Adv. Appl. Probab. **56,** 156–204 (2024)
*doi:*10.1017/apr.2023.21

BOOTSTRAP PERCOLATION IN INHOMOGENEOUS RANDOM GRAPHS

HAMED AMINI ,* *University of Florida*
NIKOLAOS FOUNTOULAKIS ,** *University of Birmingham*
KONSTANTINOS PANAGIOTOU,*** *University of Munich*

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 neighbors 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 eventually be infected and show that it also satisfies a law of large numbers.

Keywords: Bootstrap percolation; random graphs; sharp threshold

2020 Mathematics Subject Classification: Primary 05C80
Secondary 60K35

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 neighbors 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 $\mathcal{A}_f$.

Received 5 August 2019; revision received 4 March 2023.
* Postal address: Department of Industrial and Systems Engineering, University of Florida, 468 Weil Hall, Gainesville, FL 32611, USA. Email address: aminil@ufl.edu
** Postal address: School of Mathematics, University of Birmingham, Birmingham B15 2TT, UK. Email address: n.fountoulakis@bham.ac.uk
*** Postal address: Mathematical Institute, University of Munich, Theresienstr. 39, 80333 München, Germany. Email address: kpanagio@math.lmu.de

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 also has connections to the dynamics of the Ising model at zero temperature [23, 29]. A short survey of applications can be found in [1].

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 [10, 15, 24], lattices on the hyperbolic plane [32], and hypercubes [9], as well as on many models of random graphs [3, 13, 26]. In particular, consider the case where $r = 2$ and G is the two-dimensional grid with $V = [n]^2 = \{1, \ldots, n\}^2$ (i.e., a vertex becomes infected if at least two of its neighbors 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, Bollobás and Morris [11] (when G is the three-dimensional grid on $[n]^3$ and $r = 3$) and by Balogh, Bollobás, 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 such that 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 these 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 $|\mathcal{A}_0| = o(n)$, 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], in the case where the sum of the squares of the 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, namely 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., where 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 grows with the number of vertices. Power laws emerge in several contexts, ranging from ecology and economics to social networks (see e.g. the survey of Mitzenmacher [28]). Already in the late 19th century, Pareto observed a power law in the distribution of wealth within populations [30]. In a completely different context, in 1926 Lotka [27] observed a power-law distribution on the frequencies of scientists whose works had been 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 bootstrap-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], which gives a threshold function $a_c(n) = o(n)$ such that if $a(n)$ 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 this fraction exactly, and we show that as long as $a(n) = o(n)$, it does not depend on $a(n)$ itself. In the rest of this section we provide the definition of the random graph model that we consider and the statements 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 $x_n \le C y_n$ for all $n \ge N$, and we write $x_n = o(y_n)$ if $x_n / y_n \to 0$ as $n \to \infty$. We also sometimes 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* 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 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 $w_1 \le w_2 \le \cdots \le w_n$. For any $S \subseteq [n]$, set

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

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

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 in a series of papers [17–19] for fairly general choices of $\mathbf{w}$. Chung and Lu studied several typical properties of the resulting graphs, such as the average distance between two randomly chosen vertices that belong to the same component, and the component size distribution. Their 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 are referring 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

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

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 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 an $x_0 \in \mathbb{R}^+$ 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 w.h.p. The third property in Definition 1 is a minor restriction guaranteeing that no vertex has a vanishing expected degree; it is added for convenience in order to simplify several of our technical considerations.

At many points in our arguments it will be important to select vertices randomly according to their weight, i.e. so that the probability of choosing $i \in [n]$ equals $w_i / W_{[n]}(\mathbf{w})$. This is

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

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,

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

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.,

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

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 1. *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*

$$f_r\left(y; W^*_F, p\right) = 0. \tag{3}$$

Assume also that $f'_r(\hat{y}; W^*_F, p) < 0$. Then

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

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 fails to happen only if

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

and for such (rather exceptional) weight sequences we expect a different behavior. Moreover, we show (cf. Lemma 6) that if the weight sequence has a power-law distribution with exponent between 2 and 3, this case will not happen (i.e., we always have $f'_r(\hat{y}; W^*_F, p) < 0$).

Intuitively, the quantity $\hat{y}$ represents the limit of the probability that infection is passed through a random neighbor 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 neighbors. 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 neighbor 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 no longer 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. 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 $x_0 \leq x < c_3 \cdot n^{\zeta}$,

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

while $F_n(x) = 0$ 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 $c > 0$ that

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

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 typically results 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*, because of the fact that the typical distance between two vertices that belong to the same component is $O(\log\log n)$; see [18, 37].

Theorem 1 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

$$\frac{r-1}{2r-\beta+1} < \zeta \leq \frac{1}{\beta-1},$$

if $a(n) = o(a_c(n))$, 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. $|\mathcal{A}_0| = |\mathcal{A}_f|$. $\left(\text{In particular, } a_c^-(n) = a_c(n) \text{ and } a_c^+(n) = n^{1-\zeta\frac{r-\beta+2}{r-1}}.\right)$ We refine this result using the proof of Theorem 1 and determine the fraction of vertices that belong to $\mathcal{A}_f$.

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

Theorem 2. *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* $a(n) = \omega(a_c(n))$ $\left(\textit{for } \frac{r-1}{2r-\beta+1} < \zeta \leq \frac{1}{\beta-1}\right)$ *and* $a(n) = \omega(a_c^+(n))$ $\left(\textit{for } \zeta \leq \frac{r-1}{2r-\beta+1}\right)$*, then*

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

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

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

When $\beta > 3$, the regularity assumptions of Theorem 1 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. Consequently, the techniques that are used for the proof of Theorem 2 in Section 3.2 do not apply immediately but need 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, asymptotically almost surely, most vertices of weight exceeding some large constant become infected.

2.5. Outline

The proofs of Theorems 1 and 2 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 1 for finitary sequences; this is Theorem 3, stated below. The proof of Theorem 3 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 neighbors of each infected vertex. Such an exposure algorithm has also been applied in the homogeneous setting [26]. Of course, the inhomogeneous setting imposes significant obstacles. We close the paper with the proofs of some rather technical results, which transfer the condition on the derivative that appears in the statement of Theorem 1 to 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 following properties:

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

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

For $\gamma > 0$ let

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

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 the following hold:

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 Lebesgue–Stieltjes integral with respect to F.
3. The assumption $\mathbb{E}[W_F] = d < \infty$ implies that $\mathbb{P}[W_F > x] = o(1/x)$ 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 weight sequence $(\mathbf{w}(n))_{n\geq 1}$ by a sequence where most vertices have 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. 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

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

is an (ℓ, γ)-*discretization* 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 $\gamma_1, \ldots, \gamma_{p(\ell)} \in (0, 1)$ such that $\sum_{i=1}^{p_\ell} \gamma_i = 1 - h_F(\gamma)$, and there are 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 $\mathsf{C}_1(n), \ldots, \mathsf{C}_{p_\ell}(n)$, such that the following hold:

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 $n \to \infty$, the following hold:

3. For all $1 \leq i \leq p_\ell$, $n^{-1}|\mathsf{C}_i(n)| \to \gamma_i$.
4. There is an $h_F(\gamma) \leq \gamma' < h_F(\gamma) + 2W_\gamma(F)/C_\gamma(F)$ 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 $F^{(\ell,\gamma)}$.

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

The upper bounds in Items 4 and 5 are tailored to the proof of Theorem 1. Note that in the previous definition no assumption is made on the W_i, and thus $\mathbf{W}^{(\ell,\gamma)}$ might look very different from $\mathbf{w}$. The next definition quantifies when an (ℓ, γ)-discretization 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 a G-distributed random variable.

Definition 4. Let $(\mathbf{w}(n))_{n\geq 1}$ be regular and let F be its limiting distribution function. A family $\big((\mathbf{W}^{(\ell,\gamma)}(n'))_{n\geq 1}\big)_{\ell\in\mathbb{N},\gamma\in(0,1)}$ of (ℓ, γ)-discretizations of $(\mathbf{w}(n))_{n\geq 1}$ with limiting distribution functions $F^{(\ell,\gamma)}$ is called *F-convergent* if the following hold:

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

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

2. We have

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

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

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

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 4, 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 1. *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 $CL'\big(\mathbf{W}^{(\ell,\gamma)}\big)$. In this modified model the edge probabilities are proportional to the product of the weights of the vertices, except that the normalizing factor is equal not to the sum of the weights in $\mathbf{W}^{(\ell,\gamma)}$, but rather to $W_{[n]}(\mathbf{w}(n))$; that is, the edge $\{i, j\}$ is contained in $CL'\big(\mathbf{W}^{(\ell,\gamma)}\big)$ with probability

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

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

The next theorem quantifies the number of finally infected vertices when the weight sequence is a discretization of a given regular $(\mathbf{w}(n))_{n\geq 1}$. It is general enough to be used in the proof of Theorem 2 as well.

Theorem 3. *Let $(\mathbf{w}(n))_{n\geq 1}$ be regular and let F be its limiting distribution function. Let $\left(\left(\mathbf{W}^{(\ell,\gamma)}(n')\right)_{n\geq 1}\right)_{\ell\in\mathbb{N},\gamma\in(0,1)}$ be a family of (ℓ,γ)-discretizations of $(\mathbf{w}(n))_{n\geq 1}$ which is F-convergent. Moreover, assume that $f_r'(\hat{y};W_F^*,p)<0$ (cf. Theorem 1).*

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 vertices in $[n']\setminus\mathsf{C}_\gamma'(n)$ that eventually become 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 $1-o(1)$,

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

3.1. Proof of Theorem 1

Given a regular $(\mathbf{w}(n))_{n\geq 1}$, Theorem 1 follows from Theorem 3 by constructing an F-convergent family $((\mathbf{W}^{(\ell,\gamma)}(n'))_{n\geq 1})_{\ell\in\mathbb{N},\gamma\in(0,1)}$ for a certain function $n':\mathbb{N}\to\mathbb{N}$. 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 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 $\varepsilon_\ell=1/\ell$. First, 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 $i=0,\dots,t_\ell$, let y_{2i},y_{2i+1} be such that

(1) $\max\left\{\frac{1}{2}(x_{i-1}+x_i),x_i-\varepsilon_\ell\right\}<y_{2i}<x_i$;

(2) $x_i<y_{2i+1}<\min\left\{\frac{1}{2}(x_i+x_{i+1}),x_i+\varepsilon_\ell\right\}$ 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 $\mathcal{P}_\ell:=\{[y_0,y_1),\dots,[y_{2t_\ell},C_\gamma)\}$. With $p_\ell=2t_\ell+1$, for $i=1,\dots,p_\ell$, we set $I_i=[y_{i-1},y_i)$.

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 $i=1,\dots,p_\ell$ 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 $\mathsf{C}_\gamma^-=\{v_1,\dots,v_{k_-}\}$ which is disjoint from $[n]$. We identify with $[n'']$ the set $\left(\cup_{i=1}^{p_\ell}\mathsf{C}_i\right)\bigcup\mathsf{C}_\gamma^-$, through a bijective mapping $\varphi^-:\left(\cup_{i=1}^{p_\ell}\mathsf{C}_i\right)\bigcup\mathsf{C}_\gamma^-\to[n'']$. It follows that $n''=(1-h_F(\gamma)+ph_F(\gamma))n(1+o(1))$.

