THE ROYAL SWEDISH ACADEMY OF SCIENCES
INSTITUT MITTAG-LEFFLER
Auravägen 17, SE-182 60 Djursholm, Sweden
Tel. +46 8 622 05 60 Fax. +46 8 622 05 89
info@mittag-leffler.se www.mittag-leffler.se

Critical behavior in inhomogeneous random graphs

R. van der Hofstad

REPORT No. 1, 2008/2009, spring
ISSN 1103-467X
ISRN IML-R- -1-08/09- -SE+spring

Critical behavior in inhomogeneous random graphs

Remco van der Hofstad *

February 1, 2009

Abstract

We study the critical behavior of inhomogeneous random graphs where edges are present independently but with unequal edge occupation probabilities. We show that the critical behavior depends sensitively on the properties of the asymptotic degrees. Indeed, when the proportion of vertices with degree at least k is bounded above by $k^{-\tau+1}$ for some $\tau > 4$, the largest critical connected component is of order $n^{2/3}$, where n denotes the size of the graph, as on the Erdős-Rényi random graph. The restriction $\tau > 4$ corresponds to finite *third* moment of the degrees. When, the proportion of vertices with degree at least k is asymptotically equal to $ck^{-\tau+1}$ for some $\tau \in (3, 4)$, the largest critical connected component is of order $n^{(\tau-2)/(\tau-1)}$, instead.

Our results show that, for inhomogeneous random graphs with a power-law degree sequence, the critical behavior admits a transition when the third moment of the degrees turns from finite to infinite. Similar phase transitions have been shown to occur for typical distances in such random graphs when the variance of the degrees turns from finite to infinite. We present further results related to the size of the critical or scaling window, and state conjectures for this and related random graph models.

1 Introduction and results

We study the critical behavior of inhomogeneous random graphs, where edges are present independently but with unequal edge occupation probabilities. Such inhomogeneous random graphs were studied in substantial detail in the seminal paper [6], where various results have been proved, including the identification of the critical value by studying the connected component sizes in the super- and subcritical regimes.

In this paper, we study the critical behavior of such random graphs, and show that this critical behavior depends sensitively on the properties of the asymptotic degrees. When the proportion of vertices with degree at least k bounded above by $k^{-\tau+1}$ for some $\tau > 4$, the largest critical connected component is of order $n^{2/3}$, where n denotes the size of the graph, as on the Erdős-Rényi random graph [17]. The restriction on the degrees corresponds to finite *third* moment of the degrees. When, however, the proportion of vertices with degree at least k is asymptotically equal to $ck^{-\tau+1}$ for some $\tau \in (3, 4)$, the largest critical connected component is of order $n^{(\tau-2)/(\tau-1)}$, instead. Random graphs with a power-law degree sequence are sometimes called *scale free*. Our results show that, for scale-free inhomogeneous random graphs, the critical behavior admits a transition when the third

*Department of Mathematics and Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands. `rhofstad@win.tue.nl`

moment of the degrees turns from finite to infinite. Similar phase transitions have been shown to occur for typical distances in scale-free random graphs when the variance of the degrees turns from finite to infinite [10, 12, 18, 22, 23, 36, 39]. In general, such results show that the behavior in scale-free random graphs depends sensitively on the degree exponent. We present further results related to the size of the critical or scaling window, and state conjectures concerning the sub- and supercritical regimes for this model, as well as for the so-called *configuration model*, a random graph model with prescribed degrees.

1.1 Inhomogeneous random graphs: the rank-1 case

In this section, we introduce the random graph model that we shall investigate. In our models, $\mathbf{w} = \{w_j\}_{j=1}^n$ are vertex weights, and l_n is the total weight of all vertices given by

$$l_n = \sum_{j=1}^n w_j. \tag{1.1}$$

We shall mainly work with the Poissonian random graph or *Norros-Reittu random graph* [36], in which the edge probabilities are given by

$$p_{ij}^{(\mathrm{NR})} = 1 - \mathrm{e}^{-w_i w_j / l_n}. \tag{1.2}$$

More precisely, $p_{ij}^{(\mathrm{NR})}$ is the probability that edge ij is present, for $1 \leq i < j \leq n$, and the different edges are independent. We denote the Norros-Reittu random graph with vertex weights $\mathbf{w} = \{w_i\}_{i=1}^n$ by $\mathrm{NR}_n(\mathbf{w})$. In Section 2, we shall extend our results to graphs where the edge probabilities are either $p_{ij} = \max\{w_i w_j / l_n, 1\}$ (as studied by Chung and Lu in [10, 11, 12, 13, 14]) or $p_{ij} = w_i w_j / (l_n + w_i w_j)$ (as studied in [9]).

Naturally, the graph structure depends sensitively on the empirical properties of the weights. For F any distribution function, we shall take

$$w_j = [1 - F]^{-1}(j/n), \tag{1.3}$$

where $[1 - F]^{-1}$ is the generalized inverse function of $1 - F$ defined, for $u \in (0, 1)$, by

$$[1 - F]^{-1}(u) = \inf\{s : [1 - F](s) \leq u\}, \tag{1.4}$$

where, by convention, we set $[1 - F]^{-1}(1) = 0$. A simple example arises when we take

$$F(x) = \begin{cases} 0 & \text{for } x < a, \\ 1 - (a/x)^{\tau - 1} & \text{for } x \geq a, \end{cases} \tag{1.5}$$

in which case $[1 - F]^{-1}(u) = a(1/u)^{-1/(\tau-1)}$, so that $w_j = a(j/n)^{-1/(\tau-1)}$.

In the setting in (1.3), the number of vertices with degree k, which we denote by N_k, satisfies

$$N_k/n \xrightarrow{\mathbb{P}} f_k, \tag{1.6}$$

where the limiting distribution $\{f_k\}_{k=1}^\infty$ is a so-called mixed Poisson distribution given by

$$f_k = \mathbb{E}\Big[\mathrm{e}^{-W} \frac{W^k}{k!}\Big], \qquad k \geq 0, \tag{1.7}$$

i.e., conditionally on $W = w$, the distribution is Poisson with mean w, and where the random variable W has distribution function F appearing in (1.3). Since the Poisson random variable is highly concentrated, it is easy to see that the number of vertices with degree larger than k is, for large k, very close to $n[1 - F(k)]$.

In our setting, there is a giant component precisely when $\nu > 1$, where we define

$$\nu = \frac{\mathbb{E}[W^2]}{\mathbb{E}[W]}. \tag{1.8}$$

More precisely, if $\nu > 1$, then the largest connected component has $n\zeta(1 + o_{\mathbb{P}}(1))$ vertices, while if $\nu \leq 1$, the largest connected component has $o_{\mathbb{P}}(n)$ vertices. Here we write that $X_n = o_{\mathbb{P}}(b_n)$ for some sequence b_n, when X_n/b_n converges to zero in probability. See, e.g., [6, Theorem 3.1 and Section 16.4] and [11, 14, 36]. When $\nu > 1$, the rank-1 inhomogeneous random graph is called *supercritical*, when $\nu = 1$ it is called *critical*, and when $\nu < 1$, it is called *subcritical*. The aim of this paper is to study the size of the largest connected components in the critical case. In our example (1.5), we have that $\nu = a(\tau-2)/(\tau-3)$, so that it is subcritical when $a < (\tau-3)/(\tau-2)$, critical when $a = (\tau-3)/(\tau-2)$ and supercritical when $a > (\tau-3)/(\tau-2)$.

We next describe further results from the literature in the super- and subcritical regimes. In the supercritical results, when $W \geq \varepsilon$ a.s. for some $\varepsilon > 0$, then the second largest component has size $O_{\mathbb{P}}(\log n)$, as in the Erdős-Rényi random graph. Here we write that $X_n = O_{\mathbb{P}}(b_n)$ when $|X_n|/b_n$ is a tight sequence of random variables. Further, in [6, Theorem 3.12 and Section 16.4], it is stated that the largest subcritical cluster has size $O_{\mathbb{P}}(\log n)$ when W has a bounded support. When the latter is not the case, which happens for example in the *power-law case* where

$$1 - F(x) = L(x)x^{-(\tau-1)}, \qquad x \geq 0, \tag{1.9}$$

for some $\tau \geq 1$, some slowly varying function $x \mapsto L(x)$, then the largest subcritical connected component has size $cn^{1/(\tau-1)}/(1-\nu)$ (see [26]), where $cn^{1/(\tau-1)}$ corresponds to the largest degree in the graph. Thus, we can think of such clusters as basically containing the largest degree vertex with a 'few' extra vertices attached to its edges. In this paper, we aim to study the *critical case*, where $\nu = 1$, giving special attention to the power-law case in (1.9). The critical nature of percolation on various graphs has received substantial attention in the past decades. We shall discuss previous work on the largest subgraph of a graph in Section 1.3 below. The problem of connecting the critical nature of random graphs to the value of the degree power-law exponent τ is novel.

1.2 Results

Before we can state our results, we introduce some notation. We write $[n] = \{1, \ldots, n\}$ for the set of vertices. For two vertices $s, t \in [n]$, we write $s \longleftrightarrow t$ when there exists a path of occupied edges connecting s and t. By convention, we always assume that $v \longleftrightarrow v$. For $v \in [n]$, we denote the *connected component containing* v or *cluster of* v by

$$C(v) = \big\{x \in [n] \colon v \longleftrightarrow x\big\}. \tag{1.10}$$

We denote the size of $C(v)$ by $|C(v)|$. The *largest connected component* is equal to any cluster $C(v)$ for which $|C(v)|$ is maximal, so that

$$|C_{\max}| = \max\{|C(v)| \colon v \in [n]\}. \tag{1.11}$$

Note that the above definition does identify $|C_{\max}|$ uniquely, but it may not identify $C_{\max}$ uniquely. We start by investigating the critical case, i.e., the case where ν in (1.8) is equal to 1:

Theorem 1.1 (Largest critical cluster for $\tau > 4$)**.** *Fix* $\mathrm{NR}_n(\mathbf{w})$ *with* $\mathbf{w} = \{w_i\}_{i=1}^n$ *as in* (1.3)*, and assume that the distribution function* F *in* (1.3) *satisfies* $\nu = 1$*. When there exists a* $\tau > 4$ *and a constant* $c > 0$ *such that, for all large enough* $x \geq 0$*,*

$$1 - F(x) \leq c_F x^{-(\tau-1)}, \tag{1.12}$$

then there exists a constant $b > 0$ *such that for all* $\omega > 1$ *and for* n *sufficiently large,*

$$\mathbb{P}\Big(\omega^{-1} n^{2/3} \leq |C_{\max}| \leq \omega n^{2/3}\Big) \geq 1 - \frac{b}{\omega}. \tag{1.13}$$

Theorem 1.2 (Largest critical cluster for $\tau \in (3,4)$)**.** *Fix* $\mathrm{NR}_n(\mathbf{w})$ *with* $\mathbf{w} = \{w_i\}_{i=1}^n$ *as in* (1.3)*, and assume that the distribution function* F *in* (1.3) *satisfies* $\nu = 1$*. When there exist* $0 < c_1 < c_2 < \infty$ *such that, for all large enough* $x \geq 0$*,*

$$c_1 x^{-(\tau-1)} \leq 1 - F(x) \leq c_2 x^{-(\tau-1)}, \tag{1.14}$$

then there exists a constant $b > 0$ *such that for all* $\omega > 1$ *and for* n *sufficiently large,*

$$\mathbb{P}\Big(|C_{\max}| \leq \omega n^{(\tau-2)/(\tau-1)}\Big) \geq 1 - \frac{b}{\omega}. \tag{1.15}$$

When, further, there exists $0 < c_F < \infty$ *such that, as* $x \to \infty$*,*

$$1 - F(x) = c_F x^{-(\tau-1)}(1 + o(1)), \tag{1.16}$$

then also

$$\mathbb{P}\Big(|C_{\max}| \geq \omega^{-1} n^{(\tau-2)/(\tau-1)}\Big) \geq 1 - \frac{b}{\omega}. \tag{1.17}$$

We call critical behavior of random graphs of size $n^{2/3}$ as in Theorem 1.1 *random graph asymptotics*. A special case of Theorem 1.1 is the critical behavior for the Erdős-Rényi random graph, where bounds as in (1.13) have a long history (see e.g., [17], as well as [4, 29, 34, 38] and the monographs [5, 31] for the most detailed results). The Erdős-Rényi random graph corresponds to taking $w_i = c$ for all $i \in [n]$, and then ν in (1.8) equals c. Therefore, criticality corresponds to $w_i = 1$ for all $i \in [n]$. Interestingly, what Theorems 1.1 and 1.2 show is that our rank-1 inhomogeneous random graphs have random graph asymptotics when $\tau > 4$, but not when $\tau \in (3,4)$. In the latter case, the critical clusters are of order $n^{(\tau-2)/(\tau-1)}$, which is smaller than $n^{2/3}$. When $\tau \uparrow 4$, then $(\tau-2)/(\tau-1) \uparrow 2/3$, so that the two regimes match up nicely.

For the Erdős-Rényi random graph there is a tremendous amount of work on the question for which values of p, similar critical behavior is observed as for the critical value $p = 1/n$ [1, 3, 17, 29, 33, 34]. Indeed, when we take $p = (1 + \lambda n^{-1/3})/n$, the largest cluster has size $\Theta(n^{2/3})$ for every fixed $\lambda \in \mathbb{R}$, but it is $o_{\mathbb{P}}(n^{2/3})$ when $\lambda \to -\infty$, and has size $\gg n^{2/3}$ when $\lambda \gg 1$. Therefore, the values p of the form $p = (1 + \lambda n^{-1/3})/n$ are sometimes called the *critical window*. We next study the critical window in the inhomogeneous setting:

Theorem 1.3 (Critical window for $\tau > 4$)**.** *Fix* $\mathrm{NR}_n(\mathbf{w})$ *with* $\mathbf{w} = \{w_i\}_{i=1}^n$ *as in* (1.3)*, and assume that the distribution function* F *in* (1.3) *satisfies* $\nu = 1$*. Fix* $\varepsilon_n = o(1)$*, and let* $\tilde{\mathbf{w}}$ *be defined by*

$$\tilde{w}_i = (1 + \varepsilon_n) w_i. \tag{1.18}$$

Assume that (1.12) *holds for some* $\tau > 4$*, and fix* ε_n *such that* $|\varepsilon_n| \leq \Lambda n^{-1/3}$ *for some* $\Lambda > 0$*. Then there exists a constant* $b = b(\Lambda) > 0$ *such that for all* $\omega > 1$ *and for* n *sufficiently large,* $\mathrm{NR}_n(\tilde{\mathbf{w}})$ *satisfies*

$$\mathbb{P}\Big(\omega^{-1} n^{2/3} \leq |C_{\max}| \leq \omega n^{2/3}\Big) \geq 1 - \frac{b}{\omega}. \tag{1.19}$$

Theorem 1.4 (Critical window for $\tau \in (3,4)$)**.** *Fix* $\mathrm{NR}_n(\mathbf{w})$ *with* $\mathbf{w} = \{w_i\}_{i=1}^n$ *as in* (1.3)*, and assume that the distribution function* F *in* (1.3) *satisfies* $\nu = 1$*. Fix* $\varepsilon_n = o(1)$*, and let* $\tilde{\mathbf{w}}$ *be defined by*

$$\tilde{w}_i = (1 + \varepsilon_n) w_i. \tag{1.20}$$

Assume that (1.16) *holds for some* $\tau \in (3,4)$*, and fix* ε_n *such that* $|\varepsilon_n| \leq \Lambda n^{-(\tau-3)/(\tau-1)}$ *for some* $\Lambda > 0$*. Then there exists a constant* $b = b(\Lambda) > 0$ *such that for all* $\omega > 1$ *and for* n *sufficiently large,* $\mathrm{NR}_n(\tilde{\mathbf{w}})$ *satisfies*

$$\mathbb{P}\Big(\omega^{-1} n^{(\tau-2)/(\tau-1)} \leq |C_{\max}| \leq \omega n^{(\tau-2)/(\tau-1)}\Big) \geq 1 - \frac{b}{\omega}. \tag{1.21}$$

Theorems 1.3–1.4 show that the critical window has width at least $n^{-1/3}$ when $\tau > 4$ and $n^{-(\tau-3)/(\tau-1)}$ when $\tau \in (3,4)$. Below, we shall argue on a heuristic basis that the windows in Theorems 1.3–1.4 really are the critical windows.

1.3 Discussion and related results and conjectures

In this section, we discuss our results and the relevant results in the literature. We also state conjectures for sharper results and for related random graph models.

Connecting the subcritical and supercritical regimes to the critical one. We denote the forward degree of the neighbor of a uniform vertex by

$$\nu_n = \frac{\sum_{j=1}^n w_j^2}{\sum_{j=1}^n w_j}. \tag{1.22}$$

Then, in the setting of (1.3), we shall see that $\nu_n \to \nu$, where ν is defined by (1.8) (see Corollary 4.2 below). As described in more detail in Section 4.2, we can approximate the exploration of a cluster by a branching process having mean ν_n. We shall see that this branching process has *finite variance* when $\tau > 4$, but not when $\tau \in (3,4)$.

