1 - h_F(\gamma) + \gamma^+$ as $n \to \infty$. Moreover,

$$\lim_{\gamma\downarrow 0} \gamma^+ C_\gamma = \lim_{\gamma\downarrow 0}(C_\gamma \mathbb{P}[C_\gamma \le W_F \le 2C_\gamma] + 2\mathbb{E}\left[\mathbf{1}_{\{W_F \ge 2C_\gamma\}} W_F\right]) \stackrel{(5)}{=} 0. \tag{7}$$

Also, the total weight of the vertices in C_γ^+ can be bounded as follows.

$$\begin{aligned} W_{\mathsf{C}_\gamma^+}(\mathbf{W}^{(\ell,\gamma)+}) &= \sum_{j\,:\,C_\gamma \le w_j < 2C_\gamma} w_j + \sum_{j\,:\,w_j \ge 2C_\gamma} 2\lfloor \frac{w_j}{C_\gamma} \rfloor (2C_\gamma) \\ &\le \sum_{j\,:\,C_\gamma \le w_j < 2C_\gamma} w_j + 4\sum_{j\,:\,w_j \ge 2C_\gamma} w_j. \end{aligned}$$

Hence, as $n \to \infty$

$$\frac{W_{\mathsf{C}_\gamma^+}(\mathbf{W}^{(\ell,\gamma)+})}{n} \to \mathbb{E}\left[\mathbf{1}_{\{C_\gamma \le W_F < 2C_\gamma\}} W_F\right] + 4\mathbb{E}\left[\mathbf{1}_{\{W_F \ge 2C_\gamma\}} W_F\right] =: W_\gamma^+ \le 4W_\gamma \tag{8}$$

We denote by $U_n^{(\ell,\gamma)-}$ and $U_n^{(\ell,\gamma)+}$ the weight in $\mathbf{W}^{(\ell,\gamma)-}(n)$ and $\mathbf{W}^{(\ell,\gamma)+}(n)$ of a uniformly chosen vertex from $[n']$ and $[n'']$, respectively. Also, we let $F_n^{(\ell,\gamma)-}, F_n^{(\ell,\gamma)+}$ denote their distribution functions. Note that both $F_n^{(\ell,\gamma)-}, F_n^{(\ell,\gamma)+}$ converge pointwise as $n \to \infty$ to the functions $F^{(\ell,\gamma)-}, F^{(\ell,\gamma)+}$, respectively, where

–for each $i = 0, \ldots, p_\ell$ and for each $x \in I_i$ we set

$$F^{(\ell,\gamma)-}(x) := \frac{F(W_i^+)}{1 - h_F(\gamma) + p h_F(\gamma)} \quad \text{and} \quad F^{(\ell,\gamma)+}(x) = \frac{F(W_i^-)}{1 - h_F(\gamma) + \gamma^+}.$$

– for any $x \geq C_\gamma$ we have $F^{(\ell,\gamma)-}(x) = 1$ and for any $x < y_0$ we have $F^{(\ell,\gamma)-}(x) = 0, F^{(\ell,\gamma)+}(x) = 0$;

– for any $C_\gamma \leq x < 2C_\gamma$ we have

$$F^{(\ell,\gamma)+}(x) = \frac{F(x)}{1 - h_F(\gamma) + \gamma^+}, \tag{9}$$

whereas for $x \geq 2C_\gamma$ we have $F^{(\ell,\gamma)+}(x) = 1$.

We will now prove that both families $\{\mathbf{W}^{(\ell,\gamma)+}(n)\}_{\gamma\in(0,1),\ell\in\mathbb{N}}$ and $\{\mathbf{W}^{(\ell,\gamma)-}(n)\}_{\gamma\in(0,1),\ell\in\mathbb{N}}$ are F-covergent. Thus, we will verify that they satisfy both parts of Definition 3.2.

Part 1 of Definition 3.2. It will be convenient to define a probability distribution function which will be the pointwise limit of $F^{(\ell,\gamma)+}$ and $F^{(\ell,\gamma)-}$ as $\ell \to \infty$. For any $x \in [0, C_\gamma)$ we set

$$F^{(\gamma)+}(x) = \frac{F(x)}{1 - h_F(\gamma) + \gamma^+},$$

and

$$F^{(\gamma)-}(x) = \frac{F(x)}{1 - h_F(\gamma) + ph_F(\gamma)},$$

whereas $F^{(\gamma)+}(x) = F^{(\gamma)-}(x) = 0$, for $x < 0$, and $F^{(\gamma)+}(x) = F^{(\gamma)-}(x) = 1$, for $x \geq C_\gamma$. Note first that for any point $x < C_\gamma$ that is a point of continuity of F (and, thereby, of $F^{(\gamma)+}$ and $F^{(\gamma)-}$ as well), we have

$$\lim_{\ell\to\infty} F^{(\ell,\gamma)+}(x) = F^{(\gamma)+}(x)$$

and

$$\lim_{\ell\to\infty} F^{(\ell,\gamma)-}(x) = F^{(\gamma)-}(x).$$

Moreover, note for any $x > 0$ we have

$$\lim_{\gamma\downarrow 0} F^{(\gamma)+}(x), F^{(\gamma)-}(x) = F(x).$$

We will now turn to the size-biased versions of these distributions. Let $U^{(\gamma)+}$ and $U^{(\gamma)-}$ denote two random variables with probability distribution functions $F^{(\gamma)+}$ and $F^{(\gamma)-}$, respectively. Thus, as $\ell \to \infty$,

$$U^{(\ell,\gamma)+} \xrightarrow{d} U^{(\gamma)+} \text{ and } U^{(\ell,\gamma)-} \xrightarrow{d} U^{(\gamma)-} \tag{10}$$

whereas as $\gamma \downarrow 0$ we have

$$U^{(\gamma)+}, U^{(\gamma)-} \xrightarrow{d} W_F. \tag{11}$$

Claim 3.5. *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of non-negative random variables. Suppose that W is a random variable such that $X_n \xrightarrow{d} W$ as $n \to \infty$. For every $x > 0$ which is a point of continuity of the cdf of W we have*

$$\lim_{n\to\infty} \mathbb{E}(X_n \mathbf{1}_{X_n\leq x}) = \mathbb{E}(W\mathbf{1}_{W\leq x}).$$

Proof. Firstly note that $X_n\mathbf{1}_{X_n\leq x} \xrightarrow{d} W\mathbf{1}_{W\leq x}$ as $n \to \infty$. By the Skorokhod representation theorem, there is a coupling of these random variables such that $X_n\mathbf{1}_{X_n\leq x} \to W\mathbf{1}_{W\leq x}$ almost surely as $n \to \infty$. The claim now follows from the bounded convergence theorem. □

This claim implies that for any $\gamma \in (0,1)$, as $\ell \to \infty$,

$$F^{*(\ell,\gamma)+}(x) = \frac{\mathbb{E}(U^{(\ell,\gamma)+}\mathbf{1}_{U^{(\ell,\gamma)+}\le x})}{\mathbb{E}(U^{(\ell,\gamma)+})} \to \frac{\mathbb{E}(U^{(\gamma)+}\mathbf{1}_{U^{(\gamma)+}\le x})}{\mathbb{E}(U^{(\gamma)+})}, \tag{12}$$

and

$$F^{*(\ell,\gamma)-}(x) = \frac{\mathbb{E}(U^{(\ell,\gamma)-}\mathbf{1}_{U^{(\ell,\gamma)-}\le x})}{\mathbb{E}(U^{(\ell,\gamma)-})} \to \frac{\mathbb{E}(U^{(\gamma)-}\mathbf{1}_{U^{(\gamma)-}\le x})}{\mathbb{E}(U^{(\gamma)-})}. \tag{13}$$

Furthermore, we will show that $\mathbb{E}(U^{(\gamma)+}), \mathbb{E}(U^{(\gamma)-}) \to \mathbb{E}(W_F)$ as $\gamma \downarrow 0$. The proof is identical for both $U^{(\gamma)+}$ and $U^{(\gamma)-}$; so we will denote these by $U^{(\gamma)\pm}$. For $\delta > 0$, let C_δ be a continuity point of F such that $\mathbb{E}(W_F\mathbf{1}_{W_F>C_\delta}) < \delta/4$. Now, let us bound $\mathbb{E}(U^{(\gamma)\pm}\mathbf{1}_{U^{(\gamma)\pm}>C_\delta})$. Note that $\mathbb{E}(U^{(\gamma)\pm}\mathbf{1}_{U^{(\gamma)\pm}>C_\delta}) = \mathbb{E}(U^{(\gamma)\pm}\mathbf{1}_{C_\gamma \ge U^{(\gamma)\pm}>C_\delta})$, if γ is small enough. So it suffices to bound the latter term. For $C_\delta < x < C_\gamma$, we have $F^{(\gamma)\pm}(x) = \lambda^{\pm}(\gamma)\cdot F(x)$, where $\lambda^{\pm}(\gamma) \to 1$ as $\gamma \downarrow 0$. Furthermore, $F^{(\gamma)^{\pm}}(C_\gamma) - F(C_\gamma) = 1 - (1 - h_F(\gamma)) = h_F(\gamma)$. Thus,

$$\begin{aligned}\mathbb{E}(U^{(\gamma)\pm}\mathbf{1}_{C_\gamma \ge U^{(\gamma)\pm}>C_\delta}) &= \lambda^{\pm}(\gamma)\mathbb{E}(W_F\mathbf{1}_{C_\gamma \ge W_F > C_\delta}) + (1-\lambda^{\pm}(\gamma))C_\gamma \cdot h_F(\gamma) \\ &\le \lambda^{\pm}(\gamma)\mathbb{E}(W_F\mathbf{1}_{W_F>C_\delta}) + |1-\lambda^{\pm}(\gamma)|C_\gamma \cdot h_F(\gamma) < \delta/3,\end{aligned}$$

if γ is sufficiently small (using (7)). Also, by Claim 3.5 we deduce that for any γ sufficiently small

$$|\mathbb{E}(U^{(\gamma)\pm}\mathbf{1}_{U^{(\gamma)\pm}\le C_\delta}) - \mathbb{E}(W_F\mathbf{1}_{W_F\le C_\delta})| < \delta/3.$$

Combining the above we finally deduce that

$$\begin{aligned}&|\mathbb{E}(U^{(\gamma)\pm}) - \mathbb{E}(W_F)| \le \\ &\quad |\mathbb{E}(U^{(\gamma)\pm}\mathbf{1}_{U^{(\gamma)\pm}\le C_\delta}) - \mathbb{E}(W_F\mathbf{1}_{W_F\le C_\delta})| + \mathbb{E}(U^{(\gamma)\pm}\mathbf{1}_{U^{(\gamma)\pm}>C_\delta}) + \mathbb{E}(W_F\mathbf{1}_{W_F>C_\delta}) < \delta.\end{aligned} \tag{14}$$

Claim 3.5 also implies that for every $x > 0$ which is a point of continuity of F, we have as $\gamma \downarrow 0$

$$|\mathbb{E}(U^{(\gamma)\pm}\mathbf{1}_{U^{(\gamma)\pm}\le x}) - \mathbb{E}(W_F\mathbf{1}_{W_F\le x})| \downarrow 0. \tag{15}$$

So from (14),(15) and (12),(13) we deduce that for any continuity point $x \in \mathbb{R}$ and any $\delta > 0$ there exists $\gamma_0 = \gamma_0(\delta, x)$ with the property that for any $0 < \gamma < \gamma_0$ there exists ℓ_0 such that for any $\ell > \ell_0$ we have

$$|F^{*(\ell,\gamma)}(x) - F^*(x)| < \delta. \tag{16}$$

The above can now be translated into the next lemma.

Lemma 3.6. *For any bounded and continuous function $f : \mathbb{R} \to \mathbb{R}$ the following holds: for any $\delta > 0$ there exists $\gamma_0 = \gamma_0(\delta, x)$ with the property that for any $0 < \gamma < \gamma_0$ there exists ℓ_0 such that for any $\ell > \ell_0$*

$$|\mathbb{E}(f(U^{*\pm(\ell,\gamma)})) - \mathbb{E}(f(W_F^*))| < \delta.$$

Although this is a straightforward restatement of weak convergence, we give it more explicitly as $U^{*\pm(\ell,\gamma)}$ depends on two parameters ℓ and γ. It is for the sake of clarity that we state explicitly how these depend when taking the double limit.

This completes the first part of Definition 3.2 and we proceed with the second part.

Part 2 of Definition 3.2. Since $F^{(\ell,\gamma)+}$ and $F^{(\ell,\gamma)-}$ are both constant (and equal to 1) for $x \geq 2C_\gamma$ we have

$$\mathbb{E}\left(U^{(\ell,\gamma)\pm} \right) = \mathbb{E}\left(U^{(\ell,\gamma)\pm}\mathbf{1}_{U^{(\ell,\gamma)\pm}\leq 2C_\gamma} \right).$$

Furthermore,

$$\mathbb{E}\left[W_F \right] = \mathbb{E}\left(W_F\mathbf{1}_{W_F\leq 2C_\gamma} \right) + \mathbb{E}\left(W_F\mathbf{1}_{W_F> 2C_\gamma} \right)$$

Therefore,

$$\begin{aligned}\left|\mathbb{E}\left(U^{(\ell,\gamma)\pm} \right) - \mathbb{E}\left(W_F \right)\right| &= \left|\mathbb{E}\left(U^{(\ell,\gamma)\pm}\mathbf{1}_{U^{(\ell,\gamma)\pm}\leq 2C_\gamma} \right) - \mathbb{E}\left(W_F \right)\right| \\ \leq \quad & \left|\mathbb{E}\left(U^{(\ell,\gamma)\pm}\mathbf{1}_{U^{(\ell,\gamma)\pm}\leq 2C_\gamma} \right) - \mathbb{E}\left(W_F\mathbf{1}_{W_F\leq 2C_\gamma} \right)\right| + \mathbb{E}\left(W_F\mathbf{1}_{W_F> 2C_\gamma} \right).\end{aligned} \tag{17}$$

The last term converges to 0 as $\gamma \downarrow 0$ since $\mathbb{E}\left(W_F \right) < \infty$.

We will now bound the first term on the right-hand side in (17). We write $F^{(\ell,\gamma)\pm}$ for either of $F^{(\ell,\gamma)+}$ or $F^{(\ell,\gamma)-}$. Using the integration-by-parts formula for the Lebesque-Stieltjes integral we can write

$$\begin{aligned}\mathbb{E}\left(U^{(\ell,\gamma)\pm}\mathbf{1}_{U^{(\ell,\gamma)\pm}\leq 2C_\gamma} \right) &= 2C_\gamma F^{(\ell,\gamma)\pm}(2C_\gamma+) - \int_0^{2C_\gamma} F^{(\ell,\gamma)\pm}(x)dx \\ &= 2C_\gamma - \int_0^{2C_\gamma} F^{(\ell,\gamma)\pm}(x)dx.\end{aligned} \tag{18}$$

Using integration-by-parts, we also get

$$\begin{aligned}\mathbb{E}\left(W_F\mathbf{1}_{W_F\leq 2C_\gamma} \right) &= 2C_\gamma \cdot F(2C_\gamma+) - f(0)\cdot F(0-) - \int_0^{2C_\gamma} F(x)dx \\ &= 2C_\gamma \cdot (1 - \mathbb{P}\left[W_F > 2C_\gamma \right]) - \int_0^{2C_\gamma} F(x)dx.\end{aligned} \tag{19}$$

Therefore,

$$\begin{aligned}&\left|\mathbb{E}\left(U^{(\ell,\gamma)\pm}\mathbf{1}_{U^{(\ell,\gamma)\pm}\leq 2C_\gamma} \right) - \mathbb{E}\left(W_F\mathbf{1}_{W_F\leq 2C_\gamma} \right)\right| \leq \\ &2C_\gamma\mathbb{P}\left[W_F > 2C_\gamma \right] + \int_0^{C_\gamma} |F^{(\ell,\gamma)\pm}(x) - F(x)|dx + \int_{C_\gamma}^{2C_\gamma} |F^{(\ell,\gamma)\pm}(x) - F(x)|dx.\end{aligned} \tag{20}$$

We will bound $\int_0^{C_\gamma} |F^{(\ell,\gamma)+}(x) - F(x)|dx$. Firstly, we write

$$\begin{aligned}&\int_0^{C_\gamma} |F^{(\ell,\gamma)+}(x) - F(x)|dx = \sum_{i=0}^{p_\ell} \int_{y_i}^{y_{i+1}} |F^{(\ell,\gamma)+}(x) - F(x)|dx \\ &= \sum_{i=0}^{t_\ell} \int_{y_{2i}}^{y_{2i+1}} |F^{(\ell,\gamma)+}(x) - F(x)|dx + \sum_{i=0}^{t_\ell-1} \int_{y_{2i+1}}^{y_{2(i+1)}} |F^{(\ell,\gamma)+}(x) - F(x)|dx.\end{aligned}$$

In regard with the first sum of integrals, note that for any $x \in [y_{2i}, y_{2i+1})$ we have $|F(W_{2i}^-) - F(x)| \leq F(W_{2i}^+) - F(W_{2i}^-)$. Therefore, each integrand is bounded as follows:

$$\begin{aligned} |F^{(\ell,\gamma)+}(x) - F(x)| &= \left| \frac{F(W_{2i}^-)}{1 - h_F(\gamma) + \gamma^+} - F(x) \right| = \left| \frac{F(W_{2i}^-) - F(x)(1 - h_F(\gamma) + \gamma^+)}{1 - h_F(\gamma) + \gamma^+} \right| \\ &\leq \frac{|F(W_{2i}^-) - F(x)|}{1 - h_F(\gamma) + \gamma^+} + F(x) \frac{|h_F(\gamma) - \gamma^+|}{1 - h_F(\gamma) + \gamma^+} \\ &\leq \frac{F(W_{2i}^+) - F(W_{2i}^-)}{1 - h_F(\gamma) + \gamma^+} + \frac{|h_F(\gamma) - \gamma^+|}{1 - h_F(\gamma) + \gamma^+}. \end{aligned} \tag{21}$$