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

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

$$\varepsilon_j(n) = \frac{w_j(n)}{C_\gamma} - \left\lfloor \frac{w_j(n)}{C_\gamma} \right\rfloor,$$

and we set

$$R = \left\lceil 2 \sum_{j\,:\, w_j(n) \geq 2C_\gamma} \varepsilon_j(n) \right\rceil.$$

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 $2C_\gamma$:

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

Let $n' = \left| \left(\cup_{i=1}^{p_\ell} \mathsf{C}_i\right) \bigcup \mathsf{C}_\gamma^+ \right|$, and identify the set $[n']$ with the vertices in $\left(\cup_{i=1}^{p_\ell} \mathsf{C}_i\right) \bigcup \mathsf{C}_\gamma^+$, through a bijection $\varphi^+ : \left(\cup_{i=1}^{p_\ell} \mathsf{C}_i\right) \bigcup \mathsf{C}_\gamma^+ \to [n']$. We will use the symbol C_γ^+ to denote the set $[n'] \setminus \varphi^+ \left(\left(\cup_{i=1}^{p_\ell} \mathsf{C}_i\right)\right)$. 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 $W_i^- = y_{2(i-1)}$ and $W_i^+ = y_{2i-1}$; for each $j \in \mathsf{C}_i$, we set

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

For any $j \in [n''] \setminus \varphi^- \left(\cup_{i=1}^{p_\ell} \mathsf{C}_i\right)$ 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 $W_{\mathsf{C}_\gamma^-}\left(\mathbf{W}^{(\ell,\gamma)-}\right)$ denotes the total weight of these vertices, then this satisfies

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

Furthermore,

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

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

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

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

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

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

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

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

Hence, as $n \to \infty$,

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

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 $F_n^{(\ell,\gamma)-}$ and $F_n^{(\ell,\gamma)+}$ converge pointwise, as $n \to \infty$, to the functions $F^{(\ell,\gamma)-}$ and $F^{(\ell,\gamma)+}$, 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 $x \ge C_\gamma$ 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 $C_\gamma \le x < 2C_\gamma$ we have

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

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

We will now prove that both families

$$\left\{\mathbf{W}^{(\ell,\gamma)+}(n')\right\}_{\gamma \in (0,1), \ell \in \mathbb{N}} \quad \text{and} \quad \left\{\mathbf{W}^{(\ell,\gamma)-}(n'')\right\}_{\gamma \in (0,1), \ell \in \mathbb{N}}$$

are F-convergent. Thus, we will verify that they satisfy both parts of Definition 4.

Part 1 of Definition 4. 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

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

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

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 therefore 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 that 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 $\ell \to \infty$,

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

whereas as $\gamma \downarrow 0$ we have

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

Claim 4. *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 cumulative distribution function 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. First 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 $\ell \to \infty$,

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

and

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

Furthermore, we will show the following.

Lemma 2. *We have* $\lim_{\gamma\downarrow 0} \mathbb{E}\big(U^{(\gamma)+}\big) = \lim_{\gamma\downarrow 0} \mathbb{E}(U^{(\gamma)-}) = \mathbb{E}(W_F)$.

Proof. The proof is identical for both $U^{(\gamma)+}$ and $U^{(\gamma)-}$, so we will denote these by $U^{(\gamma)\pm}$. For $\delta > 0$, let $\bar{C}_\delta$ be a continuity point of F such that

$$\mathbb{E}\left(W_F\mathbf{1}_{W_F>\bar{C}_\delta}\right) < \delta/3. \tag{14}$$

By Claim 4 we deduce that for any γ sufficiently small,

$$|\mathbb{E}\left(U^{(\gamma)\pm}\mathbf{1}_{U^{(\gamma)\pm}\leq\bar{C}_\delta}\right) - \mathbb{E}\left(W_F\mathbf{1}_{W_F\leq\bar{C}_\delta}\right)| < \delta/3. \tag{15}$$

Let us bound $\mathbb{E}(U^{(\gamma)\pm}\mathbf{1}_{U^{(\gamma)\pm}>\bar{C}_\delta})$. Note that if γ is small enough, then

$$\mathbb{E}\big(U^{(\gamma)\pm}\mathbf{1}_{U^{(\gamma)\pm}>\bar{C}_\delta}\big)=\mathbb{E}\big(U^{(\gamma)\pm}\mathbf{1}_{C_\gamma\geq U^{(\gamma)\pm}>\bar{C}_\delta}\big).$$

So it suffices to bound the latter term.

Claim 5. *For any $\delta>0$, if γ is sufficiently small, then*

$$\left|\mathbb{E}\left(U^{(\gamma)\pm}\mathbf{1}_{C_\gamma\geq U^{(\gamma)\pm}>\bar{C}_\delta}\right)-\mathbb{E}\left(W_F\mathbf{1}_{C_\gamma\geq W_F>\bar{C}_\delta}\right)\right|<\delta/3.$$

Proof. For $\bar{C}_\delta<x<C_\gamma$, we have $F^{(\gamma)\pm}(x)=\lambda^\pm(\gamma)\cdot F(x)$, where $\lambda^\pm(\gamma)\to 1$ as $\gamma\downarrow 0$. Furthermore, $F^{(\gamma)^\pm}(C_\gamma)-F(C_\gamma)=1-(1-h_F(\gamma))=h_F(\gamma)$. For some integer $T>0$, consider a partition of $[0,C_\gamma]$ given by $p_0=0<p_1<\cdots<p_T=C_\gamma$, which are points of continuity of F. Taking $f(x)=x\mathbf{1}_{\bar{C}_\delta<x\leq C_\gamma}$, we write

$$\begin{aligned}&\sum_{i=1}^T f(p_i)\Big(F^{(\gamma)^\pm}(p_i)-F^{(\gamma)^\pm}(p_{i-1})\Big)=\\&f(p_T)\Big(F^{(\gamma)^\pm}(p_T)-F^{(\gamma)^\pm}(p_{T-1})\Big)+\sum_{1\leq i<T}f(p_i)\Big(F^{(\gamma)^\pm}(p_i)-F^{(\gamma)^\pm}(p_{i-1})\Big).\end{aligned}$$

The second term on the right-hand side is

$$\sum_{1\leq i<T}f(p_i)\Big(F^{(\gamma)^\pm}(p_i)-F^{(\gamma)^\pm}(p_{i-1})\Big)=\lambda^\pm(\gamma)\cdot\sum_{1\leq i<T}f(p_i)(F(p_i)-F(p_{i-1})).$$

The first term can be written as

$$\begin{aligned}&f(p_T)\Big(F^{(\gamma)^\pm}(p_T)-F^{(\gamma)^\pm}(p_{T-1})\Big)\\&=f(p_T)\Big(F^{(\gamma)^\pm}(p_T)+F(p_T)-F(p_T)-F(p_{T-1})+F(p_{T-1})-F^{(\gamma)^\pm}(p_{T-1})\Big)\\&=f(p_T)(F(p_T)-F(p_{T-1}))\\&\qquad\qquad+f(p_T)\Big(\Big(F^{(\gamma)^\pm}(p_T)-F(p_T)\Big)-\Big(F^{(\gamma)^\pm}(p_{T-1})-F(p_{T-1})\Big)\Big)\\&\overset{p_T=C_\gamma}{=}f(p_T)(F(p_T)-F(p_{T-1}))\\&\qquad\qquad+f(p_T)\left((F^{(\gamma)^\pm}(C_\gamma)-F(C_\gamma))-F(p_{T-1})(\lambda^\pm(\gamma)-1)\right)\\&\overset{F^{(\gamma)^\pm}(C_\gamma)=1,h_F(\gamma)=1-F(C_\gamma)}{=}f(p_T)(F(p_T)-F(p_{T-1}))\\&\qquad\qquad+f(p_T)\Big(h_F(\gamma)-F(p_{T-1})\Big(\lambda^\pm(\gamma)-1\Big)\Big)\\&=\lambda^\pm(\gamma)f(p_T)(F(p_T)-F(p_{T-1}))+(1-\lambda^\pm(\gamma))f(C_\gamma)(F(p_T)-F(p_{T-1}))\\&\qquad\qquad+f(C_\gamma)h_F(\gamma)-f(C_\gamma)F(p_{T-1})\left(\lambda^\pm(\gamma)-1\right).\end{aligned}$$

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

These two calculations imply that

$$\sum_{i=1}^{T} f(p_i)\left(F^{(\gamma)^{\pm}}(p_i) - F^{(\gamma)^{\pm}}(p_{i-1})\right) =$$

$$\lambda^{\pm}(\gamma)\cdot\sum_{i=1}^{T} f(p_i)(F(p_i) - F(p_{i-1})) + (1-\lambda^{\pm}(\gamma))f(C_\gamma)(F(p_T) - F(p_{T-1}))$$
$$+f(C_\gamma)h_F(\gamma) - f(C_\gamma)F(p_{T-1})\left(\lambda^{\pm}(\gamma)-1\right).$$

Taking the limit of a sequence of partitions whose mesh tends to 0, we have $p_{T-1}\uparrow p_T = C_\gamma$. Since C_γ is a point of continuity for F, we obtain

$$\mathbb{E}\left(U^{(\gamma)\pm}\mathbf{1}_{C_\gamma \ge U^{(\gamma)\pm} > \bar{C}_\delta}\right) =$$
$$\lambda^{\pm}(\gamma)\mathbb{E}\left(W_F \mathbf{1}_{C_\gamma \ge W_F > \bar{C}_\delta}\right) + C_\gamma\left(h_F(\gamma) - F(C_\gamma)\left(\lambda^{\pm}(\gamma)-1\right)\right).$$

By (5) we have that $C_\gamma h_F(\gamma)\downarrow 0$ as $\gamma\downarrow 0$. Since $|\lambda^{\pm}(\gamma)-1| = O(\gamma^+ + h_F(C_\gamma))$, by (7) we also have $C_\gamma|\lambda^{\pm}(\gamma)-1|\downarrow 0$ as $\gamma\downarrow 0$. □

Combining Claim 5 with (14) and (15), we finally deduce that given $\delta>0$, for any γ that is sufficiently small, we have the following bound:

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

□

Claim 4 also implies that for every $x>0$ which is a point of continuity of F, we have as $\gamma\downarrow 0$

$$|\mathbb{E}\left(U^{(\gamma)\pm}\mathbf{1}_{U^{(\gamma)\pm}\le x}\right) - \mathbb{E}\left(W_F\mathbf{1}_{W_F\le x}\right)|\downarrow 0. \tag{16}$$

So from Lemma 2, (16), 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

$$\left|F^{*(\ell,\gamma)}(x) - F^*(x)\right| < \delta. \tag{17}$$

The above can now be translated into the next lemma.

Lemma 3. *For any bounded and continuous function $f:\mathbb{R}\to\mathbb{R}$ the following holds: there exists a function $\gamma_f:\mathbb{R}_+\to\mathbb{R}_+$ such that $\gamma_f(x)\downarrow 0$ as $x\downarrow 0$, and moreover, for any $\delta>0$ and any $0<\gamma<\gamma_f(\delta)$ there exists ℓ_0 with the property that for any $\ell>\ell_0$*

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

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 upon each other when taking the double limit.