We first give a heuristic explanation for the critical behavior of $n^{2/3}$ in Theorem 1.1. Indeed, with $\varepsilon = \nu - 1$ and $\tau > 4$, we have that the survival probability of the branching process approximation is like $\varepsilon_n n$ when $\varepsilon_n > 0$, while the largest subcritical cluster is like ε_n^{-2} when $\varepsilon_n < 0$. This suggests that the critical behavior arises precisely when $\varepsilon_n^{-2} = n\varepsilon_n$, i.e., when $\varepsilon_n = n^{-1/3}$, and in this case, the largest connected component is $\varepsilon_n n = n^{2/3}$ as in Theorem 1.1.

We next extend this heuristic to the case where $\tau \in (3,4)$, where the picture changes completely. The results in [26] suggest that the largest subcritical cluster is like $w_1/(1-\nu) = \Theta(n^{1/(\tau-1)}/|\varepsilon_n|)$. Of course, the results in [26] only prove this when $\nu < 1$ is *fixed*, but we conjecture that it extends to all subcritical ν. In the supercritical regime, instead, the largest connected component should be like $n\rho$, where ρ is the survival probability of an infinite variance branching process. A straightforward computation shows that, when $\varepsilon > 0$ and $\varepsilon \ll 1$, we have that $\rho \sim \varepsilon^{1/(\tau-3)}$. We shall make this precise in Lemma 4.6 below. Thus, critical behavior should be characterized by taking $n^{1/(\tau-1)}/\varepsilon_n = \varepsilon_n^{1/(\tau-3)} n$, which is $\varepsilon_n = n^{-(\tau-3)/(\tau-1)}$. In this case, the largest critical cluster should be $\varepsilon_n^{1/(\tau-3)} n \sim n^{(\tau-2)/(\tau-1)}$, as in Theorem 1.2. This shows that in both cases, the subcritical and supercritical regimes connect up nicely, and it would be of interest to make these connections rigorous, particularly in the case when $\tau \in (3,4)$, by proving that in the barely subcritical regime (where $\varepsilon_n \ll n^{-(\tau-3)/(\tau-1)}$), the largest connected component is indeed $\Theta(n^{1/(\tau-1)}/|\varepsilon_n|)$, and in the barely supercritical regime (where $\varepsilon_n \gg n^{-(\tau-3)/(\tau-1)}$), the largest connected component is $\Theta(\varepsilon_n^{1/(\tau-3)} n)$. We summarize the above heuristics in the following conjecture:

Conjecture 1.5 (The off-critical regimes). *Fix* $\mathrm{NR}_n(\mathbf{w})$ *with* $\mathbf{w} = \{w_i\}_{i=1}^n$ *as in* (1.3).
(a) Assume that (1.12) *holds for some* $\tau > 4$. *Then, when* $\varepsilon_n = \nu_n - 1 \gg n^{-1/3}$,

$$|C_{\max}| = \frac{2\varepsilon_n n}{\sigma^2}(1 + o_{\mathbb{P}}(1)), \tag{1.23}$$

where

$$\sigma^2 = \frac{\mathbb{E}[W^3]}{\mathbb{E}[W]}, \tag{1.24}$$

while, when $\varepsilon_n = \nu_n - 1 \ll -n^{-1/3}$,

$$|C_{\max}| = \max\{w_1/|\varepsilon_n|, \log(n\varepsilon_n^3)/\varepsilon_n^2\}(1 + o_{\mathbb{P}}(1)). \tag{1.25}$$

(b) Assume that (1.16) *holds for some* $\tau \in (3, 4)$. *Then, when* $\varepsilon_n = \nu_n - 1 \gg n^{-(\tau-3)/(\tau-1)}$, *there exists a constant* $A > 0$ *such that*

$$|C_{\max}| = A\varepsilon_n^{1/(\tau-3)} n(1 + o_{\mathbb{P}}(1)). \tag{1.26}$$

while, when $\varepsilon_n = \nu_n - 1 \ll -n^{-(\tau-3)/(\tau-1)}$,

$$|C_{\max}| = \frac{w_1}{|\varepsilon_n|}(1 + o_{\mathbb{P}}(1)). \tag{1.27}$$

The scaling limit of cluster sizes for $\tau > 4$. Aldous [1], building on earlier work by Erdős and Rényi [17], Bollobás [3] and Łuczak [33] (see also [29, 34] and the monographs [5, 31]) proves that the ordered cluster sizes of the various critical clusters for the Erdős-Rényi random graph weakly converges to a limiting process, which can be characterized as the excursions of a standard Brownian motion with a linearly decreasing drift, ordered in their sizes. More precisely, let $\{B_t\}_{t\geq 0}$ be standard Brownian motion, and let

$$W_t = B_t - t^2/2, \qquad t \geq 0, \tag{1.28}$$

denote a Brownian motion which has a negative drift $-t$ at time $t \geq 0$. Define the reflecting version of this process by $\{R_t\}_{t\geq 0}$, so that

$$R_t = W_t - \min_{0\leq s\leq t} W_s, \qquad t \geq 0, \tag{1.29}$$

and $R_t \geq 0$ for all times. Let $\{\gamma_j\}_{j=1}^\infty$ denote the ordered excursion lengths of the process $\{R_t\}_{t\geq 0}$. Then, one of the main results of [1] is that the ordered component sizes in the Erdős-Rényi random graph with $p = 1/n$ converge in distribution to $\{\gamma_j\}_{j=1}^\infty$.

The intuition behind this result is that when we explore the clusters one by one, then this can be described in terms of a *random walk* $\{S_t\}_{t=0}^\infty$, where S_t denotes the number of active vertices at time t. At time $t = 0$, there is one active vertex, say vertex 1, and $S_0 = 1$. The vertices go from unexplored to active to inactive, the latter meaning that they are explored. In this process, when there are active vertices, we take any one of them, and let X_t denote the number of its unexplored neighbors. Then we let $S_t = S_{t-1} + X_t - 1$. For the Erdős-Rényi random graph with $p = 1/n$, and conditionally on the number of unexplored vertices at time t being N_t, X_t has a Binomial distribution with parameters $n - N_t$ and $p = 1/n$. For large t, we have that $N_t \approx t$, so that

$$X_t \approx \mathrm{Poi}(1 - t/n), \tag{1.30}$$

where $\mathrm{Poi}(\lambda)$ denotes a Poisson random variable with mean λ. We note that the cluster of vertex 1 is explored when $S_t = 0$ for the first time, and thus $|C(1)| = \min\{t : S_t = 0\}$. When there are no active vertices, on the other hand, then $S_{t-1} = 0$, and we take the inactive vertex with minimal index, and set $S_t = 1$, and we repeat the above steps. Thus, we see that the cluster sizes correspond to the times between successive visits of 0 of the process $\{S_t\}_{t=0}^{\infty}$, and we further have that the process

$$R_t^{(n)} = n^{-1/3} S_{\lceil tn^{2/3} \rceil} \tag{1.31}$$

converges in distribution to $\{R_t\}_{t \geq 0}$. Now, because of (1.30), we have that

$$S_t = \sum_{s=1}^{t} (X_s - 1) \approx \mathrm{Poi}\big(t - t^2/2n\big) - t, \tag{1.32}$$

so that, by the central limit theorem for Poisson random variables,

$$R_t^{(n)} = n^{-1/3} S_{\lceil tn^{2/3} \rceil} \approx B_t - t^2/2 = W_t. \tag{1.33}$$

The reflection in (1.29) comes from the fact that when $S_{t-1} = 0$, we restart at $S_t = 1$ rather than continuing the exploration. We thus see that the growing negative drift in (1.28) originates from the depletion of points during the exploration process.

Now, for the critical $\mathrm{NR}_n(\mathbf{w})$ with $\mathbf{w} = \{w_i\}_{i=1}^n$ as in (1.3), we can again describe the exploration by a certain random walk $\{S_t\}_{t=0}^{\infty}$, and we again see that the mean number of forward neighbors of a vertex is approximately equal to 1. Thus, a similar argument applies. However, in the use of the central limit theorem in (1.33), it is used that the variance of the number of forward neighbors of a vertex in the exploration equals 1. This variance is *bounded* when F satisfies (1.12). Denote its variance by σ^2, then, we conjecture that the ordered component sizes in the critical $\mathrm{NR}_n(\mathbf{w})$ converge in distribution to $\{\gamma_j\}_{j=1}^{\infty}$ which are the excursion lengths of the process $\{R_t\}_{t \geq 0}$ in (1.29), apart from the fact that B_t in (1.28) should be replaced with σB_t. When we use that $\{B_{at}\}_{t=0}^{\infty}$ has the same distribution as $\{\sqrt{a} B_t\}_{t=0}^{\infty}$, and using this for $a = \sigma^{2/3}$, we note that with $\{W_t^{(\sigma)}\}_{t \geq 0}$ defined by $W_t^{(\sigma)} = \sigma B_t - t^2/2$, we have that $W_{\sigma^{2/3} t}^{(\sigma)}$ has the same distribution as $\sigma^{4/3} W_t^{(1)}$. This suggests that, for $\tau > 4$, the scaling limit of the largest critical cluster converges to $\sigma^{2/3}$ times the ones for the Erdős-Rényi random graph:

Conjecture 1.6 (Weak convergence of the ordered critical clusters for $\tau > 4$)**.** *Let $|C_{(1)}| \geq |C_{(2)}| \geq \ldots$, denote the connected components of $\mathrm{NR}_n(\mathbf{w})$ with $\mathbf{w} = \{w_i\}_{i=1}^n$ as in (1.3), ordered in size. Under the assumptions of Theorem 1.1, $\big\{|C_{(j)}| n^{-2/3}\big\}_{j=1}^{k}$ converges weakly to $\sigma^{2/3}$ times the scaling limit for the critical Erdős-Rényi random graph, where σ^2 is given in (1.24).*

The scaling limit of cluster sizes for $\tau \in (3, 4)$. When $\tau \in (3, 4)$, large parts of the above discussion remain valid, however, the variance of the forward degree of a vertex obtained in the exploration process is *infinite*. Therefore, one cannot expect convergence to a process on the basis of Brownian motion to take place, but rather to a Lévy process. This suggests hat one should study excursion of Lévy process excursions. It would be interesting to see whether an explanation as the one above for the $\tau > 4$ case can be derived for the cluster sizes in Theorem 1.2. We conjecture that the ordered cluster sizes weakly convergence:

Conjecture 1.7 (Weak convergence of the ordered critical clusters for $\tau \in (3,4)$)**.** *Let* $|C_{(1)}| \geq |C_{(2)}| \geq \dots,$ *denote the connected components of* $\mathrm{NR}_n(\mathbf{w})$ *with* $\mathbf{w} = \{w_i\}_{i=1}^n$ *as in* (1.3)*, ordered in size. Assume that* $\nu = 1$ *and that* (1.16) *in Theorem* 1.2 *holds. Then* $\left\{|C_{(j)}|n^{-(\tau-2)/(\tau-1)}\right\}_{j=1}^k$ *converges weakly to a non-degenerate limit distribution.*

By Theorem 1.2, $|C_{(1)}|n^{-(\tau-2)/(\tau-1)}$ is tight, and remains strictly positive with high probability. A close inspection of the proof shows that this result can be extended to $|C_{(j)}|n^{-(\tau-2)/(\tau-1)}$ for any j.

The configuration model. Given a **degree sequence**, namely a sequence of n positive integers $\mathbf{d} = (d_1, d_2, \dots, d_n)$ with the total degree

$$l_n^{(\mathrm{CM})} = \sum_{i=1}^n d_i \tag{1.34}$$

assumed to be even, the configuration model (CM) on n vertices with degree sequence $\mathbf{d}$ is constructed as follows:

Start with n vertices and d_j stubs adjacent to vertex j. The graph is constructed by pairing up each stub to some other stub to form edges. Number the stubs from 1 to $l_n^{(\mathrm{CM})}$ in some arbitrary order. Then, at each step, two stubs (not already paired) are chosen uniformly at random among all the *free* stubs and are paired to form a single edge in the graph. These stubs are no longer free and removed from the list of free stubs. We continue with this procedure of choosing and pairing two stubs until all the stubs are paired.

By varying the degree sequence $\mathbf{d}$, one obtains random graph with various degree sequences in a similar was as how varying $\mathbf{w}$ influences the degree sequence in the rank-1 inhomogeneous random graph studied here. A first setting which produces a random graph with asymptotic degree sequences according to some distribution F arises by taking $\{d_i\}_{i=1}^n = \{D_i\}_{i=1}^n$, where $\{D_i\}_{i=1}^n$ are i.i.d. random variables with distribution F. An alternative choice is to take $\{d_i\}_{i=1}^n$ such that the number of vertices with degree k equals $\lceil nF(k)\rceil - \lceil nF(k-1)\rceil$.

The CM is not necessarily simple, i.e., it can have self-loops and multiple edges. However, when

$$\nu_n^{(\mathrm{CM})} = \frac{1}{l_n} \sum_{i=1}^n d_i(d_i - 1) \tag{1.35}$$

converges as $n \to \infty$ and $d_i = o(\sqrt{n})$ for each $i \in [n]$, then the number of self-loops and multiple converge in distribution to independent Poisson random variables (see e.g., [27] and the references therein). In [35], the phase transition of the CM was investigated, and it was shown that when

$$\nu^{(\mathrm{CM})} = \lim_{n\to\infty} \nu_n^{(\mathrm{CM})} > 1, \tag{1.36}$$

and certain conditions on the degrees are satisfied, then a giant component exists, while if $\nu^{(\mathrm{CM})} \leq 1$, then the largest connected component has size $o_{\mathbb{P}}(1)$. In [30], some of the conditions were removed. Also the *barely supercritical* regime, where $n^{1/3}(\nu_n - 1) \to \infty$, is investigated. Note that when the proportion of vertices of degree k is asymptotic to $k^{-\tau}$ (for example when $F(k) - F(k-1) = ck^{-\tau}$), then this corresponds to $\tau > 4$. However, [30] also makes a more stringent condition, namely that $\sum_{i=1}^n d_i^{4+\eta} = O(n)$ for some $\eta > 0$, which corresponds to $\tau > 5$. In [30, Remark 2.5], it is conjectured that this condition is not necessary, and that, in fact, the results should hold for $\tau > 4$. Similar results are proved in [32] under related conditions.

The results in [30, 32] suggest that the barely supercritical regime for the CM is similar to the one for the Erdős-Rényi random graph when $\tau > 4$. We strengthen this by conjecturing that the largest connected component of the CM obeys identical bounds as for the rank-1 inhomogeneous random graph studied here:

Conjecture 1.8 (The critical behavior of the configuration model). *Fix the degrees in the configuration model $\{d_i\}_{i=1}^n$ such that the number of vertices with degree k equals $\lceil nF(k)\rceil - \lceil nF(k-1)\rceil$. Then, the configuration model obeys the same asymptotics as the $\mathrm{NR}_n(\mathbf{w})$ under the same assumptions on F, with ν_n in (4.15) replaced with $\nu^{(\mathrm{CM})}$ in (1.36), and with σ^2 replaced with*

$$\left(\sigma^{(\mathrm{CM})}\right)^2 = \frac{\mathbb{E}[D(D-1)(D-2)]}{\mathbb{E}[D]}, \tag{1.37}$$

where D has distribution function F.

As discussed earlier, parts of Conjecture 1.8 have been proved in [26, 30, 32], see also Theorem 2.2 below. The parameters $\sigma^{(\mathrm{CM})}$ is quite similar to σ in (1.24) above, since, for the $\mathrm{NR}_n(\mathbf{w})$ with $\mathbf{w} = \{w_i\}_{i=1}^n$ as in (1.3), the degree of a uniform vertex has distribution close to $D = \mathrm{Poi}(W)$ (see (1.6)–(1.7)), for which

$$\mathbb{E}[D] = \mathbb{E}[W], \qquad \mathbb{E}[D(D-1)(D-2)] = \mathbb{E}[W^3], \tag{1.38}$$

so that $\sigma^{(\mathrm{CM})}$ for $D = \mathrm{Poi}(W)$ reduces to σ in (1.24). In Section 2 we elaborate on this connection. See in particular Theorem 2.2 below, where we shall show that, by asymptotic contiguity, the results in [30] also apply to the barely supercritical regime in the rank-1 inhomogeneous random graph as studied here, when $\tau > 5$.