Using this bound, we get

$$(1 - h_F(\gamma) + \gamma^+) \cdot \sum_{i=0}^{t_\ell} \int_{y_{2i}}^{y_{2i+1}} |F^{(\ell,\gamma)+}(x) - F(x)| dx \leq$$
$$\sum_{i=0}^{t_\ell} \left((F(W_i^+) - F(W_i^-)) \cdot \int_{y_{2i}}^{y_{2i+1}} dx \right) + |h_F(\gamma) - \gamma^+| \sum_{i=0}^{t_\ell} \cdot \int_{y_{2i}}^{y_{2i+1}} dx.$$

But $\int_{y_{2i}}^{y_{2i+1}} dx = y_{2i+1} - y_{2i} \leq 2\varepsilon_\ell$. So

$$\sum_{i=0}^{t_\ell} \left((F(W_i^+) - F(W_i^-)) \cdot \int_{y_{2i}}^{y_{2i+1}} dx \right) \leq 2\varepsilon_\ell \cdot \sum_{i=0}^{t_\ell} \big((F(W_i^+) - F(W_i^-)) \big) \leq 2\varepsilon_\ell.$$

Furthermore,

$$|h_F(\gamma) - \gamma^+| \sum_{i=0}^{t_\ell} \int_{y_{2i}}^{y_{2i+1}} dx \leq C_\gamma \cdot |h_F(\gamma) - \gamma^+| \leq C_\gamma \cdot |h_F(\gamma) - \gamma^+|.$$

We thus deduce that

$$\sum_{i=0}^{t_\ell} \int_{y_{2i}}^{y_{2i+1}} |F^{(\ell,\gamma)+}(x) - F(x)| dx \leq \frac{2\varepsilon_\ell + C_\gamma |h_F(\gamma) - \gamma^+|}{1 - h_F(\gamma) + \gamma^+}.$$

In regards with the second integral, note that for $x \in [y_{2i+1}, y_{2(i+1)})$ we have $|F^{(\ell,\gamma)+}(W_{2i+1}^-) - F(x)| \leq \varepsilon_\ell$. Arguing as in (21), we deduce that

$$|F^{(\ell,\gamma)+}(x) - F(x)| \leq \frac{\varepsilon_\ell + |h_F(\gamma) - \gamma^+|}{1 - h_F(\gamma) + \gamma^+}.$$

Thereby,

$$\sum_{i=0}^{t_\ell - 1} \int_{y_{2i+1}}^{y_{2(i+1)}} |F^{(\ell,\gamma)+}(x) - F(x)| dx \leq \frac{C_\gamma (\varepsilon_\ell + |h_F(\gamma) - \gamma^+|)}{1 - h_F(\gamma) + \gamma^+}.$$

We thus conclude that

$$\int_0^{C_\gamma} |F^{(\ell,\gamma)+}(x) - F(x)| dx \leq \frac{\varepsilon_\ell (C_\gamma + 2) + 2C_\gamma |h_F(\gamma) - \gamma^+|}{1 - h_F(\gamma) + \gamma^+} =: \hat{\rho}^+(\ell, \gamma). \tag{22}$$

One can similarly show that

$$\int_0^{C_\gamma} |F^{(\ell,\gamma)-}(x) - F(x)|dx \quad \leq \quad \frac{\varepsilon_\ell(C_\gamma + 2) + 2C_\gamma h_F(\gamma)}{1 - h_F(\gamma) + ph_F(\gamma)} =: \hat{\rho}^-(\ell,\gamma). \tag{23}$$

Note that $\lim_{\gamma\downarrow 0}\lim_{\ell\to\infty}\hat{\rho}^{\pm}(\ell,\gamma) = 0$.

Finally, we will argue about $\int_{C_\gamma}^{2C_\gamma}|F^{(\ell,\gamma)+}(x) - F(x)|dx$. For any $x \in [C_\gamma, 2C_\gamma]$ that is a point of continuity we have

$$\begin{aligned}|F^{(\ell,\gamma)+}(x) - F(x)| &= \left|\frac{F(x)}{1 - h_F(\gamma) + \gamma^+} - F(x)\right| = \left|\frac{F(x) - F(x)(1 - h_F(\gamma) + \gamma^+)}{1 - h_F(\gamma) + \gamma^+}\right| \\ &\leq F(x)\frac{|h_F(\gamma) - \gamma^+|}{1 - h_F(\gamma) + \gamma^+} \leq \frac{|h_F(\gamma) - \gamma^+|}{1 - h_F(\gamma) + \gamma^+}.\end{aligned}$$

Therefore:

$$\begin{aligned}&\left|\int_{C_\gamma}^{2C_\gamma}(F^{(\ell,\gamma)+}(x) - F(x))dx\right| \leq \int_{C_\gamma}^{2C_\gamma}|F^{(\ell,\gamma)+}(x) - F(x)|dx \leq \\ &\leq \frac{|h_F(\gamma) - \gamma^+|}{1 - h_F(\gamma) + \gamma^+}\cdot(2C_\gamma - C_\gamma) = C_\gamma\cdot\frac{|h_F(\gamma) - \gamma^+|}{1 - h_F(\gamma) + \gamma^+} \leq \hat{\rho}^+(\ell,\gamma).\end{aligned} \tag{24}$$

Also,

$$\begin{aligned}&\int_{C_\gamma}^{2C_\gamma}\left|F^{(\ell,\gamma)-}(x) - F(x)\right|dx = \int_{C_\gamma}^{2C_\gamma}|1 - F(x)|\,dx = \int_{C_\gamma}^{2C_\gamma}\mathbb{P}\left[\,W_F > x\,\right]dx \\ \leq\; &\int_{C_\gamma}^{\infty}\mathbb{P}\left[\,W_F \geq x\,\right]dx = \mathbb{E}\left(W_F\mathbf{1}_{W_F\geq C_\gamma}\right).\end{aligned} \tag{25}$$

The latter $\downarrow 0$ as $\gamma \downarrow 0$.

Now, we substitute the bounds of (22), (23), (24),(25) into (20), we get

$$\begin{aligned}&\left|\mathbb{E}\left(U^{(\ell,\gamma)\pm}\mathbf{1}_{U^{(\ell,\gamma)\pm}\leq 2C_\gamma}\right) - \mathbb{E}\left(W_F\mathbf{1}_{W_F\leq 2C_\gamma}\right)\right| \leq \\ &\qquad 2C_\gamma\mathbb{P}\left[\,W_F > 2C_\gamma\,\right] + 2\hat{\rho}^{\pm}(\ell,\gamma) + \mathbb{E}\left(W_F\mathbf{1}_{W_F\geq C_\gamma}\right).\end{aligned} \tag{26}$$

Using the upper bound of (26) into (17) and that $\mathbb{E}\left(W_F\mathbf{1}_{W_F>2C_\gamma}\right) \leq \mathbb{E}\left(W_F\mathbf{1}_{W_F\geq C_\gamma}\right)$ we finally get

$$\begin{aligned}\left|\mathbb{E}\left(U^{(\ell,\gamma)\pm}\right) - \mathbb{E}\left(W_F\right)\right| &\leq 2C_\gamma\mathbb{P}\left[\,W_F > 2C_\gamma\,\right] + 2\hat{\rho}^{\pm}(\ell,\gamma) + 2\mathbb{E}\left(W_F\mathbf{1}_{W_F\geq C_\gamma}\right) \\ &\leq 2\hat{\rho}^{\pm}(\ell,\gamma) + 3\mathbb{E}\left(W_F\mathbf{1}_{W_F\geq C_\gamma}\right).\end{aligned} \tag{27}$$

Another useful bound which can be deduced by (18) and (19) replacing $2C_\gamma$ with C_γ is

$$\begin{aligned}&\left|\mathbb{E}\left(U^{(\ell,\gamma)\pm}\mathbf{1}_{U^{(\ell,\gamma)\pm}\leq C_\gamma}\right) - \mathbb{E}\left(W_F\right)\right| \\ &\leq \left|\mathbb{E}\left(U^{(\ell,\gamma)\pm}\mathbf{1}_{U^{(\ell,\gamma)\pm}\leq C_\gamma}\right) - \mathbb{E}\left(W_F\mathbf{1}_{W_F\leq C_\gamma}\right)\right| + \mathbb{E}\left(W_F\mathbf{1}_{W_F>C_\gamma}\right) \\ &\leq C_\gamma\left|\mathbb{P}\left[U^{(\ell,\gamma)\pm} \leq C_\gamma\right] - \mathbb{P}\left[\,W_F \leq C_\gamma\,\right]\right| + \int_0^{C_\gamma}|F^{(\ell,\gamma)\pm}(x) - F(x)|dx + \mathbb{E}\left(W_F\mathbf{1}_{W_F>C_\gamma}\right).\end{aligned}$$

But note that $\mathbb{P}\left[\,U^{(\ell,\gamma)-} \leq C_\gamma\,\right] = 1$ whereas $\mathbb{P}\left[\,U^{(\ell,\gamma)+} \leq C_\gamma\,\right] = \frac{F(C_\gamma)}{1-h_F(\gamma)+\gamma^+}$. So

$$\left|\mathbb{P}\left[\,U^{(\ell,\gamma)+} \leq C_\gamma\,\right] - \mathbb{P}\left[\,W_F \leq C_\gamma\,\right]\right| = \frac{F(C_\gamma)|h_F(\gamma)-\gamma^+|}{1-h_F(\gamma)+\gamma^+}$$

and

$$\left|\mathbb{P}\left[\,U^{(\ell,\gamma)-} \leq C_\gamma\,\right] - \mathbb{P}\left[\,W_F \leq C_\gamma\,\right]\right| = \mathbb{P}\left[\,W_F > C_\gamma\,\right]$$

Thus,

$$\begin{aligned}
&\left|\mathbb{E}\left(\,U^{(\ell,\gamma)\pm}\mathbf{1}_{U^{(\ell,\gamma)\pm}\leq C_\gamma}\,\right) - \mathbb{E}\left(\,W_F\,\right)\right| \\
&\leq \quad C_\gamma \frac{F(C_\gamma)|h_F(\gamma)-\gamma^+|}{1-h_F(\gamma)+\gamma^+} + \int_0^{C_\gamma} |F^{(\ell,\gamma)\pm}(x) - F(x)|dx + \mathbb{E}\left(\,W_F\mathbf{1}_{W_F>C_\gamma}\,\right) \\
&\overset{(22),(23)}{\leq} \quad C_\gamma \frac{|h_F(\gamma)-\gamma^+|}{1-h_F(\gamma)+\gamma^+} + \hat{\rho}^{\pm}(\ell,\gamma) + \mathbb{E}\left(\,W_F\mathbf{1}_{W_F>C_\gamma}\,\right),
\end{aligned} \tag{28}$$

whereas

$$\left|\mathbb{E}\left(\,U^{(\ell,\gamma)+}\mathbf{1}_{U^{(\ell,\gamma)\pm}\leq C_\gamma}\,\right) - \mathbb{E}\left(\,W_F\,\right)\right| \leq C_\gamma \mathbb{P}\left[\,W_F > C_\gamma\,\right] + \hat{\rho}^{\pm}(\ell,\gamma) + \mathbb{E}\left(\,W_F\mathbf{1}_{W_F>C_\gamma}\,\right). \tag{29}$$

We conclude that (27) together with (5), (28) and (29) complete Part 2 of Definition 3.2.

3.1.2 Bounds on $|\mathcal{A}_f|$

For a subset $S \subseteq [n]$, let $\mathcal{A}_f(S)$ denote the final set of infected vertices in $CL(\mathbf{w})$ assuming that $\mathcal{A}_0 = S$. With this notation we have of course that $\mathcal{A}_f = \mathcal{A}_f(\mathcal{A}_0)$. We also set $\mathcal{A}_f^-(S)$ to be the set of infected vertices in $CL'(\mathbf{W}^{(\ell,\gamma)-})$, respectively, assuming that the initial set is $\varphi^-(S \cap \cup_{i=1}^{p_\ell} \mathsf{C}_i)$. Finally, for a subset $S \subseteq [n']$ let $\mathcal{A}_f^+(S)$ the final set of infected vertices on $CL'(\mathbf{W}^{(\ell,\gamma)+})$. We will show the following.

Claim 3.7. *Let $p \in (0,1)$. Assume that $\mathcal{A}_0$ is a random subset of $[n]$ where each vertex is included with probability p independently of any other vertex. Then there is a coupling space on which w.h.p.*

$$|\mathcal{A}_f^-(\mathcal{A}_0 \cup \mathsf{C}_\gamma^-)| \leq |\mathcal{A}_f| \leq |\mathcal{A}_f^+(\mathcal{A}_0 \cup \mathsf{C}_\gamma^+)|. \tag{30}$$

In the proof of Claim 3.7 and elsewhere we write that the probability that two vertices k and j, with $w_j(n) \leq C_\gamma$, are adjacent is equal to $w_k(n)w_j(n)/W_{[n]}$; in other words, when we apply (1) we tacitly assume that n is sufficiently large so that this ratio is less than 1.

Proof of Claim 3.7. As $\mathcal{A}_0$ is formed by including every vertex in $[n]$ independently with probability p, it follows that w.h.p. at least k_- elements of C_γ become initially infected. We identify exactly k_- of them with the set C_γ^-. Recall for each $k \in \cup_{i=1}^{p_\ell} \mathsf{C}_i$ we have $W_{\varphi^-(k)}^{(\ell,\gamma)-}(n) \leq w_k(n)$. This implies that for each pair $k, k' \in [n] \setminus \mathsf{C}_\gamma$ of distinct vertices, the probability that $\varphi^-(k)$ and $\varphi^-(k')$ are adjacent is smaller in $CL'(\mathbf{W}^{(\ell,\gamma)-})$ than the corresponding probability about k and k' in $CL(\mathbf{w})$. Hence, there is a coupling space on which

$$CL'(\mathbf{W}^{(\ell,\gamma)-}) \subseteq CL(\mathbf{w}),$$

and the first inequality in (30) follows. The second inequality follows from a slightly more involved argument. Let $j \in \mathsf{C}_\gamma$ be such that $w_j(n) \geq 2C_\gamma$ and let $k \in \cup_{i=1}^{p_\ell} \mathsf{C}_i$. The probability that k is adjacent to j in $CL(\mathbf{w})$ is equal to $w_k w_j / W_{[n]}$. Also, the probability that k is adjacent to at least one of the copies of j in $[n']$ in the random graph $CL'(\mathbf{W}^{(\ell,\gamma)+})$ is

$$1 - \left(1 - \frac{2w_k C_\gamma}{W_{[n]}}\right)^{2\lfloor w_j/C_\gamma \rfloor}.$$

Assume that we show that for n sufficiently large we have that for any $k \in \cup_{i=1}^{p_\ell} \mathsf{C}_i$ and any $j \in \mathsf{C}_\gamma$

$$\frac{w_k w_j}{W_{[n]}} \leq 1 - \left(1 - \frac{2w_{\varphi^+(k)} C_\gamma}{W_{[n]}}\right)^{2\lfloor w_j/C_\gamma \rfloor}. \tag{31}$$

Moreover, assume that every vertex in C'_γ is among those vertices that are initially infected. Now, observe that there is coupling space in which we have

$$CL(\mathbf{w})[\cup_{i=1}^{p_\ell} \mathsf{C}_i] \subseteq CL'(\mathbf{W}^{(\ell,\gamma)+})[\cup_{i=1}^{p_\ell} \mathsf{C}_i]. \tag{32}$$

This is the case, since for any $k \in \cup_{i=1}^{p_\ell} \mathsf{C}_i$ we have $w_k(n) \leq W_{\varphi^+(k)}^{(\ell,\gamma)+}(n')$. Consider a vertex $k \in \cup_{i=1}^{p_\ell} \mathsf{C}_i$ and now let $j \in \mathsf{C}_\gamma$. Now, Inequality (31) implies that the probability that k is adjacent to j in $CL(\mathbf{w})$ is at most the probability that $\varphi^+(k)$ is adjacent to at least one of the copies of j in $[n']$ within $CL'(\mathbf{W}^{(\ell,\gamma)+})$. Thereby, it follows that the number of neighbours of k in C_γ in the random graph $CL(\mathbf{w})$ is stochastically dominated by the size of the neighbourhood of k in C'_γ in the random graph $CL'(\mathbf{W}^{(\ell,\gamma)+})$. This observation together with (32) imply that

$$|\mathcal{A}_f(\mathcal{A}_0 \cup \mathsf{C}_\gamma)| \leq_{st} |\mathcal{A}_f^+(\mathcal{A}_0 \cup \mathsf{C}_\gamma^+)|.$$

But also,

$$|\mathcal{A}_f| \leq_{st} |\mathcal{A}_f(\mathcal{A}_0 \cup \mathsf{C}_\gamma)|.$$

The second stochastic inequality of the claim follows from the above two inequalities. It remains to show (31). Using the Bonferroni inequalities we have

$$1 - \left(1 - \frac{2w_{\varphi^+(k)} C_\gamma}{W_{[n]}}\right)^{2\lfloor w_j/C_\gamma \rfloor} \geq 2\lfloor \frac{w_j}{C_\gamma} \rfloor \frac{2w_{\varphi^+(k)} C_\gamma}{W_{[n]}} - 2\left(w_j/C_\gamma\right)^2 \ \frac{4w_{\varphi^+(k)}^2 C_\gamma^2}{W_{[n]}^2}. \tag{33}$$

But

$$2\lfloor \frac{w_j}{C_\gamma} \rfloor \frac{2w_{\varphi^+(k)} C_\gamma}{W_{[n]}} \geq 2\left(\frac{w_j}{C_\gamma} - 1\right) \frac{2w_{\varphi^+(k)} C_\gamma}{W_{[n]}} = 2\frac{w_j}{C_\gamma}\left(1 - \frac{C_\gamma}{w_j}\right) \frac{2w_{\varphi^+(k)} C_\gamma}{W_{[n]}} \overset{w_j/C_\gamma \geq 2}{\geq} \frac{2w_{\varphi^+(k)} w_j}{W_{[n]}}.$$

Substituting this lower bound into (33) we obtain

$$\begin{aligned} 1 - \left(1 - \frac{2w_{\varphi^+(k)} C_\gamma}{W_{[n]}}\right)^{2\lfloor \frac{w_j}{C_\gamma} \rfloor} &\geq \frac{2w_{\varphi^+(k)} w_j}{W_{[n]}} - \frac{8w_{\varphi^+(k)}^2 w_j^2}{W_{[n]}^2} = \frac{2w_{\varphi^+(k)} w_j}{W_{[n]}}\left(1 - \frac{4w_{\varphi^+(k)} w_j}{W_{[n]}}\right) \\ &> \frac{w_{\varphi^+(k)} w_j}{W_{[n]}} \geq \frac{w_k w_j}{W_{[n]}}, \end{aligned}$$

for n sufficiently large, as $w_k < C_\gamma$ and $w_j = w_j(n) = o(n)$ (uniformly for all j) but $W_{[n]} = \Theta(n)$. □

We will now apply Theorem 3.4 to the random variables that bound $|\mathcal{A}_f|$ in Claim 3.7. Theorem 3.4 implies that there exists $\gamma_2 > 0$ satisfying the following: for any $\gamma < \gamma_2$ and any $\delta \in (0,1)$ there exists an infinite set of natural numbers $\mathcal{S}^1$ such that for every $\ell \in \mathcal{S}^1$ with probability $1 - o(1)$

$$n^{-1}|\mathcal{A}_f^+(\mathcal{A}_0 \cup \mathsf{C}_\gamma^+)| \leq (1+\delta)((1-p)\mathbb{E}[\,\psi_r(W_F\hat{y})\,] + p), \tag{34}$$

and an infinite set of natural numbers $\mathcal{S}^2$ such that for every $\ell \in \mathcal{S}^2$ with probability $1 - o(1)$

$$n^{-1}|\mathcal{A}_f^-(\mathcal{A}_0 \cup \mathsf{C}_\gamma^-)| \geq (1-\delta)((1-p)\mathbb{E}[\,\psi_r(W_F\hat{y})\,] + p). \tag{35}$$

Hence, Claim 3.7 together with (34) and (35) imply the following w.h.p. bounds on the size of $\mathcal{A}_f$:

$$n^{-1}|\mathcal{A}_f| = (1 \pm \delta)((1-p)\mathbb{E}[\,\psi_r(W_F\hat{y})\,] + p),$$

whereby Theorem 2.2 follows.