This completes the first part of Definition 4. We now turn to the second part.

Part 2 of Definition 4. Since $F^{(\ell,\gamma)+}$ and $F^{(\ell,\gamma)-}$ are both constant (and equal to 1), for $x \geq 2C_\gamma$ we have

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

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,

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

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 (18). We write $F^{(\ell,\gamma)\pm}$ for either of $F^{(\ell,\gamma)+}$ and $F^{(\ell,\gamma)-}$. Using the integration-by-parts formula for the Lebesgue–Stieltjes integral, we can write

$$\begin{aligned} \mathbb{E}\left(U^{(\ell,\gamma)\pm}\mathbf{1}_{U^{(\ell,\gamma)\pm}\leq 2C_\gamma}\right) &= 2C_\gamma F^{(\ell,\gamma)\pm}(2C_\gamma+) - \int_0^{2C_\gamma} F^{(\ell,\gamma)\pm}(x)dx \\ &= 2C_\gamma - \int_0^{2C_\gamma} F^{(\ell,\gamma)\pm}(x)dx. \end{aligned} \tag{19}$$

Using integration by parts, we also get

$$\begin{aligned} \mathbb{E}\left(W_F\mathbf{1}_{W_F\leq 2C_\gamma}\right) &= 2C_\gamma\cdot F(2C_\gamma+) - 0\cdot F(0-) - \int_0^{2C_\gamma} F(x)dx \\ &= 2C_\gamma\cdot\left(1-\mathbb{P}\left[W_F>2C_\gamma\right]\right) - \int_0^{2C_\gamma} F(x)dx. \end{aligned} \tag{20}$$

Therefore,

$$\begin{aligned} &\left|\mathbb{E}\left(U^{(\ell,\gamma)\pm}\mathbf{1}_{U^{(\ell,\gamma)\pm}\leq 2C_\gamma}\right) - \mathbb{E}\left(W_F\mathbf{1}_{W_F\leq 2C_\gamma}\right)\right| \leq \\ &2C_\gamma\mathbb{P}\left[W_F>2C_\gamma\right] + \int_0^{C_\gamma} |F^{(\ell,\gamma)\pm}(x) - F(x)|dx + \int_{C_\gamma}^{2C_\gamma} |F^{(\ell,\gamma)\pm}(x) - F(x)|dx. \end{aligned} \tag{21}$$

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

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

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

For 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:

$$\begin{aligned}
|F^{(\ell,\gamma)+}(x) - F(x)| &= \left| \frac{F(W_{2i}^-)}{1 - h_F(\gamma) + \gamma^+} - F(x) \right| = \left| \frac{F(W_{2i}^-) - F(x)(1 - h_F(\gamma) + \gamma^+)}{1 - h_F(\gamma) + \gamma^+} \right| \\
&\leq \frac{|F(W_{2i}^-) - F(x)|}{1 - h_F(\gamma) + \gamma^+} + F(x) \frac{|h_F(\gamma) - \gamma^+|}{1 - h_F(\gamma) + \gamma^+} \\
&\leq \frac{F(W_{2i}^+) - F(W_{2i}^-)}{1 - h_F(\gamma) + \gamma^+} + \frac{|h_F(\gamma) - \gamma^+|}{1 - h_F(\gamma) + \gamma^+}.
\end{aligned} \tag{22}$$

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,

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

We thus deduce that

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

For the second integral, note that for $x \in [y_{2i+1}, y_{2(i+1)})$ we have

$$\left| F\big(W_{2i+1}^-\big) - F(x) \right| \leq \varepsilon_\ell.$$

Arguing as in (22), we deduce that

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

Therefore,

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

We thus conclude that

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

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

One can similarly show that

$$\int_0^{C_\gamma} |F^{(\ell,\gamma)-}(x) - F(x)|dx \le \frac{\varepsilon_\ell(C_\gamma + 2) + 2C_\gamma h_F(\gamma)}{1 - h_F(\gamma) + p h_F(\gamma)} =: \hat{\rho}^-(\ell, \gamma). \tag{24}$$

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

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

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

Therefore,

$$\begin{aligned} &\left| \int_{C_\gamma}^{2C_\gamma} (F^{(\ell,\gamma)+}(x) - F(x))dx \right| \le \int_{C_\gamma}^{2C_\gamma} |F^{(\ell,\gamma)+}(x) - F(x)|dx \le \\ &\le \frac{|h_F(\gamma) - \gamma^+|}{1 - h_F(\gamma) + \gamma^+} \cdot (2C_\gamma - C_\gamma) = C_\gamma \cdot \frac{|h_F(\gamma) - \gamma^+|}{1 - h_F(\gamma) + \gamma^+} \le \hat{\rho}^+(\ell, \gamma). \end{aligned} \tag{25}$$

Also,

$$\begin{aligned} &\int_{C_\gamma}^{2C_\gamma} \left| F^{(\ell,\gamma)-}(x) - F(x) \right| dx = \int_{C_\gamma}^{2C_\gamma} |1 - F(x)|\, dx = \int_{C_\gamma}^{2C_\gamma} \mathbb{P}\left[W_F > x\right] dx \\ &\le \int_{C_\gamma}^{\infty} \mathbb{P}\left[W_F \ge x\right] dx = \mathbb{E}\left(W_F \mathbf{1}_{W_F \ge C_\gamma}\right). \end{aligned} \tag{26}$$

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

Now, substituting the bounds of (23), (24), (25), and (26) into (21), we get

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

Using the upper bound of (27) in (18), and using the fact that $\mathbb{E}\left(W_F \mathbf{1}_{W_F > 2C_\gamma}\right) \le \mathbb{E}\left(W_F \mathbf{1}_{W_F \ge C_\gamma}\right)$, we finally get

$$\begin{aligned} \left| \mathbb{E}\left(U^{(\ell,\gamma)\pm}\right) - \mathbb{E}\left(W_F\right) \right| &\le 2C_\gamma \mathbb{P}\left[W_F > 2C_\gamma\right] + 2\hat{\rho}^{\pm}(\ell, \gamma) + 2\mathbb{E}\left(W_F \mathbf{1}_{W_F \ge C_\gamma}\right) \\ &\le 2\hat{\rho}^{\pm}(\ell, \gamma) + 3\mathbb{E}\left(W_F \mathbf{1}_{W_F \ge C_\gamma}\right). \end{aligned} \tag{28}$$

Another useful bound which can be deduced from (19) and (20), replacing $2C_\gamma$ with C_γ, is

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

But note that $\mathbb{P}\left[U^{(\ell,\gamma)-}\le C_\gamma\right]=1$, whereas

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

So

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

and

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

Thus,

$$\begin{aligned}
&\left|\mathbb{E}\left(U^{(\ell,\gamma)\pm}\mathbf{1}_{U^{(\ell,\gamma)\pm}\le C_\gamma}\right)-\mathbb{E}\left(W_F\right)\right| \\
\le\ & C_\gamma\frac{F(C_\gamma)|h_F(\gamma)-\gamma^+|}{1-h_F(\gamma)+\gamma^+}+\int_0^{C_\gamma}|F^{(\ell,\gamma)\pm}(x)-F(x)|dx+\mathbb{E}\left(W_F\mathbf{1}_{W_F>C_\gamma}\right) \\
\overset{(23),(24)}{\le} & C_\gamma\frac{|h_F(\gamma)-\gamma^+|}{1-h_F(\gamma)+\gamma^+}+\hat{\rho}^\pm(\ell,\gamma)+\mathbb{E}\left(W_F\mathbf{1}_{W_F>C_\gamma}\right),
\end{aligned}\tag{29}$$

whereas

$$\begin{aligned}
\left|\mathbb{E}\left(U^{(\ell,\gamma)+}\mathbf{1}_{U^{(\ell,\gamma)\pm}\le C_\gamma}\right)-\mathbb{E}\left(W_F\right)\right|\le& C_\gamma\mathbb{P}\left[W_F>C_\gamma\right]+\hat{\rho}^\pm(\ell,\gamma) \\
&+\mathbb{E}\left(W_F\mathbf{1}_{W_F>C_\gamma}\right).
\end{aligned}\tag{30}$$

We conclude that (28), together with (5), (29), and (30), completes Part 2 of Definition 4.

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 $CL'\big(\mathbf{W}^{(\ell,\gamma)-}\big)$, 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 $\mathcal{A}_f^+(S)$ be the final set of infected vertices on $CL'\big(\mathbf{W}^{(\ell,\gamma)+}\big)$. We will show the following.

Claim 6. *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.,*

$$\left|\mathcal{A}_f^-\left(\mathcal{A}_0\cup\mathsf{C}_\gamma^-\right)\right|\le|\mathcal{A}_f|\le\left|\mathcal{A}_f^+\left(\mathcal{A}_0\cup\mathsf{C}_\gamma^+\right)\right|.\tag{31}$$

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

In the proof of Claim 6 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 6. 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 that 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 in $CL'\big(\mathbf{W}^{(\ell,\gamma)-}\big)$ is smaller than the corresponding probability for k and k' in $CL(\mathbf{w})$. Hence, there is a coupling space on which

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

and the first inequality in (31) 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 $CL'\big(\mathbf{W}^{(\ell,\gamma)+}\big)$ is

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

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

$$\frac{w_k w_j}{W_{[n]}} \leq 1 - \left(1 - \frac{2W_{\varphi^+(k)}^{(\ell,\gamma)+} C_\gamma}{W_{[n]}}\right)^{2\lfloor w_j/C_\gamma \rfloor}. \tag{32}$$

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

$$CL(\mathbf{w})\big[\cup_{i=1}^{p_\ell} \mathsf{C}_i\big] \subseteq CL'\big(\mathbf{W}^{(\ell,\gamma)+}\big)\big[\cup_{i=1}^{p_\ell} \mathsf{C}_i\big]. \tag{33}$$

This is the case because for any $k \in \cup_{i=1}^{p_\ell} \mathsf{C}_i$ we have $w_k \leq W_{\varphi^+(k)}^{(\ell,\gamma)+}$. Consider a vertex $k \in \cup_{i=1}^{p_\ell} \mathsf{C}_i$ and now let $j \in \mathsf{C}_\gamma$. Then the inequality (32) 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 $CL'\big(\mathbf{W}^{(\ell,\gamma)+}\big)$. Therefore, it follows that the number of neighbors of k in C_γ in the random graph $CL(\mathbf{w})$ is stochastically dominated by the size of the neighborhood of k in C_γ' in the random graph $CL'\big(\mathbf{W}^{(\ell,\gamma)+}\big)$. This observation, together with (33), implies that

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

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 (32). Let us abbreviate $W_{\varphi^+(k)}^{(\ell,\gamma)+} =: W_k$. Using the Bonferroni inequalities we have

$$1 - \left(1 - \frac{2W_k C_\gamma}{W_{[n]}}\right)^{2\lfloor w_j/C_\gamma \rfloor} \geq 2\left\lfloor \frac{w_j}{C_\gamma} \right\rfloor \frac{2W_k C_\gamma}{W_{[n]}} - 2\left(w_j/C_\gamma\right)^2 \frac{4W_k^2 C_\gamma^2}{W_{[n]}^2}. \tag{34}$$

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

But

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

Substituting this lower bound into (34), we obtain

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

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 to the random variables that bound $|\mathcal{A}_f|$ in Claim 6. Theorem 3 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 $1 - o(1)$,

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

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

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

Hence, Claim 6 together with (35) and (36) imply the following w.h.p. bounds on the size of $\mathcal{A}_f$:

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

from which Theorem 1 follows.