2 Asymptotic equivalence and contiguity

The Norros-Reittu random graph model is the a special case of the so-called rank-1 case of the general inhomogeneous random graphs as studied in [6]. We start by introducing the general setting of inhomogeneous random graphs. The vertex set of the models under consideration shall be denoted by $[n] = \{1, \ldots, n\}$. We write $\mathbf{p} = \{p_{ij}\}_{1\leq i<j\leq n}$ for the edge probabilities in the graph, and we write $\mathrm{IRG}_n(\mathbf{p})$ for the inhomogeneous random graph for the probability that the edge ij is present equals p_{ij} and the events that different edges are present are independent. In [6], this setting is studied, where the edge probabilities p_{ij} are given by

$$p_{ij} = p_{ij}(\kappa) = \min\{\kappa(x_i, x_j)/n, 1\}, \tag{2.1}$$

where $\{x_i\}_{i=1}^n$ are elements of some state space $\mathcal{X}$ and $\kappa\colon \mathcal{X}\times\mathcal{X} \to [0,\infty)$ is a kernel that moderates the inhomogeneity of the graph. The rank-1 case is obtained by assuming that κ is of product form, i.e., there exists a function $\psi\colon \mathcal{X} \to [0,\infty)$ such that

$$\kappa(x, y) = \psi(x)\psi(y). \tag{2.2}$$

The rank-1 inhomogeneous random graph has attracted considerable attention in the literature, and detailed results are known about the degrees [6, 9], the phase transition [11, 14] and distances [6, 10, 12, 18, 36] in such models. We especially refer to the monographs [13, 15] for a detailed account of such results, as well as to the lecture notes [21, Chapters 6 and 9] where many properties are investigated.

We now define two random graph models that are closely related to the Norros-Reittu random graph. In the *generalized random graph model* [9], the edge probability of the edge between vertices i and j is equal to

$$p_{ij} = p_{ij}^{(\mathrm{GRG})} = \frac{w_i w_j}{l_n + w_i w_j}. \tag{2.3}$$

In the *expected degree random graph* or Chung-Lu random graph [10, 11, 12, 13, 14], the edge probabilities are given by

$$p_{ij}^{(\mathrm{CL})} = \max\{\frac{w_i w_j}{l_n}, 1\}. \tag{2.4}$$

When $\max_{i=1}^n w_i^2 \leq l_n$, we may forget about the maximum with 1 in (2.4). We shall assume $\max_{i=1}^n w_i^2 \leq l_n$ throughout this section, and denote the resulting graph by $\mathrm{CL}_n(\mathbf{w})$. The Chung-Lu model is sometimes referred to as the random graph with given expected degrees, as the expected degree is vertex i is equal to w_i. [1]

We now show that our results apply to $\mathrm{NR}_n(\mathbf{w}), \mathrm{GRG}_n(\mathbf{w})$, and $\mathrm{CL}_n(\mathbf{w})$ all at once:

Theorem 2.1 (Asymptotic equivalence [28])**.** *Fix $\mathbf{w} = \{w_j\}_{j=1}^n$ as in (1.3), and assume that F satisfies (1.12) for some $\tau > 3$. Then, the results on $\mathrm{NR}_n(\mathbf{w})$ in Theorems 1.1–1.4 hold for $\mathrm{GRG}_n(\mathbf{w})$ and $\mathrm{CL}_n(\mathbf{w})$ under the same conditions on F.*

Proof. By the results in [28], the random graphs $\mathrm{NR}_n(\mathbf{w}), \mathrm{GRG}_n(\mathbf{w})$, and $\mathrm{CL}_n(\mathbf{w})$ are asymptotically equivalent, meaning that all sequences of events have asymptotically equal probabilities, when

$$\sum_{j=1}^n w_j^3 = o(n^{3/2}). \tag{2.5}$$

In the case of (1.3), we can bound, using

$$\frac{1}{n}\sum_{i=1}^n w_i^3 \leq \frac{1}{n}\sum_{j=1}^n \big(c_F n/j\big)^{-1/(\tau-1)} = o(n^{1/2}), \tag{2.6}$$

since $1 - F(x) \leq c_F x^{-(\tau-1)}$, which implies that (by e.g., [18, (B.9)])

$$[1-F]^{-1}(u) \leq \big(c_F/u\big)^{1/(\tau-1)}. \tag{2.7}$$

□

Theorem 2.2 (The barely supercritical regime [30])**.** *When (1.12) holds for some $\tau > 5$, then the results of [30] also hold for $\mathrm{NR}_n(\mathbf{w})$, $\mathrm{GRG}_n(\mathbf{w})$ and $\mathrm{CL}_n(\mathbf{w})$. More precisely, define*

$$\varepsilon_n = \frac{1}{l_n}\sum_{j=1}^n w_j^2 - 1, \tag{2.8}$$

then, when $\varepsilon_n \to 0$ such that $n^{1/3}\varepsilon_n \to \infty$,

$$\frac{|C_{\max}|}{n\varepsilon_n} \xrightarrow{\mathbb{P}} \frac{2\mathbb{E}[W]}{\mathbb{E}[W^3]}. \tag{2.9}$$

[1] To make this precise, one needs that $w_i^2 \leq l_n$ for every i, and we add an artificial self-loop at vertex i with probability w_i^2/l_n.

Further, when $\varepsilon_n \to \nu - 1 > 0$*, then* $|C_{\max}|/n \xrightarrow{\mathbb{P}} \rho$*, where* $\rho > 0$ *is the survival probability of an appropriate branching process.*

Proof. First, by the asymptotic equivalence of $\mathrm{NR}_n(\mathbf{w})$, $\mathrm{GRG}_n(\mathbf{w})$ and $\mathrm{CL}_n(\mathbf{w})$ which is established in the proof of Theorem 2.1 using the results in [28], it suffices to prove the result for $\mathrm{GRG}_n(\mathbf{w})$. Now, it is well-known that $\mathrm{GRG}_n(\mathbf{w})$ conditioned on its degrees is uniform from all graphs with those degrees. The configuration model conditioned on not having any self-loops and multiple edges is also a uniform graph with the specified degree sequence. Thus, the two are the same. Now, when $\tau > 3$, we have that the maximal degree of the $\mathrm{GRG}_n(\mathbf{w})$ is $o(\sqrt{n})$ with high probability, and the number of vertices with degree k converges to f_k (recall (1.6)). Thus, by [27], the CM with asymptotic degree distribution $\{N_k/n\}_{k=0}^{\infty}$ has, with probability that remains strictly positive as $n \to \infty$, no self-loops and multiple edges. As a result, any limit in probability proved for the CM with this degree sequence also holds for $\mathrm{GRG}_n(\mathbf{w})$ (see [28]). □

3 Strategy of the proof

We start by describing the strategy of proof for Theorems 1.1–1.4. The proofs of all these results shall follow the same strategy. We denote by

$$Z_{\geq k} = \sum_{v=1}^{n} \mathbb{1}_{\{|C(v)|\geq k\}} \tag{3.1}$$

the number of vertices that are contained in connected components of size at least k. The random variable $Z_{\geq k}$ will be used to prove the asymptotics of $|C_{\max}|$. This can be understood by noting that

$$|C_{\max}| = \max\{k : Z_{\geq k} \geq k\}, \tag{3.2}$$

which allows us to prove bounds on $|C_{\max}|$ by investigating $Z_{\geq k}$ for appropriately chosen values of k. This strategy has been successfully applied in several related settings, such as percolation on the torus in general dimension [8] as well as for percolation on high-dimensional tori [7, 20, 25]. This is the first time that this methodology is applied to an *inhomogeneous* setting.

Proposition 3.1 (An upper bound on the largest critical cluster)**.** *Suppose that there exists an* $a_1 > 0$ *such that for all* $k \geq n^{\delta/(1+\delta)}$ *and for* V *a uniform vertex in* $[n]$*, the bound*

$$\mathbb{P}(|C(V)| \geq k) \leq a_1\big(k^{-1/\delta} + \big(\varepsilon_n \vee n^{-\alpha/(\delta+1)}\big)^{1/\alpha}\big) \tag{3.3}$$

holds, where

$$|\varepsilon_n| \leq \Lambda n^{-\alpha/(\delta+1)}. \tag{3.4}$$

Then, there exists a $b_1 = b_1(\Lambda) > 0$ *such that, for all* $\omega \geq 1$*,*

$$\mathbb{P}\big(|C_{\max}| \geq \omega n^{\delta/(1+\delta)}\big) \leq \frac{b_1}{\omega}. \tag{3.5}$$

Proof. We use the first moment method or Markov inequality, to bound

$$\mathbb{P}(|C_{\max}| \geq k) = \mathbb{P}(Z_{\geq k} \geq k) \leq \frac{1}{k}\mathbb{E}[Z_{\geq k}] = \frac{n}{k}\mathbb{P}(|C(V)| \geq k), \tag{3.6}$$

where $V \in [n]$ is a uniform vertex. Thus, we need to bound $\mathbb{P}(|C(V)| \geq k)$ for an appropriately chosen $k = k_n$. We use (3.3), so that

$$\begin{aligned}
\mathbb{P}(|C_{\max}| \geq k) &\leq \frac{a_1 n}{k}\Big(k^{-1/\delta} + \big(\varepsilon_n \vee n^{-\alpha/(\delta+1)}\big)^{1/\alpha}\Big) \\
&\leq a_1\big(\omega^{-(1+1/\delta)} + (n^{1/(\delta+1)}|\varepsilon_n|^{1/\alpha} \vee 1)\omega^{-1}\big) \\
&\leq a_1\big(\omega^{-(1+1/\delta)} + (1+\Lambda)\omega^{-1}\big), \qquad (3.7)
\end{aligned}$$

when $k = k_n = \omega n^{1/(1+1/\delta)} = \omega n^{\delta/(1+\delta)}$, and where we have used (3.4). This completes the proof of Theorem 3.1, with $b_1 = a_1(2+\Lambda)$. □

Proposition 3.1 immediately yields that in order to prove an upper bound on $|C_{\max}|$, it suffices to prove an upper bound on the cluster tails of a uniform vertex. In order to prove a matching lower bound on $|C_{\max}|$, we shall use the second moment method, for which we need to give a bound on the variance of $Z_{\geq k}$, in which we make use of the notation

$$\chi_{\geq k}(\mathbf{p}) = \mathbb{E}[|C(V)|\mathbb{1}_{\{|C(V)|\geq k\}}], \qquad (3.8)$$

where we recall that V is a uniform vertex in $[n]$, and where we recall that $\mathbf{p} = \{p_{ij}\}_{1\leq i<j\leq n}$ denote the edge probabilities of the inhomogeneous random graph. Then the main variance estimate on $Z_{\geq k}$ is as follows:

Proposition 3.2 (A variance estimate for $Z_{\geq k}$)**.** *For any inhomogeneous random graph with edge probabilities $\mathbf{p} = \{p_{ij}\}_{1\leq i<j\leq n}$, every n and $k \in [n]$,*

$$\mathrm{Var}(Z_{\geq k}) \leq n\chi_{\geq k}(\mathbf{p}). \qquad (3.9)$$

Proof. We use that

$$\mathrm{Var}(Z_{\geq k}) = \sum_{i,j=1}^{n} \big[\mathbb{P}(|C(i)| \geq k, |C(j)| \geq k) - \mathbb{P}(|C(i)| \geq k)\mathbb{P}(|C(j)| \geq k)\big]. \qquad (3.10)$$

We split the probability $\mathbb{P}(|C(i)| \geq k, |C(j)| \geq k)$, depending on whether $i \longleftrightarrow j$ or not, i.e., we split

$$\begin{aligned}
\mathbb{P}(|C(i)| \geq k, |C(j)| \geq k) = \mathbb{P}(|C(i)| \geq k, |C(j)| \geq k, i \longleftrightarrow j) \\
+ \mathbb{P}(|C(i)| \geq k, |C(j)| \geq k, i \not\longleftrightarrow j). \qquad (3.11)
\end{aligned}$$

We can bound

$$\mathbb{P}(|C(i)| \geq k, |C(j)| \geq k, i \not\longleftrightarrow j) \leq \mathbb{P}\Big(\{|C(i)| \geq k\} \circ \{|C(j)| \geq k\}\Big), \qquad (3.12)$$

where, for two increasing events E and F, we write $E \circ F$ to denote the event that E and F occur disjointly, i.e., that there exists a (random) set of edges K such that we can see that E occurs by only inspecting the edges in K and that F occurs by only inspecting the edges in K^c. Then, the BK-inequality [2, 19] states that

$$\mathbb{P}(E \circ F) \leq \mathbb{P}(E)\mathbb{P}(F). \qquad (3.13)$$

applying this to (3.12), we obtain that

$$\mathbb{P}(|C(i)| \geq k, |C(j)| \geq k, i \not\longleftrightarrow j) \leq \mathbb{P}(|C(i)| \geq k)\mathbb{P}(|C(j)| \geq k). \qquad (3.14)$$

Therefore,

$$\mathrm{Var}(Z_{\geq k}) \leq \sum_{i,j=1}^{n} \mathbb{P}(|C(i)| \geq k, |C(j)| \geq k, i \longleftrightarrow j), \tag{3.15}$$

and we arrive at the fact that

$$\begin{aligned}\mathrm{Var}(Z_{\geq k}) &\leq \sum_{i,j=1}^{n} \mathbb{P}(|C(i)| \geq k, |C(j)| \geq k, i \longleftrightarrow j) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathbb{E}\big[\mathbb{1}_{\{|C(i)|\geq k\}}\mathbb{1}_{\{j\in C(i)\}}\big] \\ &= \sum_{i=1}^{n} \mathbb{E}\Big[\mathbb{1}_{\{|C(i)|\geq k\}}\sum_{j=1}^{n}\mathbb{1}_{\{j\in C(i)\}}\Big]. \end{aligned}\tag{3.16}$$

Since $\sum_j \mathbb{1}_{\{j\in C(i)\}} = |C(i)|$, we thus arrive at

$$\mathrm{Var}(Z_{\geq k}) \leq \sum_{i=1}^{n} \mathbb{E}[|C(i)|\mathbb{1}_{\{|C(i)|\geq k\}}] = n\mathbb{E}[|C(V)|\mathbb{1}_{\{|C(V)|\geq k\}}] = n\chi_{\geq k}(\mathbf{p}). \tag{3.17}$$

□

Proposition 3.3 (A lower bound on the largest critical cluster)**.** *Suppose that there exists $a_2 > 0$ and $K > 0$ such that for all $k \leq n^{\delta/(1+\delta)}$ and for V a uniform vertex in $[n]$,*

$$\mathbb{P}(|C(V)| \geq k) \geq \frac{a_2}{k^{1/\delta}}, \tag{3.18}$$

while

$$\mathbb{E}[|C(V)|] \leq Kn^{(\delta-1)/(\delta+1)}, \tag{3.19}$$

then there exists a $b_2 > 0$ such that, for all $\omega \geq 1$,

$$\mathbb{P}\big(|C_{\max}| \leq \omega^{-1}n^{\delta/(1+\delta)}\big) \leq \frac{b_2}{\omega^{2/\delta}}. \tag{3.20}$$

We can intuitively understand (3.19) as follows. Clearly,

$$\mathbb{E}[|C(V)|] = \sum_{k=1}^{\infty} \mathbb{P}(|C(V)| \geq k). \tag{3.21}$$

By Proposition 3.1, we see that $|C_{\max}| \leq \omega n^{\delta/(1+\delta)}$ with high probability. This suggests that we may restrict the sum in (3.21) to $k \leq n^{\delta/(1+\delta)}$. Then using (3.3) and (3.18), we are lead to

$$\mathbb{E}[|C(V)|] \approx \sum_{k=1}^{n^{\delta/(1+\delta)}} k^{-1/\delta} \sim n^{(\delta-1)/(\delta+1)}, \tag{3.22}$$

where $\approx$ denotes a asymptotic relation with an uncontrolled error term.

Equation (3.19) shows that this is indeed the right order of magnitude for the upper bound. It is trivial to use (3.18) to prove a lower bound of the same order. Thus, we can think of (3.19) as yielding the right asympotics for the expected cluster size.