3.2 Proof of Theorem 2.4

Let us assume that $\mathcal{A}_0$ is randomly selected, including each vertex independently with probability $a(n)/n$, where $a(n) \gg a_c(n)$ but $a(n) = o(n)$ (cf. Theorem 2.4 for the definition of the function $a_c(n)$). For $\varepsilon \in (0,1)$ let $\mathcal{A}_0^{(\varepsilon)}$ denote a random subset of $[n]$ where each vertex is included independently with probability ε. If n is large enough, then $\mathcal{A}_0$ can be coupled with $\mathcal{A}_0^{(\varepsilon)}$, that is, there is a coupling space in which $\mathcal{A}_0 \subseteq \mathcal{A}_0^{(\varepsilon)}$. The following stochastic upper bound can be deduced as in Claim 3.7.

Claim 3.8. *For any $\varepsilon \in (0,1)$ and any $\gamma > 0$, if n is large enough, then*

$$|\mathcal{A}_f| \leq_{st} |\mathcal{A}_f(\mathcal{A}_0^{(\varepsilon)} \cup \mathsf{C}_\gamma)| \leq_{st} |\mathcal{A}_f^+(\mathcal{A}_0^{(\varepsilon)} \cup \mathsf{C}_\gamma^+)|.$$

We will now deduce a stochastic lower bound on $|\mathcal{A}_f|$. For $C > 0$, let $\mathcal{K}_C$ denote the set of vertices having weight at least C in $\mathbf{w}$. In [6] the first two authors prove that if $\varepsilon \in (0,1)$ is sufficiently small and $\mathcal{A}_0$ is selected as above, then at least a $(1-\varepsilon)$-fraction of the vertices of $\mathcal{K}_C$ become infected if we consider a bootstrap percolation process on $CL(\mathbf{w})$ with activation threshold r where the vertices in $[n] \setminus \mathcal{K}_C$ are assumed to be "frozen", that is, they never get infected.

Lemma 3.9 (Proposition 3.7 [6]). *There exists an $\varepsilon_0 = \varepsilon_0(\beta, c_1, c_2) > 0$ such that for any positive $\varepsilon < \varepsilon_0$ there exists $C = C(c_1, c_2, \beta, \varepsilon, r) > 0$ for which the following holds. Assume that $\mathcal{A}_0$ is as above and consider a bootstrap percolation process on $CL(\mathbf{w})$ with activation threshold $r \geq 2$ and the set $\mathcal{A}_0$ as the initial set, with the restriction that the vertices in $[n] \setminus \{\mathcal{K}_C \cup \mathcal{A}_0\}$ never become infected. Then at least $(1-\varepsilon)|\mathcal{K}_C|$ vertices of $\mathcal{K}_C$ become infected with probability $1 - o(1)$.*

Lemma 3.9 implies that for any $\varepsilon > 0$ that is sufficiently small there exists $C = C(c_1, c_2, \beta, \varepsilon, r) > 0$ such that with probability $1 - o(1)$ at least $(1-\varepsilon)|\mathcal{K}_C|$ vertices of $\mathcal{K}_C$ will be infected in $CL(\mathbf{w})$, assuming that the vertices in $[n] \setminus \{\mathcal{K}_C \cup \mathcal{A}_0\}$ never become infected. Let $\mathcal{E}_{C,\varepsilon,n}$ denote this event and, if it is realised, we let $\mathcal{K}_{C,\varepsilon}$ denote a subset of $\lfloor (1-\varepsilon)|\mathcal{K}_C| \rfloor =: k$ vertices in $\mathcal{K}_C$ that become infected chosen in some particular way (for example, the k lexicographically smallest vertices). Hence, the following holds.

Claim 3.10. *For any $C > 0$ and any $\varepsilon \in (0,1)$, there is a coupling such that if $\mathcal{E}_{C,\varepsilon,n}$ is realised, then we have*

$$\mathcal{A}_f(\mathcal{K}_{C,\varepsilon}) \subseteq \mathcal{A}_f.$$

Let $\gamma \in F([0,\infty))$ be such that $C_\gamma = C$, where $C = C(\varepsilon)$ is as in Lemma 3.9. (Under the assumptions of Theorem 2.4, F is continuous (cf. Definition 2.3), and therefore $h_F(\gamma) = \gamma$.)

Consider a set of vertices $\{v_1, \ldots, v_k\}$ which is disjoint from $[n]$. We define a sequence $\tilde{\mathbf{W}}^{(\ell,\gamma)-}$ on $(\cup_{i=1}^{p_\ell} \mathsf{C}_i) \bigcup \{v_1, \ldots, v_k\}$ as follows. For every $j \in \mathsf{C}_i$, with $i = 1, \ldots, p_\ell$, we have $\tilde{W}_j^{(\ell,\gamma)-} = W_j^{(\ell,\gamma)-}$, whereas for every $j = 1, \ldots, k$ we let $\tilde{W}_{v_j}^{(\ell,\gamma)-} = C_\gamma$. We let n_- be the number of vertices of the sequence $\tilde{\mathbf{W}}^{(\ell,\gamma)-}$, that is, the size of $(\cup_{i=1}^{p_\ell} \mathsf{C}_i) \bigcup \{v_1, \ldots, v_k\}$. Since $k = (1-\varepsilon)\gamma n(1+o(1))$, this satisfies $n_- = ((1-\gamma) + \gamma(1-\varepsilon))n(1+o(1)) = (1-\gamma\varepsilon)n(1+o(1))$. Hence, for large n we have $n_- < n$. We identify the vertices in $\{v_1, \ldots, v_k\}$ with the lexicographically k first vertices in C_γ and we denote both subsets by $\mathsf{C}_{\gamma,k}$. Setting $\tilde{W}_\gamma^- := (1-\varepsilon)\gamma C_\gamma$, the weight of these vertices is $n\tilde{W}_\gamma^-(1+o(1))$, since each of them has weight equal to C_γ.

The weight sequence $\tilde{\mathbf{W}}^{(\ell,\gamma)-}$ gives rise to a probability distribution which is the limiting probability distribution of the weight of a uniformly chosen vertex from $[n_-]$. We let $\tilde{U}^{(\ell,\gamma)-}$ be a random variable which follows this distribution and let $\tilde{W}_F^{(\ell,\gamma)-}$ denote a random variable which follows the $\tilde{U}^{(\ell,\gamma)-}$ size-biased distribution. The definition of $\tilde{\mathbf{W}}^{(\ell,\gamma)-}$ yields

$$\mathbb{P}\left[\tilde{U}^{(\ell,\gamma)-} = W_i^{(\ell,\gamma)-}\right] = \frac{\gamma_i}{1-\gamma\varepsilon}, \text{ and } \mathbb{P}\left[\tilde{U}^{(\ell,\gamma)-} = C_\gamma\right] = \frac{(1-\varepsilon)\gamma}{1-\gamma\varepsilon}.$$

As we did in Section 3.1.1 for the sequence $\{\mathbf{W}^{(\ell,\gamma)-}(n)\}_{\gamma\in(0,1),\ell\in\mathbb{N}}$, one can show that $\tilde{\mathbf{W}}^{(\ell,\gamma)-}$ is an F-convergent weight sequence. We omit the proof.

Let $\hat{\mathcal{A}}_f(\mathsf{C}_{\gamma,k})$ be the final set of infected vertices in $CL(\mathbf{w})$ assuming that the initial set is $\mathsf{C}_{\gamma,k}$ and moreover no vertices in $\mathsf{C}_\gamma \setminus \mathsf{C}_{\gamma,k}$ ever become infected. Hence, on the event $\mathcal{E}_{C_\gamma,\varepsilon,n}$ we have

$$|\hat{\mathcal{A}}_f(\mathsf{C}_{\gamma,k})| \leq_{st} |\mathcal{A}_f(\mathcal{K}_{C_\gamma,\varepsilon})|.$$

(The symbol $\leq_{st}$ denotes stochastic domination.) But the assumption that no vertices in $\mathsf{C}_\gamma \setminus \mathsf{C}_{\gamma,k}$ ever become active amounts to a bootstrap percolation process on $CL'(\tilde{\mathbf{W}}^{(\ell,\gamma)-})$ with activation threshold equal to r. Let $\tilde{\mathcal{A}}_f(S)$ denote the final set under the assumption that the initial set is $S \subseteq [n']$. Since $CL'(\tilde{\mathbf{W}}^{(\ell,\gamma)-}) \subseteq CL(\mathbf{w})$ on a certain coupling space we have

$$|\tilde{\mathcal{A}}_f(\mathsf{C}_{\gamma,k})| \leq_{st} |\hat{\mathcal{A}}_f(\mathsf{C}_{\gamma,k})|.$$

Therefore

$$|\tilde{\mathcal{A}}_f(\mathsf{C}_{\gamma,k})| \leq_{st} |\mathcal{A}_f(\mathcal{K}_{C_\gamma,\varepsilon})|.$$

This together with Claim 3.10 imply the following stochastic lower bound on $|\mathcal{A}_f|$.

Claim 3.11. *For any $\gamma, \varepsilon \in (0,1)$, if $\mathcal{E}_{C_\gamma,\varepsilon,n}$ is realised, then*

$$|\tilde{\mathcal{A}}_f(\mathsf{C}_{\gamma,k})| \leq_{st} |\mathcal{A}_f|.$$

We will now apply Theorem 3.4 to the random variables that bound $|\mathcal{A}_f|$ in Claims 3.8 and 3.11. Let $\hat{y}_\varepsilon^+, \hat{y}$ be the smallest positive solutions of

$$y = (1-\varepsilon)\,\mathbb{E}\left[\psi_r\left(W_F^* y\right)\right] + \varepsilon,$$

and

$$y = \mathbb{E}\left[\, \psi_r\left(W_F^* y\right)\right],$$

respectively.

For $\varepsilon < \varepsilon_0$ let C be as in Lemma 3.9 and let $\gamma < \gamma_2$ (cf. Theorem 3.4) be such that $C = C_\gamma$. Theorem 3.4 implies that for any $\delta \in (0,1)$ there exists an infinite set of natural numbers $\mathcal{S}^1$ such that for every $\ell \in \mathcal{S}^1$ with probability $1 - o(1)$

$$\frac{|\mathcal{A}_f^+(\mathcal{A}_0^{(\varepsilon)} \cup \mathsf{C}_\gamma^+)|}{n} \le (1+\delta)((1-\varepsilon)\mathbb{E}\left[\,\psi_r(W_F \hat{y}_\varepsilon^+)\,\right] + \varepsilon), \tag{36}$$

and an infinite set of natural numbers $\mathcal{S}^2$ such that for every $\ell \in \mathcal{S}^2$ with probability $1 - o(1)$

$$\frac{|\tilde{\mathcal{A}}_f(\mathsf{C}_{\gamma,k})|}{n} \ge (1-\delta)\mathbb{E}\left[\,\psi_r(W_F \hat{y})\,\right] \tag{37}$$

Hence, Claims 3.8 and 3.11 together with (36) and (37) imply that w.h.p.

$$\frac{|\mathcal{A}_f|}{n} \le (1+\delta)((1-\varepsilon)\mathbb{E}\left[\,\psi_r(W_F \hat{y}_\varepsilon^+)\,\right] + \varepsilon),$$

and

$$\frac{|\mathcal{A}_f|}{n} \ge (1-\delta)\mathbb{E}\left[\,\psi_r(W_F \hat{y})\,\right].$$

But $y_\varepsilon^+ \to \hat{y}$ as $\varepsilon \to 0$ and Theorem 2.4 follows.

4 Proof of Theorem 3.4

In this section we will give the proof of Theorem 3.4. At the moment our analysis does not depend on the parameters ℓ, γ and, to simplify notation, we will drop the superscript (ℓ, γ). For $j = 0, \ldots, r-1$, we denote by $\mathsf{C}_{i,j}$ the subset of C_i which consists of those vertices of C_i which have j infected neighbours. We also denote by $\mathsf{C}_{i,r}$ the subset of C_i containing all those vertices that are infected, that is, they have *at least* r infected neighbours or are initially infected.

We will determine the size of the final set of infected vertices exposing *sequentially* the neighbours of each infected vertex and keeping track of the number of infected neighbours an uninfected vertex has. In other words, we will be keeping track of the size of the sets $\mathsf{C}_{i,j}$. This method of exposure has also been applied in the analysis in [26]. However, the inhomogeneity in the present context bears additional difficulties as the evolutions of the sets $\mathsf{C}_{i,j}$ are interdependent.

The *sequential* exposure proceeds as follows. For $i = 1, \ldots, p_\ell$ and $j = 0, \ldots, r-1$, let $\mathsf{C}_{i,j}(t)$ denote set of vertices which have j infected neighbours after the execution of the tth step . We also denote by $\mathsf{C}_{i,r}(t)$ the set of all those vertices that have *at least* r infected neighbours after tth step.

Here $\mathsf{C}_{i,j}(0)$ denotes the set $\mathsf{C}_{i,j}$ before the beginning of the execution. Furthermore, let $\mathsf{U}(t)$ denote the set of infected *unexposed* vertices after the execution of the tth step, with $\mathsf{U}(0)$ denoting the set of infected vertices before the beginning of the process.

At step $t \ge 1$, if $\mathsf{U}(t-1)$ is non-empty,

i. choose a vertex v uniformly at random from $\mathsf{U}(t-1)$;

ii. expose the neighbours v in the set $\bigcup_{i=1}^{p_\ell} \cup_{j=0}^{r-1} \mathsf{C}_{i,j}(t-1)$;

iii. set $\mathsf{U}(t) := \mathsf{U}(t-1) \setminus \{v\}$.

The above set of steps is repeated for as long as the set U is non-empty. The exposure of the neighbours of v can be alternatively thought of as a random assignment of a mark to each vertex of $\bigcup_{i=1}^{p_\ell} \cup_{j=0}^{r-1} \mathsf{C}_{i,j}(t-1)$ independently of every other vertex; if a vertex in $\mathsf{C}_{i,j}(t-1)$ receives such a mark, then it is moved to $\mathsf{C}_{i,j+1}(t)$. Hence, during the execution of the tth step each vertex in $\mathsf{C}_{i,j}(t-1)$ either remains a member of $\mathsf{C}_{i,j}(t)$ or it is moved to $\mathsf{C}_{i,j+1}(t)$.