3.2. Proof of Theorem 2

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. 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 6.

Claim 7. *For any $\varepsilon \in (0, 1)$ and any $\gamma > 0$, if n is large enough, then*

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

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

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

$\mathcal{K}_C$ 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 4. ([6, Proposition 3.7].) *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 4 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; if it is realized, 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). Then the following holds.

Claim 8. *For any* $C > 0$ *and any* $\varepsilon \in (0, 1)$, *there is a coupling such that if* $\mathcal{E}_{C,\varepsilon,n}$ *is realized, 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 4. (Under the assumptions of Theorem 2, F is continuous (cf. Definition 2), 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 $\left(\cup_{i=1}^{p_\ell} \mathsf{C}_i\right) \bigcup \{v_1, \ldots, v_k\}$ 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 $\left(\cup_{i=1}^{p_\ell} \mathsf{C}_i\right) \bigcup \{v_1, \ldots, v_k\}$. 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 k lexicographically 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

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

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

(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 $CL'\big(\tilde{\mathbf{W}}^{(\ell,\gamma)-}\big)$ 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 $CL'\big(\tilde{\mathbf{W}}^{(\ell,\gamma)-}\big) \subseteq CL(\mathbf{w})$ on a certain coupling space we have

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

Therefore

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

This together with Claim 8 implies the following stochastic lower bound on $|\mathcal{A}_f|$.

Claim 9. *For any γ, $\varepsilon \in (0, 1)$, if $\mathcal{E}_{C_\gamma,\varepsilon,n}$ is realized, then*

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

We will now apply Theorem 3 to the random variables that bound $|\mathcal{A}_f|$ in Claims 7 and 9. 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 4 and let $\gamma < \gamma_2$ (cf. Theorem 3) be such that $C = C_\gamma$. Theorem 3 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 $1 - o(1)$,

$$\frac{\big|\mathcal{A}_f^+\big(\mathcal{A}_0^{(\varepsilon)} \cup \mathsf{C}_\gamma^+\big)\big|}{n} \leq (1+\delta)\big((1-\varepsilon)\mathbb{E}\left[\psi_r\big(W_F \hat{y}_\varepsilon^+\big)\right] + \varepsilon\big), \tag{37}$$

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

$$\frac{\big|\tilde{\mathcal{A}}_f(\mathsf{C}_{\gamma,k})\big|}{n} \geq (1-\delta)\mathbb{E}\left[\psi_r(W_F \hat{y})\right]. \tag{38}$$

Hence, Claims 7 and 9 together with (37) and (38) imply that w.h.p.

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

and

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

But $y_\varepsilon^+ \to \hat{y}$ as $\varepsilon \to 0$, and Theorem 2 follows.

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

4. Proof of Theorem 3

In this section we will give the proof of Theorem 3. At the moment our analysis does not depend on the parameters ℓ, γ, so, 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 neighbors. 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 neighbors or are initially infected.

We will determine the size of the final set of infected vertices by exposing *sequentially* the neighbors of each infected vertex and keeping track of the number of infected neighbors that 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 introduces 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 the set of vertices which have j infected neighbors 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 neighbors after the 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 $t \geq 1$, if $\mathsf{U}(t-1)$ is non-empty,

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

(ii) expose the neighbors 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 neighbors of v can alternatively be 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 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 $\mathbf{c}(t) = \big(u(t), w_{\mathsf{U}}(t), (c_{i,j}(t))_{i,j}\big)$. 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

$$\begin{aligned}\mathbb{E}\left[c_{i,0}(t) - c_{i,0}(t-1) \mid \mathbf{c}(t-1)\right] &= -c_{i,0}(t-1) \sum_{v \in \mathsf{U}(t-1)} \frac{W_i w_v}{W_{[n]}} \frac{1}{u(t-1)} \\ &= -c_{i,0}(t-1) \frac{W_i}{W_{[n]}} \frac{w_{\mathsf{U}}(t-1)}{u(t-1)}.\end{aligned} \tag{39}$$

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 $c_{i,j-1}$.

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

For $i=1, \ldots, p_\ell$ and $0<j<r$ we have

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

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

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

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

$$
\begin{aligned}
&\mathbb{E}\left[w_{\mathsf{U}}(t)-w_{\mathsf{U}}(t-1) \mid \mathbf{c}(t-1)\right] \\
&= -\frac{w_{\mathsf{U}}(t-1)}{u(t-1)} + \sum_{i=1}^{p_\ell} W_i c_{i,r-1}(t-1) \sum_{v\in \mathsf{U}(t-1)} \frac{W_i w_v}{W_{[n]}} \frac{1}{u(t-1)} \\
&= -\frac{w_{\mathsf{U}}(t-1)}{u(t-1)} + \frac{w_{\mathsf{U}}(t-1)}{u(t-1)} \sum_{i=1}^{p_\ell} \frac{W_i^2}{W_{[n]}} c_{i,r-1}(t-1).
\end{aligned} \tag{42}
$$

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 ν, $\mu_{\mathsf{U}} : [0, \infty) \to \mathbb{R}$ 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 determines the functions $\gamma_{i,j}$ is as follows:

$$
\begin{aligned}
\frac{d\gamma_{i,0}}{d\tau} &= -\gamma_{i,0}(\tau)\frac{W_i}{d}G(\tau), \\
\frac{d\gamma_{i,j}}{d\tau} &= \left(\gamma_{i,j-1}(\tau)-\gamma_{i,j}(\tau)\right)\frac{W_i}{d}G(\tau), \qquad 1\le j\le r-1.
\end{aligned} \tag{43}
$$

The continuous counterparts of (41) and (42) are

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

and

$$\frac{d\mu_{\mathsf{U}}}{d\tau} = -G(\tau) + G(\tau)\sum_{i=1}^{p_\ell} \frac{W_i^2}{d}\,\gamma_{i,r-1}(\tau). \tag{45}$$

The initial conditions are

$$\begin{aligned}
&\nu(0) = p(1-h_F(\gamma)) + \gamma', \ \text{ for } p \in [0,1) \text{ (recall that } p \text{ is the initial infection rate)},\\
&\mu_{\mathsf{U}}(0) = W'_\gamma + p\sum_{i=1}^{p_\ell} W_i\gamma_i,\\
&\gamma_{i,0}(0) = (1-p)\gamma_i,\\
&\gamma_{i,j}(0) = 0, \ \text{ for } j = 1, \ldots, r-1.
\end{aligned} \tag{46}$$

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

Proposition 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 $1 \le j \le r-1$,

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

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. Therefore, and using the expression for $\gamma_{i,0}$, we obtain

$$\exp\left(\int_0^s a(\rho)d\rho\right) = \frac{\gamma_{i,0}(s)}{\gamma_{i,0}(0)}. \tag{47}$$

Hence

$$\begin{aligned}
&\int_0^\tau b(s)\exp\left(-\int_0^s a(\rho)d\rho\right)ds = \\
&\qquad \frac{W_i}{d(j-1)!}\int_0^\tau \gamma_{i,0}(s)\log^{j-1}\left(\frac{\gamma_{i,0}(0)}{\gamma_{i,0}(s)}\right)G(s)\,\frac{\gamma_{i,0}(0)}{\gamma_{i,0}(s)}ds\\
&\qquad = \gamma_{i,0}(0)\,\frac{W_i}{d(j-1)!}\int_0^\tau \gamma_{i,0}(s)\log^{j-1}\left(\frac{\gamma_{i,0}(0)}{\gamma_{i,0}(s)}\right)\frac{G(s)}{\gamma_{i,0}(s)}ds
\end{aligned}$$

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

$$
\begin{aligned}
&= -\frac{\gamma_{i,0}(0)}{(j-1)!}\int_0^\tau \frac{1}{\gamma_{i,0}(s)}\log^{j-1}\left(\frac{\gamma_{i,0}(0)}{\gamma_{i,0}(s)}\right)\left(-\gamma_{i,0}(s)\frac{W_i}{d}G(s)\right)ds \\
&\overset{(43)}{=} -\frac{1}{(j-1)!}\int_0^\tau \frac{\gamma_{i,0}(0)}{\gamma_{i,0}(s)}\log^{j-1}\left(\frac{\gamma_{i,0}(0)}{\gamma_{i,0}(s)}\right)\left(\frac{d\gamma_{i,0}}{ds}\right)ds \\
&= -\frac{\gamma_{i,0}(0)}{(j-1)!}\int_0^\tau \frac{\gamma_{i,0}(0)}{\gamma_{i,0}(s)}\log^{j-1}\left(\frac{\gamma_{i,0}(0)}{\gamma_{i,0}(s)}\right)d\left(\frac{\gamma_{i,0}}{\gamma_{i,0}(0)}\right) \\
&\overset{(x=\gamma_{i,0}/\gamma_{i,0}(0))}{=} -\frac{\gamma_{i,0}(0)}{(j-1)!}\int_1^{\gamma_{i,0}(\tau)/\gamma_{i,0}(0)} \frac{1}{x}\log^{j-1}\left(\frac{1}{x}\right)dx \\
&= (-1)^{j-1}\frac{\gamma_{i,0}(0)}{(j-1)!}\int_{\gamma_{i,0}(\tau)/\gamma_{i,0}(0)}^1 \frac{\log^{j-1}(x)}{x}dx.
\end{aligned} \tag{48}
$$

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

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

Therefore, the last integral in (48) 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 (48), we obtain

$$
\int_0^\tau b(s)\exp\left(-\int_0^s a(\rho)d\rho\right)ds = \frac{\gamma_{i,0}(0)}{j!}\log^j\left(\frac{\gamma_{i,0}(0)}{\gamma_{i,0}(\tau)}\right). \tag{49}
$$

Combining (47) and (49), 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 $1 \leq i \leq p_\ell$, and integrate (44) in order to deduce the expressions for ν and μ_{U}.