Proof of Proposition 3.3. We use the Chebychev inequality, as well as

$$\{|C_{\max}| < k\} = \{Z_{\geq k} = 0\}, \tag{3.23}$$

to obtain that

$$\mathbb{P}\big(|C_{\max}| < \omega^{-1} n^{\delta/(1+\delta)}\big) = \mathbb{P}_1\big(Z_{\geq \omega^{-1} n^{\delta/(1+\delta)}} = 0\big) \leq \frac{\mathrm{Var}\big(Z_{\geq \omega^{-1} n^{\delta/(1+\delta)}}\big)}{\mathbb{E}[Z_{\geq \omega^{-1} n^{\delta/(1+\delta)}}]^2}. \tag{3.24}$$

By (3.18), we have that

$$\mathbb{E}[Z_{\geq \omega^{-1} n^{\delta/(1+\delta)}}] = n\mathbb{P}(|C(V)| \geq \omega^{-1} n^{\delta/(1+\delta)}) \geq \frac{na_2}{n^{1/(1+\delta)}} = a_2 \omega^{1/\delta} n^{\delta/(\delta+1)}. \tag{3.25}$$

Also, by Proposition 3.2, with $k_n = \omega^{-1} n^{2/3}$,

$$\mathrm{Var}(Z_{\geq \omega^{-1} n^{\delta/(\delta+1)}}) \leq n\chi_{\geq \omega^{-1} n^{\delta/(\delta+1)}}(\mathbf{p}) \leq K n^{1+(\delta-1)/(\delta+1)} = K n^{2\delta/(\delta+1)}. \tag{3.26}$$

Substituting (3.24)–(3.26), we obtain, for n sufficiently large,

$$\mathbb{P}_1\big(|C_{\max}| < \omega^{-1} n^{\delta/(1+\delta)}\big) \leq \frac{K n^{2\delta/(\delta+1)}}{a_1^2 \omega^{2/\delta} n^{2\delta/(\delta+1)}} = \frac{K}{a_1^2 \omega^{2/\delta}}. \tag{3.27}$$

This completes the proof of Proposition 3.3. □

4 Preliminaries

In this section, we derive preliminary results needed to prove the bounds on cluster tails and expected cluster sizes that we shall need in order to apply Propositions 3.1–3.3. We start in Section 4.1 by analyzing sums involving powers of $\{w_j\}_{j=1}^n$, and in Section 4.2 we describe a beautiful connection between branching processes and clusters in the Norros-Reittu model due to Norros and Reittu in [36].

4.1 The weight of a random vertex W_n

In this section, we investigate the weight of a uniform vertex in $[n]$, which we denote by W_n. For this, we first note that

$$[1-F]^{-1}(1-u) = F^{-1}(u) = \inf\{x : F(x) \geq u\}, \tag{4.1}$$

which, in particular, implies that $[1-F]^{-1}(U)$ has distribution function F when U is uniform on $(0,1)$. This implies that W_n is a random variable with distribution function F_n given by

$$F_n(x) = \mathbb{P}(W_n \leq x) = \frac{1}{n}\big(\lfloor nF(x) \rfloor + 1\big) \wedge 1. \tag{4.2}$$

Indeed, we note that

$$\begin{aligned}
\sum_{j=1}^n \mathbb{1}_{\{w_j \leq x\}} &= \sum_{j=1}^n \mathbb{1}_{\{[1-F]^{-1}(\frac{j}{n}) \leq x\}} = \sum_{i=0}^{n-1} \mathbb{1}_{\{[1-F]^{-1}(1-\frac{i}{n}) \leq x\}} \\
&= \sum_{i=0}^{n-1} \mathbb{1}_{\{F^{-1}(\frac{i}{n}) \leq x\}} = \sum_{i=0}^{n-1} \mathbb{1}_{\{\frac{i}{n} \leq F(x)\}} = \min\{n, \lfloor nF(x) \rfloor + 1\},
\end{aligned} \tag{4.3}$$

where we write $j = n - i$ in the second equality and use (4.1) in the third equality. This proves (4.2). Note that $F_n(x) \geq F(x)$, which shows that W_n is stochastically dominated by W, so that, in particular, for *increasing* functions $x \mapsto h(x)$,

$$\frac{1}{n}\sum_{j=1}^{n} h(w_j) \leq \mathbb{E}[h(W)]. \tag{4.4}$$

In the sequel, we shall repeatedly rely on the following lemma in order to bound expectations of W_n:

Lemma 4.1 (Expectations of W_n)**.** *Let W have distribution function F and assume that (1.12) holds for some $\tau > 3$. Let $x \mapsto h(x)$ be a differentiable function with $h(0) = 0$, and such that*

$$\int_0^\infty |h'(x)|[1 - F(x)]dx < \infty. \tag{4.5}$$

Then,

$$\Big|\mathbb{E}[h(W_n)] - \mathbb{E}[h(W)]\Big| \leq \int_{an^{1/(\tau-1)}}^\infty |h'(x)|[1 - F(x)]dx + \frac{1}{n}\int_0^{an^{1/(\tau-1)}} |h'(x)|dx. \tag{4.6}$$

Proof. We need to investigate $\mathbb{E}[h(W)]$ for a random variable W. For this, we write, using that $h(0) = 0$,

$$\begin{aligned}\mathbb{E}[h(W)] &= \mathbb{E}\Big[\int_0^W h'(x)dx\Big] = \mathbb{E}\Big[\int_0^\infty h'(x)\mathbb{1}_{\{x<W\}}dx\Big] \\ &= \int_0^\infty h'(x)\mathbb{E}\Big[\mathbb{1}_{\{x<W\}}\Big]dx = \int_0^\infty h'(x)[1 - F(x)]dx,\end{aligned} \tag{4.7}$$

where we use Fubini, which is allowed since $|h'(x)|[1 - F(x)]$ is integrable by assumption. Because of this representation, we have that

$$\mathbb{E}[h(W_n)] - \mathbb{E}[h(W)] = \int_0^\infty h'(x)[F(x) - F_n(x)]dx. \tag{4.8}$$

Now, when (1.12) holds, then

$$F_n(an^{1/(\tau-1)}) = 1 \tag{4.9}$$

for some $a > 0$. Thus,

$$\Big|\mathbb{E}[h(W_n)] - \mathbb{E}[h(W)]\Big| \leq \int_{an^{1/(\tau-1)}}^\infty |h'(x)|[1 - F(x)]dx + \int_0^{an^{1/(\tau-1)}} |h'(x)|[F_n(x) - F(x)]dx. \tag{4.10}$$

We finally use that

$$F_n(x) - F(x) \leq 1/n \tag{4.11}$$

to arrive at

$$\Big|\mathbb{E}[h(W_n)] - \mathbb{E}[h(W)]\Big| \leq \int_{an^{1/(\tau-1)}}^\infty |h'(x)|[1 - F(x)]dx + \frac{1}{n}\int_0^{an^{1/(\tau-1)}} |h'(x)|dx. \tag{4.12}$$

□

Corollary 4.2 (Bounds on characteristic function and mean degrees)**.** *Let W have distribution function F and assume that* (1.12) *holds for some $\tau > 3$. Let*

$$\varphi_n(t) = \frac{\mathbb{E}[W_n \mathrm{e}^{it((1+\varepsilon_n)W_n-1)}]}{\mathbb{E}[W_n]}, \qquad \varphi(t) = \frac{\mathbb{E}[W\mathrm{e}^{it(W-1)}]}{\mathbb{E}[W]}. \tag{4.13}$$

Then,

$$|\varphi_n(t) - \varphi(t)| \leq cn^{-\frac{\tau-2}{\tau-1}} + c|t|(n^{-\frac{\tau-3}{\tau-1}} + |\varepsilon_n|). \tag{4.14}$$

Further, with

$$\tilde{\nu}_n = (1+\varepsilon_n)\frac{\mathbb{E}[W_n^2]}{\mathbb{E}[W_n]}, \tag{4.15}$$

and assume that (1.12) *holds for some $\tau > 3$. Then, with ν in* (1.8)*,*

$$|\tilde{\nu}_n - \nu| \leq c(|\varepsilon_n| + n^{-\frac{\tau-3}{\tau-1}}). \tag{4.16}$$

Proof. We first take $\varepsilon_n = 0$, and split

$$\varphi_n(t) - \varphi(t) = \frac{\varphi(t)}{\mathbb{E}[W]}\Big(\big(\mathbb{E}[W_n]\big)^{-1} - \frac{1}{\mathbb{E}[W]}\Big) + \frac{\varphi(t)}{\mathbb{E}[W_n]}\Big(\mathbb{E}[W_n\mathrm{e}^{it(W_n-1)}] - \mathbb{E}[W\mathrm{e}^{it(W-1)}]\Big). \tag{4.17}$$

Lemma 4.1 applied to $h(x) = x$, yields

$$\big|\mathbb{E}[W_n] - \mathbb{E}[W]\big| \leq cn^{-\frac{\tau-2}{\tau-1}}. \tag{4.18}$$

We next apply Lemma 4.1 to $h(x) = x\mathrm{e}^{it(x-1)}$, for which we compute

$$|h'(x)| = |\mathrm{e}^{it(x-1)} + itx\mathrm{e}^{it(x-1)}| \leq 1 + x|t|. \tag{4.19}$$

Therefore,

$$\begin{aligned}\Big|\mathbb{E}[W_n\mathrm{e}^{it(W_n-1)}] - \mathbb{E}[W\mathrm{e}^{it(W-1)}]\Big| &\leq \int_{an^{1/(\tau-1)}}^{\infty}(1+x|t|)[1-F(x)]dx + \frac{1}{n}\int_0^{an^{1/(\tau-1)}}(1+x|t|)dx \\ &\leq cn^{-\frac{\tau-2}{\tau-1}} + c|t|n^{-\frac{\tau-3}{\tau-1}}.\end{aligned} \tag{4.20}$$

Together, these two estimates prove the claim for $\varepsilon_n = 0$. When $\varepsilon_n \neq 0$, then we use that

$$\begin{aligned}\frac{\mathbb{E}[W_n\mathrm{e}^{it((1+\varepsilon_n)W_n-1)}]}{\mathbb{E}[W_n]} - \frac{\mathbb{E}[W_n\mathrm{e}^{it(W_n-1)}]}{\mathbb{E}[W_n]} &= \frac{\mathbb{E}\big[W_n\mathrm{e}^{it(W_n-1)}(\mathrm{e}^{it\varepsilon_n W_n}-1)\big]}{\mathbb{E}[W_n]} \\ &= O(|\varepsilon_n||t|\mathbb{E}[W_n^2]/\mathbb{E}[W_n]) = O(|\varepsilon_n||t|).\end{aligned} \tag{4.21}$$

The proof of (4.16) is similar. □

Lemma 4.3 (Bounds and asymptotics of moments of W_n)**.** *Let W_n have distribution function F_n in* (4.2)*.*

(i) Assume that (1.12) *holds for some $\tau > 3$, and let $a < \tau - 1$. Then, for x sufficiently large, there exists a $C = C(a, \tau)$ such that, uniformly in n,*

$$\mathbb{E}[W_n^a \mathbb{1}_{\{W_n \geq x\}}] \leq Cx^{a+1-\tau}. \tag{4.22}$$

(ii) Assume that (1.16) holds for some $\tau > 3$, and let $a > \tau - 1$. Then, there exists $C_1 = C_1(a,\tau)$ and $C_2 = C_2(a,\tau)$ such that, uniformly in n,

$$C_1\big(x \wedge n^{1/(\tau-1)}\big)^{a+1-\tau} \leq \mathbb{E}[W_n^a \mathbb{1}_{\{W_n \leq x\}}] \leq C_2 x^{a+1-\tau}, \tag{4.23}$$

where, in the lower bound, we write, for $x, y \in \mathbb{R}$, $x \wedge y = \min\{x, y\}$.

Proof. (i) When $a < \tau - 1$, the expectation is finite. We rewrite the integral, using partial integration, as

$$\begin{aligned}\mathbb{E}[W_n^a \mathbb{1}_{\{W_n \geq x\}}] &= \int_x^\infty w^a dF_n(w) = a^{-1} \int_0^\infty v^{a-1}[1 - F_n(x \vee v)]dv \\ &= x^a[1 - F_n](x) + a^{-1} \int_x^\infty v^{a-1}[1 - F_n(v)]dv,\end{aligned} \tag{4.24}$$

where $x \vee v = \max\{x, v\}$. Now, $1 - F_n(x) \leq 1 - F(x)$, so that we may replace the F_n by F in an upper bound. When (1.12) holds for some $\tau > 3$, then we can further bound this as

$$\mathbb{E}[W_n^a \mathbb{1}_{\{W \geq x\}}] \leq cx^{a+1-\tau} + Ca^{-1} \int_x^\infty w^{a-\tau} dw \leq Cx^{a+1-\tau}, \tag{4.25}$$

where the value of the constant C changes from line to line.
(ii) We again rewrite, using partial integration,

$$\begin{aligned}\mathbb{E}[W_n^a \mathbb{1}_{\{W_n \leq x\}}] &= \int_0^x w^a dF(w) = a^{-1} \int_0^x v^{a-1}[F_n(x) - F_n(v)]dv \\ &\leq a^{-1} \int_0^x v^{a-1}[1 - F_n(v)]dv \leq a^{-1} \int_0^x v^{a-1}[1 - F_n(v)]dv \\ &\leq ca^{-1} \int_0^x v^{a-\tau} dv \leq Cx^{a+1-\tau}.\end{aligned} \tag{4.26}$$

For the lower bound, we first assume that $x \leq n^{1/(\tau-1)}$ and use that

$$\begin{aligned}\mathbb{E}[W_n^a \mathbb{1}_{\{W_n \leq x\}}] &= a^{-1} \int_0^x v^{a-1}[F_n(x) - F_n(v)]dv \geq a^{-1} \int_0^{\varepsilon x} v^{a-1}[F_n(x) - F_n(v)]dv \\ &= a^{-1} \int_0^{x/2} v^{a-1}[1 - F_n(v)]dv - a^{-1}[1 - F_n(x)] \int_0^{x/2} v^{a-1} dv \\ &\geq a^{-1} \int_0^{\varepsilon x} v^{a-1}[1 - F(v)]dv - a^{-1} \int_0^{\varepsilon x} v^{a-1}/n dv - a^{-1}[1 - F(x)] \int_0^{\varepsilon x} v^{a-1} dv \\ &\geq C(\varepsilon x)^{a+1-\tau} - C(\varepsilon x)^a/n - Cx^{-(\tau-1)}(\varepsilon x)^a \\ &= C\varepsilon^{a+1-\tau} x^{a+1-\tau}\Big(1 - \varepsilon^{\tau-1}\big(x/n^{1/(\tau-1)}\big)^{\tau-1} - \varepsilon^{\tau-1}\Big) \geq C_2 x^{a+1-\tau},\end{aligned} \tag{4.27}$$

when we take $\varepsilon > 0$ sufficiently small, and we use that $x \leq n^{1/(\tau-1)}$. When $x \geq n^{1/(\tau-1)}$, then we can use that $w_1 \geq cn^{1/(\tau-1)}$, so that

$$\mathbb{E}[W_n^a \mathbb{1}_{\{W_n \leq x\}}] \geq w_1^a/n \geq (cn^{1/(\tau-1)})^a/n = C_2\big(n^{1/(\tau-1)}\big)^{a+1-\tau}. \tag{4.28}$$

□

4.2 Connection to mixed Poisson branching processes

In this section, we discuss the relation between our Poissonian random graph and mixed Poisson branching processes due to Norros and Reittu [36].

Stochastic domination of neighborhoods by a branching process. We shall dominate the cluster of a vertex in the Norros-Reittu model by the total progeny of a two-stage branching processes with mixed Poissonian offspring.

In order to describe this relation, we consider the neighborhood shells of a uniformly chosen vertex $V \in [n]$, i.e., all vertices on a fixed graph distance of vertex V. More precisely,

$$\partial\mathcal{N}_0 = \{V\} \qquad \text{and} \qquad \partial\mathcal{N}_l = \{1 \le j \le n : d(V,j) = l\}, \tag{4.29}$$

where $d(i,j)$ denotes the graph distance between vertices i and j in $\mathrm{NR}_n(\mathbf{w})$, i.e., the minimum number of edges in a path between the vertices i and j. Define the set of vertices reachable in at most j steps from vertex V by

$$\mathcal{N}_l = \{1 \le j \le n : d(V,j) \le l\} = \bigcup_{k=0}^{l} \partial\mathcal{N}_k. \tag{4.30}$$

The main idea is that we can explicitly view the neighborhood sizes $\{|\mathcal{N}_l|\}_{l=0}^{\infty}$ as a marked Poisson branching process where repeated vertices are thinned. The NR-process is a marked two-stage branching process denoted by $\{Z_l, \mathbf{M}\}_{l\ge 0}$, where Z_l denotes the number of individuals of generation l, and where the vector

$$\mathbf{M} = (M_{l,1}, M_{l,2}, \dots, M_{l,Z_l}) \in [n]^{Z_l}, \tag{4.31}$$

denotes the marks of the individuals in generation l. These marks shall label to which vertex in $\mathrm{NR}_n(\mathbf{w})$ an individual in the branching process corresponds. We now give a more precise definition of the NR-process and describe its connection with $\mathrm{NR}_n(\mathbf{w})$.