4.1 Conditional Expected Evolution

Let $c_{i,j}$ denote the size of the set $\mathsf{C}_{i,j}$ for all $i = 1, \ldots, p_\ell$ and $j = 0, \ldots, r-1$. Our equations will also incorporate the size of U at time $t-1$, which we denote by $u(t-1)$, as well as the total weight of vertices in $\mathsf{U}(t-1)$, which we denote by $w_{\mathsf{U}}(t-1)$. For these values of i and j we let $\mathbf{c}(t) = (u(t), w_{\mathsf{U}}(t), (c_{i,j}(t))_{i,j})$. This vector determines the state of the process after step t. We will now give the expected change of $c_{i,j}$ during the execution of step t, conditional on $\mathbf{c}(t-1)$. If step t is to be executed, it is necessary to have $u(t-1) > 0$, which we will assume to be the case. We begin with $c_{i,0}$, for $i = 1, \ldots, p_\ell$, having

$$\begin{aligned}\mathbb{E}\left[\, c_{i,0}(t) - c_{i,0}(t-1) \mid \mathbf{c}(t-1) \,\right] &= -c_{i,0}(t-1) \sum_{v \in \mathsf{U}(t-1)} \frac{W_i w_v}{W_{[n]}} \, \frac{1}{u(t-1)} \\ &= -c_{i,0}(t-1) \, \frac{W_i}{W_{[n]}} \, \frac{w_{\mathsf{U}}(t-1)}{u(t-1)}.\end{aligned} \tag{38}$$

The evolution of $c_{i,j}$ for $0 < j < r$ involves a term that accounts for the "losses" from the set $c_{i,j}$ as well as a term which describes the expected "gain" from the set $c_{i,j-1}$. For $i = 1, \ldots, p_\ell$ and $0 < j < r$ we have

$$\begin{aligned}&\mathbb{E}\left[\, c_{i,j}(t) - c_{i,j}(t-1) \mid \mathbf{c}(t-1) \,\right] \\ =&c_{i,j-1}(t-1) \sum_{v \in \mathsf{U}(t-1)} \frac{W_i w_v}{W_{[n]}} \, \frac{1}{u(t-1)} - c_{i,j}(t-1) \sum_{v \in \mathsf{U}(t-1)} \frac{W_i w_v}{W_{[n]}} \, \frac{1}{u(t-1)} \\ =&(c_{i,j-1}(t-1) - c_{i,j}(t-1)) \, \frac{W_i}{W_{[n]}} \, \frac{w_{\mathsf{U}}(t-1)}{u(t-1)}.\end{aligned} \tag{39}$$

Finally, we will need to describe the expected change in the size of U during step t. In this case, *one* vertex is removed from $\mathsf{U}(t-1)$, but additional vertices may arrive from the sets $\mathsf{C}_{i,r-1}(t-1)$. More specifically, we write

$$\begin{aligned}\mathbb{E}\left[\, u(t) - u(t-1) \mid \mathbf{c}(t-1) \,\right] &= -1 + \sum_{i=1}^{p_\ell} c_{i,r-1}(t-1) \sum_{v \in \mathsf{U}(t-1)} \frac{W_i w_v}{W_{[n]}} \, \frac{1}{u(t-1)} \\ &= -1 + \frac{w_{\mathsf{U}}(t-1)}{u(t-1)} \sum_{i=1}^{p_\ell} \frac{W_i}{W_{[n]}} \, c_{i,r-1}(t-1).\end{aligned} \tag{40}$$

Similarly, the expected change in the weight of U during step t is as follows:

$$
\begin{aligned}
&\mathbb{E}\left[w_{\mathsf{U}}(t) - w_{\mathsf{U}}(t-1) \mid \mathbf{c}(t-1) \right] \\
&= - \frac{w_{\mathsf{U}}(t-1)}{u(t-1)} + \sum_{i=1}^{p_\ell} W_i c_{i,r-1}(t-1) \sum_{v \in \mathsf{U}(t-1)} \frac{W_i w_v}{W_{[n]}} \frac{1}{u(t-1)} \\
&= - \frac{w_{\mathsf{U}}(t-1)}{u(t-1)} + \frac{w_{\mathsf{U}}(t-1)}{u(t-1)} \sum_{i=1}^{p_\ell} \frac{W_i^2}{W_{[n]}} c_{i,r-1}(t-1).
\end{aligned}
\tag{41}
$$

4.2 Continuous Approximation

The above quantities will be approximated by the solution of a system of ordinary differential equations. We will consider a collection of continuous differentiable functions $\gamma_{i,j} : [0,\infty) \to \mathbb{R}$, for all $i = 1,\ldots,p_\ell$ and $j = 0,\ldots,r-1$, through which we will approximate the quantities $c_{i,j}$. To be more precise, $\gamma_{i,j}$ will be shown to be close to $c_{i,j}/n$. Moreover, u and w_{U} will be approximated through the continuous differentiable functions $\nu, \mu_{\mathsf{U}} : [0,\infty) \to \mathbb{R}$ in a similar way. We will also use another continuous function $G : [0,\infty) \to \mathbb{R}$ which will approximate the ratio w_{U}/u; note that this is the average weight of the set of infected unexposed vertices.

The system of differential equations that determine the functions $\gamma_{i,j}$ is as follows:

$$
\begin{aligned}
\frac{d\gamma_{i,0}}{d\tau} &= -\gamma_{i,0}(\tau) \frac{W_i}{d} G(\tau), \\
\frac{d\gamma_{i,j}}{d\tau} &= \left(\gamma_{i,j-1}(\tau) - \gamma_{i,j}(\tau)\right) \frac{W_i}{d} G(\tau), \qquad 1 \le j \le r-1.
\end{aligned}
\tag{42}
$$

The continuous counterparts of (40) and (41) are

$$
\frac{d\nu}{d\tau} = -1 + G(\tau) \sum_{i=1}^{p_\ell} \frac{W_i}{d} \gamma_{i,r-1}(\tau), \tag{43}
$$

and

$$
\frac{d\mu_{\mathsf{U}}}{d\tau} = -G(\tau) + G(\tau) \sum_{i=1}^{p_\ell} \frac{W_i^2}{d} \gamma_{i,r-1}(\tau). \tag{44}
$$

The initial conditions are

$$
\begin{aligned}
\nu(0) &= p(1 - h_F(\gamma)) + \gamma', \text{ for } p \in [0,1) \text{ (recall that } p \text{ is the initial infection rate)}, \\
\mu_{\mathsf{U}}(0) &= W'_\gamma + p \sum_{i=1}^{p_\ell} W_i \gamma_i, \\
\gamma_{i,0}(0) &= (1-p)\gamma_i, \\
\gamma_{i,j}(0) &= 0, \text{ for } j = 1,\ldots,r-1.
\end{aligned}
\tag{45}
$$

In the following proposition, we will express the formal solution of the above system in terms of $\gamma_{i,0}(\tau)$.

Proposition 4.1. *With $I(\tau) = \int_0^\tau G(s)ds$, we have*

$$
\gamma_{i,0}(\tau) = \gamma_{i,0}(0) \exp\left(-W_i I(\tau)/d\right).
$$

Moreover, for $1 \leq j \leq r-1$

$$\gamma_{i,j}(\tau) = \frac{\gamma_{i,0}(\tau)}{j!} \log^j \left(\frac{\gamma_{i,0}(0)}{\gamma_{i,0}(\tau)} \right).$$

Proof. The expression for $\gamma_{i,0}(\tau)$ can be obtained through separation of variables – we omit the details. The remaining expressions will be obtained by induction. Let us consider the differential equation for $\gamma_{i,j}$, where $0 < j < r$, assuming that we have derived the expression for $\gamma_{i,j-1}$. This differential equation is a first order ordinary differential equation of the form $y'(\tau) = a(\tau)y(\tau) + b(\tau)$ with initial condition $y(0) = 0$. Its general solution is equal to

$$y(\tau) = \exp\left(\int_0^\tau a(s)ds \right) \cdot \int_0^\tau b(s) \exp\left(-\int_0^s a(\rho)d\rho \right) ds.$$

Here, we have

$$a(\tau) = -\frac{W_i}{d} G(\tau), \; b(\tau) = \gamma_{i,j-1}(\tau) \frac{W_i}{d} G(\tau) = \frac{W_i}{d} \; \frac{\gamma_{i,0}(\tau)}{(j-1)!} \; \log^{j-1} \left(\frac{\gamma_{i,0}(0)}{\gamma_{i,0}(\tau)} \right) G(\tau),$$

by the induction hypothesis. Thereby and using the expression for $\gamma_{i,0}$ we obtain

$$\exp\left(\int_0^s a(\rho)d\rho \right) = \frac{\gamma_{i,0}(s)}{\gamma_{i,0}(0)}. \tag{46}$$

Hence

$$\begin{aligned}
&\int_0^\tau b(s) \exp\left(-\int_0^s a(\rho)d\rho \right) ds = \\
&\qquad \frac{W_i}{d(j-1)!} \int_0^\tau \gamma_{i,0}(s) \log^{j-1} \left(\frac{\gamma_{i,0}(0)}{\gamma_{i,0}(s)} \right) \; G(s) \; \frac{\gamma_{i,0}(0)}{\gamma_{i,0}(s)} ds \\
&\qquad = \gamma_{i,0}(0) \; \frac{W_i}{d(j-1)!} \int_0^\tau \gamma_{i,0}(s) \log^{j-1} \left(\frac{\gamma_{i,0}(0)}{\gamma_{i,0}(s)} \right) \; \frac{G(s)}{\gamma_{i,0}(s)} ds \\
&\qquad = -\frac{\gamma_{i,0}(0)}{(j-1)!} \int_0^\tau \frac{1}{\gamma_{i,0}(s)} \log^{j-1} \left(\frac{\gamma_{i,0}(0)}{\gamma_{i,0}(s)} \right) \left(-\gamma_{i,0}(s) \frac{W_i}{d} G(s) \right) ds \\
&\qquad \overset{(42)}{=} -\frac{1}{(j-1)!} \int_0^\tau \frac{\gamma_{i,0}(0)}{\gamma_{i,0}(s)} \log^{j-1} \left(\frac{\gamma_{i,0}(0)}{\gamma_{i,0}(s)} \right) \left(\frac{d\gamma_{i,0}}{ds} \right) ds \\
&\qquad = -\frac{\gamma_{i,0}(0)}{(j-1)!} \int_0^\tau \frac{\gamma_{i,0}(0)}{\gamma_{i,0}(s)} \log^{j-1} \left(\frac{\gamma_{i,0}(0)}{\gamma_{i,0}(s)} \right) d \left(\frac{\gamma_{i,0}}{\gamma_{i,0}(0)} \right) \\
&\qquad \overset{(x=\gamma_{i,0}/\gamma_{i,0}(0))}{=} -\frac{\gamma_{i,0}(0)}{(j-1)!} \int_1^{\gamma_{i,0}(\tau)/\gamma_{i,0}(0)} \frac{1}{x} \log^{j-1} \left(\frac{1}{x} \right) dx \\
&\qquad = (-1)^{j-1} \frac{\gamma_{i,0}(0)}{(j-1)!} \int_{\gamma_{i,0}(\tau)/\gamma_{i,0}(0)}^1 \frac{\log^{j-1}(x)}{x} dx.
\end{aligned} \tag{47}$$

For $j = 1$, the last integral equals $\log(\gamma_{i,0}(0)/\gamma_{i,0}(\tau))$. For $j \geq 2$, it can be calculated using integration by parts.

$$\int \frac{\log^{j-1}(x)}{x} dx = \int (\log(x))' \log^{j-1}(x) dx = \log^j(x) - (j-1) \int \frac{\log^{j-1}(x)}{x} dx,$$

which yields

$$\int \frac{\log^{j-1}(x)}{x} dx = \frac{\log^j(x)}{j}.$$

Thereby, the last integral in (47) is

$$\int_{\gamma_{i,0}(\tau)/\gamma_{i,0}(0)}^{1} \frac{\log^{j-1}(x)}{x} dx = -\frac{1}{j} \ \log^j \left(\frac{\gamma_{i,0}(\tau)}{\gamma_{i,0}(0)}\right) = \frac{(-1)^{j+1}}{j} \ \log^j \left(\frac{\gamma_{i,0}(0)}{\gamma_{i,0}(\tau)}\right).$$

Substituting this into (47) we obtain:

$$\int_0^\tau b(s) \exp\left(-\int_0^s a(\rho) d\rho\right) ds = \frac{\gamma_{i,0}(0)}{j!} \ \log^j \left(\frac{\gamma_{i,0}(0)}{\gamma_{i,0}(\tau)}\right). \tag{48}$$

Combining (46) and (48), we have

$$\gamma_{i,j}(\tau) = \frac{\gamma_{i,0}(\tau)}{j!} \ \log^j \left(\frac{\gamma_{i,0}(0)}{\gamma_{i,0}(\tau)}\right).$$

□

In the sequel we will use the expressions for $\gamma_{i,r-1}$, where $1 \le i \le p_\ell$, and integrate (43) in order to deduce the expressions for ν and μ_{U}.

Proposition 4.2. *We have*

$$\nu(\tau) = p \ (1 - h_F(\gamma)) + \gamma' - \tau + (1-p) \sum_{i=1}^{p_\ell} \gamma_i \mathbb{P}\left[\mathsf{Po}\left(\frac{W_i}{d} I(\tau)\right) \ge r\right]$$

and

$$\mu_{\mathsf{U}}(\tau) = W'_\gamma + p \sum_{i=1}^{p_\ell} W_i \gamma_i - I(\tau) + (1-p) \sum_{i=1}^{p_\ell} W_i \gamma_i \mathbb{P}\left[\mathsf{Po}\left(\frac{W_i}{d} I(\tau)\right) \ge r\right].$$

Proof. Applying Proposition 4.1 to (43) yiels

$$\frac{d\nu}{d\tau} = -1 + G(\tau) \sum_{i=1}^{p_\ell} \frac{W_i}{d} \ \frac{\gamma_{i,0}(\tau)}{(r-1)!} \log^{r-1}\left(\frac{\gamma_{i,0}(0)}{\gamma_{i,0}(\tau)}\right).$$

By integrating this expression we obtain

$$\begin{aligned} \nu(\tau) &= \nu(0) - \tau + \frac{1}{(r-1)!} \sum_{i=1}^{p_\ell} \int_0^\tau \frac{W_i}{d} \ \gamma_{i,0}(s) G(s) \log^{r-1}\left(\frac{\gamma_{i,0}(0)}{\gamma_{i,0}(s)}\right) ds \\ &\overset{(42)}{=} \nu(0) - \tau - \frac{1}{(r-1)!} \sum_{i=1}^{p_\ell} \int_0^\tau \left(\frac{d\gamma_{i,0}}{ds}\right) \log^{r-1}\left(\frac{\gamma_{i,0}(0)}{\gamma_{i,0}(s)}\right) ds \qquad (49) \\ &= \nu(0) - \tau - \frac{1}{(r-1)!} \sum_{i=1}^{p_\ell} \gamma_{i,0}(0) \int_1^{\gamma_{i,0}(\tau)/\gamma_{i,0}(0)} \log^{r-1}\left(\frac{1}{x}\right) dx. \end{aligned}$$

We calculate the last integral substituting y for $1/x$ and using integration by parts. We have

$$\int \log^{r-1}\left(\frac{1}{x}\right) dx = -\int \frac{\log^{r-1}(y)}{y^2} dy = \int \left(\frac{1}{y}\right)' \log^{r-1}(y) dy$$
$$= \frac{\log^{r-1}(y)}{y} - (r-1)\int \frac{\log^{r-2}(y)}{y^2} dy.$$

As $\int \frac{1}{y^2} dy = -\frac{1}{y}$, dividing and multiplying by $(r-1)!$, we obtain

$$\int \log^{r-1}\left(\frac{1}{x}\right) dx = \frac{(r-1)!}{y} \sum_{i=0}^{r-1} \frac{\log^i(y)}{i!},$$

where $y = 1/x$. Thereby, for all $i = 1, \ldots, p_\ell$ we have

$$\int_1^{\gamma_{i,0}(\tau)/\gamma_{i,0}(0)} \log^{r-1}\left(\frac{1}{x}\right) dx = (r-1)!\left(\frac{\gamma_{i,0}(\tau)}{\gamma_{i,0}(0)} \sum_{i=0}^{r-1} \frac{1}{i!} \log^i\left(\frac{\gamma_{i,0}(0)}{\gamma_{i,0}(\tau)}\right) - 1\right).$$

Substituting the above into (49) we obtain

$$\nu(\tau) = \nu(0) - \tau + \sum_{i=1}^{p_\ell} \gamma_{i,0}(0) \left(1 - \frac{\gamma_{i,0}(\tau)}{\gamma_{i,0}(0)} \sum_{j=0}^{r-1} \frac{1}{j!} \log^j\left(\frac{\gamma_{i,0}(0)}{\gamma_{i,0}(\tau)}\right)\right)$$

Observe now that the expression in brackets is equal to the probability that a Poisson distributed random variable with parameter $\log(\gamma_{i,0}(0)/\gamma_{i,0}(\tau))$ is at least r. But by Proposition 4.1, we have

$$\log\left(\frac{\gamma_{i,0}(0)}{\gamma_{i,0}(\tau)}\right) = \frac{W_i}{d} I(\tau).$$

Also, recall that by (45) $\gamma_{i,0}(0) = (1-p)\gamma_i$, for each $i = 1, \ldots, p_\ell$, and $\nu(0) = p\,(1-\gamma) + \gamma'$. Hence

$$\nu(\tau) = p\,(1-\gamma) + \gamma' - \tau + (1-p)\sum_{i=1}^{\ell} \gamma_i \mathbb{P}\left[\mathsf{Po}\left(\frac{W_i}{d} I(\tau)\right) \geq r\right].$$

The expression of μ_{U} is obtained along the same lines and we omit its proof. □

4.3 Wormald's Theorem

We summarize here the method introduced by Wormald in [38, 39] for the analysis of a discrete random process by using differential equations. Recall that a function $f(u_1, ..., u_{b+1})$ satisfies a Lipschitz condition in a domain $D \subseteq \mathbb{R}^{b+1}$ if there is a constant $L > 0$ such that

$$|f(u_1, ..., u_{b+1}) - f(v_1, ..., v_{b+1})| \leq L \max_{1 \leq i \leq b+1} |u_i - v_i|$$

for all $(u_1, ..., u_{b+1}), (v_1, ..., v_{b+1}) \in D$. For variables $Y_1, ..., Y_b$, the *stopping time* $T_D(Y_1, ..., Y_b)$ is defined to be the minimum t such that

$$(t/n; Y_1(t)/n, ..., Y_b(t)/n) \notin D.$$

This is written as T_D when $Y_1, ..., Y_b$ are understood from the context.