Proposition 2. *We have*

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

and

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

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

Proof. Applying Proposition 1 to (44) yields

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

By integrating this expression we obtain

$$\begin{aligned} \nu(\tau) &= \nu(0) - \tau + \frac{1}{(r-1)!} \sum_{i=1}^{p_\ell} \int_0^\tau \frac{W_i}{d} \gamma_{i,0}(s) G(s) \log^{r-1}\left(\frac{\gamma_{i,0}(0)}{\gamma_{i,0}(s)}\right) ds \\ &\overset{(43)}{=} \nu(0) - \tau - \frac{1}{(r-1)!} \sum_{i=1}^{p_\ell} \int_0^\tau \left(\frac{d\gamma_{i,0}}{ds}\right) \log^{r-1}\left(\frac{\gamma_{i,0}(0)}{\gamma_{i,0}(s)}\right) ds \\ &= \nu(0) - \tau - \frac{1}{(r-1)!} \sum_{i=1}^{p_\ell} \gamma_{i,0}(0) \int_1^{\gamma_{i,0}(\tau)/\gamma_{i,0}(0)} \log^{r-1}\left(\frac{1}{x}\right) dx. \end{aligned} \tag{50}$$

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

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

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 $y = 1/x$. Therefore, 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 (50), 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 $\log\left(\gamma_{i,0}(0)/\gamma_{i,0}(\tau)\right)$ is at least r. But by Proposition 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 (46), $\gamma_{i,0}(0) = (1-p)\gamma_i$, 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 for μ_{U} is obtained along the same lines, and we omit its proof. □

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

4.3. Wormald's theorem

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

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

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

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

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

Theorem 10. *([38].) Let $b, n \in \mathbb{N}$. For $1 \leq j \leq b$, suppose that $Y_j^{(n)}(t)$ is a sequence of real-valued random variables such that $0 \leq Y_j^{(n)} \leq Cn$ for some constant $C > 0$. Let H_t be the history up to time t, i.e., the sequence $\{Y_j^{(n)}(k),\ 0 \leq j \leq b,\ 0 \leq k \leq t\}$. Suppose also that for some bounded connected open set $D \subseteq \mathbb{R}^{b+1}$ containing the intersection of $\{(t, z_1, \ldots, z_b) : t \geq 0\}$ with some neighborhood of*

$$\left\{(0, z_1, \ldots, z_b) : \mathbb{P}\left[Y_j^{(n)}(0) = z_j n,\ 1 \leq j \leq b\right] \neq 0 \text{ for some } n\right\},$$

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 $l \leq b$ 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 $l \leq b$ and uniformly over all $t < T_D$,*
$$\mathbb{E}[Y_l^{(n)}(t+1) - Y_l^{(n)}(t) | H_t] = f_l(t/n, Y_1^{(n)}(t)/n, \ldots, Y_b^{(n)}(t)/n) + o(1).$$
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 $(0, \hat{z}_1, \ldots, \hat{z}_b) \in D$, the system of differential equations*
$$\frac{dz_l}{ds} = f_l(s, z_1, \ldots, z_l), \qquad l = 1, \ldots, b,$$
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 $0 \leq t \leq \min\{\sigma n, T_D\}$ 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.

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

4.4. Proof of Theorem 3

We will apply Theorem 10 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., $\{\nu,\, \mu_{\mathsf{U}},\, (\gamma_{i,j})_{i=1,\ldots,p_\ell, j=0,\ldots,r-1}\}$.

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

$$D_\epsilon = \left\{(\tau,\, \nu,\, \mu_{\mathsf{U}},\, (\gamma_{i,j})_{i,j}) \in \mathbb{R}^{b+1} \mid -\epsilon < \tau < 1,\ 0 < \frac{\mu_{\mathsf{U}}}{\nu} < 2C_\gamma,\ -\epsilon < \gamma_{i,j} < \gamma_i + \epsilon,\right.$$
$$\left. \epsilon < \mu_{\mathsf{U}} < W'_\gamma + \sum_{i=1}^{p_\ell} W_i\gamma_i \right\}.$$

We now apply the last part (b) of Theorem 10. Note that the 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 (39)–(42). 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

$$\begin{aligned} u(t) &= n\nu(t/n) + o_p(n),\\ w_{\mathsf{U}}(t) &= n\mu_{\mathsf{U}}(t/n) + o_p(n),\\ c_{i,j}(t) &= n\gamma_{i,j}(t/n) + o_p(n), \ \text{ for all } i = 1, \ldots, p_\ell,\ j = 0, \ldots, r-1. \end{aligned} \tag{51}$$

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 $2C_\gamma$, implies that at round T_{D_ϵ} we have $w_{\mathsf{U}}(T_{D_\epsilon})/n \le \epsilon$, but $w_{\mathsf{U}}(T_{D_\epsilon} - 1)/n > \epsilon$.

First, 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 (51), w.h.p.

$$\mu_{\mathsf{U}}\left(\frac{T_{D_\epsilon} - 1}{n}\right) < 2\epsilon. \tag{52}$$

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

$$u\left(T_{D_\epsilon} - 1\right)/n \le \frac{1.5\epsilon}{W_0}. \tag{53}$$

Therefore, (51) again implies that w.h.p.

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

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

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 2 implies that

$$\left| \frac{T_{D_\epsilon} - 1}{n} - \alpha\left(\frac{1}{d} I\left(\frac{T_{D_\epsilon} - 1}{n}\right)\right)\right| \le \frac{2\epsilon}{W_0}. \tag{54}$$

Let $\hat{\tau}^{(\ell,\gamma)}$ denote the minimum $\tau > 0$ such that $\mu_{\mathsf{U}}(\tau) = 0$. By Lemma 5 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

$$\left|\alpha(\hat{y}_{\ell,\gamma}) - \left(p + (1-p)\mathbb{E}(\psi_r(W_F \hat{y}))\right)\right| < \delta, \tag{55}$$

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 (52), from the continuity of the function μ_{U}, we deduce that there exists $\delta_1 = \delta_1(\epsilon) > 0$ such that, for n large enough,

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

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

$$\left|\alpha\left(\frac{1}{d} I\left(\hat{\tau}^{(\ell,\gamma)}\right)\right) - \alpha\left(\frac{1}{d} I\left(\frac{T_{D_\epsilon} - 1}{n}\right)\right)\right| < f(\delta_1). \tag{57}$$

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

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

whereby $\hat{y}_{\ell,\gamma} = \lim_{\tau \uparrow \hat{\tau}^{(\ell,\gamma)}} x(\tau)$. Thus the triangle inequality together with (54), (55), and (57) implies that for any $\gamma < c_1$, any $\delta \in (0, 1)$, and any $\ell \in \mathcal{S}$, w.h.p.

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

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 will be complete if we show that the process will finish soon after T_{D_ϵ}.

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

4.4.1. *The end of the process.* We will show that w.h.p. 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 subcritical. The subcriticality is encompassed by the following: there exists $\kappa_0 < 1$ such that with probability $1 - o(1)$,

$$\sum_{i=1}^{p_\ell} \frac{W_i^2}{W_{[n]}} c_{i,r-1}\left(\frac{T_{D_\epsilon}-1}{n}\right) < \kappa_0 < 1. \tag{58}$$

We start this section by proving this. First, let us observe that using the expression for μ_{U} from Proposition 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

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

But also by (45) we can write

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

So in particular, for $\tau = (T_{D_\epsilon} - 1)/n$, we have $G((T_{D_\epsilon} - 1)/n) > 0$ and therefore (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 5 and Proposition 4 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

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

and, moreover, $\hat{y}_{\ell,\gamma} < 1$. (Note that $\hat{y} < 1$ by its definition.) By Claim 15 below, the family $\left\{f_r^{(\ell,\gamma)\prime}(x)\right\}_{\ell \geq \ell'_1}$ restricted to [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 (56), 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 (51) imply that there exists $\kappa_0 < 1$ such that (58) holds with probability $1 - o(1)$.

From step T_{D_ϵ} onward, we will provide a stochastic upper bound on U by a set $\widehat{\mathsf{U}}$, whose size is an essentially subcritical multi-type branching process. In particular, the expression in (58) will dominate the principal eigenvalue of the expected progeny matrix of this branching process. In this process, rather than exposing the vertices of $\widehat{\mathsf{U}}$ one at a time, we will expose all of their neighbors simultaneously in each round. We let $\widehat{\mathsf{U}}(s)$ be the set $\widehat{\mathsf{U}}$ after s rounds.

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

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 $\widehat{\mathsf{U}}(0) = \mathsf{U}(t_0)$, $\widehat{\mathsf{C}}_{i,r-1}(0) = \mathsf{C}_{i,r-1}(t_0)$, and $\widehat{\mathsf{C}}_{i,<r-1}(0) = \cup_{k=2}^{r} \mathsf{C}_{i,r-k}(t_0)$, for all $i = 1, \ldots, p_\ell$. 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 coordinates are the sizes of the sets $\widehat{\mathsf{U}}_i(s)$. A vertex $v \in \widehat{\mathsf{U}}_j(s)$ (for $j \in \{1, \ldots, p_\ell\}$) 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 vertex in $\widehat{\mathsf{U}}(s)$.

Furthermore, if a vertex in $\widehat{\mathsf{C}}_{i,<r-1}(s)$ is adjacent to exactly one vertex in $\widehat{\mathsf{U}}(s)$, then a vertex is moved into $\widehat{\mathsf{C}}_{i,r-1}$. In this process the set $\widehat{\mathsf{C}}_{i,r-1}$ can only gain, not lose, vertices. 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 only once* (it could be a child of any adjacent vertex in $\widehat{\mathsf{U}}(s)$); hence it is included in $\widehat{\mathsf{U}}(s+1)$ only once. 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 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 because 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 $\mathbb{E}\left(\overline{u}_{s+1}^T | \mathcal{H}_s\right) \le \overline{u}_s^T A_s$ (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 $\hat{c}_{j,r-1}(s) \ge \hat{c}_{j,r-1}(0) = c_{i,r-1}\left(\frac{T_{D_\epsilon}-1}{n}\right)$. Thus for $s > 0$ we have

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

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

Note also that since $\hat{c}_{j,r-1}(s)$, $\hat{u}(s)$, $\hat{c}_{j,<r-1}(s) \leq n$, and $W_{[n]} = \Theta(n)$, we have the bound $\rho_s \leq D$ for some $D > 0$, which depends on γ, and for all $s \geq 0$.

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 (53),

$$\hat{u}(0)\frac{4C_\gamma^2}{W_{[n]}}\hat{c}_{j,<r-1}(0) \leq \hat{u}(0)\frac{4C_\gamma^2}{W_{[n]}}n < \hat{u}(0)\frac{5C_\gamma^2}{dn}n \overset{(53)}{\leq} \epsilon\frac{15C_\gamma^2}{2d}n.$$

Hence, combining this with (58), we deduce that if ϵ is small enough, then ρ_0 is smaller than 1 and in fact is bounded away from 1.

Let $\lambda_i := W_i / \sum_j W_j$ and set $\xi := [\lambda_1, \ldots, \lambda_{p_\ell}]^T$. 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 11. *Conditional on $\mathcal{H}_s$, with probability at least $1 - 1/n^2$, we have $Z_s = 0$ or*

$$Z_{s+1} \leq \rho_s Z_s + Z_s^{1/2} \log^{3/2} n. \tag{59}$$

Proof. In the following we condition on $\mathcal{H}_s$, which is suppressed from the notation. Assume that $Z_s > 0$. Note that Z_{s+1} is a weighted sum of Bernoulli-distributed random variables, where the weights are bounded. More specifically,

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

This expansion also shows that $Z_s > 0$ implies that $Z_s > c$, for some $c > 0$ which does not depend on s (or n).

We will appeal to Talagrand's inequality (see for example Theorem 2.29 in [25]).