We define $Z_0 = 1$ and take $M_{0,1}$ uniformly from the set $[n]$, corresponding to the choice of A_1, which is uniformly over all the vertices. The offspring of an individual with mark $m \in [n]$ is as follows: the total number of children has a Poisson distribution with parameter w_m, of which, for each $i \in [n]$, a Poisson distributed number with parameter

$$\frac{w_i w_m}{l_n}, \tag{4.32}$$

bears mark i, independently of the other individuals. Since

$$\sum_{i=1}^{n} \frac{w_i w_m}{l_n} = \frac{w_m}{l_n} \sum_{i=1}^{n} w_i = w_m, \tag{4.33}$$

and sums of independent Poisson random variables are again Poissonian, we may take the number of children with different marks mutually independent. As a result of this definition, the marks of the children of an individual in $\{Z_l, \mathbf{M}\}_{l\ge 0}$ can be seen as independent realizations of a random variable M, with distribution

$$\mathbb{P}(M = m) = \frac{w_m}{l_n}, \quad 1 \le m \le n, \tag{4.34}$$

and, consequently,

$$\mathbb{P}(w_M \le x) = \sum_{m=1}^{n} \mathbb{1}_{\{w_m \le x\}} \mathbb{P}(M = m) = \frac{1}{l_n} \sum_{m=1}^{n} w_m \mathbb{1}_{\{w_m \le x\}} = \mathbb{P}(W_n^* \le x), \tag{4.35}$$

where, for a non-negative random variable W, we let W^* denote its size-biased distribution given by

$$\mathbb{P}(W^* \le x) = \frac{\mathbb{E}[W \mathbb{1}_{\{W \le x\}}]}{\mathbb{E}[W]}. \tag{4.36}$$

The above shows that size-biasing naturally arises in the context of inhomogeneous random graphs.

For the definition of the $\underline{\text{NR}}$-process, we start with a copy of the NR-process $\{Z_l, \mathbf{M}\}_{l \ge 0}$, and reduce this process generation by generation, i.e., in the order

$$M_{0,1}, M_{1,1}, M_{1,2}, \dots, M_{1,Z_1}, M_{2,1}, \dots \tag{4.37}$$

by discarding each individual and all its descendants whose mark has been chosen before, i.e., if $M_{i_1,j_1} = M_{i_2,j_2}$ for some i_1, j_1 which has appeared before i_2, j_2, then the individual corresponding to i_2, j_2, as well as its entire offspring, is erased. The process obtained in this way is called the $\underline{\text{NR}}$-process and is denoted by the sequence $\{\underline{Z}_l, \underline{\mathbf{M}}_l\}_{l \ge 0}$. One of the main results of [36, Proposition 3.1] is the fact that the distribution of $\{\partial \mathcal{N}_l\}_{l \ge 0}$ is equal to that of :

Proposition 4.4 (Neighborhoods are a thinned marked branching process)**.** *Let $\{\underline{Z}_l, \underline{\mathbf{M}}_l\}_{l \ge 0}$ be the $\underline{\text{NR}}$-process and let $\underline{\mathbf{M}}_l$ be the set of marks in the l^{th} generation, then the sequence of sets $\{\underline{\mathbf{M}}_l\}_{l \ge 0}$ has the same distribution as the sequence $\{\partial \mathcal{N}_l\}_{l \ge 0}$ given by* (4.29).

As a consequence of Proposition 4.4, we can couple the $\underline{\text{NR}}$-process to the neighborhood shells of a uniformly chosen vertex $V \in [n]$, i.e., all vertices on a fixed graph distance of V, see (4.29) and note that $V \sim M_{0,1}$. Thus, using Proposition 4.4, we can couple the expansion of the neighborhood shells and the $\underline{\text{NR}}$-process in such a way that

$$\underline{\mathbf{M}}_l = \partial \mathcal{N}_l \qquad \text{and} \qquad \underline{Z}_l = |\partial \mathcal{N}_l|, \quad l \ge 0. \tag{4.38}$$

Furthermore, we see that an individual with mark m in the $\underline{\text{NR}}$-process is identified with vertex m in the graph $\text{NR}_n(\mathbf{w})$ whose capacity is w_m.

For given weights $\{w_i\}_{i=1}^n$, we now describe the distribution of the marked Poisson process. For this, we note that since the marks are mutually *independent*, the marked Poisson process is a branching process, see [36, Proposition 3.2] and the discussion proceeding it. The offspring distribution $f^{(n)}$ of Z_1, i.e., the first generation of $\{Z_l\}_{l \ge 0}$, is given by

$$f_k^{(n)} = \mathbb{P}\big(\text{Poi}(W_n) = k\big) = \frac{1}{n} \sum_{m=1}^{n} \mathrm{e}^{-w_m} \frac{w_m^k}{k!}, \tag{4.39}$$

for $k \ge 0$. Recall that individuals in the second and further generations have a random mark distributed as an independent copy of M given by (4.34). Hence, if we denote the offspring distribution of the second and further generations by $g^{(n)}$, then we obtain

$$g_k^{(n)} = \mathbb{P}\big(\text{Poi}(W_n^*) = k\big) = \frac{1}{l_n} \sum_{m=1}^{n} \mathrm{e}^{-w_m} \frac{w_m^{k+1}}{k!}, \tag{4.40}$$

for $k \ge 0$. This describes a stochastic upper bound on the neighborhood shells of a uniform vertex $V \in [n]$ in the Norros-Reittu model $\text{NR}_n(\mathbf{w})$ in terms of a normal branching process.

Otter-Dwass formula for the branching process total progeny. Our proofs make crucial use of the *Otter-Dwass formula*, which describes the distribution of the total progeny of a branching process. This result is sometimes called the *random-walk hitting time theorem*, see [16] for the special case when the branching process starts with a single individual and [37] for the more general case. See [24] for a simple proof based on induction. The Otter-Dwass formula is an extremely useful result to study the cluster sizes in random graphs.

Lemma 4.5 (Otter-Dwass formula)**.** *Let $X_1, X_2, X_3, \ldots$ be i.i.d. random variables distributed as Z. Let $\mathbb{P}_m$ denote the Galton-Watson process measure started from m initial individuals. For all $n \in \mathbb{N}$,*

$$\mathbb{P}_m(T = k) = \frac{m}{k}\mathbb{P}(\sum_{i=1}^{k} X_i = k - m). \tag{4.41}$$

The survival probability of near-critical mixed Poisson branching processes. We shall also need bounds on the survival probability of near-critical mixed Poisson branching processes.

Lemma 4.6 (Survival probability of near-critical mixed Poisson branching processes.)**.** *Let ρ_n be the survival probability of a mixed Poisson branching process with mixing distribution W_n^*. Assume that $\varepsilon_n = \mathbb{E}[W_n^*] - 1 = \nu_n - 1 \geq 0$ and $\varepsilon_n = o(1)$. When (1.12) holds for some $\tau > 4$, then there exists a constant $c > 0$ such that*

$$\rho_n \leq c\varepsilon_n. \tag{4.42}$$

When (1.16) holds for some $\tau \in (3, 4)$, then

$$\rho_n \leq c\big(\varepsilon_n^{1/(\tau-3)} \vee n^{-1/(\tau-1)}\big). \tag{4.43}$$

Proof. We use that the survival probability ρ of a branching process with offspring distribution X satisfies

$$1 - \rho = \mathbb{E}[(1-\rho)^X]. \tag{4.44}$$

In our case, X has a mixed Poisson distribution with mixing distribution W_n^*, so that

$$1 - \rho_n = \mathbb{E}[(1-\rho_n)^X] = \mathbb{E}[\mathrm{e}^{-\rho_n W_n^*}]. \tag{4.45}$$

Now, we use that $\mathrm{e}^{-x} \geq 1 - x$ when $x \geq 1/2$ and $\mathrm{e}^{-x} \geq 1 - x + x^2/4$ when $x \leq 1/2$, to arrive at

$$1 - \rho_n \geq 1 - \rho_n\mathbb{E}[W_n^*] + \frac{\rho_n^2}{4}\mathbb{E}\big[(W_n^*)^2\mathbb{1}_{\{\rho_n W_n^* \leq 1/2\}}\big]. \tag{4.46}$$

Rearranging terms and using that $\mathbb{E}[W_n^*] = \nu_n$, we obtain

$$\rho_n\mathbb{E}\big[(W_n^*)^2\mathbb{1}_{\{\rho_n W_n^* \leq 1/2\}}\big] \leq 4(\nu_n - 1) = 4\varepsilon_n. \tag{4.47}$$

Now, when (1.12) holds for some $\tau > 4$, then

$$\mathbb{E}\big[(W_n^*)^2\mathbb{1}_{\{\rho_n W_n^* \leq 1/2\}}\big] = \mathbb{E}[(W^*)^2](1 + o(1)) = \frac{\mathbb{E}[W^3]}{\mathbb{E}[W]}(1 + o(1)), \tag{4.48}$$

so that $\rho_n \leq c\varepsilon_n$ for some constant $c > 0$. When, on the other hand, (1.16) holds for some $\tau \in (3, 4)$, then, either $\rho_n \leq 2n^{-1/(\tau-1)}$, or $\rho_n \geq 2n^{-1/(\tau-1)}$, in which case we may apply Lemma 4.3(ii) to obtain

$$\mathbb{E}\big[(W_n^*)^2\mathbb{1}_{\{\rho_n W_n^* \leq 1/2\}}\big] \geq c\rho_n^{\tau-4}, \tag{4.49}$$

so that

$$c\rho_n^{\tau-3} \leq 4\varepsilon_n, \tag{4.50}$$

which proves (4.43). This proves that $\rho_n \leq c\big(\varepsilon_n^{1/(\tau-3)} \vee n^{-1/(\tau-1)}\big)$. □

5 An upper bound on the cluster tail

In this section, we shall prove an upper bound on the tail probabilities of critical clusters. In its statement, we denote

$$\delta = \begin{cases} 2 & \text{for } \tau > 4, \\ \tau - 2 & \text{for } \tau \in (3,4), \end{cases} \qquad \alpha = \begin{cases} 1 & \text{for } \tau > 4, \\ \tau - 3 & \text{for } \tau \in (3,4). \end{cases} \tag{5.1}$$

We shall prove the upper bounds in Theorems 1.1 –1.4 all at once.

Proposition 5.1 (An upper bound on the cluster tail)**.** *Fix* $\mathrm{NR}_n(\mathbf{w})$ *with* $\mathbf{w} = \{w_i\}_{i=1}^n$ *as in* (1.3)*, and assume that the distribution function* F *in* (1.3) *satisfies* $\nu = 1$*. Fix* $\varepsilon_n = o(1)$*, and let* $\tilde{\mathbf{w}}$ *be defined as in* (1.20)*. Assume that* (1.12) *holds for some* $\tau > 4$*, or that* (1.14) *holds for some* $\tau \in (3,4)$*, and fix* ε_n *such that* $|\varepsilon_n| \leq \Lambda n^{-\alpha/(\delta+1)}$ *for some* $\Lambda > 0$*, where we recall the definitions of* δ *and* α *in* (5.1)*. Then, there exists an* $a_1 > 0$ *such that for all* $k \geq 1$ *and for* V *a uniform vertex in* $[n]$*, there exists a constant* $a_1 > 0$ *such that*

$$\mathbb{P}(|C(V)| \geq k) \leq a_1\big(k^{-1/\delta} + \big(\varepsilon_n \vee n^{-\alpha/(\delta+1)}\big)^{1/\alpha}\big). \tag{5.2}$$

By Proposition 3.1, Proposition 5.1 proves the upper bounds on $|C_{\max}|$ in Theorems 1.1 –1.4. The remainder of this section will be devoted to the proof of Proposition 5.1.

Dominating the two-stage branching process by an ordinary branching process. Using the description of the neighborhood shells in Proposition 4.4, we arrive at the bound, valid for all k,

$$\mathbb{P}(|C(V)| \geq k) \leq \mathbb{P}(T^{(2)} \geq k), \tag{5.3}$$

where $T^{(2)}$ is the total progeny of the two-stage branching process described below (4.40). Unfortunately, the Otter-Dwass formula (Lemma 4.5) is not valid as is for two-stage branching processes, and we first establish that, for every $k \geq 0$,

$$\mathbb{P}(T^{(2)} \geq k) \leq \mathbb{P}(T \geq k), \tag{5.4}$$

where T is the total progeny of a mixed Poisson branching process with offspring distribution $g^{(n)} = \{g_k^{(n)}\}_{k=0}^\infty$ defined in (4.40). The bound in (5.4) is equivalent to the fact that $T^{(2)} \preceq T$, where $X \preceq Y$ means that X is stochastically smaller than Y. Since the distributions of $T^{(2)}$ and T agree except for the offspring of the root, we have that $T^{(2)} \preceq T$ follows when $Z_1^{\mathrm{un}} \preceq Z_1^{\mathrm{sb}}$, where Z_1^{un} has mixed Poisson distribution with mixing distribution W_n, and where Z_1^{sb} has a mixed Poisson distribution with mixing random variable W_n^*. For two mixed Poisson random variables X, Y with mixing random variables W_X and W_Y, respectively, $X \preceq Y$ follows when $W_X \preceq W_Y$. The proof of (5.4) is completed by noting that, for any non-negative random variable W, and for W^* its size-biased version, we have $W \preceq W^*$.

The total progeny of our mixed Poisson branching process. By (5.3)–(5.4),

$$\begin{aligned} \mathbb{P}(|C(V)| \geq k) \leq \mathbb{P}(T \geq k) &= \mathbb{P}(T = \infty) + \sum_{l=k}^\infty \mathbb{P}(T = l) \\ &= \mathbb{P}(T = \infty) + \sum_{l=k}^\infty \frac{1}{l}\mathbb{P}\Big(\sum_{i=1}^l X_i = l-1\Big), \end{aligned} \tag{5.5}$$

where the last formula follows from Lemma 4.5 for $m = 1$, and where $\{X_i\}_{i=1}^{\infty}$ is an i.i.d. sequence with a mixed Poisson distribution with mixing random variable W_n^*. In the following lemma, we shall investigate $\mathbb{P}(\sum_{i=1}^{l} X_i = l-1)$:

Proposition 5.2 (Upper bound on probability mass function of $\sum_{i=1}^{l} X_i$). *Let $\{X_i\}_{i=1}^{\infty}$ be an i.i.d. sequence with a mixed Poisson distribution with mixing random variable $\tilde{W}_n^* = (1+\varepsilon_n)W_n^*$, where W_n^* is defined in (4.35). Under the assumptions of Proposition 5.1, there exists an $\tilde{a}_1 > 0$ such that for all $l \geq n^{\delta/(1+\delta)}$, such that, for n sufficiently large,*

$$\mathbb{P}\Big(\sum_{i=1}^{l} X_i = l-1\Big) \leq \tilde{a}_1\Big(l^{-1/\delta} + \big(n^{(\tau-4)/2(\tau-1)} \wedge 1\big)l^{-1/2}\Big). \tag{5.6}$$

where $\delta > 0$ is defined in (5.1).

Proof. We rewrite, using the Fourier inversion theorem, and writing $\phi_n(t) = \mathbb{E}[\mathrm{e}^{itX_1}]$ for the characteristic function of the random variables $\{X_i\}_{i=1}^{\infty}$,

$$\mathbb{P}\Big(\sum_{i=1}^{l} X_i = l-1\Big) = \int_{[-\pi,\pi]} \mathrm{e}^{-i(l-1)t}\phi_n(t)^l \frac{dt}{2\pi}, \tag{5.7}$$

so that

$$\mathbb{P}\Big(\sum_{i=1}^{l} X_i = l-1\Big) \leq \int_{[-\pi,\pi]} |\phi_n(t)|^l \frac{dt}{2\pi}, \tag{5.8}$$

Since X_1 has a mixed Poisson distribution with mixing random variable $\tilde{W}_n^*$, we obtain

$$\phi_n(t) = \mathbb{E}[\mathrm{e}^{itX_1}] = \mathbb{E}[\mathrm{e}^{(\mathrm{e}^{it}-1)\tilde{W}_n^*}]. \tag{5.9}$$

By dominated convergence and the weak convergence of $\tilde{W}_n^*$ to W^*, for every $t \in [-\pi,\pi]$,

$$\lim_{n\to\infty} \phi_n(t) = \phi(t) = \mathbb{E}[\mathrm{e}^{W^*(\mathrm{e}^{it}-1)}]. \tag{5.10}$$

Since, further,

$$|\phi_n'(t)| = \big|\mathbb{E}\big[W_n^*\mathrm{e}^{it}\mathrm{e}^{(\mathrm{e}^{it}-1)\tilde{W}_n^*}\big]\big| \leq \mathbb{E}[\tilde{W}_n^*] = \tilde{\nu}_n = (1+\varepsilon_n)\nu_n = 1+o(1), \tag{5.11}$$

which is uniformly bounded, the convergence in (5.10) is uniform for all $t \in [-\pi,\pi]$. Finally, a mixed Poisson random variable for which the mixing distribution is not degenerated at 0 satisfies that for every $\eta > 0$, there exists $\varepsilon > 0$ such that $|\phi(t)| < 1-2\varepsilon$ for all $|t| > \eta$. Therefore, uniformly for sufficiently large n, for every $\eta > 0$, there exists $\varepsilon > 0$ such that $|\phi_n(t)| < 1-\varepsilon$ for all $|t| > \eta$. Thus,

$$\int_{[-\pi,\pi]} |\phi_n(t)|^l \frac{dt}{2\pi} \leq (1-\varepsilon)^l + \int_{[-\eta,\eta]} |\phi_n(t)|^l \frac{dt}{2\pi}. \tag{5.12}$$

We start by deriving the bound when $\tau > 4$, by bounding

$$|\phi_n(t)| \leq \mathbb{E}[\mathrm{e}^{-\tilde{W}_n^*[1-\cos(t)]}]. \tag{5.13}$$

Now using that, uniformly for $t \in [-\pi,\pi]$, there exists an $a > 0$ such that

$$1-\cos(t) \geq at^2, \tag{5.14}$$

and, for $x \leq 1$, $\mathrm{e}^{-x} \leq 1 - x/2$, we arrive at

$$\begin{aligned} |\phi_n(t)| &\leq \mathbb{E}[\mathrm{e}^{-a\tilde{W}_n^* t^2}] \leq \mathbb{E}\big[(1 - a\tilde{W}_n^* t^2/2)\mathbb{1}_{\{a\tilde{W}_n^* t^2 \leq 1\}}\big] + \mathbb{E}\big[\mathbb{1}_{\{a\tilde{W}_n^* t^2 > 1\}}\big] \\ &= 1 - at^2\mathbb{E}[\tilde{W}_n^*] + \mathbb{E}\big[\mathbb{1}_{\{a\tilde{W}_n^* t^2 > 1\}}(1 + a\tilde{W}_n^* t^2/2)\big]. \end{aligned} \tag{5.15}$$

Further bounding, using Lemma 4.3 and $\tau > 4$,

$$\mathbb{E}\big[\mathbb{1}_{\{a\tilde{W}_n^* t^2 > 1\}}(1 + a\tilde{W}_n^* t^2/2)\big] = o(t^2), \tag{5.16}$$

we finally obtain that, uniformly for $t \in [-\eta, \eta]$, there exists a $b > 0$ such that

$$|\phi_n(t)| \leq 1 - bt^2. \tag{5.17}$$

Thus, there exists a constant $a_2 > 0$ such that

$$\int_{[-\pi,\pi]} |\phi_n(t)|^l \frac{dt}{2\pi} \leq (1-\varepsilon)^l + \int_{[-\eta,\eta]} (1 - bt^2)^l \frac{dt}{2\pi} \leq \frac{a_2}{l^{1/2}}, \tag{5.18}$$

which proves (5.6) for $\delta = 2$ for all $\tau > 3$.