Theorem 4.3 ([39]). *Let $b, n \in \mathbb{N}$. For $1 \le j \le b$, suppose that $Y_j^{(n)}(t)$ is a sequence of real-valued random variables such that $0 \le Y_j^{(n)} \le Cn$ for some constant $C > 0$. Let H_t be the history up to time t, i.e., the sequence $\{Y_j^{(n)}(k),\ 0 \le j \le b,\ 0 \le k \le t\}$. Suppose also that for some bounded connected open set $D \subseteq \mathbb{R}^{b+1}$ containing the intersection of $\{(t, z_1, ..., z_b) : t \ge 0\}$ with some neighborhood of*

$$\left\{(0, z_1, ..., z_b) : \mathbb{P}(Y_j^{(n)}(0) = z_j n, 1 \le j \le b) \ne 0 \text{ for some } n\right\},$$

the following three conditions are satisfied:

1. (Boundedness). *For some function $\omega = \omega(n)$ and $\lambda = \lambda(n)$ with $\lambda^4 \log n < \omega < n^{2/3}/\lambda$ and $\lambda \to \infty$ as $n \to \infty$, for all $l \le b$ and uniformly for all $t < T_D$,*

$$\mathbb{P}\left(|Y_l^{(n)}(t+1) - Y_l^{(n)}(t)| > \frac{\sqrt{\omega}}{\lambda^2 \sqrt{\log n}} \mid H_t\right) = o(n^{-3});$$

2. (Trend). *For all $l \le b$ and uniformly over all $t < T_D$,*

$$\mathbb{E}[Y_l^{(n)}(t+1) - Y_l^{(n)}(t) | H_t] = f_l(t/n, Y_1^{(n)}(t)/n, ..., Y_b^{(n)}(t)/n) + o(1);$$

3. (Lipschitz). *For each l the function f_l is continuous and satisfies a Lipschitz condition on D with all Lipschitz constants uniformly bounded.*

Then the following hold.

(a) *For $(0, \hat{z}_1, ..., \hat{z}_b) \in D$, the system of differential equations*

$$\frac{dz_l}{ds} = f_l(s, z_1, ..., z_l), \quad l = 1, ..., b,$$

has a unique solution in D, $z_l : \mathbb{R} \to \mathbb{R}$ for $l = 1, \ldots, b$, which passes through $z_l(0) = \hat{z}_l$, $l = 1, \ldots, b,$, and which extends to points arbitrarily close to the boundary of D.

(b) *We have*

$$Y_l^{(n)}(t) = n z_l(t/n) + o_p(n)$$

uniformly for $0 \le t \le \min\{\sigma n, T_D\}$ and for each l. Here $z_l(t)$ is the solution in (a) with $\hat{z}_l = Y_l^{(n)}(0)/n$, and $\sigma = \sigma_D(n)$ is the supremum of those s to which the solution can be extended.

4.4 Proof of Theorem 3.4

We will apply Theorem 4.3 to show that the trajectory of $\{u(t), w_{\mathsf{U}}(t), (c_{i,j}(t))_{1 \le i \le p_\ell, 0 \le j \le r-1}\}$ throughout the algorithm is w.h.p. close to the solution of the deterministic equations suggested by these equations, i.e., $\{\nu, \mu_{\mathsf{U}}, (\gamma_{i,j})_{i=1,...,p_\ell, j=0,...,r-1}\}$.

We set $b = rp_\ell + 2$. For $\epsilon > 0$, we define

$$D_\epsilon = \{(\tau, \nu, \mu_{\mathsf{U}}, (\gamma_{i,j})_{i,j}) \in \mathbb{R}^{b+1} \mid -\epsilon < \tau < 1,\ 0 < \frac{\mu_{\mathsf{U}}}{\nu} < 2C_\gamma,\ -\epsilon < \gamma_{i,j} < \gamma_i + \epsilon,$$

$$\epsilon < \mu_{\mathsf{U}} < W'_\gamma + \sum_{i=1}^{p_\ell} W_i \gamma_i\},$$

We now apply the last part (b) of Theorem 4.3. Note that Boundedness and Trend hypotheses are verified for $t < T_{D_\epsilon}$. More specifically, the Boundedness hypothesis follows since the changes in the quantities $u(t), w_{\mathsf{U}}(t), c_{i,j}(t)$ are bounded by a constant multiple of the maximum degree of the random graph. But since the maximum weight is bounded, we may choose, for example, $\lambda = n^{1/8}$ and $\omega = n^{25/48}$, and show that the maximum degree is bounded by $\sqrt{\omega}/(\lambda^2 \log n) = n^{1/96}/\log n$ with probability $1 - o(n^{-3})$. The Trend hypothesis is verified by (38)–(41). By the assumption that $0 < \frac{\mu_{\mathsf{U}}}{\nu} < 2C_\gamma$, the Lipschitz condition is also verified. Hence, for $0 \le t \le \min\{\sigma_D n, T_{D_\epsilon}\}$, we have

$$\begin{aligned} u(t) &= n\nu(t/n) + o_p(n), \\ w_{\mathsf{U}}(t) &= n\mu_{\mathsf{U}}(t/n) + o_p(n), \\ c_{i,j}(t) &= n\gamma_{i,j}(t/n) + o_p(n), \text{ for all } i = 1, \dots, p_\ell, \ j = 0, \dots, r-1. \end{aligned} \tag{50}$$

This gives us the convergence up to the point where the solution leaves D_ϵ. Observe that the definition of the domain D_ϵ together with the fact that the maximum weight is bounded by $2C_\gamma$ imply that at round T_{D_ϵ} we have $w_{\mathsf{U}}(T_{D_\epsilon})/n \le \epsilon$, but $w_{\mathsf{U}}(T_{D_\epsilon} - 1)/n > \epsilon$.

Firstly, we will bound $|\mathcal{A}_f(T_{D_\epsilon})|$. Observe that $T_{D_\epsilon} = |\mathcal{A}_f(T_{D_\epsilon})|$ as exactly one vertex is removed at each step. Also, as we noted above $w_{\mathsf{U}}(T_{D_\epsilon})/n < \epsilon$, but $w_{\mathsf{U}}(T_{D_\epsilon} - 1)/n \ge \epsilon$. Since the maximum degree is $o_p(n)$ and the weights are bounded, w.h.p. we have

$$\epsilon \le w_{\mathsf{U}}(T_{D_\epsilon} - 1)/n \le 1.5\epsilon.$$

Hence, by (50) w.h.p.

$$\mu_{\mathsf{U}}\left(\frac{T_{D_\epsilon} - 1}{n}\right) < 2\epsilon. \tag{51}$$

Also, as the mimimum weight is bounded from below by W_0, the bound on w_{U} implies that

$$u\left(T_{D_\epsilon} - 1\right)/n \le \frac{1.5\epsilon}{W_0}. \tag{52}$$

Therefore, (50) again implies that w.h.p.

$$\nu\left(\frac{T_{D_\epsilon} - 1}{n}\right) \le \frac{2\epsilon}{W_0}.$$

Let

$$\alpha(y) := p\,\left(1 - h_F(\gamma)\right) + \gamma' + (1-p)\sum_{i=1}^{p_\ell} \gamma_i \psi_r\left(W_i y\right).$$

The first part of Proposition 4.2 implies that

$$\left|\frac{T_{D_\epsilon} - 1}{n} - \alpha\left(\frac{1}{d} I\left(\frac{T_{D_\epsilon} - 1}{n}\right)\right)\right| \le \frac{2\epsilon}{W_0}. \tag{53}$$

Let $\hat{\tau}^{(\ell,\gamma)}$ denote the minimum $\tau > 0$ such that $\mu_{\mathsf{U}}(\tau) = 0$. By Lemma 4.6 below there exists $c_1 > 0$ with the property that for any $\gamma < c_1$ and any $\delta \in (0,1)$ there exists an infinite set of positive integers $\mathcal{S}$ such that when $\ell \in \mathcal{S}$, it holds that

$$|\alpha(\hat{y}_{\ell,\gamma}) - (p + (1-p)\mathbb{E}(\psi_r(W_F \hat{y})))| < \delta, \tag{54}$$

where $\hat{y}_{\ell,\gamma}$ is the smallest positive root of

$$y = \frac{W'_\gamma}{d} + p\frac{1}{d}\sum_{i=1}^{p_\ell} W_i\gamma_i + (1-p)\sum_{i=1}^{p_\ell}\frac{W_i\gamma_i}{d}\psi_r(W_i y).$$

Its existence is implied by the continuity of $I(\tau)$ and $\alpha(y)$. By (51), the continuity of the function μ_{U} we deduce that there exists $\delta_1 = \delta_1(\epsilon) > 0$ such that, for n large enough,

$$\left|\frac{T_{D_\epsilon} - 1}{n} - \hat{\tau}^{(\ell,\gamma)}\right| < \delta_1. \tag{55}$$

Now, let $I\left(\hat{\tau}^{(\ell,\gamma)}\right) = \lim_{\tau\uparrow\hat{\tau}^{(\ell,\gamma)}} I(\tau)$. The continuity of I and α implies that there exists an increasing function $f : (0,1) \to (0,1)$, (depending on ℓ and γ) such that $f(x) \downarrow 0$ as $x \downarrow 0$ and

$$\left|\alpha\left(\frac{1}{d}I\left(\hat{\tau}^{(\ell,\gamma)}\right)\right) - \alpha\left(\frac{1}{d}I\left(\frac{T_{D_\epsilon}-1}{n}\right)\right)\right| < f(\delta_1). \tag{56}$$

Let us set $x = x(\tau) = I(\tau)/d$. Since $\mu_{\mathsf{U}}(\hat{\tau}^{(\ell,\gamma)}) = 0$, this implies that

$$\begin{aligned}\frac{I(\hat{\tau}^{(\ell,\gamma)})}{d} &= \frac{W'_\gamma}{d} + p\,\frac{1}{d}\sum_{i=1}^{p_\ell} W_i\gamma_i + (1-p)\sum_{i=1}^{p_\ell}\frac{W_i\gamma_i}{d}\mathbb{P}\left[\mathsf{Po}\left(\frac{W_i}{d}I(\hat{\tau}^{(\ell,\gamma)})\right) \geq r\right] \\ &= \frac{W'_\gamma}{d} + p\,\frac{1}{d}\sum_{i=1}^{p_\ell} W_i\gamma_i + (1-p)\sum_{i=1}^{p_\ell}\frac{W_i\gamma_i}{d}\psi_r\left(\frac{W_i}{d}I(\hat{\tau}^{(\ell,\gamma)})\right),\end{aligned}$$

whereby $\hat{y}_{\ell,\gamma} = \lim_{\tau\uparrow\hat{\tau}^{(\ell,\gamma)}} x(\tau)$. Thus the triangle inequality together with (53), (54) and (56) imply that for any $\gamma < c_1$, any $\delta \in (0,1)$ and any $\ell \in \mathcal{S}$ w.h.p.

$$\left|n^{-1}|\mathcal{A}_f(T_{D_\epsilon})| - \alpha(\hat{y}_{\ell,\gamma})\right| < \frac{\epsilon}{C_\gamma} + \delta + f(\delta_1).$$

Recall that $f(\delta_1)$ can become arbitrarily small if we make ϵ small enough. Therefore, the right-hand side of the above can become as small as we please. The proof of Theorem 3.4 will be complete, if we show that the process will finish soon after T_{D_ϵ}.

4.4.1 The end of the process

We will show that with high probability only a small fraction of vertices are added after T_{D_ϵ}. From now on, we start exposing the edges incident to all vertices of U simultaneously. Hence, we change the time scaling. Informally, each round will be approximated by a generation of a multi-type branching process which is sub-critical. The sub-criticality is encompassed by the following: there exists $\kappa_0 < 1$ such that with probability $1 - o(1)$

$$\sum_{i=1}^{p_\ell}\frac{W_i^2}{W_{[n]}}c_{i,r-1}\left(\frac{T_{D_\epsilon}-1}{n}\right) < \kappa_0 < 1. \tag{57}$$

We start this section by proving this. First, let us observe that using the expression for μ_{U} from Proposition 4.2 and the chain rule, for any $\tau < T_{D_\epsilon}$ we can write

$$\mu'_{\mathsf{U}}(\tau) = G(\tau)f_r^{(\ell,\gamma)\prime}(I(\tau)/d)$$

where

$$f_r^{(\ell,\gamma)}(x) = \frac{W'_\gamma}{d} + p\ \frac{1}{d}\sum_{i=1}^{p_\ell} W_i\gamma_i - x + (1-p)\ \frac{1}{d}\sum_{i=1}^{p_\ell} W_i\gamma_i \mathbb{P}\left[\, \mathsf{Po}\left(W_i x\right) \ge r\,\right].$$

But also by (44) we can write

$$\mu'_{\mathsf{U}}(\tau) = G(\tau)\left(-1 + \sum_{i=1}^{p_\ell} \frac{W_i^2}{d}\gamma_{i,r-1}\left(\tau\right)\right).$$

So, in particular for $\tau = (T_{D_\epsilon} - 1)/n$, we have $G((T_{D_\epsilon} - 1)/n) > 0$ and thereby (with $x(\tau) = I(\tau)/d$)

$$f_r^{(\ell,\gamma)\prime}\left(x\left(\frac{T_{D_\epsilon}-1}{n}\right)\right) = -1 + \sum_{i=1}^{p_\ell}\frac{W_i^2}{d}\gamma_{i,r-1}\left(\frac{T_{D_\epsilon}-1}{n}\right).$$

But by Lemma 4.6 and Proposition 4.8 below, for any $\delta \in (0,1)$ there exists $c_1 > 0$ with the property that for any $\gamma < c_1$ there exists an infinite set of positive integers $\mathcal{S}$ such that when $\ell \in \mathcal{S}$, it holds that

$$f_r^{(\ell,\gamma)\prime}(\hat{y}_{\ell,\gamma}) < f'_r(\hat{y}; W_F^*, p) + \delta,$$

and, moreover, $\hat{y}_{\ell,\gamma} < 1$. (Note that $\hat{y} < 1$ by its definition.) By Claim 4.11 below the family $\{f_r^{(\ell,\gamma)\prime}(x)\}_{\ell \ge \ell'_1}$ rescstricted on $[0,1]$ for $\ell'_1 = \ell'_1(\gamma)$ is equicontinuous provided that $\gamma < c'_4$. Hence, for any $\gamma < c_1 \wedge c'_4$ and any ℓ sufficiently large in $\mathcal{S}$ using (55), we conclude that if ε is sufficiently small we have

$$f_r^{(\ell,\gamma)\prime}\left(x\left(\frac{T_{D_\epsilon}-1}{n}\right)\right) < f'_r(\hat{y}; W_F^*, p) + 2\delta.$$

We select δ small enough so that the right-hand side of the above is negative. This and (50) imply that there exists $\kappa_0 < 1$ such that (57) holds with probability $1 - o(1)$.

From step T_{D_ϵ} onwards, we will provide a stochastic upper bound on U by a set $\widehat{\mathsf{U}}$, whose size is a essentially subcritical multi-type branching process. In particular, the expression in (57) will dominate the principal eigenvalue of the expected progeny matrix of this branching process. In this process we will not expose the vertices of $\widehat{\mathsf{U}}$ in a one-at-a-time fashion, but we will expose the neighbours of all of them simultaneously in each round. We let $\widehat{\mathsf{U}}(s)$ be the set $\widehat{\mathsf{U}}$ after s rounds.

Hence, $\widehat{\mathsf{U}}(s)$ will be the sth generation of this process, which resembles a multi-type branching process. We will keep track of the size of $\widehat{\mathsf{U}}$ through a functional which is well-known in the theory of multi-type branching processes to give rise to a supermartingale. Let us proceed with the details of this argument.

We set $t_0 := T_{D_\epsilon} - 1$. Let $\widehat{\mathsf{U}}(0) = \mathsf{U}(t_0)$ and $\widehat{\mathsf{C}}_{i,<r-1}(0) = \cup_{k=2}^r \mathsf{C}_{i,r-k}(t_0)$, for all $i = 1, \ldots, p_\ell$ and $k = 1, \ldots, r-1$. Let $\widehat{\mathsf{U}}_i(s)$ denote the subset of $\widehat{\mathsf{U}}(s)$ which consists of those vertices that have weight W_i and let $u_i(s) := |\widehat{\mathsf{U}}_i(s)|$ – we say that these vertices are of type i. We set $\hat{c}_{i,<r-1}(s) = |\widehat{\mathsf{C}}_{i,<r-1}(s)|$ and $\hat{c}_{i,r-1}(s) = |\widehat{\mathsf{C}}_{i,r-1}(s)|$. Let $\overline{u}_s = [u_1(s), \ldots, u_{p_\ell}(s)]^T$ be the vector whose co-ordinates are the sizes of the sets $\widehat{\mathsf{U}}_i(s)$. A vertex $v \in \widehat{\mathsf{U}}_j(s)$ can "give birth" to vertices of type i (i.e., of weight W_i). These may be vertices from the set $\widehat{\mathsf{C}}_{i,r-1}(s)$ or from the set $\widehat{\mathsf{C}}_{i,<r-1}(s)$. If v becomes adjacent to a vertex in $\widehat{\mathsf{C}}_{i,r-1}(s)$, then a child of v is

produced. Similarly, we say that a vertex in $\widehat{\mathsf{C}}_{i,<r-1}(s)$ produces a child of v, if it is adjacent to v and to *some other* vertex in $\widehat{\mathsf{U}}(s)$. In that sense, a vertex in $\widehat{\mathsf{C}}_{i,r-1} \cup \widehat{\mathsf{C}}_{i,<r-1}$ may be responsible for the birth of a child of more than one vertices in $\widehat{\mathsf{U}}(s)$.

Furthermore, if a vertex in $\widehat{\mathsf{C}}_{i,<r-1}(s)$ is adjacent to exaclty one vertex in $\widehat{\mathsf{U}}(s)$, then a vertex is added into $\widehat{\mathsf{C}}_{i,r-1}$. In this process the set $\widehat{\mathsf{C}}_{i,r-1}$ does not lose vertices but may only gain. Clearly, $|\widehat{\mathsf{U}}(s)|$ is a stochastic upper bound on U.