Theorem 12. *Let $Z_1, \ldots, Z_N$ be independent random variables taking values in some sets $\Lambda_1, \ldots, \Lambda_N$, respectively. Suppose that $X = f(Z_1, \ldots, Z_N)$, where $f : \Lambda_1 \times \cdots \times \Lambda_N \to \mathbb{R}$ is a function which satisfies the following conditions:*

(i) *There are constants c_k, for $k = 1, \ldots, N$, such that if $z, z' \in \Lambda_1 \times \cdots \times \Lambda_N$ differ only in their kth coordinates, z_k and z'_k respectively, then $|f(z) - f(z')| \leq c_k$.*

(ii) *There exists a function $\psi : \mathbb{R} \to \mathbb{R}$ such that if $z \in \Lambda_1 \times \cdots \times \Lambda_N$ satisfies $f(z) \geq r$, then there is a witness set $J \subseteq \{1, \ldots, N\}$ with $\sum_{k \in J} c_k^2 \leq \psi(r)$ such that any $y \in \Lambda_1 \times \cdots \times \Lambda_N$ with $y_k = z_k$, for $k \in J$, also has $f(y) \geq r$.*

If m is a median of X, then for every $t \geq 0$,

$$\mathbb{P}(X \leq m - t) \leq 2e^{-t^2/4\psi(m)}$$

and

$$\mathbb{P}(X \geq m + t) \leq 2e^{-t^2/4\psi(m+t)}.$$

We will apply Theorem 12 to the random variable Z_{s+1}. First, note that Z_{s+1} is a function of independent Bernoulli random variables, which correspond to the (potential) edges that are incident to $\widehat{\mathsf{U}}(s)$. Let us write the random variable Z_{s+1} more explicitly as a function on the sample space which consists of all subsets of $\cup_{j=1}^{p_\ell}\left(\widehat{\mathsf{C}}_{j,r-1}(s)\cup\widehat{\mathsf{C}}_{j,<r-1}(s)\right)\times\widehat{\mathsf{U}}(s)$. More specifically, for any subset

$$E\subseteq\cup_{j=1}^{p_\ell}\left(\widehat{\mathsf{C}}_{j,r-1}(s)\cup\widehat{\mathsf{C}}_{j,<r-1}(s)\right)\times\widehat{\mathsf{U}}(s)$$

we can write

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

where the $\mathbf{1}_{\mathcal{E}}(E)$ denotes the indicator function of the event $\{E:E\in\mathcal{E}\}$. For $v\in\widehat{\mathsf{C}}_{j,r-1}(s)\cup\widehat{\mathsf{C}}_{j,<r-1}(s)$, if we change any pair of vertices between v and $\widehat{\mathsf{U}}(s)$ (add it if it is not an edge, or remove it if it is), then Z_{s+1} may change by at most λ_j.

Now, for any subset of edges $E\subseteq\cup_{j=1}^{p_\ell}\left(\widehat{\mathsf{C}}_{j,r-1}(s)\cup\widehat{\mathsf{C}}_{j,<r-1}(s)\right)\times\widehat{\mathsf{U}}(s)$ and a subset $J\subseteq\cup_{j=1}^{p_\ell}\left(\widehat{\mathsf{C}}_{j,r-1}(s)\cup\widehat{\mathsf{C}}_{j,<r-1}(s)\right)$, let

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

Suppose that $Z_{s+1}(E)\geq x$. Let $J^*=J^*(E)\subseteq\cup_{j=1}^{p_\ell}\left(\widehat{\mathsf{C}}_{j,r-1}(s)\cup\widehat{\mathsf{C}}_{j,<r-1}(s)\right)$ be a minimal subset which is also a minimizer of $Z_{s+1}^{J}(E)$ subject to $Z_{s+1}^{J}(E)\geq x$. (Note that a minimizer may not be minimal, as it may include a vertex v such that $\mathbf{1}_{d_{\widehat{\mathsf{U}}(s)}(v)\geq 1}(E)$, $\mathbf{1}_{d_{\widehat{\mathsf{U}}(s)}(v)\geq 2}(E)=0$.) Observe that the minimality of J^* implies that $Z_{s+1}^{J^*}(E)<x+\lambda_{max}$, where $\lambda_{\max}=\max\{\lambda_j\}_{j=1,\dots,p_\ell}$.

We select a set of edges $E_{J^*}\subseteq E$ as follows: for any $v\in\widehat{\mathsf{C}}_{j,r-1}(s)\cap J^*$, we add to E_{J^*} one of the edges in E that are incident to v, if there are any such; for any $v\in\widehat{\mathsf{C}}_{j,<r-1}(s)\cap J^*$, we add to E_{J^*} two of the edges in E that are incident to v, if there are at least two such edges. Hence, any subset of edges $E'\subseteq\cup_{j=1}^{p_\ell}\left(\widehat{\mathsf{C}}_{j,r-1}(s)\cup\widehat{\mathsf{C}}_{j,<r-1}(s)\right)\times\widehat{\mathsf{U}}(s)$ such that $E_{J^*}\subseteq E'$ also satisfies $Z_{s+1}(E')\geq x$.

For any $e\in\left(\widehat{\mathsf{C}}_{j,r-1}(s)\cup\widehat{\mathsf{C}}_{j,<r-1}(s)\right)\times\widehat{\mathsf{U}}(s)$, we set $\lambda_e=\lambda_j$. Consider now $\sum_{e\in E_{J^*}}\lambda_e^2$. Since $\lambda_e\leq 1$, we have the bound

$$\begin{aligned}\sum_{e\in E_{J^*}}\lambda_e^2&\leq\sum_{j=1}^{p_\ell}\left(\sum_{e\in\left(\widehat{\mathsf{C}}_{j,r-1}(s)\times\widehat{\mathsf{U}}(s)\right)\cap E_{J^*}}\lambda_e+\sum_{e\in\left(\widehat{\mathsf{C}}_{j,<r-1}(s)\times\widehat{\mathsf{U}}(s)\right)\cap E_{J^*}}\lambda_e\right)\\&=\sum_{j=1}^{p_\ell}\left(\sum_{e\in\left(\widehat{\mathsf{C}}_{j,r-1}(s)\times\widehat{\mathsf{U}}(s)\right)\cap E_{J^*}}\lambda_j+\sum_{e\in\left(\widehat{\mathsf{C}}_{j,<r-1}(s)\times\widehat{\mathsf{U}}(s)\right)\cap E_{J^*}}\lambda_j\right)\\&=\sum_{j=1}^{p_\ell}\left(\sum_{v\in\widehat{\mathsf{C}}_{j,r-1}(s)\cap J^*}\lambda_j+2\sum_{v\in\widehat{\mathsf{C}}_{j,<r-1}(s)\cap J^*}\lambda_j\right)\\&\leq 2Z_{s+1}^{J^*}\leq 2(x+\lambda_{\max}).\end{aligned}$$

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

Hence, we can apply Theorem 12, taking $\psi(x) = 2(x + \lambda_{\max})$ (conditional on $\mathcal{H}_s$ which we suppress). With $m(Z_{s+1})$ being a median of Z_{s+1}, Talagrand's inequality yields

$$\mathbb{P}\left[Z_{s+1} \geq m(Z_{s+1}) + \frac{1}{2} Z_s^{1/2} \log^{3/2} n\right] \leq 2e^{-\frac{Z_s \log^3 n}{32\left(m(Z_{s+1}) + \frac{1}{2} Z_s^{1/2} \log^{3/2} n + \lambda_{\max}\right)}}. \tag{60}$$

Since $\psi(x)$ is increasing with respect to x and Z_{s+1} takes only non-negative values, using an argument similar to that of [25, pp. 41–42] we have

$$\begin{aligned}
&|\mathbb{E}\,(Z_{s+1}) - m(Z_{s+1})| \leq \mathbb{E}\,(|Z_{s+1} - m(Z_{s+1})|) = \int_0^\infty \mathbb{P}\left[|Z_{s+1} - m(Z_{s+1})| > t\right] dt \\
&\leq \int_0^{m(Z_{s+1})} 4e^{-\frac{t^2}{4\psi(2m(Z_{s+1}))}} dt + \int_{m(Z_{s+1})}^\infty 2e^{-\frac{t^2}{4\psi(m(Z_{s+1})+t)}} dt \\
&\leq \int_0^{m(Z_{s+1})} 4e^{-\frac{t^2}{16m(Z_{s+1})+8\lambda_{\max}}} dt + \int_{m(Z_{s+1})}^\infty 2e^{-\frac{t^2}{16t+8\lambda_{\max}}} dt \\
&\leq 8\sqrt{\pi(m(Z_{s+1}) + \lambda_{\max})} + 32,
\end{aligned}$$

and using that $m(Z_{s+1}) \leq 2m(Z_{s+1})\mathbb{P}\left[Z_{s+1} \geq m(Z_{s+1})\right] \leq 2\mathbb{E}\,(Z_{s+1})$, we obtain

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

Hence, for n large enough, using that $\mathbb{E}\,(Z_{s+1}) \leq \rho_s Z_s$, we have

$$\begin{aligned}
&\mathbb{P}\left[Z_{s+1} \geq \mathbb{E}\,(Z_{s+1}) + Z_s^{1/2} \log^{3/2} n\right] \\
&\leq \mathbb{P}\left[Z_{s+1} \geq m(Z_{s+1}) - O(\mathbb{E}\,(Z_{s+1})^{1/2}) + Z_s^{1/2} \log^{3/2} n\right] \\
&\leq \mathbb{P}\left[Z_{s+1} \geq m(Z_{s+1}) + \frac{1}{2} Z_s^{1/2} \log^{3/2} n\right].
\end{aligned}$$

So by (60) we conclude (using that $m(Z_{s+1}) \leq 2\mathbb{E}\,(Z_{s+1}) \leq 2\rho_s Z_s$) that

$$\begin{aligned}
&\mathbb{P}\left[Z_{s+1} \geq \mathbb{E}\,(Z_{s+1}) + Z_s^{1/2} \log^{3/2} n\right] \\
&\leq 2e^{-\frac{Z_s \log^3 n}{32(m(Z_{s+1}) + \frac{1}{2} Z_s^{1/2} \log^{3/2} n + \lambda_{\max})}} \leq 2e^{-\frac{Z_s \log^3 n}{32(2\rho_s Z_s + \frac{1}{2} Z_s^{1/2} \log^{3/2} n + \lambda_{\max})}} = e^{-\Omega(\log^{3/2} n)},
\end{aligned}$$

since ρ_s and $\lambda_{\max}$ are bounded uniformly over all s and Z_s is bounded away from 0, if $Z_s > 0$. □

We denote the above event ($Z_s = 0$ or (59) holds) by $\mathcal{E}_s$. If $Z_s > \log^6 n$ and $\mathcal{E}_s$ is realized, we have

$$\begin{aligned}
Z_{s+1} &\leq \rho_s Z_s \left(1 + \frac{Z_s^{1/2} \log^{3/2} n}{\rho_s Z_s}\right) \overset{\rho_s \geq \kappa_0'}{\leq} \rho_s Z_s \left(1 + \frac{\log^{3/2} n}{\kappa_0' Z_s^{1/2}}\right) \\
&\overset{Z_s > \log^6 n}{\leq} \rho_s Z_s \left(1 + \frac{\log^{3/2} n}{\kappa_0' \log^3 n}\right) = \rho_s Z_s \left(1 + \frac{1}{\kappa_0' \log^{3/2} n}\right).
\end{aligned} \tag{61}$$