In order to prove (5.6) for $\delta = \tau - 2 < 2$ for $\tau \in (3, 4)$, we have to obtain a sharper bound on $|\phi_n(t)|$. For this, we identify

$$\phi_n(t) = \mathrm{Re}(\phi_n(t)) + i\mathrm{Im}(\phi_n(t)), \tag{5.19}$$

where

$$\mathrm{Re}(\phi_n(t)) = \mathbb{E}\big[\cos(\tilde{W}_n^* \sin(t))\mathrm{e}^{-\tilde{W}_n^*[1-\cos(t)]}\big], \tag{5.20}$$

$$\mathrm{Im}(\phi_n(t)) = \mathbb{E}\big[\sin(\tilde{W}_n^* \sin(t))\mathrm{e}^{-\tilde{W}_n^*[1-\cos(t)]}\big], \tag{5.21}$$

so that

$$|\phi_n(t)|^2 = \mathrm{Re}(\phi_n(t))^2 + \mathrm{Im}(\phi_n(t))^2. \tag{5.22}$$

We start by upper bounding $|\mathrm{Im}(\phi_n(t))|$, by using that for all $t \in \mathbb{R}$,

$$|\sin(t)| \leq |t|, \tag{5.23}$$

so that, since $\tilde{\nu}_n = 1 + o(1)$,

$$|\mathrm{Im}(\phi_n(t))| \leq |t|\mathbb{E}[\tilde{W}_n^*] = |t|(1 + o(1)). \tag{5.24}$$

Further,

$$\mathrm{Re}(\phi_n(t)) = 1 - \mathbb{E}[1 - \cos(\tilde{W}_n^* \sin(t))] + \mathbb{E}\big[\cos(\tilde{W}_n^* \sin(t))[\mathrm{e}^{-\tilde{W}_n^*[1-\cos(t)]} - 1]\big]. \tag{5.25}$$

By the uniform convergence in (5.10) and the fact that, for $\eta > 0$ small enough, $\mathrm{Re}(\phi(t)) \geq 0$, we only need to derive an upper bound on $\mathrm{Re}(\phi_n(t))$ rather than on $|\mathrm{Re}(\phi_n(t))|$. For this, we use that $1 - \mathrm{e}^{-x} \leq x$ and $1 - \cos(t) \leq t^2/2$, to bound

$$\big|\mathbb{E}\big[\cos(\tilde{W}_n^* \sin(t))[\mathrm{e}^{-\tilde{W}_n^*[1-\cos(t)]} - 1]\big]\big| \leq \mathbb{E}\big[1 - \mathrm{e}^{-\tilde{W}_n^*[1-\cos(t)]}\big] \leq [1 - \cos(t)]\mathbb{E}[\tilde{W}_n^*] \leq \tilde{\nu}_n t^2/2. \tag{5.26}$$

Further, using (5.14) whenever $\tilde{W}_n^*|t| \leq 1$, so that also $\tilde{W}_n^*|\sin(t)| \leq \tilde{W}_n^*|t| \leq 1$, and $1 - \cos(\tilde{W}_n^* \sin(t)) \geq 0$ otherwise, we obtain

$$\mathrm{Re}(\phi_n(t)) \leq 1 - a\sin(t)^2\mathbb{E}\big[\big(\tilde{W}_n^*\big)^2\mathbb{1}_{\{\tilde{W}_n^*|t|\leq 1\}}\big] + \tilde{\nu}_n t^2/2 = 1 - at^2\frac{\mathbb{E}\big[W_n^3\mathbb{1}_{\{W_n|t|\leq 1\}}\big]}{\mathbb{E}[W_n]} + \tilde{\nu}_n t^2/2. \quad (5.27)$$

By Lemma 4.3, we have that

$$\mathbb{E}\big[W_n^3\mathbb{1}_{\{W_n|t|\leq 1\}}\big] \geq C_1\big(|t| \vee n^{-1/(\tau-1)}\big)^{\tau-4}. \quad (5.28)$$

Combining (5.27) with (5.28), we obtain that, uniformly in $|t| \leq \eta$ for some small enough $\eta > 0$,

$$\mathrm{Re}(\phi_n(t)) \leq \begin{cases} 1 - 2a_{\mathrm{ub}}|t|^{\tau-2} & \text{for } |t| \geq n^{-1/(\tau-1)}, \\ 1 - 2a_{\mathrm{ub}}t^2 n^{(4-\tau)/(\tau-1)} & \text{for } |t| \leq n^{-1/(\tau-1)}. \end{cases} \quad (5.29)$$

which, combined with (5.22) and (5.24) shows that, for $|t| \leq \eta$ and $\eta > 0$ sufficiently small,

$$|\phi_n(t)| \leq \begin{cases} \mathrm{e}^{-a_{\mathrm{ub}}t^{2-\tau}/2} & \text{for } |t| \geq n^{-1/(\tau-1)}, \\ \mathrm{e}^{-a_{\mathrm{ub}}t^2 n^{(4-\tau)/(\tau-1)}/2} & \text{for } |t| \leq n^{-1/(\tau-1)}. \end{cases} \quad (5.30)$$

Thus, there exists a constant $\tilde{a}_1 > 0$ such that

$$\begin{aligned}\int_{[-\pi,\pi]} |\phi_n(t)|^l \frac{dt}{2\pi} &\leq (1-\varepsilon)^l + \int_{[-\eta,\eta]} \mathrm{e}^{-la_{\mathrm{ub}}|t|^{2-\tau}/2}dt + \int_{[-n^{-1/(\tau-1)},n^{-1/(\tau-1)}]} \mathrm{e}^{-la_{\mathrm{ub}}t^2n^{(4-\tau)/(\tau-1)}/2}dt \\ &\leq \frac{\tilde{a}_1}{l^{1/(\tau-2)}} + \frac{\tilde{a}_1 n^{(\tau-4)/2(\tau-1)}}{\sqrt{l}} = \tilde{a}_1\big(l^{-1/\delta} + n^{(\tau-4)/2(\tau-1)}l^{-1/2}\big), \end{aligned} \quad (5.31)$$

which proves (5.6) for $\delta = \tau - 2$ for when $\tau \in (3, 4)$. □

Now we are ready to prove Proposition 5.1:

Proof of Proposition 5.1. By (5.5) and Lemma 4.6,

$$\begin{aligned}\mathbb{P}(|C(V)| \geq k) &\leq c(\varepsilon_n^{1/\alpha} \vee n^{-1/(\tau-1)}) + \tilde{a}_1\sum_{l=k}^{\infty}\frac{1}{l^{(\delta+1)/\delta}} + \tilde{a}_1\sum_{l=k}^{\infty} l^{-3/2} \\ &\leq c\big(\varepsilon_n \vee n^{\alpha/(\tau-1)}\big)^{1/\alpha} + \frac{\tilde{a}_1\delta}{k^{1/\delta}} + \big(n^{(\tau-4)/2(\tau-1)} \wedge 1\big)k^{-1/2}, \end{aligned} \quad (5.32)$$

the final term being absent for $\tau > 4$. The proof is completed by noting that, for $k \geq n^{\delta/(\delta+1)} = n^{(\tau-2)/(\tau-1)}$

$$n^{(\tau-4)/2(\tau-1)}k^{-1/2} \leq n^{(\tau-4)/2(\tau-1)}n^{(\tau-2)/2(\tau-1)} = n^{-1/(\tau-1)}. \quad (5.33)$$

Thus, the last term in (5.32) can be incorporated into the first term, for the appropriate choice of a_1. This proves the claim in (5.2). □

5.1 An upper bound on the expected cluster size

We now slightly extend the above computation to prove a bound on the expected cluster size, which we shall need in order to apply Proposition 3.3 (recall (3.19)).

Proposition 5.3 (An upper bound on the expected cluster size)**.** *Assuming that* $\tilde{\nu}_n \leq 1 - cn^{-\alpha/(\delta+1)}$*, there exists a* $K > 0$ *such that for* V *a uniform vertex in* $[n]$*,*

$$\mathbb{E}[|C(V)|] \leq Kn^{\alpha/(\delta+1)}. \tag{5.34}$$

Proposition 5.3 proves (3.19) in Proposition 3.3 when we note that, by (5.1), $\alpha = \delta - 1$.

Proof. We note that by (5.3) and (5.4),

$$\mathbb{E}[|C(V)|] \leq \mathbb{E}[T^{(2)}] \leq \mathbb{E}[T] = \frac{1}{1-\tilde{\nu}_n} \leq c^{-1}n^{\alpha/(\delta+1)} = Kn^{\alpha/(\delta+1)}, \tag{5.35}$$

when $K = 1/c$. This completes the proof of Proposition 5.3. □

6 A lower bound on the cluster tail

In this section, we prove a lower bound on the cluster tail. We shall pick $\tilde{w}_i$ as in (1.18), and assume the bounds on ε_n in Theorem 1.3 for $\tau > 4$ and the assumptions of Theorem 1.4 for $\tau \in (3, 4)$. The main result is the following proposition:

Proposition 6.1 (A lower bound on the cluster tail)**.** *Fix* $\mathrm{NR}_n(\mathbf{w})$ *with* $\mathbf{w} = \{w_i\}_{i=1}^n$ *as in* (1.3)*, and assume that the distribution function* F *in* (1.3) *satisfies* $\nu = 1$*. Fix* $\varepsilon_n = o(1)$*, and let* $\tilde{\mathbf{w}}$ *be defined as in* (1.20)*. Assume that* (1.12) *holds for some* $\tau > 4$*, or that* (1.16) *holds for some* $\tau \in (3, 4)$*, and fix* ε_n *such that* $|\varepsilon_n| \leq \Lambda n^{-\alpha/(\delta+1)}$ *for some* $\Lambda > 0$*, where we recall the definitions of* δ *and* α *in* (5.1)*. Then, there exists an* $a_1 > 0$ *such that for all* $k \geq 1$ *and for* V *a uniform vertex in* $[n]$*, there exists a constant* $a_2 > 0$ *such that*

$$\mathbb{P}(|C(V)| \geq k) \geq \frac{a_2}{k^{1/\delta}}. \tag{6.1}$$

The key ingredient in the proof of Proposition 6.1 is again the coupling to branching processes. Indeed, let $T^{(2)}$ denote the total progeny of the branching process $\{Z_l\}_{l=0}^{\infty}$ defined in Section 4.2. Note the explicit coupling between the cluster size $|C(V)|$ and T described there. We can then bound

$$\mathbb{P}(|C(V)| \geq k) \geq \mathbb{P}(T^{(2)} \geq 2k, |C(V)| \geq k) = \mathbb{P}(T^{(2)} \geq 2k) - \mathbb{P}(T^{(2)} \geq 2k, |C(V)| < k). \tag{6.2}$$

The following lemmas contain bounds on both contributions:

Lemma 6.2 (Lower bound tail total progeny)**.** *Under the assumptions of Proposition 6.1, there exists a constant* $a_2 > 0$ *such that*

$$\mathbb{P}(T^{(2)} \geq k) \geq \frac{2a_2}{k^{1/\delta}}. \tag{6.3}$$

Lemma 6.3 (Upper bound cluster tail coupling)**.** *Under the assumptions of Theorem 1.1 for* $\tau > 4$ *and the assumptions of Theorem 1.2 for* $\tau \in (3, 4)$*, for all* $k \leq \varepsilon n^{(\delta-1)/(\delta+1)}$*, there exists constants* $c, p > 0$ *such that*

$$\mathbb{P}(T^{(2)} \geq 2k, |C(V)| < k) \leq \frac{c\varepsilon^p}{k^{1/\delta}}. \tag{6.4}$$

Proof of Proposition 6.1 subject to Lemmas 6.2-6.3. Recall (6.2), and substitute the bounds in Lemmas 6.2-6.3 to conclude that, for all $\varepsilon > 0$ sufficiently small,

$$\mathbb{P}(|C(V)| \geq k) \geq \frac{2a_2}{(2k)^{1/\delta}} - \frac{c\varepsilon^p}{k^{1/\delta}} \geq \frac{a_2}{k^{1/\delta}}, \tag{6.5}$$

where $\varepsilon > 0$ is so small that $2^{1-1/\delta}a_2 - c\varepsilon^p \geq a_2$. This is possible, since $\delta > 1$. □

Proof of Lemma 6.2. We start by noting that

$$\mathbb{P}(T^{(2)} \geq k) \geq \mathbb{P}(T^{(2)} \geq k, Z_1 = 1) = \mathbb{P}(T \geq k-1)\mathbb{P}(Z_1 = 1) \geq \mathbb{P}(T \geq k)\mathbb{P}(Z_1 = 1), \tag{6.6}$$

where T is a standard branching process with offspring distribution $g^{(n)}$ in (4.40). Note that, by (4.39),

$$\mathbb{P}(Z_1 = 1) = f_1^{(n)} = \mathbb{E}[W_n e^{-W_n}] = \mathbb{E}[W \mathrm{e}^{-W}] + o(1), \tag{6.7}$$

which remains strictly positive. Thus, it suffices to prove a lower bound on $\mathbb{P}(T \geq k-1)$. For this, we bound

$$\begin{aligned}\mathbb{P}(T \geq k) &\geq \sum_{l=k}^{\infty} \mathbb{P}(T = l) = \sum_{l=k-1}^{\infty} \frac{1}{l}\mathbb{P}\big(\sum_{i=1}^{l} X_i = l-1\big)\\ &\geq \sum_{l=k}^{\infty} \frac{1}{l}\mathbb{P}\big(\sum_{i=1}^{l} X_i = l-1\big) \geq \sum_{l=k}^{2k} \frac{1}{l}\mathbb{P}\big(\sum_{i=1}^{l} X_i = l-1\big).\end{aligned} \tag{6.8}$$

We start by studying the case $\tau > 4$. Denote $Y_l = W^*_{n,l}$, where $\{W^*_{n,s}\}_{s=1}^{\infty}$ are i.i.d. copies of the random variable W^*_n. Then, conditionally on $\{W^*_{n,s}\}_{s=1}^{\infty}$, the random variable $\sum_{i=1}^{l} X_i$ has a Poisson distribution with parameter $l\bar{Y}_l$, where $\bar{Y}_l = \frac{1}{l}\sum_{i=1}^{l} Y_l$. Thus,

$$\mathbb{P}\big(\sum_{i=1}^{l} X_i = l-1\big) = \mathbb{E}\Big[\mathbb{P}\big(\sum_{i=1}^{l} X_i = l-1 \mid \bar{Y}_l\big)\Big] = \mathbb{E}\Big[\frac{(l\bar{Y}_l)^{l-1}}{(l-1)!}\mathrm{e}^{-l\bar{Y}_l}\Big] = \mathbb{E}\Big[\frac{1}{\bar{Y}_l}\frac{(l\bar{Y}_l)^{l}}{l!}\mathrm{e}^{-l\bar{Y}_l}\Big]. \tag{6.9}$$