If a vertex is a child of more than one vertex, we assume that it is *born more than once*; it is included in $\widehat{\mathsf{U}}(s+1)$ as many times as it is born. In fact, the former case is much more likely than the latter. The expected number of those children that are born out of $\mathsf{C}_{i,r-1}(s)$ is bounded by $\frac{W_j W_i}{W_{[n]}} c_{i,r-1}(s)$. The expected number of the vertices of type i that originate from $\widehat{\mathsf{C}}_{i,<r-1}$ is bounded by $\hat{c}_{i,<r-1}(s)\frac{W_j W_i}{W_{[n]}}\left(|\widehat{\mathsf{U}}(s)|(2C_\gamma)^2/W_{[n]}\right)$. This is the case as the factor $|\widehat{\mathsf{U}}(s)|(2C_\gamma)^2/W_{[n]}$ bounds from above the probability that a given vertex in $\widehat{\mathsf{C}}_{i,<r-1}(s)$ is adjacent to some other vertex in $\widehat{\mathsf{U}}(s)$.

Now, if we let A_s be the $p_\ell \times p_\ell$ matrix whose ij entry is the expected number of children of type i that a vertex of type j has, then $\mathbb{E}\left(\overline{u}_{s+1}^T|\mathcal{H}_s\right) \leq \overline{u}_s^T A_s$ (the inequality is meant to be pointwise), where $\mathcal{H}_s$ is the sub-σ-algebra generated by the history of the process up to round s. One can view the matrix A_s as the expected progeny matrix of a multi-type branching process, where the expected number of children of type j that a vertex of type i gives birth to is at most

$$A_s[i,j] := \frac{W_i W_j}{W_{[n]}} a_j(s), \text{ where } a_j(s) := \hat{c}_{j,r-1}(s) + \hat{u}(s)\frac{4C_\gamma^2}{W_{[n]}}\hat{c}_{j,<r-1}(s).$$

Throughout this section, we will be working with this upper bound, which comes from a stochastic upper bound on the process. It is not hard to see that the vector $[W_1, \ldots, W_{p_\ell}]^T$ is a right eigenvector of A_s, with

$$\sum_{i=1}^{p_\ell} \frac{W_i^2}{W_{[n]}} a_i(s) =: \rho_s$$

being the corresponding eigenvalue. In fact, this is the unique positive eigenvalue of A_s. Since $\hat{c}_{j,r-1}(s)$ does not decrease, we have $\hat{c}_{j,r-1}(s) \geq \hat{c}_{j,r-1}(0) = c_{i,r-1}\left(\frac{T_{D_\epsilon}-1}{n}\right)$. Thus for $s > 0$ we have

$$\rho_s \geq \sum_{i=1}^{p_\ell} \frac{W_i^2}{W_{[n]}} c_{i,r-1}\left(\frac{T_{D_\epsilon}-1}{n}\right) > \kappa_0' > 0,$$

for some constant κ_0' and for any n sufficiently large.

For $s = 0$, it is not hard to see that ρ_0 is less than and bounded away from 1, if we choose ϵ small enough. Indeed, by (52)

$$\hat{u}(0)\frac{4C_\gamma^2}{W_{[n]}}\hat{c}_{j,<r-1}(0) \leq \hat{u}(0)\frac{4C_\gamma^2}{W_{[n]}}n < \hat{u}(0)\frac{5C_\gamma^2}{dn}n \overset{(52)}{\leq} \epsilon\frac{15C_\gamma}{2d}n.$$

Hence, together with (57) we deduce that if ϵ is small enough, then ρ_0 is smaller than 1 and, in fact, it is bounded away from 1.

Let $\lambda_i := W_i / \sum_j W_j$ and set $\xi := [\lambda_1, \dots, \lambda_{p_\ell}]^T$. Clearly, this is also a right eigenvector of A_t. Consider now the random variable $Z_s = (\xi, \overline{u}_s)$, where $(\cdot,\cdot)$ is the usual *dot* product. Therefore,

$$\mathbb{E}\left(Z_{s+1} | \mathcal{H}_s \right) \leq \rho_s Z_s.$$

Claim 4.4. *Conditional on $\mathcal{H}_s$, with probability at least $1 - 1/n^2$ we have*

$$Z_{s+1} \leq \rho_s Z_s + Z_s^{1/2} \log n.$$

Proof of Claim 4.4. Note that Z_{s+1} is a weighted sum of Bernoulli random variables, where the weights are bounded. More specifically,

$$Z_{s+1} = \sum_{j=1}^{p_\ell} \lambda_j \left(\sum_{v \in \mathsf{C}_{j,r-1}(s)} \mathbf{1}_{d_{\widehat{\mathsf{U}}(s)}(v) \geq 1} + \sum_{v \in \widehat{\mathsf{C}}_{j,<r-1}(s)} \mathbf{1}_{d_{\widehat{\mathsf{U}}(s)}(v) \geq 2} \right).$$

We will appeal to Talagrand's inequality (see for example Theorem 2.29 in [25]). Firstly, note that Z_{s+1} is a function of independent Bernoulli random variables, which correspond to the (potential) edges that are incident to $\widehat{\mathsf{U}}(s)$. If we change any one of them, then Z_{s+1} will change accordingly by at most 1 (as all the λ_js are at most 1). Furthermore, if $Z_{s+1} \geq x$, for some $x \geq 0$, then there is a collection of edges whose presence witnesses this fact. Let E_x be such a subset having size $\lceil x \rceil$. Let $\lambda_{\min} = \min_i \{\lambda_i\}$. Hence, we can apply Theorem 2.29 from [25] taking $\psi(x) = (x+1)/\lambda_{\min}$ (as $\sum_{e \in E_x} 1^2 \leq (x+1)/\lambda_{\min}$), with $m(Z_{s+1})$ being the median of Z_{t+1}; Talagrand's inequality yields

$$\mathbb{P}\left[Z_{s+1} \geq m(Z_{s+1}) + \frac{1}{2} Z_s^{1/2} \log n \right] \leq 2e^{-\frac{Z_s \log^2 n}{4(m(Z_{s+1}) + Z_s^{1/2} \log n + 1)/\lambda_{\min}}}. \tag{58}$$

Since $\psi(x)$ is proportional to x and Z_{s+1} takes only non-negative integer values, (using an argument similar to that on pages 41–42 in [25]) it follows that

$$|\mathbb{E}\left(Z_{s+1} \right) - m(Z_{s+1})| = O(\mathbb{E}\left(Z_{s+1} \right)^{1/2}).$$

Hence, for n large enough

$$\mathbb{P}\left[Z_{s+1} \geq \mathbb{E}\left(Z_{s+1} \right) + Z_s^{1/2} \log n \right] \leq \mathbb{P}\left[Z_{t+1} \geq m(Z_{s+1}) - O(\mathbb{E}\left(Z_{s+1} \right)^{1/2}) + Z_s^{1/2} \log n \right]$$
$$\leq \mathbb{P}\left[Z_{s+1} \geq m(Z_{s+1}) + \frac{1}{2} Z_s^{1/2} \log n \right].$$

So by (58) we conclude (using that $m(Z_{s+1}) \leq 2\mathbb{E}\left(Z_{s+1} \right) \leq 2\rho_s Z_s$) that

$$\mathbb{P}\left[Z_{s+1} \geq \mathbb{E}\left(Z_{s+1} \right) + Z_s^{1/2} \log n \right]$$
$$\leq 2e^{-\frac{Z_s \log^2 n}{4(m(Z_{s+1}) + Z_s^{1/2} \log n + 1)/\lambda_{\min}}} \leq 2e^{-\frac{Z_s \log^2 n}{4(2\rho_s Z_s + Z_s^{1/2} \log n + 1)/\lambda_{\min}}} = e^{-\Omega(\log^2 n)}.$$

□

We denote the above event by $\mathcal{E}_s$. If $Z_s > \log^6 n$ and $\mathcal{E}_s$ is realised, we have

$$\begin{aligned} Z_{s+1} &\leq \rho_s Z_s \left(1 + \frac{Z_s^{1/2}\log n}{\rho_s Z_s}\right) \overset{\rho_s \geq \kappa_0'}{\leq} \rho_s Z_s \left(1 + \frac{\log n}{\kappa_0' Z_s^{1/2}}\right) \overset{Z_s > \log^6 n}{\leq} \rho_s Z_s \left(1 + \frac{\log n}{\kappa_0' \log^3 n}\right) \\ &= \rho_s Z_s \left(1 + \frac{1}{\kappa_0' \log^2 n}\right). \end{aligned} \tag{59}$$

In a multi-type branching process, the variable Z_s/ρ^s, where ρ is the largest positive eigenvalue of the progeny matrix, is a martingale (see for example Theorem 4 in Chapter V.6 of [8]). Here, we use this fact only approximately, since the progeny matrix changes as the process evolves. Nevertheless, it does not change immensely, and we are able to control the increase of the eigenvalue ρ_s. Let us now make this precise.

By (57), the largest positive eigenvalue of A_0 is bounded by a constant $\rho_0 < 1$, with probability $1 - o(1)$. Recall that $\lambda_{\min} = \min_i\{\lambda_i\}$. For any $s \geq 0$, let

$$\mathcal{D}_s := \left\{ \sum_{j=1}^{p_\ell} \sum_{v \in \widehat{\mathsf{U}}(s)} d_{\widehat{\mathsf{C}}_{j,<r-1}(s)}(v) < \max\left\{ \frac{10C_\gamma^2}{\lambda_{\min} d} Z_s, \log^6 n \right\} \right\}.$$

Claim 4.5. *For any $s \geq 0$ we have $\mathbb{P}[\mathcal{D}_s] = 1 - o(1/n^2)$.*

Proof of Claim 4.5. The random variable $\sum_{j=1}^{p_\ell} \sum_{v \in \widehat{\mathsf{U}}(s)} d_{\widehat{\mathsf{C}}_{j,<r-1}(s)}(v)$ is stochastically bounded from above by $\sum_{v \in \widehat{\mathsf{U}}(s)} X_v$, where the X_vs are i.i.d. random variables that are distributed as $\mathrm{Bin}(n, (2C_\gamma)^2/W_{[n]})$. The expected value of this sum bounded by $\frac{5C_\gamma^2}{d} u(t)$ for large n. Also, $u(s) \leq Z_s/\lambda_{\min}$, as $Z_s = (\xi, \overline{u}_s) = \sum_i \lambda_i u_i(s) \geq \lambda_{\min} \sum_i u_i(s)$. So the expectation is at most $\frac{5C_\gamma^2}{\lambda_{\min} d} Z_s$. The claim follows from a standard Chernoff bound on the binomial distribution (as the sum of binomial is itself binomially distributed). □

Let $B_s := \max\left\{ \frac{10C_\gamma^2}{\lambda_{\min} d} Z_s, \log^6 n \right\}$.

On the event $\mathcal{D}_s$, the total degree of the vertices in $\widehat{\mathsf{U}}(s)$ into the set $\widehat{\mathsf{C}}_{i,<r-1}(s)$ bounds the number of vertices that enter into the set $\widehat{\mathsf{C}}_{i,r-1}$. Hence, on the event $\mathcal{D}_s$, we have

$$\hat{c}_{i,r-1}(s+1) \leq \hat{c}_{i,r-1}(s) + B_s.$$

Furthermore, for large n

$$u(s+1)\frac{4C_\gamma^2}{W_{[n]}} \hat{c}_{i,<r-1}(s+1) \leq u(s+1)\frac{4C_\gamma^2}{W_{[n]}} n \leq u(s+1)\frac{5C_\gamma^2}{dn} n = u(s+1)\frac{5C_\gamma^2}{d} \leq Z_{s+1}\frac{5C_\gamma^2}{\lambda_{\min} d}.$$

Also, on $\mathcal{E}_s$ we have $Z_{s+1} \leq \beta_1 Z_s$, for some constant $\beta_1 > 0$, if $\frac{10C_\gamma^2}{\lambda_{\min} d} Z_s \geq \log^6 n$; otherwise, since ρ_s is uniformly bounded bounded by some constant over all $s > 0$ we have $Z_{s+1} \leq$

$\beta_2 \log^6 n$ for some constant $\beta_2 > 0$. Therefore, on $\mathcal{D}_s \cap \mathcal{E}_s$ we have

$$\begin{aligned} \sum_{i=1}^{p_\ell} \frac{W_i^2}{W_{[n]}} a_i(s+1) &\leq \sum_{i=1}^{p_\ell} \frac{W_i^2}{W_{[n]}} \left(\hat{c}_{i,r-1}(s) + B_s + Z_{s+1} \frac{5C_\gamma^2}{\lambda_{\min} d} \right) \\ &\leq \sum_{i=1}^{p_\ell} \frac{W_i^2}{W_{[n]}} \hat{c}_{i,r-1}(s) + \beta \frac{B_s}{n} + \beta' \frac{Z_{s+1}}{n} \\ &\leq \rho_s + \beta \frac{B_s}{n} + \beta' \frac{Z_{s+1}}{n}, \end{aligned}$$

for some constants $\beta, \beta' > 0$ and any n. Furthermore, for some other constant $\beta'' > 0$,

$$B_s \leq \beta''(Z_s + \log^6 n).$$

Therefore, for some $\gamma > 0$, we finally obtain

$$\rho_{s+1} \leq \rho_s + \frac{1}{n} \left(\gamma (Z_s + Z_{s+1}) + \beta'' \log^6 n \right). \tag{60}$$

Let $\lambda_n = 1 + \frac{1}{\kappa_0' \log^2 n}$ and $T_0 = 2\lceil \log_{1/\tau} n \rceil$. We use induction in order to show that for every $\delta > 0$ there exists $\varepsilon > 0$ such that if $Z_0/n < \varepsilon$, then for all $s \leq T_0$

$$\rho_s \leq \rho_0 + \gamma \frac{Z_0}{n} \left(2 \sum_{k=0}^{s-1} (\rho_0 + \delta)^k \lambda_n^k + (\rho_0 + \delta)^s \lambda_n^s \right) + (4\gamma\gamma' + \beta'') \frac{\log^6 n}{n} \sum_{k=1}^{s-1} k + 2\gamma\gamma' \frac{\log^6 n}{n} s. \tag{61}$$

Now, observe that for n sufficiently large $(\rho_0 + \delta)\lambda_n < \rho_0' < 1$ From this inequality, we deduce that for every $s \leq T_0$, we have

$$\begin{aligned} \rho_s &\leq \rho_0 + \frac{2\gamma Z_0}{n} \sum_{k=0}^{\infty} \rho_0'^k + O\left(\frac{\log^8 n}{n} \right) = \rho_0 + \frac{2\gamma Z_0}{n} \frac{1}{1 - \rho_0'} + O\left(\frac{\log^8 n}{n} \right) \\ &\overset{Z_0/n < \varepsilon}{\leq} \rho_0 + 2\gamma\varepsilon \frac{1}{1 - \rho_0'} + O\left(\frac{\log^8 n}{n} \right) < \rho_0 + \delta < 1, \end{aligned} \tag{62}$$

provided that $\varepsilon > 0$ is small enough.

By (59), on the event $\cap_{s'=1}^{T_0} \{\mathcal{E}_{s'} \cap \mathcal{D}_{s'}\}$ for any $s < T_0$ we have

$$Z_s \leq \rho_s Z_{s-1} \lambda_n + \gamma' \log^6 n,$$

for some constant $\gamma' > 0$. Repeating this we get

$$\begin{aligned} Z_s &\leq \rho_{s-1}\rho_{s-2} Z_{s-2} \lambda_n^2 + \gamma'(\lambda_n + 1) \log^6 n \\ &\vdots \\ &\leq Z_0 \lambda_n^s \prod_{i=1}^{s-1} \rho_{s-i} + \gamma' \log^6 n \sum_{i=0}^{s-1} \lambda_n^i \\ &\leq Z_0 \lambda_n^s \prod_{i=0}^{s-1} \rho_{s-i} + 2s\gamma' \log^6 n, \end{aligned}$$

where in the last inequality we used $\lambda_n^i \le 2$ for n sufficiently large, uniformly over $i \le T_0$. By the inductive hypothesis $\rho_{s-i} \le \rho_0 + \delta$, for all $i \le s$. We thus deduce that

$$Z_s \le Z_0(\rho_0+\delta)^s \lambda_n^s + 2s\gamma' \log^6 n. \tag{63}$$

Substituting (61) into (60) and using (63), we obtain

$$\begin{aligned}
\rho_{s+1} \quad &\le \quad \rho_0 + \frac{\gamma Z_0}{n}\left(2\sum_{k=0}^{s-1}(\rho_0+\delta)^k\lambda_n^k + (\rho_0+\delta)^s\lambda_n^s\right) + (4\gamma\gamma' + \beta'')\frac{\log^6 n}{n}\sum_{k=1}^{s-1} k \\
&\quad +2\gamma\gamma'\frac{\log^6 n}{n}s + \beta''\frac{\log^6 n}{n} \\
&\quad +\frac{1}{n}\gamma\left(Z_0(\rho_0+\delta)^s\lambda_n^s + 2s\gamma'\log^6 n + Z_0(\rho_0+\delta)^{s+1}\lambda_n^{s+1} + 2(s+1)\gamma'\log^6 n\right) \\
&\le \quad \rho_0 + \frac{\gamma Z_0}{n}\left(2\sum_{k=0}^{s}(\rho_0+\delta)^k\lambda_n^k + (\rho_0+\delta)^{s+1}\lambda_n^{s+1}\right) + (4\gamma\gamma' + \beta'')\frac{\log^6 n}{n}\sum_{k=1}^{s} k \\
&\quad +2\gamma\gamma'\frac{\log^6 n}{n}(s+1),
\end{aligned}$$

where in the last inequality we used that $\beta''\frac{\log^6 n}{n} \le \beta''\frac{\log^6 n}{n}s$.