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

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 (58), the largest positive eigenvalue of A_0 is bounded by a constant $\rho_0 < 1$, with probability $1 - o(1)$. We set $\lambda_{\min} = \min\{\lambda_i\}_{i=1,\ldots,p_\ell}$. 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 13. *For any $s \geq 0$ we have $\mathbb{P}[\mathcal{D}_s] = 1 - o(1/n^2)$.*

Proof of Claim 13. 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_v are independent and identically distributed random variables that are distributed as $\mathrm{Bin}(n, (2C_\gamma)^2/W_{[n]})$. The expected value of this sum is bounded by $\frac{5C_\gamma^2}{d} u(t)$ for large n. Also, $u(s) \leq Z_s/\lambda_{\min}$, as $Z_s = (\xi, \bar{u}_s) = \sum_i \lambda_i u_i(s) \geq \lambda_{\min} \sum_i u_i(s)$. 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 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 β, $\beta' > 0$ 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

$$\rho_{s+1} \leq \rho_s + \frac{1}{n} \left(\gamma (Z_s + Z_{s+1}) + \beta'' \log^6 n \right). \tag{62}$$

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

$$\rho_s \leq \rho_0 + \gamma \frac{Z_0}{n} \left(2 \sum_{k=0}^{s-1} (\rho_0 + \delta)^k \lambda_n^k + (\rho_0 + \delta)^s \lambda_n^s \right) + (4\gamma\gamma' + \beta'') \frac{\log^6 n}{n} \sum_{k=1}^{s-1} k + 2\gamma\gamma' \frac{\log^6 n}{n} s. \tag{63}$$

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

$$\begin{aligned}\rho_s &\leq \rho_0 + \frac{2\gamma Z_0}{n} \sum_{k=0}^{\infty} \rho_0'^k + O\left(\frac{\log^8 n}{n} \right) = \rho_0 + \frac{2\gamma Z_0}{n} \frac{1}{1-\rho_0'} + O\left(\frac{\log^8 n}{n} \right) \\ &\overset{Z_0/n<\varepsilon}{\leq} \rho_0 + 2\gamma\varepsilon \frac{1}{1-\rho_0'} + O\left(\frac{\log^8 n}{n} \right) < \rho_0 + \delta < 1,\end{aligned} \tag{64}$$

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

By (61), 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

$$Z_s \le Z_0(\rho_0+\delta)^s \lambda_n^s + 2s\gamma' \log^6 n. \tag{65}$$

Substituting (63) into (62) and using (65), we obtain

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

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 11 and 13 imply that

$$\mathbb{P}\left[\cap_{s\le T_0}\{\mathcal{E}_s \cap \mathcal{D}_s\}\right] = 1 - O(\log n/n^2). \tag{66}$$

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

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

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,

$$\begin{aligned}
\mathbb{E}\left(Z_{T_0}\right) &\le \mathbb{E}\left(\mathbb{E}\left(Z_{T_0} \mid \mathcal{H}_{T_0-1}\right)\mathbf{1}_{\mathcal{S}_{T_0-1}}\right) + O(\log n/n) \\
&\overset{(64)}{\le} \mathbb{E}\left((\rho_0+\delta)Z_{T_0-1}\right) + O(\log n/n).
\end{aligned}$$

Repeating this, we get

$$\mathbb{E}\left(Z_{T_0}\right) \le (\rho_0+\delta)^{T_0}\mathbb{E}\left(Z_0\right) + O(T_0\log n/n) \overset{\mathbb{E}(Z_0)=O(n)}{\le} O(1/n + T_0/n) = o(1). \tag{67}$$

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

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

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

Lemma 5. *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) $f_r^{(\ell,\gamma)\prime}\left(\hat{y}_{\ell,\gamma}\right) < f_r'(\hat{y}; W_F^*, p) + \delta$, *and*

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

Proof. Using Definition 3, we can express the γ_i in terms of the γ_i': $\gamma_i = (1 - h_F(\gamma) + \gamma')\gamma_i'$. The expression for $f_r^{(\ell,\gamma)}$ yields the following:

$$\begin{aligned} f_r^{(\ell,\gamma)}(x) &= \frac{W'_\gamma}{d} + p\,\frac{1}{d}\sum_{i=1}^{p_\ell} W_i\gamma_i - x + (1-p)\,\frac{1}{d}\sum_{i=1}^{p_\ell} W_i\gamma_i\mathbb{P}\left[\mathsf{Po}\left(W_i x\right) \geq r\right] \\ &= \frac{W'_\gamma}{d} + (1 - h_F(\gamma) + \gamma')\left(p\,\frac{\hat{d}^{(\ell,\gamma)}}{d} + (1-p)\,\frac{d^{(\ell,\gamma)}}{d}\sum_{i=1}^{p_\ell}\frac{W_i\gamma_i'}{d^{(\ell,\gamma)}}\mathbb{P}\left[\mathsf{Po}\left(W_i x\right) \geq r\right]\right) - x, \end{aligned} \tag{68}$$

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 (68) becomes

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

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

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

$$\begin{aligned} f_r^{(\ell,\gamma)\prime}(x) &= -1 + c(\gamma)(1-p)\,\frac{d^{(\ell,\gamma)}}{d}\int_0^{C_\gamma} ye^{-yx}\frac{(yx)^{r-1}}{(r-1)!}dF^{*(\ell,\gamma)}(y) \\ &= -1 + c(\gamma)(1-p)\,\frac{d^{(\ell,\gamma)}}{d}\,\frac{r}{x}\int_0^{C_\gamma} e^{-yx}\frac{(yx)^{r}}{r!}dF^{*(\ell,\gamma)}(y). \end{aligned} \tag{69}$$

Similarly, we can write

$$\alpha(x) = p(1-h_F(\gamma)) + \gamma' + c(\gamma)(1-p)\int_0^{C_\gamma} \psi_r\,(yx)\,dF^{(\ell,\gamma)}(y). \tag{70}$$

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 3. *Let $f:[0,\infty)\to[0,\infty)$ be a bounded function which is everywhere differentiable. Also, let $y_1\in\mathbb{R}$. 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. *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 1. *Let $f:[0,\infty)\to[0,\infty)$ be a bounded function which is everywhere differentiable. For any $\delta>0$ and 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 (69) and (70), 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 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 they have bounded derivatives.

By the second part of Definition 4 and the fact that $c(\gamma)\to 1$ as $\gamma\downarrow 0$, we have

$$c(\gamma)\left|\frac{d^{(\ell,\gamma)}}{d}-1\right| < \delta \tag{71}$$

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

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 4.

Claim 14. *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 $\gamma\to 0$ implies that if γ is small enough and ℓ is large enough, we have

$$\left|p^{(\ell,\gamma)}-p\right|<\delta \quad \text{and} \quad \left|p(1-h_F(\gamma))+\gamma'-p\right|<\delta. \tag{72}$$

Both parts of the lemma now follow from Corollary 1 together with (71) and (72). We now proceed with the proofs of Propositions 3 and 4.

Proof of Proposition 3. The proof of this proposition will proceed in two steps. First 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

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

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

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

Having proved these inequalities, we deduce that

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

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

$$\left|\int_0^{C_\gamma} f(yy_2)dF^{(\ell,\gamma)}(y)-\int_0^{C_\gamma} f(yy_1)dF(y)\right|<\delta/2. \tag{75}$$

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

$$\left|\int_{C_\gamma}^{\infty} f(yy_1)dF(y)\right|<\delta/2. \tag{76}$$

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

As before, from (75) and (76) we deduce

$$\left| \int_0^{C_\gamma} f(yy_2) dF^{(\ell,\gamma)}(y) - \mathbb{E}\left[f(W_F y_1) \right] \right|$$
$$\leq \left| \int_0^{C_\gamma} f(yy_2) dF^{(\ell,\gamma)}(y) - \int_0^{C_\gamma} f(yy_1) dF(y) \right| + \left| \int_{C_\gamma}^{\infty} f(yy_1) dF(y) \right| \stackrel{(75),(76)}{<} \delta.$$

We proceed with the proofs of (73) and (74)—the proofs of (75) and (76) 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 (73). We begin with the specification of δ'. We let δ' be such that whenever $|y_1 - y_2| < \delta'$, we have

$$|f(xy_1) - f(xy_2)| < \delta/4 \tag{77}$$

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

$$\begin{aligned}
&\left| \int_0^{C_\gamma} f(xy_2) dF^{*(\ell,\gamma)}(x) - \int_0^{C_\gamma} f(xy_1) dF^*(x) \right| \\
&\leq \int_0^{C_\gamma} |f(xy_2) - f(xy_1)| dF^{*(\ell,\gamma)}(x) \\
&\qquad + \left| \int_0^{C_\gamma} f(xy_1) dF^{*(\ell,\gamma)}(x) - \int_0^{C_\gamma} f(xy_1) dF^*(x) \right| \\
&\stackrel{(77)}{\leq} \delta/4 + \left| \int_0^{C_\gamma} f(xy_1) dF^{*(\ell,\gamma)}(x) - \int_0^{C_\gamma} f(xy_1) dF^*(x) \right|.
\end{aligned} \tag{78}$$

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 (17), as the latter implies that $F^{*(\ell,\gamma)}$ converges weakly to F^* as $\gamma \downarrow 0$ and $\ell \to \infty$. Since f has been assumed to be bounded and continuous, by Lemma 3 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 (74), 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 1, 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 (78), we can deduce (73). □

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

We now proceed with the proof of (74)

Proof of (74). Assume that $|f(x)| < b$ for any $x \in \mathbb{R}$. Hence we have

$$\left| \int_{C_\gamma}^{\infty} f(yy_1) dF^*(y) \right| < b\mathbb{E}\left[\mathbf{1}_{W_F^* \geq C_\gamma} \right]. \tag{79}$$

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 $\mathbb{E}\left[W_F\right] < \infty$, 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 (74) follows from (79). □

Proof of Proposition 4. We consider the functions $f_r^{(\ell,\gamma)}(x)$ restricted to the unit interval [0,1].