Now, by Stirling's formula, $l! \leq 3l^{l+1/2}\mathrm{e}^{-l}$, so that

$$\mathbb{P}(\sum_{i=1}^{l} X_i = l-1) \geq \frac{1}{3\sqrt{l}}\mathbb{E}\Big[\frac{1}{\bar{Y}_l}\big(e\bar{Y}_l\big)^{l}\mathrm{e}^{-l\bar{Y}_l}\Big] \geq \frac{1}{3(1-\eta)l^{3/2}}\mathbb{E}\Big[\mathrm{e}^{-lI(\bar{Y}_l)}\mathbb{1}_{\{\bar{Y}_l\in[1-\eta,1+\eta]\}}\Big]. \tag{6.10}$$

where we denote

$$I(\lambda) = \lambda - 1 - \log \lambda. \tag{6.11}$$

By a Taylor expansion, we see that, for all $\lambda \in [1-\eta, 1+\eta]$,

$$I(\lambda) = \frac{1}{2}(\lambda-1)^2 + O(|\lambda-1|^3) \leq (\lambda-1)^2, \tag{6.12}$$

so that we arrive at

$$\begin{aligned}\mathbb{P}(\sum_{i=1}^{l} X_i = l-1) &\geq \frac{1}{3(1-\eta)l^{3/2}}\mathbb{E}\Big[\mathrm{e}^{-l(\bar{Y}_l-1)^2}\mathbb{1}_{\{\bar{Y}_l\in[1-\eta,1+\eta]\}}\Big]\\ &= \frac{1}{3(1-\eta)l^{3/2}}\Big(\mathbb{E}\big[\mathrm{e}^{-l(\bar{Y}_l-1)^2}\big] - \mathrm{e}^{-\eta^2 l}\Big).\end{aligned} \tag{6.13}$$

By Jensen's inequality,

$$\mathbb{P}\Big(\sum_{i=1}^{l} X_i = l-1\Big) \geq \frac{1}{3(1-\eta)l^{3/2}}\Big(\mathrm{e}^{-l\mathbb{E}[(\bar{Y}_l-1)^2]} - \mathrm{e}^{-\eta^2 l}\Big), \tag{6.14}$$

and we compute

$$l\mathbb{E}[(\bar{Y}_l - 1)^2] = \frac{1}{l}\mathrm{Var}\Big(\sum_{i=1}^{l} Y_l\Big) + l\big(\mathbb{E}[Y_1]-1\big)^2 = \sigma_n^2 + l(\tilde{\nu}_n - \nu)^2, \tag{6.15}$$

where

$$\sigma_n^2 = \mathrm{Var}(Y_1) = (1+\varepsilon_n)^2\Big(\mathbb{E}[(W_n^*)^2] - \nu_n^2\Big) \to \sigma^2 = \mathbb{E}[(W^*)^2] - \nu^2 = \frac{\mathbb{E}[W^3]}{\mathbb{E}[W]} - \nu^2 < \infty, \tag{6.16}$$

since $\tau > 4$. By Corollary 4.2, and for $l \leq 2k \leq 2\varepsilon n^{2/3}$,

$$l(\tilde{\nu}_n - \nu)^2 \leq c\varepsilon n^{2/3}|\varepsilon_n|^2 + 2\varepsilon n^{2/3} n^{-2(\tau-3)/(\tau-1)} \leq 2c\varepsilon, \tag{6.17}$$

so that

$$\mathbb{P}\Big(\sum_{i=1}^{l} X_i = l-1\Big) \geq \frac{1}{3(1-\eta)l^{3/2}}\big(\mathrm{e}^{-\sigma^2 - 2c\varepsilon} + o(1)\big). \tag{6.18}$$

Thus,

$$\mathbb{P}(T \geq k) \geq \sum_{l=k}^{2k} \frac{1}{3(1-\eta)l^{3/2}}\big(\mathrm{e}^{-\sigma^2 - 2c\varepsilon} + o(1)\big) \geq \frac{\tilde{a}_2}{k^{1/2}}, \tag{6.19}$$

which, together with (6.6) and (6.7), completes the proof for $\tau > 4$.

We continue with the proof for $\tau \in (3,4)$. We shall follow a large part of the analysis for the upper bound in Proposition 5.2, and, thus, the proof will be somewhat more sketchy.

Recall (5.7), to get

$$\mathbb{P}\Big(\sum_{i=1}^{l} X_i = l-1\Big) = \int_{[-\pi,\pi]} \mathrm{e}^{it}\big(\mathrm{e}^{-it}\phi_n(t)\big)^l \frac{dt}{2\pi}. \tag{6.20}$$

It is not hard to see that, by the arguments in the proof of Proposition 5.2 (in particular, recall (5.30)),

$$\int_{[-\pi,\pi]} (\mathrm{e}^{it} - 1)\big(\mathrm{e}^{-it}\phi_n(t)\big)^l \frac{dt}{2\pi} \leq c l^{-2/(\tau-2)}, \tag{6.21}$$

and, with $K_l = K l^{-1/(\tau-2)}$,

$$\int_{[-\pi,\pi]\setminus[-K_l,K_l]} \big(\mathrm{e}^{-it}\phi_n(t)\big)^l \frac{dt}{2\pi} \geq -\mathrm{e}^{-cK^{\tau-2}} l^{-1/(\tau-2)}, \tag{6.22}$$

so we are left to give a lower bound to of the form $cl^{-1/(\tau-2)}$ for

$$\int_{[-K_l,K_l]} \big(\mathrm{e}^{-it}\phi_n(t)\big)^l \frac{dt}{2\pi}. \tag{6.23}$$

Note that, by the uniform convergence of $\mathrm{e}^{-it}\phi_n(t)$ to $\mathrm{e}^{-it}\phi(t)$, we have that $\mathrm{e}^{-it}\phi_n(t) \to 1$ uniformly for $t \in [-K_l, K_l]$.

We shall first show that

$$\mathrm{e}^{-it}\phi_n(t) = \varphi_n(t) + O(t^2), \tag{6.24}$$

where $\varphi_n(t)$ is defined in (4.13). For this, we use that

$$\mathrm{Re}(\mathrm{e}^{-it}\phi_n(t)) = \mathbb{E}\big[\cos(W_n^*\sin(t) - t)\mathrm{e}^{-W_n^*[1-\cos(t)]}\big], \tag{6.25}$$

$$\mathrm{Im}(\mathrm{e}^{-it}\phi_n(t)) = \mathbb{E}\big[\sin(W_n^*\sin(t) - t)\mathrm{e}^{-W_n^*[1-\cos(t)]}\big], \tag{6.26}$$

Using that

$$|\mathrm{e}^{-W_n^*[1-\cos(t)]} - 1| \leq cW_n^*[1-\cos(t)] \leq cW_n^* t^2/2, \tag{6.27}$$

we see that we can replace $\mathrm{e}^{-W_n^*[1-\cos(t)]}$ by 1 in (6.25), at the expense of an error term of size $O(t^2)$. Further, using that

$$|\cos(x+y) - \cos(x)| \leq |x|, \qquad |\sin(x+y) - \sin(x)| \leq |x|, \qquad |\sin(t) - t| \leq |t|^3/3, \tag{6.28}$$

we see that

$$|\cos(W_n^*\sin(t) - t) - \cos((W_n^* - 1)t)| \leq W_n^*|\sin(t) - t| \leq W_n^*|t|^3/3, \tag{6.29}$$

and similarly $|\sin(w_j\sin(t) - t) - \sin((w_j - 1)t| \leq w_j|t|^3/6$. This proves (6.24), and implies that

$$\begin{aligned}\int_{[-K_l,K_l]} \mathrm{Re}\big(\mathrm{e}^{-it}\phi_n(t)\big)^l \frac{dt}{2\pi} &= \int_{[-K_l,K_l]} \varphi_n(t)^l \frac{dt}{2\pi} + O\big(l\int_{[-K_l,K_l]} |t|^2 dt\big)\\ &= \int_{[-K_l,K_l]} \mathrm{Re}\big(\varphi_n(t)^l\big)\frac{dt}{2\pi} + O(lK_l^3)\\ &= \int_{[-K_l,K_l]} \mathrm{Re}\big(\varphi_n(t)^l\big)\frac{dt}{2\pi} + O(Kl^{1-3/(\tau-2)}),\end{aligned} \tag{6.30}$$

where $l^{1-3/(\tau-2)} = l^{-(5-\tau)/(tau-2)} = o(l^{-1/(\tau-2)})$ for $\tau \in (3,4)$. Thus, we shall be done if there exists a constant $C > 0$, which is independent from K, such that

$$\int_{[-K_l,K_l]} \mathrm{Re}\big(\varphi_n(t)^l\big)\frac{dt}{2\pi} = l^{-1/(\tau-2)}\int_{[-K,K]} \mathrm{Re}\big(\varphi_n(tl^{1/(\tau-2)})^l\big)\frac{dt}{2\pi} \geq Cl^{-1/(\tau-2)}. \tag{6.31}$$

To prove (6.31), we rewrite

$$\varphi_n(t)^l = \varphi(t)^l\Big(1 + \frac{\varphi_n(t) - \varphi(t)}{\varphi(t)}\Big)^l, \tag{6.32}$$

so that

$$\mathrm{Re}\big(\varphi_n(t)^l\big) = \mathrm{Re}\Big(\varphi(t)^l\Big)\mathrm{Re}\Big(1 + \frac{\varphi_n(t) - \varphi(t)}{\varphi(t)}\Big)^l - \mathrm{Im}\Big(\varphi(t)^l\Big)\mathrm{Im}\Big(1 + \frac{\varphi_n(t) - \varphi(t)}{\varphi(t)}\Big)^l. \tag{6.33}$$

When $t \to 0$,

$$\varphi(t) = \mathrm{e}^{-c|t|^{\tau-2}(1+o(1))}, \tag{6.34}$$

since W^* is in the domain of attraction of a stable distribution when (1.16) holds. Thus, pointwise in t as $l \to \infty$,

$$\mathrm{Re}\Big(\varphi(tl^{1/(\tau-2)})^l\Big) = \mathrm{Re}\Big(\mathrm{e}^{-c|t|^{\tau-2}(1+o(1))}\Big) = \mathrm{e}^{-c|t|^{\tau-2}}(1+o(1)), \tag{6.35}$$

and, pointwise in t as $l \to \infty$,

$$\mathrm{Im}\Big(\varphi(tl^{1/(\tau-2)})^l\Big) = o(1). \tag{6.36}$$

Therefore, also using that

$$\Big|\mathrm{Im}\Big(\varphi(tl^{1/(\tau-2)})^l\Big)\mathrm{Im}\Big(1 + \frac{\varphi_n(tl^{1/(\tau-2)}) - \varphi(tl^{1/(\tau-2)})}{\varphi(tl^{1/(\tau-2)})}\Big)^l\Big| \leq \big|\varphi_n(tl^{1/(\tau-2)})^l\big| \leq \mathrm{e}^{-c|t|^{\tau-2}}, \tag{6.37}$$

which is integrable, the dominated convergence theorem gives that, for every $K > 0$,

$$\int_{[-K,K]} \mathrm{Im}\Big(\varphi(t)^l\Big)\mathrm{Im}\Big(1 + \frac{\varphi_n(t) - \varphi(t)}{\varphi(t)}\Big)^l \frac{dt}{2\pi} = o(1). \tag{6.38}$$

Moreover, by Corollary 4.2 and the fact that $|\varphi(t)| \geq 1/2$ for all $|t| \leq K_l$,

$$\begin{aligned}\mathrm{Re}\Big(1 + \frac{\varphi_n(t) - \varphi(t)}{\varphi(t)}\Big)^l &\geq (1 - 2|\varphi_n(t) - \varphi(t)|)^l\\ &\geq (1 - 2lcn^{-\frac{\tau-2}{\tau-1}} + 2cl|t|\big((|\varepsilon_n| + n^{-\frac{\tau-3}{\tau-1}}\big)\\ &\geq 1 - 2c(ln^{-\frac{\tau-2}{\tau-1}}) - 2c|t|l^{1/(\tau-2)}(ln^{-\frac{\tau-2}{\tau-1}})^{(\tau-3)/(\tau-2)}.\end{aligned} \tag{6.39}$$

Now we use that $l \leq 2k \leq 2\varepsilon n^{(\tau-2)/(\tau-1)}$ and $|\varepsilon_n| \leq \Lambda n^{-\frac{\tau-3}{\tau-1}}$, so that

$$\mathrm{Re}\Big(1 + \frac{\varphi_n(t) - \varphi(t)}{\varphi(t)}\Big)^l \geq 1 - 4c\varepsilon - 4c\varepsilon^{(\tau-3)/(\tau-2)}(1+\Lambda)|t|l^{1/(\tau-2)}. \tag{6.40}$$

In particular, this implies that, when $4c\varepsilon(1+\Lambda) + 4c\varepsilon^{(\tau-3)/(\tau-2)}K \leq 1$, and for all $|t| \leq K_l$,

$$\mathrm{Re}\Big(\varphi(t)^l\Big)\mathrm{Re}\Big(1 + \frac{\varphi_n(t) - \varphi(t)}{\varphi(t)}\Big)^l \geq 0. \tag{6.41}$$

Therefore, we arrive at the claim that there exists $C = C(\varepsilon, \Lambda) > 0$ such that, as $l \to \infty$,

$$\int_{[-K,K]} \mathrm{Re}\big(\varphi_n(tl^{1/(\tau-2)})^l\big)\frac{dt}{2\pi} \geq \int_{[-K,K]} \mathrm{e}^{-c|t|^{\tau-2}}\Big(1 - 4c\varepsilon - 4c\varepsilon^{(\tau-3)/(\tau-2)}|t|l^{1/(\tau-2)}\Big)dt + o(1) \geq C, \tag{6.42}$$

when we choose $\varepsilon > 0$ sufficiently small. This completes the proof of (6.31). □

Proof of Lemma 6.3. We write

$$\mathbb{P}(T^{(2)} \geq 2k, |C(V)| < k) = \sum_{t=1}^{k-1} \mathbb{P}(T^{(2)} \geq 2k, |C(V)| = t). \tag{6.43}$$

When $|C(V)| = t$, but $T^{(2)} \geq 2k > t$, then we must have that the total progeny of the thinned vertices is at least k. Denote the set of vertices which are thinned by time t by N_t, so that, by the description in Section 4.2, the set consists of those vertices which are drawn at least twice before time t. For $j \in [n]$, let $M_j(t)$ denote the number of times the vertex j is drawn in the first t draws. Thus, $j \in N_t$ precisely when $M_j(t) \geq 2$. Then we write

$$\mathbb{P}\big(T^{(2)} \geq 2k, |C(V)| = t\big) \leq \mathbb{P}\Big(|C(V)| = t, \sum_{j \in N_t} \tilde{T}_{j,t} \geq k\Big), \tag{6.44}$$

where

$$\tilde{T}_{j,t} = \sum_{s=1}^{M_j(t)-1} T_{j,s}, \tag{6.45}$$

and $\{T_{j,s}\}_{s=1}^{\infty}$ is an i.i.d. sequence of random variables where $T_{j,1}$ is the total progeny of a branching process which has a Poisson distribution with parameter w_j in the first generation, and offspring distribution $g^{(n)}$ in all later generations. Now, clearly N_t and $M_j(t)$ are non-decreasing in time, so that

$$\mathbb{P}\big(T^{(2)} \geq 2k, |C(V)| = t\big) \leq \mathbb{P}\Big(|C(V)| = t, \sum_{j \in N_k} \tilde{T}_{j,k} \geq k\Big), \tag{6.46}$$

and we arrive at

$$\mathbb{P}\big(T^{(2)} \geq 2k, |C(V)| < k\big) \leq \mathbb{P}\Big(T^{(2)} \geq 2k, \sum_{j \in N_k} \tilde{T}_{j,k} \geq k\Big). \tag{6.47}$$