Set $\tau = \rho_0 + \delta$ and recall that $T_0 = 2\lceil \log_{1/\tau} n\rceil$. Claims 4.4 and 4.5 imply that

$$\mathbb{P}\left[\cap_{s\le T_0}\{\mathcal{E}_s \cap \mathcal{D}_s\}\right] = 1 - O(\log n/n^2). \tag{64}$$

For any $S \in \mathbb{N}$, we let $\mathcal{S}_S = \cap_{s\le S}\{\mathcal{E}_s \cap \mathcal{D}_s\}$ and note that if $S < S'$, then $\mathcal{S}_{S'} \subset \mathcal{S}_S$.

Using the tower property of the conditional expectation, we write

$$\mathbb{E}\left(Z_{T_0}\right) \quad = \quad \mathbb{E}\left(\mathbb{E}\left(Z_{T_0} \mid \mathcal{H}_{T_0-1}\right)\right) = \mathbb{E}\left(\mathbb{E}\left(Z_{T_0} \mid \mathcal{H}_{T_0-1}\right)\left(\mathbf{1}_{\mathcal{S}_{T_0-1}} + \mathbf{1}_{\mathcal{S}^c_{T_0-1}}\right)\right).$$

But $Z_{T_0} = O(n)$, whereby

$$\mathbb{E}\left(\mathbb{E}\left(Z_{T_0} \mid \mathcal{H}_{T_0-1}\right)\mathbf{1}_{\mathcal{S}^c_{T_0-1}}\right) = O(n)\mathbb{E}\left(\mathbf{1}_{\mathcal{S}^c_{T_0-1}}\right) = O(\log n/n).$$

Therefore,

$$\begin{aligned}
\mathbb{E}\left(Z_{T_0}\right) \quad &\le \quad \mathbb{E}\left(\mathbb{E}\left(Z_{T_0} \mid \mathcal{H}_{T_0-1}\right)\mathbf{1}_{\mathcal{S}_{T_0-1}}\right) + O(\log n/n) \\
&\overset{(62)}{\le} \quad \mathbb{E}\left((\rho_0+\delta)Z_{T_0-1}\right) + O(\log n/n).
\end{aligned}$$

Repeating this, we get

$$\mathbb{E}\left(Z_{T_0}\right) \le (\rho_0+\delta)^{T_0}\mathbb{E}\left(Z_0\right) + O(T_0\log n/n) \overset{\mathbb{E}(Z_0)=O(n)}{\le} O(1/n + T_0/n) = o(1). \tag{65}$$

Therefore, $\mathbb{P}\left[\widehat{\mathsf{U}}(T_0) \ne \varnothing\right] = o(1)$.

4.5 Auxiliary lemmas

Recall that $\hat{\tau}^{(\ell,\gamma)}$ denotes the minimum $\tau > 0$ such that $\mu_{\mathsf{U}}(\tau) = 0$. Recall also that $\hat{y}$ is the smallest positive solution of $f_r(y; W_F^*, p) = 0$ and that we have assumed that $f_r'(\hat{y}; W_F^*, p) < 0$. Recall that for $\gamma \in (0,1)$ and $\ell \in \mathbb{N}$ we set

$$f_r^{(\ell,\gamma)}(x) = \frac{W_\gamma'}{d} + p\ \frac{1}{d}\sum_{i=1}^{p_\ell} W_i\gamma_i - x + (1-p)\ \frac{1}{d}\sum_{i=1}^{p_\ell} W_i\gamma_i \mathbb{P}\left[\,\mathsf{Po}\left(W_i x\right) \geq r\,\right].$$

Also, recall that

$$\alpha(y) := p\ (1 - h_F(\gamma)) + \gamma' + (1-p)\sum_{i=1}^{p_\ell} \gamma_i \psi_r\left(W_i y\right).$$

The following lemma shows that if γ is taken small enough and ℓ is a large positive integer, then $\alpha(\hat{y}_{\ell,\gamma})$ and $f_r^{(\ell,\gamma)}(\hat{y}_{\ell,\gamma})$ can be approximated by the corresponding functions of $\hat{y}$.

Lemma 4.6. *For any $\delta > 0$, there exists c_1 such that for any $\gamma < c_1$, there exists a subsequence $\{\ell_k\}_{k\in\mathbb{N}}$ with the property that for every $\ell \in \{\ell_k\}_{k\in\mathbb{N}}$:*

1. $f_r^{(\ell,\gamma)\prime}(\hat{y}_{\ell,\gamma}) < f_r'(\hat{y}; W_F^*, p) + \delta$;
2. $|\alpha(\hat{y}_{\ell,\gamma}) - (p + (1-p)\mathbb{E}(\psi_r(W_F\hat{y})))| < \delta$.

Proof. Using Definition 3.1, we can express the γ_is in terms of the γ_i's: $\gamma_i = (1 - h_F(\gamma) + \gamma')\gamma_i'$. The expression for $f_r^{(\ell,\gamma)}$ yields the following:

$$\begin{aligned} f_r^{(\ell,\gamma)}(x) &= \frac{W_\gamma'}{d} + p\ \frac{1}{d}\sum_{i=1}^{p_\ell} W_i\gamma_i - x + (1-p)\ \frac{1}{d}\sum_{i=1}^{p_\ell} W_i\gamma_i \mathbb{P}\left[\,\mathsf{Po}\left(W_i x\right) \geq r\,\right] \\ &= \frac{W_\gamma'}{d} + (1 - h_F(\gamma) + \gamma')\left(p\ \frac{\hat{d}^{(\ell,\gamma)}}{d} + (1-p)\ \frac{d^{(\ell,\gamma)}}{d}\sum_{i=1}^{p_\ell} \frac{W_i\gamma_i'}{d^{(\ell,\gamma)}}\mathbb{P}\left[\,\mathsf{Po}\left(W_i x\right) \geq r\,\right]\right) - x, \end{aligned} \tag{66}$$

where $d^{(\ell,\gamma)} = \int_0^\infty x dF^{(\ell,\gamma)}(x)$ and $\hat{d}^{(\ell,\gamma)} = \int_0^{C_\gamma} x dF^{(\ell,\gamma)}(x) = \sum_{i=1}^{p_\ell} W_i\gamma_i'$. Hence, the second sum in the above expression can be rewritten as

$$\sum_{i=1}^{p_\ell} \frac{W_i\gamma_i'}{d^{(\ell,\gamma)}}\mathbb{P}\left[\,\mathsf{Po}\left(W_i x\right) \geq r\,\right] = \int_0^{C_\gamma} \psi_r\left(yx\right) dF^{*(\ell,\gamma)}(y),$$

where $F^{*(\ell,\gamma)}$ is the distribution function of the $U^{(\ell,\gamma)}$ size-biased distribution.

We set $c(\gamma) = 1 - h_F(\gamma) + \gamma'$ and write $p^{(\ell,\gamma)} = \frac{W_\gamma'}{dc(\gamma)} + p\ \frac{\hat{d}^{(\ell,\gamma)}}{d}$. The expression in (66) becomes

$$f_r^{(\ell,\gamma)}(x) = c(\gamma)\left(p^{(\ell,\gamma)} + (1-p)\ \frac{d^{(\ell,\gamma)}}{d}\ \int_0^{C_\gamma} \psi_r\left(yx\right) dF^{*(\ell,\gamma)}(y)\right) - x.$$

Hence, the derivative of $f_r^{(\ell,\gamma)}(x)$ with respect to x is

$$\begin{aligned} f_r^{(\ell,\gamma)\prime}(x) &= -1 + c(\gamma)(1-p)\ \frac{d^{(\ell,\gamma)}}{d}\ \int_0^{C_\gamma} ye^{-yx}\frac{(yx)^{r-1}}{(r-1)!}dF^{*(\ell,\gamma)}(y) \\ &= -1 + c(\gamma)(1-p)\ \frac{d^{(\ell,\gamma)}}{d}\ \frac{r}{x}\ \int_0^{C_\gamma} e^{-yx}\frac{(yx)^r}{r!}dF^{*(\ell,\gamma)}(y). \end{aligned} \tag{67}$$

Similarly, we can write

$$\alpha(x) = p(1 - h_F(\gamma)) + \gamma' + c(\gamma)(1-p) \int_0^{C_\gamma} \psi_r(yx)\, dF^{(\ell,\gamma)}(y). \tag{68}$$

For real numbers y and $\delta > 0$, let $B(y;\delta)$ denote the open ball of radius δ around y. We show the following result.

Proposition 4.7. *Let $f : [0,\infty) \to [0,\infty)$ be a bounded function which is everywhere differentiable. Let also $y_1 \in \mathbb{R}$. For any $\delta > 0$ there exists $c_2 = c_2(\delta)$ with the property that for any $\gamma < c_2$, there exist $\ell_0 = \ell_0(\delta,\gamma) > 0$ and $\delta' = \delta'(\delta,\gamma)$ such that for any $\ell > \ell_0$ and any $y_2 \in B(y_1;\delta')$,*

$$\left| \int_0^{C_\gamma} f(yy_2) dF^{*(\ell,\gamma)}(y) - \mathbb{E}\left(f(W_F^* y_1)\right) \right| < \delta,$$

and

$$\left| \int_0^{C_\gamma} f(yy_2) dF^{(\ell,\gamma)}(y) - \mathbb{E}\left(f(W_F y_1)\right) \right| < \delta.$$

We will further show that $\hat{y}_{\ell,\gamma}$ is close to $\hat{y}$ over a subsequence $\{\ell_k\}_{k\in\mathbb{N}}$.

Proposition 4.8. *There exists a $c_3 > 0$ such that for all $\gamma < c_3$ and any $\delta' > 0$ there exists a subsequence $\{\ell_k\}_{k\in\mathbb{N}}$ such that $\hat{y}_{\ell_k,\gamma} \in B(\hat{y};\delta')$.*

The above two propositions yield the following:

Corollary 4.9. *Let $f : [0,\infty) \to [0,\infty)$ be a bounded function which is everywhere differentiable. For any $\delta > 0$, any $\gamma < c_2 \wedge c_3$, there exists a subsequence $\{\ell_k\}_{k\in\mathbb{N}}$ such that*

$$\left| \int_0^{C_\gamma} f(y\hat{y}_{\ell_k,\gamma}) dF^{*(\ell_k,\gamma)}(y) - \mathbb{E}\left(f(W_F^* \hat{y})\right) \right| < \delta,$$

and

$$\left| \int_0^{C_\gamma} f(y\hat{y}_{\ell_k,\gamma}) dF^{(\ell_k,\gamma)}(y) - \mathbb{E}\left(f(W_F \hat{y})\right) \right| < \delta.$$

The two statements of the lemma can be deduced from (67) and (68), if we let $f(x)$ be $\psi_r(x)$ in the former case, and $e^{-x}\frac{x^r}{r!}$ in the latter. Note that the choice of the subsequence is determined through Proposition 4.8 and can be the same for both choices of $f(x)$. Observe that both functions are bounded (by 1), they are differentiable everywhere in $\mathbb{R}$ and have bounded derivatives.

By the second part of Definition 3.2 and the fact that $c(\gamma) \to 1$ as $\gamma \downarrow 0$ we have

$$c(\gamma) \left| \frac{d^{(\ell,\gamma)}}{d} - 1 \right| < \delta, \tag{69}$$

for any γ that is small enough and any ℓ that is large enough. We will show now that $p^{(\ell,\gamma)}$ is close to p. We will need the following claim, which is a direct consequence of the second part of Definition 3.2.

Claim 4.10. *There is a function $r : (0,1) \to (0,1)$ such that $r(\gamma) \to 0$ as $\gamma \downarrow 0$ with the following property. For any $\gamma \in (0,1)$ there exists $\ell_1(\gamma)$ such that for any $\ell > \ell_1(\gamma)$*

$$\left| \hat{d}^{(\ell,\gamma)} - d \right| < r(\gamma).$$

The above claim together with the fact that $W'_\gamma \to 0$ as $\gamma \to 0$ imply that if γ is small enough and ℓ is large enough, we have

$$\left|p^{(\ell,\gamma)} - p\right| < \delta \text{ and } \left|p(1 - h_F(\gamma)) + \gamma' - p\right| < \delta. \tag{70}$$

Both parts of the lemma now follow from Corollary 4.9 together with (69) and (70). We now proceed with the proofs of Propositions 4.7 and 4.8.

Proof of Proposition 4.7. The proof of this proposition will proceed in two steps. Firstly, we will show that for any $\gamma < 1$ there exist $\delta' = \delta'(\delta, \gamma)$ and $\ell'_0 = \ell'_0(\delta, \gamma)$ such that for any $y_2 \in B(y_1; \delta')$ and $\ell > \ell'_0$ we have

$$\left|\int_0^{C_\gamma} f(yy_2)dF^{*(\ell,\gamma)}(y) - \int_0^{C_\gamma} f(yy_1)dF^*(y)\right| < \delta/2. \tag{71}$$

The proposition will follow if we show that there exists $c'_2 = c'_2(\delta)$ such that for any $\gamma < c'_2$ it holds that

$$\left|\int_{C_\gamma}^{\infty} f(yy_1)dF^*(y)\right| < \delta/2. \tag{72}$$

Having proved these inequalities, we deduce that

$$\begin{aligned}
&\left|\int_0^{C_\gamma} f(yy_2)dF^{*(\ell,\gamma)}(y) - \mathbb{E}\left[\, f(W_F^* y_1)\,\right]\right| \\
\leq &\left|\int_0^{C_\gamma} f(yy_2)dF^{*(\ell,\gamma)}(y) - \int_0^{C_\gamma} f(yy_1)dF^*(y)\right| + \left|\int_{C_\gamma}^{\infty} f(yy_1)dF^*(y)\right| \overset{(71),(72)}{<} \delta.
\end{aligned}$$

The proof for the case of $F^{(\ell,\gamma)}$ proceeds along the same lines. We can show that for any $\gamma < 1$ there exist $\delta' = \delta'(\delta, \gamma)$ and $\ell''_0 = \ell''_0(\delta, \gamma)$ such that for any $y_2 \in B(y_1; \delta')$ and $\ell > \ell''_0$ we have

$$\left|\int_0^{C_\gamma} f(yy_2)dF^{(\ell,\gamma)}(y) - \int_0^{C_\gamma} f(yy_1)dF(y)\right| < \delta/2. \tag{73}$$

Then we show that there exists $c''_2 = c''_2(\delta)$ such that for any $\gamma < c''_2$ it holds that

$$\left|\int_{C_\gamma}^{\infty} f(yy_1)dF(y)\right| < \delta/2. \tag{74}$$

As before, from (73) and (74) we deduce

$$\begin{aligned}
&\left|\int_0^{C_\gamma} f(yy_2)dF^{(\ell,\gamma)}(y) - \mathbb{E}\left[\, f(W_F y_1)\,\right]\right| \\
\leq &\left|\int_0^{C_\gamma} f(yy_2)dF^{(\ell,\gamma)}(y) - \int_0^{C_\gamma} f(yy_1)dF(y)\right| + \left|\int_{C_\gamma}^{\infty} f(yy_1)dF(y)\right| \overset{(73),(74)}{<} \delta.
\end{aligned}$$

We proceed with the proofs of (71) and (72) – the proofs of (73) and (74) are very similar (in fact, simpler) and are omitted. The lemma will follow if we take $c_2 = c'_2 \wedge c''_2$ and $\ell_0 = \ell'_0 \vee \ell''_0$.

Proof of (71). We begin with the specification of δ'. We let δ' be such that whenever $|y_1 - y_2| < \delta'$ we have

$$|f(xy_1) - f(xy_2)| < \delta/4, \tag{75}$$

for any $x \in [0, C_\gamma]$. This choice of δ' is possible since f is continuous and therefore uniformly continuous in any closed interval. Consider $y_2 \in B(y_1; \delta')$. We then have

$$\begin{aligned}
&\left| \int_0^{C_\gamma} f(xy_2) dF^{*(\ell,\gamma)}(x) - \int_0^{C_\gamma} f(xy_1) dF^*(x) \right| \\
&\leq \int_0^{C_\gamma} |f(xy_2) - f(xy_1)| dF^{*(\ell,\gamma)}(x) \\
&\qquad + \left| \int_0^{C_\gamma} f(xy_1) dF^{*(\ell,\gamma)}(x) - \int_0^{C_\gamma} f(xy_1) dF^*(x) \right| \\
&\overset{(75)}{\leq} \delta/4 + \left| \int_0^{C_\gamma} f(xy_1) dF^{*(\ell,\gamma)}(x) - \int_0^{C_\gamma} f(xy_1) dF^*(x) \right|.
\end{aligned} \tag{76}$$

We will argue that the second expression is also bounded from above by $\delta/4$ when γ is small enough and ℓ is large enough. This follows from (16) as the latter implies that $F^{*(\ell,\gamma)}$ converges weakly to F^* as $\ell \to \infty$ and $\gamma \downarrow 0$ and $\ell \to \infty$. Since f has been assumed to be bounded and continuous, by Lemma 3.6 there exists c_1 such that for any $0 < \gamma < c_1$ and any $\ell > \ell_1(\gamma)$ we have

$$\left| \int_0^{\infty} f(xy_1) dF^{*(\ell,\gamma)}(x) - \mathbb{E}(f(W_F^* y_1)) \right| < \delta/8.$$

Furthermore, by (72) if γ is sufficiently small we have

$$\left| \mathbb{E}(f(W_F^* y_1) \mathbf{1}_{W_F^* \geq C_\gamma}) \right| < \delta/8.$$

Also, by Lemma 3.3, for any γ sufficiently small and any ℓ sufficiently large

$$\left| \int_{C_\gamma}^{\infty} f(xy_1) dF^{*(\ell,\gamma)}(x) \right| < \delta/8.$$

Therefore, for any such γ and any ℓ sufficiently large we get

$$\left| \int_0^{C_\gamma} f(xy_1) dF^{*(\ell,\gamma)}(x) - \mathbb{E}(f(W_F^* y_1) \mathbf{1}_{W_F^* < C_\gamma}) \right| < \delta/4.$$

Substituting this bound into (76), we can deduce (71).

□

We now proceed with the proof of (72).