Claim 15. *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 also holds 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 15. 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 4). 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 4, Part 2),

$$\left| \frac{d^{(\ell,\gamma)}}{d} - 1 \right| < \rho(\gamma)/d. \tag{80}$$

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,

$$\left| \int_0^{C_\gamma} \psi_r\left(yx_1\right) dF^{*(\ell,\gamma)}(y) - \int_0^{C_\gamma} \psi_r\left(yx_2\right) dF^{*(\ell,\gamma)}(y) \right| < \frac{d\varepsilon}{2\rho}. \tag{81}$$

Therefore, 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

$$\begin{aligned}
&|f_r^{(\ell,\gamma)}(x_1) - f_r^{(\ell,\gamma)}(x_2)| \\
&\le |x_1 - x_2| + \frac{d^{(\ell,\gamma)}}{d} \left| \int_0^{C_\gamma} \psi_r\,(yx_1)\, dF^{*(\ell,\gamma)(y)} - \int_0^{C_\gamma} \psi_r\,(yx_2)\, dF^{*(\ell,\gamma)(y)} \right| \\
&\overset{(80),(81)}{\le} \frac{\delta}{C_\gamma} + \frac{d\varepsilon}{2\rho}\frac{\rho}{d} \le \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \le \varepsilon.
\end{aligned}$$

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

By the Arzelà–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,

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

By the L_∞-convergence of the family

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

restricted to [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

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

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 values of k, 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 3. Consider $\gamma < c_2$. Then there exists j_0 such that for $j > j_0$ we have

$$\left| \int_0^{C_\gamma} \psi_r(y\hat{y}_{\ell_{k_j},\gamma}) dF^{*(\ell_{k_j},\gamma)}(y) - \mathbb{E}\left(\psi_r(W_F^*\hat{y}_\gamma)\right) \right| < \delta/3. \tag{82}$$

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

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

by Part 2 of Definition 4 and by Claim 14. Hence

$$\begin{aligned}\left|c(\gamma)\frac{\hat{d}^{(\ell_{k_j},\gamma)}}{d}-1\right| &\leq |c(\gamma)-1|\ \frac{d^{(\ell_{k_j},\gamma)}}{d}+\left|\frac{\hat{d}^{(\ell_{k_j},\gamma)}}{d}-1\right|\\ &\leq |c(\gamma)-1|+|c(\gamma)-1|\left|\frac{d^{(\ell_{k_j},\gamma)}}{d}-1\right|+\left|\frac{\hat{d}^{(\ell_{k_j},\gamma)}}{d}-1\right|\leq 3\delta/9=\delta/3.\end{aligned} \tag{83}$$

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

$$\left|c(\gamma)\frac{d^{(\ell_{k_j},\gamma)}}{d}-1\right|\leq\frac{\delta}{3}. \tag{84}$$

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

$$\begin{aligned}\hat{f}_r(\hat{y}_\gamma)&=\hat{f}_r(\hat{y}_\gamma)-f_r^{(\ell_{k_j},\gamma)}(\hat{y}_{\ell_{k_j},\gamma})\\ &\leq p\left|c(\gamma)\frac{\hat{d}^{(\ell_{k_j},\gamma)}}{d}-1\right|\\ &\qquad+(1-p)\left|c(\gamma)\frac{d^{(\ell_{k_j},\gamma)}}{d}\int_0^{C_\gamma}\ \psi_r(y\hat{y}_{\ell_{k_j},\gamma})dF^{*(\ell_{k_j},\gamma)}(y)-\mathbb{E}\left(\psi_r(W_F^*\hat{y}_\gamma)\right)\right|\\ &\leq p\left|c(\gamma)\frac{\hat{d}^{(\ell_{k_j},\gamma)}}{d}-1\right|+(1-p)\left|\left(c(\gamma)\frac{d^{(\ell_{k_j},\gamma)}}{d}-1\right)\int_0^{C_\gamma}\ \psi_r(y\hat{y}_{\ell_{k_j},\gamma})dF^{*(\ell_{k_j},\gamma)}(y)\right|\\ &\qquad+(1-p)\left|\int_0^{C_\gamma}\ \psi_r(y\hat{y}_{\ell_{k_j},\gamma})dF^{*(\ell_{k_j},\gamma)}(y)-\mathbb{E}\left(\psi_r(W_F^*\hat{y}_\gamma)\right)\right|\\ &\overset{(83),(84),(82)}{\leq}\ \frac{\delta}{3}+\frac{\delta}{3}+\frac{\delta}{3}=\delta.\end{aligned}$$

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 a contradiction, as $\hat{y}$ is the smallest positive root of $f_r(x;W_F^*,p)=0$. □

https://doi.org/10.1017/apr.2023.21 Published online by Cambridge University Press

The following lemma shows that if the weight sequence has a 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 1 is always satisfied.

Lemma 6. *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

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

To prove the claim it is thus sufficient to argue that

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

In turn, it suffices to prove that

$$r\mathbb{E}\left[e^{-W_F^*\hat{y}}\frac{\left(W_F^*\hat{y}\right)^r}{r!}\right] < \mathbb{E}\left[\psi_r(W_F^*\hat{y})\right]. \tag{85}$$

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

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

which is equivalent to (85). 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 $j \in \{r, r+1\}$,

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

Therefore, 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 to the integral of the left-hand side, we obtain

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

□

Acknowledgements

We wish to thank the anonymous referees for their valuable comments and suggestions to improve the presentation of the paper.

Funding information

N. Fountoulakis was partially supported by the EPSRC grant EP/K019740/1.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] Adler, J. and Lev, U. (2003). Bootstrap percolation: visualizations and applications. *Brazilian J. Phys.* **33**, 641–644.
[2] Albert, R. and Barabási, A. (2002). Statistical mechanics of complex networks. *Rev. Modern Phys.* **74**, 47–97.
[3] Amini, H. (2010). Bootstrap percolation and diffusion in random graphs with given vertex degrees. *Electron. J. Combinatorics* **17**, article no. R25.
[4] Amini, H. (2010). Bootstrap percolation in living neural networks. *J. Statist. Phys.* **141**, 459–475.
[5] Amini, H., Cont, R. and Minca, A. (2016). Resilience to contagion in financial networks. *Math. Finance* **26**, 329–365.
[6] Amini, H. and Fountoulakis, N. (2014). Bootstrap percolation in power-law random graphs. *J. Statist. Phys.* **155**, 72–92.
[7] Amini, H. and Minca, A. (2016). Inhomogeneous financial networks and contagious links. *Operat. Res.* **64**, 1109–1120.
[8] Athreya, K. and Ney, P. (1972). *Branching Processes*. Springer, Berlin, Heidelberg.
[9] Balogh, J. and Bollobás, B. (2006). Bootstrap percolation on the hypercube. *Prob. Theory Relat. Fields* **134**, 624–648.
[10] Balogh, J., Bollobás, B., Duminil-Copin, H. and Morris, R. (2012). The sharp threshold for bootstrap percolation in all dimensions. *Trans. Amer. Math. Soc.* **36**, 2667–2701.
[11] Balogh, J., Bollobás, B. and Morris, R. (2009). Bootstrap percolation in three dimensions. *Ann. Prob.* **37**, 1329–1380.
[12] Balogh, J., Peres, Y. and Pete, G. (2006). Bootstrap percolation on infinite trees and non-amenable groups. *Combinatorics Prob. Comput.* **15**, 715–730.
[13] Balogh, J. and Pittel, B. G. (2007). Bootstrap percolation on the random regular graph. *Random Structures Algorithms* **30**, 257–286.
[14] Bollobás, B., Janson, S. and Riordan, O. (2007). The phase transition in inhomogeneous random graphs. *Random Structures Algorithms* **31**, 3–122.
[15] Cerf, R. and Manzo, F. (2002). The threshold regime of finite volume bootstrap percolation. *Stoch. Process. Appl.* **101**, 69–82.
[16] Chalupa, J., Leath, P. L. and Reich, G. R. (1979). Bootstrap percolation on a Bethe lattice. *J. Phys. C* **12**, L31–L35.
[17] Chung, F. and Lu, L. (2002). Connected components in random graphs with given expected degree sequences. *Ann. Combinatorics* **6**, 125–145.

[18] CHUNG, F. AND LU, L. (2003). The average distance in a random graph with given expected degrees. *Internet Math.* **1**, 91–113.

[19] CHUNG, F., LU, L. AND VU, V. (2004). The spectra of random graphs with given expected degrees. *Internet Math.* **1**, 257–275.

[20] DETERING, N., MEYER-BRANDIS, T. AND PANAGIOTOU, K. (2019). Bootstrap percolation in directed inhomogeneous random graphs. *Electron. J. Combinatorics* **26**, article no. P3.12.

[21] DETERING, N., MEYER-BRANDIS, T., PANAGIOTOU, K. AND RITTER, D. (2019). Managing default contagion in inhomogeneous financial networks. *SIAM J. Financial Math.* **10**, 578–614.

[22] FONTES, L. AND SCHONMANN, R. (2008). Bootstrap percolation on homogeneous trees has 2 phase transitions. *J. Statist. Phys.* **132**, 839–861.

[23] FONTES, L. R., SCHONMANN, R. H. AND SIDORAVICIUS, V. (2002). Stretched exponential fixation in stochastic Ising models at zero temperature. *Commun. Math. Phys.* **228**, 495–518.

[24] HOLROYD, A. E. (2003). Sharp metastability threshold for two-dimensional bootstrap percolation. *Prob. Theory Relat. Fields* **125**, 195–224.

[25] JANSON, S., ŁUCZAK, T. AND RUCIŃSKI, A. (2000). *Random Graphs*. John Wiley, New York.

[26] JANSON, S., ŁUCZAK, T., TUROVA, T. AND VALLIER, T. (2012). Bootstrap percolation on the random graph $G_{n,p}$. *Ann. Appl. Prob.* **22**, 1989–2047.

[27] LOTKA, A. J. (1926). The frequency distribution of scientific productivity. *J. Washington Acad. Sci.* **16**, 317–323.

[28] MITZENMACHER, M. (2004). A brief history of generative models for power law and lognormal distributions. *Internet Math.* **1**, 226–251.

[29] MORRIS, R. (2009). Zero-temperature Glauber dynamics on $\mathbb{Z}^d$. *Prob. Theory Relat. Fields* **149**, 417–434.

[30] PARETO, V. (1896). *Cours d'Économie Politique*. Dronz, Geneva.

[31] SABHAPANDIT, S., DHAR, D. AND SHUKLA, P. (2002). Hysteresis in the random-field Ising model and bootstrap percolation. *Phys. Rev. Lett.* **88**, article no. 197202.

[32] SAUSSET, F., TONINELLI, C., BIROLI, G. AND TARJUS, G. (2010). Bootstrap percolation and kinetically constrained models on hyperbolic lattices. *J. Statist. Phys.* **138**, 411–430.

[33] SÖDERBERG, B. (2002). General formalism for inhomogeneous random graphs. *Phys. Rev. E* **66**, article no. 066121.

[34] TLUSTY, T. AND ECKMANN, J. (2009). Remarks on bootstrap percolation in metric networks. *J. Phys. A* **42**, article no. 205004.

[35] TONINELLI, C., BIROLI, G. AND FISHER, D. S. (2006). Jamming percolation and glass transitions in lattice models. *Phys. Rev. Lett.* **96**, article no. 035702.

[36] TORRISI, G., GARETTO, M. AND LEONARDI, E. (2018). *Bootstrap percolation on the stochastic block model*. Preprint. Available at https://arxiv.org/abs/1812.09107.

[37] VAN DER HOFSTAD, R. (2016). *Random Graphs and Complex Networks*, Vol. **1**. Cambridge University Press.

[38] WORMALD, N. (1999). The differential equation method for random graph processes and greedy algorithms. In *Lectures on Approximation and Randomized Algorithms*, eds M. Karonski and H.-J. Prömel, Polish Scientific Publishers PWN, pp. 73–155.

[39] WORMALD, N. C. (1995). Differential equations for random processes and random graphs. *Ann. Appl. Prob.* **5**, 1217–1235.