Now, we split the event $\sum_{j \in N_k} \tilde{T}_{j,k} \geq k$ into the event where all $\tilde{T}_{j,k} \leq k$ and the event where there exists a $j \in N_k$ such that $\tilde{T}_{j,k} \geq k$, to arrive at

$$\begin{aligned}\mathbb{P}\big(T^{(2)} \geq 2k, |C(V)| < k\big) &\leq \sum_{j=1}^{n} \mathbb{P}(T^{(2)} \geq 2k, j \in N_k, \tilde{T}_{j,k} \geq k) + \mathbb{P}(T^{(2)} \geq 2k, \sum_{j \in N_k} \tilde{T}_{j,k} \mathbb{1}_{\{\tilde{T}_{j,k} \leq k\}} \geq k) \\ &\leq \sum_{j=1}^{n} \mathbb{P}(T^{(2)} \geq 2k, j \in N_k, \tilde{T}_{j,k} \geq k) + \frac{1}{k}\mathbb{E}\Big[\mathbb{1}_{\{T^{(2)} \geq 2k\}} \sum_{j \in N_k} \tilde{T}_{j,k} \mathbb{1}_{\{\tilde{T}_{j,k} \leq k\}}\Big],\end{aligned}$$

the last bound by the Markov inequality. Thus, we are lead to

$$\begin{aligned}&\mathbb{P}\big(T^{(2)} \geq 2k, |C(V)| < k\big) \\ &\quad \leq \sum_{j=1}^{n} \Big[\mathbb{P}\big(T^{(2)} \geq 2k, j \in N_k, \tilde{T}_{j,k} \geq k\big) + \frac{1}{k}\sum_{l=1}^{k} \mathbb{P}\big(T^{(2)} \geq 2k, j \in N_k, \tilde{T}_{j,k} \geq l\big)\Big].\end{aligned} \tag{6.48}$$

Now, we make use of the following lemma:

Lemma 6.4 (Upper bound on tail probabilities of random sums)**.** *Suppose that $\{T_l\}_{l=1}^{\infty}$ are i.i.d. random variables, with*

$$\mathbb{P}(T_l \geq k) \leq K k^{-1/\delta}, \tag{6.49}$$

for some constants c_l and some $\delta > 1$. Then, when M is independent from $\{T_l\}_{l=1}^{\infty}$, there exists a c_δ such that

$$\mathbb{P}(\sum_{l=1}^{M} T_l \geq k) \leq c_\delta K \mathbb{E}[M] k^{-1/\delta}. \tag{6.50}$$

Proof. We perform similar bounds as in (6.48) to see that

$$\begin{aligned}\mathbb{P}(\sum_{l=1}^{M} T_l \geq k) &\leq \mathbb{P}(\exists j \in [M] : T_j \geq k) + \mathbb{P}(\sum_{l=1}^{M} T_l \mathbb{1}_{\{T_l \leq k\}} \geq k)\\ &\leq \mathbb{E}[M]\mathbb{P}(T_1 \geq k) + \frac{1}{k}\mathbb{E}\big[\sum_{l=1}^{M} T_l \mathbb{1}_{\{T_l \leq k\}}\big]\\ &\leq K\mathbb{E}[M]k^{-1/\delta} + \frac{1}{k}\mathbb{E}[M]\sum_{l=1}^{k}\mathbb{P}(T_1 \geq l)\\ &\leq K\mathbb{E}[M]\Big(k^{-1/\delta} + \sum_{l=1}^{k} l^{-1/\delta}\Big) \leq c_\delta K\mathbb{E}[M]k^{-1/\delta}. \end{aligned} \tag{6.51}$$

□

Note that $T_{j,1} = 1 + \sum_{l=1}^{P_j} T_j$, where P_j has a Poisson distribution with parameter w_j. Thus, by Lemma 6.4, we have

$$\mathbb{P}(T_{j,1} \geq l) = \mathbb{P}\Big(\sum_{l=1}^{P_j} T_j \geq l-1\Big) \leq c_\delta a_1 \mathbb{E}[P_j] l^{-1/\delta} = \frac{c_\delta a_1 w_j}{l^{1/\delta}}. \tag{6.52}$$

We next again apply Lemma 6.4 to $\tilde{T}_{j,k}$ in (6.45), where M is $M_j(k)-1$ conditioned on $M_j(k) \geq 2$. It is not hard to see that, when $k \leq \varepsilon n^{(\tau-2)/(\tau-1)}$,

$$\mathbb{E}[M_j(k) - 1 \mid M_j(k) \geq 2] \leq c. \tag{6.53}$$

When j is not so large, so that w_j is quite large, the factor w_j appearing in (6.52) can be quite harmful. Thus, for j such that $w_j \geq k^{1/\delta}$, we simply bound the restrictions on $\tilde{T}_{j,k}$ away, so that

$$\begin{aligned}\mathbb{P}\big(T^{(2)} \geq 2k, |C(V)| < k\big) &\leq \sum_{j=1}^{n}\mathbb{P}\big(T^{(2)} \geq 2k, j \in N_k, \tilde{T}_{j,k} \geq k\big) + \mathbb{P}\big(T^{(2)} \geq 2k, \sum_{j\in N_k}\tilde{T}_{j,k}\mathbb{1}_{\{\tilde{T}_{j,k}\leq k\}} \geq k\big)\\ &\leq 2\sum_{j=1}^{n}\mathbb{P}\big(T^{(2)} \geq 2k, j \in N_k\big)\mathbb{1}_{\{w_j \geq k^{1/\delta}\}}\\ &\quad + \sum_{j=1}^{n}\mathbb{1}_{\{w_j < k^{1/\delta}\}}\mathbb{P}\big(T^{(2)} \geq 2k, j \in N_k, \tilde{T}_{j,k} \geq k\big)\\ &\quad + \sum_{j=1}^{n}\mathbb{1}_{\{w_j < k^{1/\delta}\}}\frac{1}{k}\sum_{l=1}^{k}\mathbb{P}\big(T^{(2)} \geq 2k, j \in N_k, \tilde{T}_{j,k} \geq l\big). \end{aligned} \tag{6.54}$$

On the second and third sum, we apply Lemma 6.4, to arrive at

$$\mathbb{P}\big(T^{(2)} \geq 2k, |C(V)| < k\big) \leq cc_\delta a_1 \sum_{j=1}^{n}\big(\frac{w_j}{k^{1/\delta}} \wedge 1\big)\Big[\mathbb{P}\big(T^{(2)} \geq 2k, j \in N_k\big) + \frac{1}{k}\sum_{l=1}^{k}\mathbb{P}\big(T^{(2)} \geq 2k, j \in N_k\big)\Big], \tag{6.55}$$

where we note that, without loss of generality, we may assume that $cc_\delta a_1 \geq 1$. Performing the sum over l leads to

$$\mathbb{P}\big(T^{(2)} \geq 2k, |C(V)| < k\big) \leq cc_\delta a_1 \sum_{j=1}^{n} \mathbb{P}\big(T^{(2)} \geq 2k, j \in N_k\big)\big(\frac{w_j}{k^{1/\delta}} \wedge 1\big). \tag{6.56}$$

When $j \in N_k$, there must be $k_1 < k_2 \leq k$ such that $M_{k_1} = M_{k_2} = j$. We may assume that $M_l \neq j$ when $l \in [k_2] \setminus \{k_1, k_2\}$, and we note that when $T^{(2)} \geq 2k$, then we must have that $T^{(2)} \geq k_1$, so that

$$\begin{aligned}
\mathbb{P}\big(T^{(2)} \geq 2k, |C(V)| < k\big) &\leq cc_\delta a_1 \sum_{j=1}^{n} \sum_{1\leq k_1<k_2\leq k} \mathbb{P}(T^{(2)} \geq k_1, M_{k_1} = M_{k_2} = j)\big(\frac{w_j}{k^{1/\delta}} \wedge 1\big) \\
&\leq cc_\delta a_1 \sum_{j=1}^{n} \big(\frac{w_j}{l_n}\big)^2 a_2 k_1^{-1/\delta}\big(\frac{w_j}{k^{1/\delta}} \wedge 1\big) \\
&\leq cc_\delta a_1 \frac{k^{2-1/\delta}}{n} \mathbb{E}\Big[W_n^*\big(\frac{W_n^*}{k^{1/\delta}} \wedge 1\big)\Big].
\end{aligned} \tag{6.57}$$

When $\tau > 4$, and we use that $\delta = 2$ and $k \leq \varepsilon n^{2/3}$, then we can bound the above by

$$\mathbb{P}\big(T^{(2)} \geq 2k, |C(V)| < k\big) \leq cc_\delta a_1 \frac{k^{2-2/\delta}}{n} \mathbb{E}[(W^*)^2] = C_\delta \frac{k}{n} = C_\delta \frac{k^{3/2}}{n} k^{-1/2} \leq C_\delta \varepsilon^{3/2} k^{-1/2}, \tag{6.58}$$

so that (6.4) follows with $p = 3/2$.

When $\tau \in (3, 4)$, then we need to be more careful. For this, we use Lemma 4.3, now using that $k \leq \varepsilon n^{(\tau-2)/(\tau-1)}$, $\delta = \tau - 2$ and $\mathbb{E}[h(W_n^*)] = \mathbb{E}[W_n h(W_n)]/\mathbb{E}[W_n]$, to obtain

$$\begin{aligned}
\mathbb{P}\big(T^{(2)} \geq 2k, |C(V)| < k\big) &\leq cc_\delta a_1 \frac{k^{2-2/\delta}}{n} \mathbb{E}\Big[(W_n^*)^2 \mathbb{1}_{\{W_n^* \leq k^{1/\delta}\}}\Big] + cc_\delta a_1 \frac{k^{2-1/\delta}}{n} \mathbb{E}\Big[W_n^* \mathbb{1}_{\{W_n^* > k^{1/\delta}\}}\Big] \\
&\leq C_\delta \Big(\frac{k^{(\tau-1)/(\tau-2)}}{n}\Big) k^{-1/(\tau-2)} \leq C_\delta \varepsilon^{(\tau-1)/(\tau-2)} k^{-1/(\tau-2)},
\end{aligned} \tag{6.59}$$

so that (6.4) follows with $p = (\tau - 1)/(\tau - 2) > 1$. □

7 Proof of the main results

Upper bounds in Theorems 1.1–1.4. The upper bounds follow immediately from Propositions 3.1 and 5.1, when we recall the definition of δ and α in (5.1), so that (3.4) is the same as $|\varepsilon_n| \leq \Lambda n^{-1/3}$ for $\tau > 4$, as assumed in Theorem 1.3, and $|\varepsilon_n| \leq \Lambda n^{-(\tau-3)/(\tau-2)}$ for $\tau \in (3, 4)$, as assumed in Theorem 1.4.

Lower bounds in Theorems 1.1–1.4. For the lower bounds, we note that there is an obvious monotonicity in the weights, so that the cluster for $\tilde{w}_i$ as in (1.18) with $\tilde{\varepsilon}_n = -\Lambda n^{-\alpha/(\delta+1)}$ is stochastically smaller than the one for ε_n with $|\varepsilon_n| \leq \Lambda n^{-\alpha/(\delta+1)}$. Now we pick $\Lambda > 0$ so large that $\tilde{\nu}_n = (1 + \tilde{\varepsilon}_n)\nu_n \leq 1 - cn^{-\alpha/(\delta+1)}$, which is possible since $\nu_n - 1 \leq cn^{-\frac{\tau-3}{\tau-1}} \leq cn^{-\frac{\alpha}{\delta+1}}$ by Corollary 4.2 and (5.1). Then, we make use of Proposition 3.3, and check that its assumptions are satisfied due to Propositions 5.3 and 6.1, since $\alpha = \delta - 1$.

This completes the proofs of Theorems 1.1–1.4. □

Acknowledgements. This work was supported in part by the Netherlands Organisation for Scientific Research (NWO). The author gratefully acknowledges the hospitality of the Mittag-Leffler Institute (MLI), where part of this work was carried out, and the lively discussions at the MLI with Svante Janson, Tomasz Łuczak, Michal Karoński and Tatyana Turova where the heuristic connection between the subcritical and supercritical regimes discussed in Section 1.3 was derived, and which led to Theorems 1.3–1.4, as well as Conjectures 1.5–1.8.

References

[1] D. Aldous. Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab.*, **25**(2):812–854, (1997).

[2] J. van den Berg and H. Kesten. Inequalities with applications to percolation and reliability. *J. Appl. Prob.*, **22**:556–569, (1985).

[3] B. Bollobás. The evolution of random graphs. *Trans. Amer. Math. Soc.*, **286**(1):257–274, (1984).

[4] B. Bollobás. The evolution of sparse graphs. In *Graph theory and combinatorics (Cambridge, 1983)*, pages 35–57. Academic Press, London, (1984).

[5] B. Bollobás. *Random graphs*, volume **73** of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, (2001).

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

[7] C. Borgs, J. Chayes, R. van der Hofstad, G. Slade, and J. Spencer. Random subgraphs of finite graphs. I. The scaling window under the triangle condition. *Random Structures Algorithms*, **27**(2):137–184, (2005).

[8] C. Borgs, J. T. Chayes, H. Kesten, and J. Spencer. Uniform boundedness of critical crossing probabilities implies hyperscaling. *Random Structures Algorithms*, **15**(3-4):368–413, (1999). Statistical physics methods in discrete probability, combinatorics, and theoretical computer science (Princeton, NJ, 1997).

[9] T. Britton, M. Deijfen, and A. Martin-Löf. Generating simple random graphs with prescribed degree distribution. *J. Stat. Phys.*, **124**(6):1377–1397, (2006).

[10] F. Chung and L. Lu. The average distances in random graphs with given expected degrees. *Proc. Natl. Acad. Sci. USA*, **99**(25):15879–15882 (electronic), (2002).

[11] F. Chung and L. Lu. Connected components in random graphs with given expected degree sequences. *Ann. Comb.*, **6**(2):125–145, (2002).

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

[13] F. Chung and L. Lu. *Complex graphs and networks*, volume **107** of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC, (2006).

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

[15] R. Durrett. *Random graph dynamics.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, (2007).

[16] M. Dwass. The total progeny in a branching process and a related random walk. *J. Appl. Prob.*, **6**:682–686, (1969).

[17] P. Erdős and A. Rényi. On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, **5**:17–61, (1960).

[18] H. van den Esker, R. van der Hofstad, and G. Hooghiemstra. Universality for the distance in finite variance random graphs. *J. Stat. Phys.*, **133**(1):169–202, (2008).

[19] G. Grimmett. *Percolation.* Springer, Berlin, 2nd edition, (1999).

[20] M. Heydenreich and R. van der Hofstad. Random graph asymptotics on high-dimensional tori. *Comm. Math. Phys.*, **270**(2):335–358, (2007).

[21] R. van der Hofstad. Random graphs and complex networks. In preparation, see `http://www.win.tue.nl/∼rhofstad/NotesRGCN.pdf`, (2009).

[22] R. van der Hofstad, G. Hooghiemstra, and P. Van Mieghem. Distances in random graphs with finite variance degrees. *Random Structures Algorithms*, **27**(1):76–123, (2005).

[23] R. van der Hofstad, G. Hooghiemstra, and D. Znamenski. Distances in random graphs with finite mean and infinite variance degrees. *Electron. J. Probab.*, **12**(25):703–766 (electronic), (2007).

[24] R. van der Hofstad and M. Keane. An elementary proof of the hitting time theorem. *Amer. Math. Monthly*, **115**(8):753–756, 2008.

[25] R. van der Hofstad and M. Luczak. Random subgraphs of the 2D Hamming graph: The supercritical phase. Preprint (2006).

[26] S. Janson. The largest component in a subcritical random graph with a power law degree distribution. *Ann. Appl. Probab.*, **18**(4):1651–1668, (2008).

[27] S. Janson. The probability that a random multigraph is simple. Preprint (2006).

[28] S. Janson. Asymptotic equivalence and contiguity of some random graphs. Preprint (2008).

[29] S. Janson, D.E. Knuth, T. Łuczak, and B. Pittel. The birth of the giant component. *Random Structures Algorithms*, **4**(3):231–358, (1993). With an introduction by the editors.

[30] S. Janson and M. Luczak. A new approach to the giant component problem. Preprint (2007).

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

[32] M. Kang and T. G. Seierstad. The critical phase for random graphs with a given degree sequence. *Combin. Probab. Comput.*, **17**(1):67–86, (2008).

[33] T. Łuczak. Component behavior near the critical point of the random graph process. *Random Structures Algorithms*, **1**(3):287–310, (1990).

[34] T. Łuczak, B. Pittel, and J. Wierman. The structure of a random graph at the point of the phase transition. *Trans. Amer. Math. Soc.*, **341**(2):721–748, (1994).

[35] M. Molloy and B. Reed. A critical point for random graphs with a given degree sequence. In *Proceedings of the Sixth International Seminar on Random Graphs and Probabilistic Methods in Combinatorics and Computer Science, "Random Graphs '93" (Poznań, 1993)*, volume **6**, pages 161–179, (1995).

[36] I. Norros and H. Reittu. On a conditionally Poissonian graph process. *Adv. in Appl. Probab.*, **38**(1):59–75, (2006).

[37] R. Otter. The multiplicative process. *Ann. Math. Statist.*, **20**:206–224, (1949).

[38] B. Pittel. On the largest component of the random graph at a nearcritical stage. *J. Combin. Theory Ser. B*, **82**(2):237–269, (2001).

[39] H. Reittu and I. Norros. On the power law random graph model of massive data networks. *Performance Evaluation*, **55**(1-2):3–23, (2004).