Proof of (72). Assume that $|f(x)| < b$ for any $x \in \mathbb{R}$. Hence we have

$$\left| \int_{C_\gamma}^{\infty} f(yy_1) dF^*(y) \right| < b\mathbb{E}\left[\mathbf{1}_{W_F^* \geq C_\gamma} \right]. \tag{77}$$

Now, observe that

$$\mathbb{E}\left[\mathbf{1}_{W_F^* \geq C_\gamma} \right] = \frac{\mathbb{E}\left[W_F \mathbf{1}_{W_F \geq C_\gamma} \right]}{\mathbb{E}\left[W_F \right]}.$$

Since $\mathbb{E}[W_F] < \infty$, the latter is at most $\delta/(2b)$, if $\gamma > 0$ is small enough.

Therefore,

$$\mathbb{E}\left[\mathbf{1}_{W_F^* \geq C_\gamma}\right] \leq \frac{\delta}{2b},$$

and (72) follows from (77). □

□

Proof of Proposition 4.8. We consider the functions $f_r^{(\ell,\gamma)}(x)$ restricted on the unit interval $[0,1]$.

Claim 4.11. *There exists $c_4 > 0$ such that for any $\gamma < c_4$ the family*

$$\left\{f_r^{(\ell,\gamma)}(x)\right\}_{\ell > \ell_1},$$

for some $\ell_1 = \ell_1(\gamma)$, is equicontinuous. The analogous statement holds also for $\left\{f_r^{(\ell,\gamma)\prime}(x)\right\}_{\ell > \ell_1'}$ for some $\ell_1' = \ell_1'(\gamma)$ and some other constant $c_4' > 0$ (this is used in Section 4.4.1).

Proof of Claim 4.11. Let $\varepsilon \in (0,1)$ and let c_4 be such that for any $\gamma < c_4$ we have $1/C_\gamma < \varepsilon/2$. Recall that $\{\mathbf{W}^{(\ell,\gamma)}\}_{\gamma \in (0,1), \ell \in \mathbb{N}}$ is F-convergent (cf. Definition 3.2). So there exists a function $\rho : (0,1) \to (0,1)$ satisfying $\rho(\gamma) \downarrow 0$ as $\gamma \downarrow 0$, such that for any γ and for any ℓ sufficiently large (cf. Definition 3.2 Part 2)

$$\left|\frac{d^{(\ell,\gamma)}}{d} - 1\right| < \rho(\gamma)/d. \tag{78}$$

The function $\psi_r(y)$ is uniformly continuous on the closed interval $[0, C_\gamma]$. Hence there exists $\delta \in (0,1)$ such that for any $x_1, x_2 \in [0,1]$ with $|x_1 - x_2| < \delta/C_\gamma$ we have $|\psi_r(wx_1) - \psi_r(wx_2)| < d\varepsilon/(2\rho)$. Thus,

$$\left|\int_0^{C_\gamma} \psi_r(yx_1)\, dF^{*(\ell,\gamma)}(y) - \int_0^{C_\gamma} \psi_r(yx_2)\, dF^{*(\ell,\gamma)}(y)\right| < \frac{d\varepsilon}{2\rho}. \tag{79}$$

Thereby, for any $\gamma < c_4$ and ℓ sufficiently large, if $x_1, x_2 \in [0,1]$ are such that $|x_1 - x_2| < \delta/C_\gamma$, then

$$\begin{aligned} |f_r^{(\ell,\gamma)}(x_1) - f_r^{(\ell,\gamma)}(x_2)| &\leq |x_1 - x_2| + \frac{d^{(\ell,\gamma)}}{d}\left|\int_0^{C_\gamma} \psi_r(yx_1)\, dF^{*(\ell,\gamma)(y)} - \int_0^{C_\gamma} \psi_r(yx_2)\, dF^{*(\ell,\gamma)(y)}\right| \\ &\overset{(78),(79)}{\leq} \frac{\delta}{C_\gamma} + \frac{d\varepsilon}{2\rho}\,\frac{\rho}{d} \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \leq \varepsilon. \end{aligned}$$

The proof for the family $\left\{f_r^{(\ell,\gamma)\prime}(x)\right\}$ is similar and we omit it. □

By the Arzelá-Ascoli Theorem, there exists a subsequence $\{\ell_k\}_{k \in \mathbb{N}}$ such that

$$\left\{f_r^{(\ell_k,\gamma)}(x)\right\}_{k \in \mathbb{N}}$$

is convergent in the L_∞-norm on the space of all continuous real-valued functions on $[0,1]$.

Now, recall that $\hat{y}$ is the smallest positive root of $f_r(y; W_F^*, p) = 0$ and, moreover, $f_r'(\hat{y}; W_F^*, p) < 0$. Also, $\hat{y} < 1$, since, by its definition, $\hat{y} = (1-p)\mathbb{E}[\psi_r(W_F^* \hat{y})] + p < 1$. Hence, there exists $\delta_0 > 0$ such that $\hat{y} + \delta_0 < 1$ and, furthermore,

$$f_r(\hat{y} + \delta_0; W_F^*, p) < 0 \text{ and}$$
$$f_r(\hat{y} - \delta_0; W_F^*, p) > 0.$$

By the the L_∞-covergence of the family $\left\{ f_r^{(\ell_k,\gamma)}(x) \right\}_{k \in \mathbb{N}}$ restricted on $[0,1]$, we deduce that there exists $\ell_1 = \ell_1(\delta_0, \gamma)$ with the property that for any k such that $\ell_k > \ell_1$ we have

$$f_r^{(\ell_k,\gamma)}(\hat{y} + \delta_0) < 0 \text{ and}$$
$$f_r^{(\ell_k,\gamma)}(\hat{y} - \delta_0) > 0.$$

In turn, this implies that for any such k there exists a root of $f_r^{(\ell_k,\gamma)}(x)$ in $B(\hat{y}; \delta_0)$.

To conclude the proof of the proposition, we need to show that for all but finitely many ks there is no positive root of $f_r^{(\ell_k,\gamma)}$ in the interval $[0, \hat{y} - \delta_0]$. Assume, for the sake of contradiction, that there exists a sub-subsequence $\{\ell_{k_i}\}_{i \in \mathbb{N}}$ such that $\hat{y}_{\ell_{k_i},\gamma} \in [0, \hat{y} - \delta_0]$. By the sequential compactness of this interval, we deduce that there is a further sub-subsequence $\{\ell_{k_j}\}_{j \in \mathbb{N}}$ over which

$$\hat{y}_{\ell_{k_j},\gamma} \to \hat{y}_\gamma,$$

as $j \to \infty$, for some $\hat{y}_\gamma \in [0, \hat{y} - \delta_0]$.

Let $\delta \in (0,1)$ and let $c_2 = c_2(\delta)$ be as in Proposition 4.7. Consider $\gamma < c_2$. Then there exists j_0 such that for $j > j_0$ we have

$$\left| \int_0^{C_\gamma} \psi_r(y \hat{y}_{\ell_{k_j},\gamma}) dF^{*(\ell_{k_j},\gamma)}(y) - \mathbb{E}\left(\psi_r(W_F^* \hat{y}_\gamma)\right) \right| < \delta/3. \tag{80}$$

Assume that γ is small enough so that

$$|c(\gamma) - 1|, \; \rho(\gamma)/d, \; r(\gamma)/d < \delta/9.$$

Moreover, assume that j_0 is large enough so that for $j > j_0$ we have

$$\left| \frac{d^{(\ell_{k_j},\gamma)}}{d} - 1 \right| < \rho(\gamma)/d \text{ and } \left| \frac{\hat{d}^{(\ell_{k_j},\gamma)}}{d} - 1 \right| < r(\gamma)/d,$$

by Definition 3.2 Part 2 and Claim 4.10. Hence

$$\begin{aligned} \left| c(\gamma) \frac{\hat{d}^{(\ell_{k_j},\gamma)}}{d} - 1 \right| &\leq |c(\gamma) - 1| \frac{d^{(\ell_{k_j},\gamma)}}{d} + \left| \frac{\hat{d}^{(\ell_{k_j},\gamma)}}{d} - 1 \right| \\ &\leq |c(\gamma) - 1| + |c(\gamma) - 1| \left| \frac{d^{(\ell_{k_j},\gamma)}}{d} - 1 \right| + \left| \frac{\hat{d}^{(\ell_{k_j},\gamma)}}{d} - 1 \right| \leq 3\delta/9 = \delta/3. \end{aligned} \tag{81}$$

Similarly, we can show that for γ small enough and j large enough

$$\left| c(\gamma) \frac{d^{(\ell_{k_j},\gamma)}}{d} - 1 \right| \leq \frac{\delta}{3}. \tag{82}$$

Now, consider the function $\hat{f}_r(x) := f_r(x; W_F^*, p) + W'_\gamma/d$. Since $f_r^{(\ell_{k_j},\gamma)}(\hat{y}_{\ell_{k_j},\gamma}) = 0$, we can write

$$\begin{aligned}
\hat{f}_r(\hat{y}_\gamma) &= \hat{f}_r(\hat{y}_\gamma) - f_r^{(\ell_{k_j},\gamma)}(\hat{y}_{\ell_{k_j},\gamma}) \\
&\leq p \left| c(\gamma) \frac{\hat{d}^{(\ell_{k_j},\gamma)}}{d} - 1 \right| + (1-p) \left| c(\gamma) \frac{d^{(\ell_{k_j},\gamma)}}{d} \int_0^{C_\gamma} \psi_r(y \hat{y}_{\ell_{k_j},\gamma}) dF^{*(\ell_{k_j},\gamma)}(y) - \mathbb{E}\left(\psi_r(W_F^* \hat{y}_\gamma)\right) \right| \\
&\leq p \left| c(\gamma) \frac{\hat{d}^{(\ell_{k_j},\gamma)}}{d} - 1 \right| + (1-p) \left| \left(c(\gamma) \frac{d^{(\ell_{k_j},\gamma)}}{d} - 1 \right) \int_0^{C_\gamma} \psi_r(y \hat{y}_{\ell_{k_j},\gamma}) dF^{*(\ell_{k_j},\gamma)}(y) \right| \\
&\quad + (1-p) \left| \int_0^{C_\gamma} \psi_r(y \hat{y}_{\ell_{k_j},\gamma}) dF^{*(\ell_{k_j},\gamma)}(y) - \mathbb{E}\left(\psi_r(W_F^* \hat{y}_\gamma)\right) \right| \\
&\overset{(81),(82),(80)}{\leq} \frac{\delta}{3} + \frac{\delta}{3} + \frac{\delta}{3} = \delta.
\end{aligned}$$

Since δ is arbitrary, it follows that

$$\hat{f}_r(\hat{y}_\gamma) := f_r(\hat{y}_\gamma; W_F^*, p) + W'_\gamma = 0,$$

whereby $f_r(\hat{y}_\gamma; W_F^*, p) < 0$. Recall also that $f_r(\hat{y} - \delta_0; W_F^*, p) > 0$. The continuity of f_r implies that there is a root in $(0, \hat{y} - \delta_0)$. But this leads to contradiction as $\hat{y}$ is the smallest positive root of $f_r(x; W_F^*, p) = 0$. □

□

The following lemma shows that if the weight sequence has power law distribution with exponent between 2 and 3, then the condition on the derivative of $f_r(x; W_F^*, p)$ that appears in the statement of Theorem 2.2 is always satisfied.

Lemma 4.12. *Assume that $(\mathbf{w}(n))_{n \geq 1}$ follows a power law with exponent $\beta \in (2,3)$. Then $f'_r(\hat{y}; W_F^*, p) < 0$.*

Proof. From the definition of f we obtain that

$$f'_r(x; W_F^*, p) = -1 + (1-p) \frac{r}{x} \mathbb{E}\left[e^{-W_F^* x} \frac{(W_F^* x)^r}{r!} \right].$$

To show the claim it is thus sufficient to argue that

$$(1-p) r \mathbb{E}\left[e^{-W_F^* \hat{y}} \frac{(W_F^* \hat{y})^r}{r!} \right] < \hat{y} = p + (1-p) \mathbb{E}\left[\psi_r(W_F^* \hat{y}) \right].$$

In turn, it suffices to prove that

$$r \mathbb{E}\left[e^{-W_F^* \hat{y}} \frac{(W_F^* \hat{y})^r}{r!} \right] < \mathbb{E}\left[\psi_r(W_F^* \hat{y}) \right]. \tag{83}$$

We set $p_r(x) = e^{-x} x^r / r!$. Furthermore, we set $g(x) := \mathbb{E}\left[p_r(W_F^* x) \right]$ and $f(x) := \mathbb{E}\left[\psi_r\left(W_F^* x\right) \right]$. Then we claim that

$$f(x) > r g(x) \quad \text{for any } x \in (0,1],$$

which is equivalent to (83). To see the claim, we will consider the difference $f(x) - rg(x)$ and show that it is increasing with respect to x; the statement then follows from $f(0) - rg(0) = 0$. The derivative with respect to x is

$$\begin{aligned}(f(x) - rg(x))' &= \mathbb{E}\left[W_F^* p_{r-1}\left(W_F^* x\right)\right] + r\left(\mathbb{E}\left[W_F^* p_r\left(W_F^* x\right)\right] - \mathbb{E}\left[W_F^* p_{r-1}\left(W_F^* x\right)\right]\right) \\ &= -\frac{r(r-1)}{x}\,\mathbb{E}\left[p_r\left(W_F^* x\right)\right] + \frac{r(r+1)}{x}\,\mathbb{E}\left[p_{r+1}\left(W_F^* x\right)\right] \\ &= \frac{r}{x}\,\left(-(r-1)\,\mathbb{E}\left[p_r\left(W_F^* x\right)\right] + (r+1)\,\mathbb{E}\left[p_{r+1}\left(W_F^* x\right)\right]\right).\end{aligned}$$

Hence, it suffices to show that

$$(r+1)\,\mathbb{E}\left[p_{r+1}\left(W_F^* x\right)\right] > (r-1)\,\mathbb{E}\left[p_r\left(W_F^* x\right)\right],$$

for $x \in (0,1]$. Note that the probability density function of W_F^* is $(\beta-1)cw^{-\beta+1}$, for $w > x_0$; otherwise it is equal to 0. So we obtain for $j \in \{r, r+1\}$

$$\mathbb{E}\left[p_j\left(W_F^* x\right)\right] = (\beta-1)c\int_{x_0}^{\infty} e^{-wx}\,\frac{(wx)^j}{j!}\,w^{-\beta+1}dw \overset{(z=wx)}{=} (\beta-1)\frac{x^{\beta-2}}{j!}c\int_{x_0}^{\infty} e^{-z}z^{j-\beta+1}dz.$$

Thereby, it suffices to show that

$$\int_{x_0}^{\infty} e^{-z}z^{r-\beta+2}dz > (r-1)\int_{x_0}^{\infty} e^{-z}z^{r-\beta+1}dz.$$

Applying integration by parts on the integral of the left-hand side we obtain

$$\begin{aligned}\int_{x_0}^{\infty} e^{-z}z^{r-\beta+2}dz &= e^{-x_0}x_0^{r-\beta+2} + (r-\beta+2)\int_{x_0}^{\infty} e^{-z}z^{r-\beta+1}dz \\ &> (r-\beta+2)\int_{x_0}^{\infty} e^{-z}z^{r-\beta+1}dz \overset{(\beta<3)}{>} (r-1)\int_{x_0}^{\infty} e^{-z}z^{r-\beta+1}dz.\end{aligned}$$

□

References

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

[2] R. Albert and A. Barabási. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1):47–97, 2002.

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

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

[5] H. Amini, R. Cont, and A. Minca. Resilience to contagion in financial networks. *Mathematical finance*, 26(2):329–365, 2016.

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

[7] H. Amini and A. Minca. Inhomogeneous financial networks and contagious links. *Operations Research*, 64(5):1109–1120, 2016.

[8] K.B. Athreya and P.E. Ney. *Branching processes*. Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen. Springer Verlag, 1972.

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

[10] 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.

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

[12] 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.

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

[14] B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Structures & Algorithms*, 31(1):3–122, 2007.

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

[16] 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.

[17] F. Chung and L. Lu. Connected components in random graphs with given expected degree sequences. *Annals of Combinatorics*, 6:125–145, 2002.

[18] F. Chung and L. Lu. The average distance in a random graph with given expected degrees. *Internet Mathematics*, 1(1):91–113, 2003.

[19] F. Chung, L. Lu, and V. Vu. The spectra of random graphs with given expected degrees. *Internet Mathematics*, 1(3):257–275, 2004.

[20] N. Detering, T. Meyer-Brandis, and K. Panagiotou. Bootstrap percolation in directed inhomogeneous random graphs. *Electronic Journal of Combinatorics*, P3.12, 2019.

[21] N. Detering, T. Meyer-Brandis, K. Panagiotou, and D. Ritter. Managing default contagion in inhomogeneous financial networks. *SIAM Journal of Financial Mathematics*, 10(2):578–614, 2019.

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

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

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

[26] 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.

[27] A. J. Lotka. The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences*, 16:317–323, 1926.

[28] M. Mitzenmacher. A brief history of generative models for power law and lognormal distributions. *Internet Mathematics*, 1:226–251, 2004.

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

[30] V. Pareto. *Cours d'Économie Politique*. Dronz, Geneva Switzerland, 1896.

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

[32] 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.

[33] B. Söderberg. General formalism for inhomogeneous random graphs. *Physical Review E*, 66:066121, 2002.

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

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

[36] G.L. Torrisi, M. Garetto, and E. Leonardi. Bootstrap percolation on the stochastic block model. arXiv:1812.09107, 53 pages.

[37] R. van der Hofstad. *Random Graphs and Complex Networks*, volume 1. Cambridge university press, 2016.

[38] N. C. Wormald. Differential equations for random processes and random graphs. *The Annals of Applied Probability*, 5(4):1217 – 1235, 1995.

[39] N.C. Wormald. The differential equation method for random graph processes and greedy algorithms. In *Lectures on Approximation and Randomization Algorithms (M. Karonski and H.-J. Prömel, eds.)*, 1999.